Self Case Study - 1

Customer Relationship Prediction

In [6]:

```
import pickle
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


from sklearn.linear_model import SGDClassifier
from sklearn.model_selection import GridSearchCV
from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier

from sklearn.metrics import roc_auc_score,roc_curve
from prettytable import PrettyTable

from mlxtend.classifier import StackingCVClassifier
from sklearn.svm import SVC
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/externals/six.py:31: FutureWarning: The module is deprec
ated in version 0.21 and will be removed in version 0.23 since we've dropped support for Python 2.7. Pl
ease rely on the official version of six (https://pypi.org/project/six/).
  "(https://pypi.org/project/six/).", FutureWarning)
```

Loading Data

Appetency

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Appetency/standardized freq encoding/X_tes
t_appetency_poly.pickle', 'rb') as handle:
    X_test_appetency = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Appetency/standardized freq encoding/X_tes
t_appetency_vanilla.pickle', 'rb') as handle:
    X_test_appetency_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Appetency/standardized freq encoding/X_tra
in_appetency_poly.pickle', 'rb') as handle:
    X_train_appetency = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Appetency/standardized freq encoding/X_tra
in_appetency_vanilla.pickle', 'rb') as handle:
    X_train_appetency_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/Appetency/y_test_appetency.pickle', 'r
b') as handle:
```

```
    y_test_appetency = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/Appetency/y_train_appetency.pickle', '
rb') as handle:
    y_train_appetency = pickle.load(handle)
```

In [ ]:

```
y_test_appetency = y_test_appetency.Appetency.values
```

In [ ]:

```
y_train_appetency = y_train_appetency.Appetency.values
```

Churn

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Churn/X_test_churn_poly.pickle', 'rb') as
handle:
    X_test_churn = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Churn/X_test_churn_vanilla.pickle', 'rb')
as handle:
    X_test_churn_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Churn/X_train_churn_poly.pickle', 'rb') as
handle:
    X_train_churn = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/Churn/X_train_churn_vanilla.pickle', 'rb')
as handle:
    X_train_churn_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/Churn/y_test_churn.pickle', 'rb') as h
andle:
    y_test_churn = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/Churn/y_train_churn.pickle', 'rb') as
handle:
    y_train_churn = pickle.load(handle)
```

In [ ]:

```
y_test_churn = y_test_churn.Churn.values
```

In [ ]:

```
y_train_churn = y_train_churn.Churn.values
```

Upselling

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/upselling/X_test_upselling_poly.pickle', 'rb') as handle:
    X_test_upselling = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/upselling/X_test_upselling_vanilla.pickle', 'rb') as handle:
    X_test_upselling_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/upselling/X_train_upselling_poly.pickle', 'rb') as handle:
    X_train_upselling = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Modeling/upselling/X_train_upselling_vanilla.pickle', 'rb') as handle:
    X_train_upselling_vanilla = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/upselling/y_test_upselling.pickle', 'rb') as handle:
    y_test_upselling = pickle.load(handle)
```

In [ ]:

```
with open('/content/drive/MyDrive/Case Study 1/Data/Feature engg/upselling/y_train_upselling.pickle', 'rb') as handle:
    y_train_upselling = pickle.load(handle)
```

In [ ]:

```
y_test_upselling = y_test_upselling.Upselling.values
```

In [ ]:

```
y_train_upselling = y_train_upselling.Upselling.values
```

Models

Vanilla (without feature engg.)

Appetency

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_appetency_vanilla, y_train_appetency)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  15 out of  15 | elapsed:   11.9s finished

-- Epoch 1
Norm: 0.80, NNZs: 72, Bias: -0.444277, T: 40000, Avg. loss: 0.654479
Total training time: 0.02 seconds.
-- Epoch 2
Norm: 0.72, NNZs: 72, Bias: -0.455214, T: 80000, Avg. loss: 0.573790
Total training time: 0.05 seconds.
-- Epoch 3
Norm: 0.73, NNZs: 72, Bias: -0.447964, T: 120000, Avg. loss: 0.564914
Total training time: 0.07 seconds.
-- Epoch 4
Norm: 0.71, NNZs: 72, Bias: -0.455829, T: 160000, Avg. loss: 0.572673
Total training time: 0.10 seconds.
-- Epoch 5
Norm: 0.71, NNZs: 72, Bias: -0.456041, T: 200000, Avg. loss: 0.566742
Total training time: 0.13 seconds.
-- Epoch 6
Norm: 0.71, NNZs: 72, Bias: -0.458073, T: 240000, Avg. loss: 0.567396
Total training time: 0.15 seconds.
-- Epoch 7
Norm: 0.70, NNZs: 72, Bias: -0.460697, T: 280000, Avg. loss: 0.569265
Total training time: 0.18 seconds.
-- Epoch 8
Norm: 0.70, NNZs: 72, Bias: -0.459879, T: 320000, Avg. loss: 0.565467
Total training time: 0.20 seconds.
Convergence after 8 epochs took 0.20 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=None,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test_
0	0.743698	0.127860	0.015140	0.003199	0.001	{'alpha': 0.001}	0.739125	0.783613
1	0.398957	0.030549	0.010637	0.000093	0.01	{'alpha': 0.01}	0.766228	0.810251
2	0.348170	0.040166	0.010320	0.000104	0.1	{'alpha': 0.1}	0.762538	0.808816
3	2.963274	3.864577	0.009592	0.002109	1	{'alpha': 1}	0.693727	0.789863
4	0.495732	0.484616	0.011110	0.001149	10	{'alpha': 10}	0.693279	0.759432

In [ ]:

```
clf = classifier.best_estimator_
# clf = SGDClassifier(loss = 'log',alpha = 0.001, class_weight= 'balanced', n_jobs = -1)
```

In [ ]:

```
clf.fit(X_train_appetency_vanilla,y_train_appetency)
```

```
-- Epoch 1
Norm: 0.80, NNZs: 72, Bias: -0.444277, T: 40000, Avg. loss: 0.654479
Total training time: 0.02 seconds.
-- Epoch 2
Norm: 0.72, NNZs: 72, Bias: -0.455214, T: 80000, Avg. loss: 0.573790
Total training time: 0.05 seconds.
-- Epoch 3
Norm: 0.73, NNZs: 72, Bias: -0.447964, T: 120000, Avg. loss: 0.564914
Total training time: 0.07 seconds.
-- Epoch 4
Norm: 0.71, NNZs: 72, Bias: -0.455829, T: 160000, Avg. loss: 0.572673
Total training time: 0.10 seconds.
-- Epoch 5
Norm: 0.71, NNZs: 72, Bias: -0.456041, T: 200000, Avg. loss: 0.566742
Total training time: 0.12 seconds.
-- Epoch 6
Norm: 0.71, NNZs: 72, Bias: -0.458073, T: 240000, Avg. loss: 0.567396
Total training time: 0.15 seconds.
-- Epoch 7
Norm: 0.70, NNZs: 72, Bias: -0.460697, T: 280000, Avg. loss: 0.569265
Total training time: 0.17 seconds.
-- Epoch 8
Norm: 0.70, NNZs: 72, Bias: -0.459879, T: 320000, Avg. loss: 0.565467
Total training time: 0.20 seconds.
Convergence after 8 epochs took 0.20 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.1, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=None, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency_vanilla)[:,1]
```

```
y_test_appetency_pred = clf.predict_proba(X_test_appetency_vanilla)[:,1]
```

In [ ]:

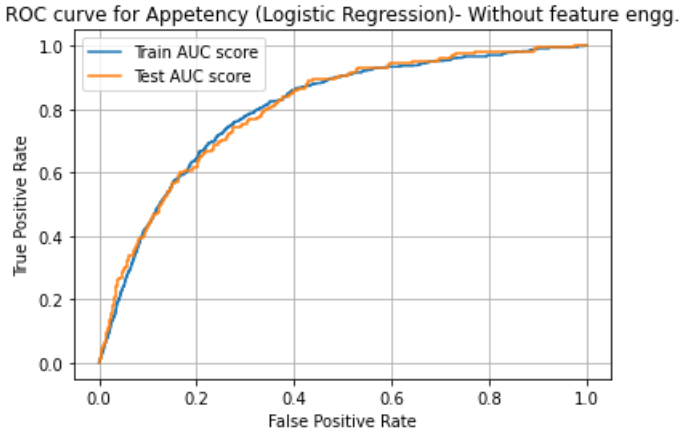
```
lr_train_auc_score_appetency_vanilla = roc_auc_score(y_train_appetency, y_train_appetency_pred)
lr_test_auc_score_appetency_vanilla = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Logistic Regression)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
lr_appetency_vanilla_score = ['Logistic Regression (Appetency Vanilla)', lr_train_auc_score_appetency_v
anilla, lr_test_auc_score_appetency_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+----------------------------------------+--------------------+--------------------+
|                 Model                  |     Train AUC      |      Test AUC      |
+----------------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency Vanilla) | 0.8004914800871239 | 0.8022376961601908 |
+----------------------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced',n_jobs = -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_appetency_vanilla,y_train_appetency)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:  2.8min
[Parallel(n_jobs=-1)]: Done  90 out of  90 | elapsed:  8.6min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	0.577689	0.017091	0.120537	0.005795	3	10	{'max_dep 3, 'n_estima 10}
1	0.747663	0.010782	0.117489	0.003201	3	20	{'max_dep 3, 'n_estima

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	20} pa
2	1.959984	0.150075	0.164925	0.049753	3	50	{'max_dep 3, 'n_estima 50}
3	3.754670	0.146095	0.305361	0.047536	3	100	{'max_dep 3, 'n_estima 100}
4	7.822746	0.019682	0.472720	0.042278	3	200	{'max_dep 3, 'n_estima 200}
5	19.139384	1.137449	0.967741	0.050304	3	500	{'max_dep 3, 'n_estima 500}
6	0.760469	0.054644	0.131657	0.005672	5	10	{'max_dep 5, 'n_estima 10}
7	1.353080	0.027494	0.134040	0.002144	5	20	{'max_dep 5, 'n_estima 20}
8	2.925882	0.182161	0.162985	0.050864	5	50	{'max_dep 5, 'n_estima 50}
9	5.725605	0.061788	0.307917	0.052442	5	100	{'max_dep 5, 'n_estima 100}
10	11.653933	0.162458	0.470370	0.053613	5	200	{'max_dep 5, 'n_estima 200}
11	28.857828	1.224374	1.005098	0.047616	5	500	{'max_dep 5, 'n_estima 500}
12	0.933711	0.041565	0.134806	0.005199	7	10	{'max_dep 7, 'n_estima 10}
13	1.665884	0.047921	0.128274	0.003703	7	20	{'max_dep 7, 'n_estima 20}
14	3.995880	0.048363	0.229677	0.008247	7	50	{'max_dep 7, 'n_estima 50}
15	7.495989	0.136508	0.339149	0.001469	7	100	{'max_dep 7, 'n_estima 100}
							{'max_dep 7,

16	14.645817 mean_fit_time	0.049791 std_fit_time	0.575541 mean_score_time	0.046617 std_score_time	7 param_max_depth	200 param_n_estimators	'n_estima 200}
17	35.707886	1.078612	1.302021	0.044760	7	500	{'max_dep 7, 'n_estima 500}
18	1.104814	0.006780	0.142469	0.003577	10	10	{'max_dep 10, 'n_estima 10}
19	1.935635	0.061368	0.131008	0.007008	10	20	{'max_dep 10, 'n_estima 20}
20	4.461021	0.104792	0.235305	0.005927	10	50	{'max_dep 10, 'n_estima 50}
21	8.498910	0.297254	0.333159	0.005986	10	100	{'max_dep 10, 'n_estima 100}
22	17.446327	0.362201	0.530856	0.004049	10	200	{'max_dep 10, 'n_estima 200}
23	42.532655	1.607236	1.301788	0.045002	10	500	{'max_dep 10, 'n_estima 500}
24	1.128788	0.054695	0.135320	0.002182	15	10	{'max_dep 15, 'n_estima 10}
25	2.081094	0.178547	0.134203	0.001945	15	20	{'max_dep 15, 'n_estima 20}
26	5.014886	0.102696	0.230513	0.006635	15	50	{'max_dep 15, 'n_estima 50}
27	9.305908	0.111638	0.374444	0.048345	15	100	{'max_dep 15, 'n_estima 100}
28	18.886215	0.646416	0.637262	0.003429	15	200	{'max_dep 15, 'n_estima 200}
29	41.784392	8.394775	1.197567	0.355466	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator
```

In [ ]:

```
clf.fit(X_train_appetency_vanilla,y_train_appetency)
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=5, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=500,
                       n_jobs=-1, oob_score=False, random_state=None, verbose=0,
                       warm_start=False)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency_vanilla)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency_vanilla)[:,1]
```

In [ ]:

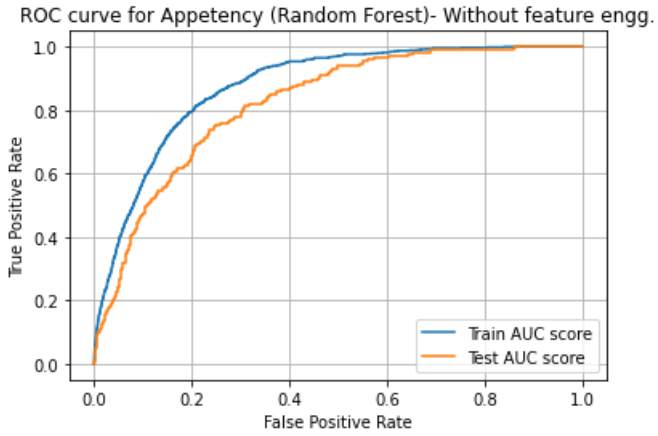
```
rf_train_auc_score_appetency_vanilla = roc_auc_score(y_train_appetency, y_train_appetency_pred)
rf_test_auc_score_appetency_vanilla = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Random Forest)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_appetency_vanilla_score = ['Random Forest (Appetency Vanilla)', rf_train_auc_score_appetency_vanilla
, rf_test_auc_score_appetency_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_appetency_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+----------------------------------+--------------------+--------------------+
|               Model              |     Train AUC      |      Test AUC      |
+----------------------------------+--------------------+--------------------+
| Random Forest (Appetency Vanilla) | 0.8760928373360423 | 0.8213869803856969 |
+----------------------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_appetency, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,250,500], 'max_depth' : [1,2,3,4]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_appetency_vanilla,y_train_appetency)
```

```
Fitting 3 folds for each of 24 candidates, totalling 72 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:  2.9min
[Parallel(n_jobs=-1)]: Done  72 out of  72 | elapsed:  7.5min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=55.17977528089887,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=-1,
             param_grid={'max_depth': [1, 2, 3, 4],
                         'n_estimators': [10, 20, 50, 100, 250, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	0.876116	0.135697	0.031803	0.007528	1	10	{'max_dep 1, 'n_estima 10}
1	1.075858	0.020172	0.027183	0.000263	1	20	{'max_dep 1, 'n_estima 20}
2	2.228308	0.021386	0.034595	0.001537	1	50	{'max_dep 1, 'n_estima 50}
3	4.192808	0.021484	0.045907	0.002745	1	100	{'max_dep 1, 'n_estima 100}
4	9.911119	0.067871	0.082553	0.005665	1	250	{'max_dep 1, 'n_estima 250}
5	19.810298	0.183523	0.140996	0.003850	1	500	{'max_dep 1, 'n_estima 500}
6	0.892716	0.003522	0.034711	0.011239	2	10	{'max_dep 2, 'n_estima 10}
7	1.521835	0.002415	0.030937	0.000264	2	20	{'max_dep 2, 'n_estima 20}
8	3.429620	0.034347	0.041643	0.000153	2	50	{'max_dep 2, 'n_estima 50}
9	6.513378	0.057886	0.065422	0.000718	2	100	{'max_dep 2, 'n_estima 100}
10	15.865312	0.024442	0.135420	0.000622	2	250	{'max_dep 2, 'n_estima 250}
11	31.407983	0.105696	0.232030	0.006502	2	500	{'max_dep 2, 'n_estima 500}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
12	1.162541	0.010920	0.027455	0.000164	3	10	{'max_dep 3, 'n_estima 10}
13	2.063547	0.023115	0.033161	0.000629	3	20	{'max_dep 3, 'n_estima 20}
14	4.745597	0.027727	0.051811	0.000447	3	50	{'max_dep 3, 'n_estima 50}
15	9.093669	0.046172	0.085147	0.004901	3	100	{'max_dep 3, 'n_estima 100}
16	22.164728	0.046537	0.169076	0.001842	3	250	{'max_dep 3, 'n_estima 250}
17	43.696494	0.019642	0.320809	0.003504	3	500	{'max_dep 3, 'n_estima 500}
18	1.455560	0.005047	0.029286	0.000211	4	10	{'max_dep 4, 'n_estima 10}
19	2.650553	0.018028	0.037057	0.000253	4	20	{'max_dep 4, 'n_estima 20}
20	6.175600	0.058528	0.059964	0.000263	4	50	{'max_dep 4, 'n_estima 50}
21	11.830695	0.072867	0.101354	0.001566	4	100	{'max_dep 4, 'n_estima 100}
22	28.734066	0.053072	0.221566	0.009719	4	250	{'max_dep 4, 'n_estima 250}
23	52.631518	6.171264	0.385890	0.039205	4	500	{'max_dep 4, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_appetency_vanilla,y_train_appetency)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=3,
              min_child_weight=1, missing=None, n_estimators=20, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=55.17977528089887,
              seed=None, silent=None, subsample=1, verbosity=1)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency_vanilla)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency_vanilla)[:,1]
```

In [ ]:

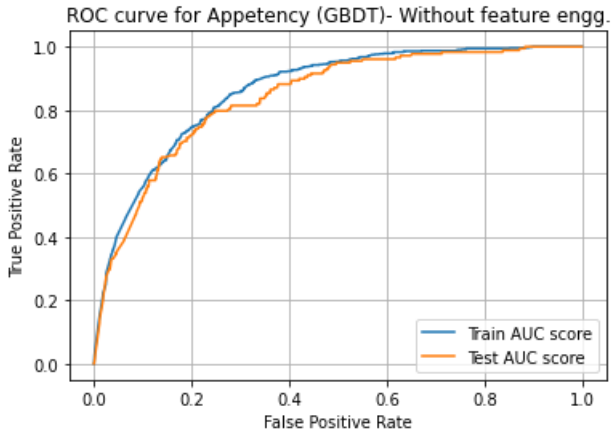
```
gbdt_train_auc_score_appetency_vanilla = roc_auc_score(y_train_appetency, y_train_appetency_pred)
gbdt_test_auc_score_appetency_vanilla = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (GBDT)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
gbdt_appetency_vanilla_score = ['GBDT (Appetency Vanilla)', gbdt_train_auc_score_appetency_vanilla, gbd
t_test_auc_score_appetency_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_appetency_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------------------+--------------------+--------------------+
|          Model           |     Train AUC      |      Test AUC      |
+--------------------------+--------------------+--------------------+
| GBDT (Appetency Vanilla) | 0.8614878188496815 | 0.8419659260682851 |
+--------------------------+--------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_vanilla_score)
score_table.add_row(rf_appetency_vanilla_score)
score_table.add_row(gbdt_appetency_vanilla_score)

print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency Vanilla) | 0.8004914800871239 | 0.8022376961601908 |
|    Random Forest (Appetency Vanilla)    | 0.8760928373360423 | 0.8213869803856969 |
|         GBDT (Appetency Vanilla)        | 0.8614878188496815 | 0.8419659260682851 |
+-----------------------------------------+--------------------+--------------------+
```

Obsetvation:

- **GBDT is performing best on Appetency vanilla Dataset**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.1, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 500, max_depth= 5, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_appetency, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 20, max_depth= 3, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.001,0.01,0.1,1,10]}
```

In [ ]:

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_appetency_vanilla,y_train_appetency)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits

[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 13.5min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
                                                                       eta0=0.0,
                                                                       fit_intercept=True,
                                                                       l1_ratio=0.15,
                                                                       learning_rate='optimal',
                                                                       loss='log',
                                                                       max_iter=1000,
                                                                       n_iter_no_change=5,
                                                                       n_jobs=-1,
                                                                       penalty='l2',
                                                                       power_t=0.5,
                                                                       random_state=None,
                                                                       shuffle=True,
                                                                       t...
                                                                          validation_fraction=0.1,
                                                                          verbose=0,
                                                                          warm_start=False),
                                            shuffle=True,
                                            store_train_meta_features=False,
                                            stratify=True, use_clones=True,
                                            use_features_in_secondary=False,
                                            use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	51.702154	1.044708	0.638183	0.004568	0.001	{'meta_classifier__al 0.001}
1	53.594777	3.878958	0.640064	0.002612	0.01	{'meta_classifier__al 0.01}
2	51.503189	0.317352	0.641331	0.002452	0.1	{'meta_classifier__al 0.1}
3	52.055218	0.804770	0.643383	0.007215	1	{'meta_classifier__al 1}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	{'meta_classifier__al
4	51.594383	0.169867	0.642029	0.004483	10	10}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_appetency_vanilla,y_train_appetency)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1, average=False,
                                                class_weight='balanced',
                                                early_stopping=False,
                                                epsilon=0.1, eta0=0.0,
                                                fit_intercept=True,
                                                l1_ratio=0.15,
                                                learning_rate='optimal',
                                                loss='log', max_iter=1000,
                                                n_iter_no_change=5, n_jobs=-1,
                                                penalty='l2', power_t=0.5,
                                                random_state=None, shuffle=True,
                                                tol=0.001,
                                                validation_fraction=0.1,
                                                verbose=0, w...
                                                   fit_intercept=True,
                                                   l1_ratio=0.15,
                                                   learning_rate='optimal',
                                                   loss='log', max_iter=1000,
                                                   n_iter_no_change=5,
                                                   n_jobs=-1, penalty='l2',
                                                   power_t=0.5,
                                                   random_state=None,
                                                   shuffle=True, tol=0.001,
                                                   validation_fraction=0.1,
                                                   verbose=0,
                                                   warm_start=False),
                          shuffle=True, store_train_meta_features=False,
                          stratify=True, use_clones=True,
                          use_features_in_secondary=False, use_probas=True,
                          verbose=0)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency_vanilla)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency_vanilla)[:,1]
```

In [ ]:

```
stack_train_auc_score_appetency_vanilla = roc_auc_score(y_train_appetency, y_train_appetency_pred)
stack_test_auc_score_appetency_vanilla = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

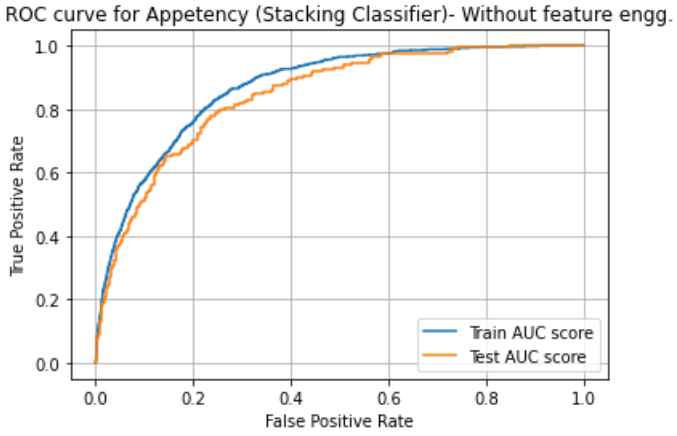
In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Stacking Classifier)- Without feature engg.')
```

```
plt.title('ROC curve for Appetency (Stacking Classifier)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_appetency_vanilla_score = ['Stacking Classifier (Appetency Vanilla)', stack_train_auc_score_appet
ency_vanilla, stack_test_auc_score_appetency_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(stack_appetency_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
| Stacking Classifier (Appetency Vanilla) | 0.8689946497086338 | 0.8438709020566075 |
+-----------------------------------------+--------------------+--------------------+
```

Churn

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_churn_vanilla, y_train_churn)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  15 out of  15 | elapsed:    5.1s finished
```

```
-- Epoch 1
Norm: 0.41, NNZs: 72, Bias: -0.130897, T: 40000, Avg. loss: 0.679042
Total training time: 0.03 seconds.
-- Epoch 2
Norm: 0.41, NNZs: 72, Bias: -0.139987, T: 80000, Avg. loss: 0.645103
Total training time: 0.05 seconds.
-- Epoch 3
Norm: 0.40, NNZs: 72, Bias: -0.141419, T: 120000, Avg. loss: 0.641798
Total training time: 0.08 seconds.
-- Epoch 4
Norm: 0.40, NNZs: 72, Bias: -0.146122, T: 160000, Avg. loss: 0.645662
Total training time: 0.11 seconds.
-- Epoch 5
Norm: 0.40, NNZs: 72, Bias: -0.145821, T: 200000, Avg. loss: 0.641147
Total training time: 0.13 seconds.
-- Epoch 6
Norm: 0.40, NNZs: 72, Bias: -0.147054, T: 240000, Avg. loss: 0.643310
Total training time: 0.15 seconds.
-- Epoch 7
Norm: 0.40, NNZs: 72, Bias: -0.147142, T: 280000, Avg. loss: 0.642164
Total training time: 0.18 seconds.
-- Epoch 8
Norm: 0.40, NNZs: 72, Bias: -0.147237, T: 320000, Avg. loss: 0.641698
Total training time: 0.21 seconds.
Convergence after 8 epochs took 0.21 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=None,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
0	0.676439	0.198540	0.012283	0.001276	0.001	{'alpha': 0.001}	0.669284	0.669949
1	0.419627	0.016270	0.019135	0.005877	0.01	{'alpha': 0.01}	0.674878	0.677014
2	0.236400	0.030160	0.010769	0.000197	0.1	{'alpha': 0.1}	0.675645	0.678588

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
3	0.603823	0.315538	0.010592	0.000151	1	{'alpha': 1}	0.667071	0.661127
4	0.507201	0.241104	0.009424	0.001984	10	{'alpha': 10}	0.652559	0.642734

In [ ]:

```
clf = classifier.best_estimator_
```

In [ ]:

```
# clf = SGDClassifier(loss = 'log',alpha = 0.001, class_weight= 'balanced', n_jobs = -1)
```

In [ ]:

```
clf.fit(X_train_churn_vanilla,y_train_churn)
```

```
-- Epoch 1
Norm: 0.41, NNZs: 72, Bias: -0.130897, T: 40000, Avg. loss: 0.679042
Total training time: 0.03 seconds.
-- Epoch 2
Norm: 0.41, NNZs: 72, Bias: -0.139987, T: 80000, Avg. loss: 0.645103
Total training time: 0.05 seconds.
-- Epoch 3
Norm: 0.40, NNZs: 72, Bias: -0.141419, T: 120000, Avg. loss: 0.641798
Total training time: 0.08 seconds.
-- Epoch 4
Norm: 0.40, NNZs: 72, Bias: -0.146122, T: 160000, Avg. loss: 0.645662
Total training time: 0.10 seconds.
-- Epoch 5
Norm: 0.40, NNZs: 72, Bias: -0.145821, T: 200000, Avg. loss: 0.641147
Total training time: 0.13 seconds.
-- Epoch 6
Norm: 0.40, NNZs: 72, Bias: -0.147054, T: 240000, Avg. loss: 0.643310
Total training time: 0.15 seconds.
-- Epoch 7
Norm: 0.40, NNZs: 72, Bias: -0.147142, T: 280000, Avg. loss: 0.642164
Total training time: 0.18 seconds.
-- Epoch 8
Norm: 0.40, NNZs: 72, Bias: -0.147237, T: 320000, Avg. loss: 0.641698
Total training time: 0.20 seconds.
Convergence after 8 epochs took 0.20 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.1, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=None, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn_vanilla)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn_vanilla)[:,1]
```

In [ ]:

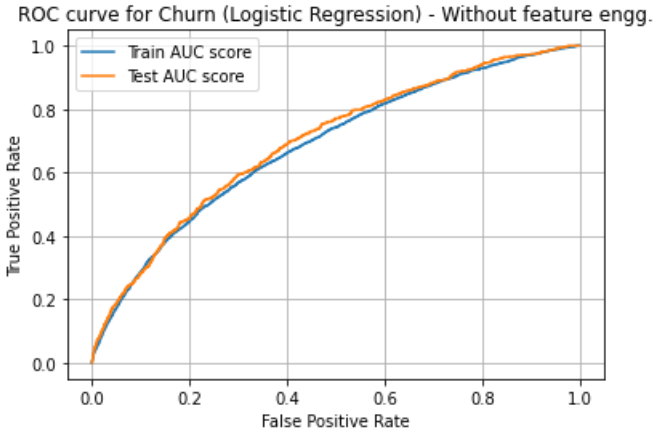
```
lr_train_auc_score_churn_vanilla = roc_auc_score(y_train_churn, y_train_churn_pred)
lr_test_auc_score_churn_vanilla = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Logistic Regression) - Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
lr_churn_vanilla_score = ['Logistic Regression (Churn Vanilla)', lr_train_auc_score_churn_vanilla, lr_t
est_auc_score_churn_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+-------------------------------------+--------------------+--------------------+
|                Model                |     Train AUC      |      Test AUC      |
+-------------------------------------+--------------------+--------------------+
| Logistic Regression (Churn Vanilla) | 0.6842109623015381 | 0.6967547407503685 |
+-------------------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced', n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= Tr
ue)
```

In [ ]:

```
classifier.fit(X_train_churn_vanilla,y_train_churn)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  90 out of  90 | elapsed: 10.4min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	0.396319	0.025482	0.110572	0.000393	3	10	{'max_dep 3, 'n_estima 10}
1	0.546118	0.057228	0.112581	0.001149	3	20	{'max_dep 3, 'n_estima 20}
2	1.232434	0.013550	0.111576	0.000880	3	50	{'max_dep 3, 'n_estima 50}
3	2.275039	0.033231	0.144148	0.047133	3	100	{'max_dep 3, 'n_estima 100}
4	4.261907	0.069236	0.211096	0.000402	3	200	{'max_dep 3, 'n_estima

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	200} pa
5	10.571623	0.070665	0.512264	0.000880	3	500	{'max_dep 3, 'n_estima 500}
6	0.464924	0.009638	0.110606	0.000443	5	10	{'max_dep 5, 'n_estima 10}
7	0.760910	0.005566	0.110416	0.000255	5	20	{'max_dep 5, 'n_estima 20}
8	1.728446	0.033643	0.110235	0.000244	5	50	{'max_dep 5, 'n_estima 50}
9	3.210008	0.029865	0.110127	0.000190	5	100	{'max_dep 5, 'n_estima 100}
10	6.255930	0.016779	0.212312	0.001758	5	200	{'max_dep 5, 'n_estima 200}
11	15.419491	0.049116	0.511094	0.000185	5	500	{'max_dep 5, 'n_estima 500}
12	0.558663	0.006362	0.112522	0.002985	7	10	{'max_dep 7, 'n_estima 10}
13	0.969508	0.004588	0.110376	0.000263	7	20	{'max_dep 7, 'n_estima 20}
14	2.185903	0.048921	0.110226	0.000221	7	50	{'max_dep 7, 'n_estima 50}
15	4.260006	0.041757	0.210688	0.000184	7	100	{'max_dep 7, 'n_estima 100}
16	8.333255	0.016622	0.311884	0.001289	7	200	{'max_dep 7, 'n_estima 200}
17	20.459204	0.069260	0.613690	0.003485	7	500	{'max_dep 7, 'n_estima 500}
18	0.672152	0.006262	0.110517	0.000319	10	10	{'max_dep 10, 'n_estima 10}
							{'max_dep 10

19	1.175808	0.018490	0.111336	0.000710	10	20	'n_estima
	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
							20}
20	2.767316	0.033562	0.111684	0.001223	10	50	{'max_dep 10, 'n_estima 50}
21	5.365231	0.004451	0.210415	0.000380	10	100	{'max_dep 10, 'n_estima 100}
22	10.594088	0.034498	0.310662	0.000298	10	200	{'max_dep 10, 'n_estima 200}
23	26.217347	0.129530	0.711457	0.000225	10	500	{'max_dep 10, 'n_estima 500}
24	0.761506	0.048730	0.113498	0.003198	15	10	{'max_dep 15, 'n_estima 10}
25	1.467881	0.011142	0.110368	0.000337	15	20	{'max_dep 15, 'n_estima 20}
26	3.414927	0.023102	0.110325	0.000102	15	50	{'max_dep 15, 'n_estima 50}
27	6.605994	0.001808	0.210960	0.000291	15	100	{'max_dep 15, 'n_estima 100}
28	13.185388	0.017115	0.411067	0.000207	15	200	{'max_dep 15, 'n_estima 200}
29	32.270094	0.312875	0.813185	0.002084	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn_vanilla,y_train_churn)
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=7, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=500,
```

```
                       n_jobs=-1, oob_score=False, random_state=None, verbose=0,
                       warm_start=False)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn_vanilla)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn_vanilla)[:,1]
```

In [ ]:

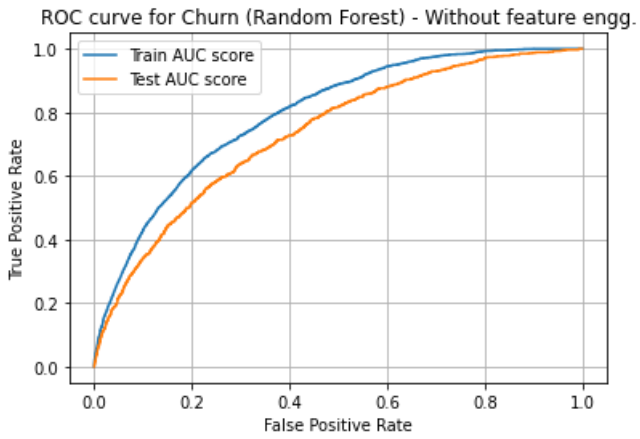
```
rf_train_auc_score_churn_vanilla = roc_auc_score(y_train_churn, y_train_churn_pred)
rf_test_auc_score_churn_vanilla = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Random Forest) - Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_churn_vanilla_score = ['Random Forest (Churn Vanilla)', rf_train_auc_score_churn_vanilla, rf_test_au
c_score_churn_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_churn_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+------------------------------+--------------------+--------------------+
|             Model            |     Train AUC      |      Test AUC      |
+------------------------------+--------------------+--------------------+
| Random Forest (Churn Vanilla) | 0.7946363330829114 | 0.7351618321589403 |
+------------------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_churn, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,300,500], 'max_depth' : [1,2,3,4,5]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc',n_jobs= -1, cv = 3, verbose=1, return_trai
n_score= True)
```

In [ ]:

```
classifier.fit(X_train_churn_vanilla,y_train_churn)
```

```
Fitting 3 folds for each of 35 candidates, totalling 105 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:  3.1min
[Parallel(n_jobs=-1)]: Done 105 out of 105 | elapsed: 13.8min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=12.614703880190605,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=-1,
             param_grid={'max_depth': [1, 2, 3, 4, 5],
                         'n_estimators': [10, 20, 50, 100, 200, 300, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
							{'max_der

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	{'max_dep
0	0.781127	0.049295	0.025006	0.000521	1	10	1, pa
							'n_estima
							10}
1	1.052067	0.007456	0.028445	0.002963	1	20	{'max_dep
							1,
							'n_estima
							20}
2	2.210635	0.021363	0.033299	0.000558	1	50	{'max_dep
							1,
							'n_estima
							50}
3	4.150005	0.023554	0.048046	0.002020	1	100	{'max_dep
							1,
							'n_estima
							100}
4	8.019980	0.014970	0.070252	0.000357	1	200	{'max_dep
							1,
							'n_estima
							200}
5	11.810273	0.036733	0.095261	0.000509	1	300	{'max_dep
							1,
							'n_estima
							300}
6	19.544491	0.116012	0.147299	0.004658	1	500	{'max_dep
							1,
							'n_estima
							500}
7	0.892777	0.003783	0.026502	0.000337	2	10	{'max_dep
							2,
							'n_estima
							10}
8	1.518763	0.000317	0.029926	0.000239	2	20	{'max_dep
							2,
							'n_estima
							20}
9	3.397994	0.015242	0.045805	0.004175	2	50	{'max_dep
							2,
							'n_estima
							50}
10	6.528205	0.045987	0.065704	0.000323	2	100	{'max_dep
							2,
							'n_estima
							100}
11	12.833131	0.033458	0.109805	0.001030	2	200	{'max_dep
							2,
							'n_estima
							200}
12	18.885443	0.042568	0.150644	0.002721	2	300	{'max_dep
							2,
							'n_estima
							300}
13	31.333681	0.129898	0.230218	0.005920	2	500	{'max_dep
							2,
							'n_estima
							500}
14	1.175054	0.003400	0.028022	0.000098	3	10	{'max_dep
							3,
							'n_estima
							10}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
15	2.063443	0.012211	0.033814	0.001118	3	20	{'max_dep 3, 'n_estima 20}
16	4.740690	0.043230	0.051481	0.000333	3	50	{'max_dep 3, 'n_estima 50}
17	9.120365	0.101596	0.083756	0.002607	3	100	{'max_dep 3, 'n_estima 100}
18	17.921890	0.179613	0.140725	0.001860	3	200	{'max_dep 3, 'n_estima 200}
19	26.530840	0.136606	0.193792	0.001923	3	300	{'max_dep 3, 'n_estima 300}
20	43.886382	0.312167	0.308696	0.012696	3	500	{'max_dep 3, 'n_estima 500}
21	1.467247	0.011388	0.029151	0.000066	4	10	{'max_dep 4, 'n_estima 10}
22	2.638001	0.018523	0.036163	0.000247	4	20	{'max_dep 4, 'n_estima 20}
23	6.083824	0.027310	0.061612	0.000589	4	50	{'max_dep 4, 'n_estima 50}
24	11.742836	0.018710	0.101354	0.001438	4	100	{'max_dep 4, 'n_estima 100}
25	23.054617	0.029312	0.171491	0.002207	4	200	{'max_dep 4, 'n_estima 200}
26	34.400711	0.256632	0.245659	0.004781	4	300	{'max_dep 4, 'n_estima 300}
27	57.151117	0.230591	0.390237	0.005214	4	500	{'max_dep 4, 'n_estima 500}
28	1.797335	0.013839	0.032799	0.000421	5	10	{'max_dep 5, 'n_estima 10}
29	3.295058	0.018161	0.042553	0.000795	5	20	{'max_dep 5,

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	'n_estima pa 20}
30	7.673005	0.068674	0.071281	0.000375	5	50	{'max_dep 5, 'n_estima 50}
31	14.700492	0.061496	0.119962	0.000207	5	100	{'max_dep 5, 'n_estima 100}
32	28.744477	0.178194	0.208921	0.007359	5	200	{'max_dep 5, 'n_estima 200}
33	42.882012	0.380348	0.305286	0.002361	5	300	{'max_dep 5, 'n_estima 300}
34	65.484928	7.980656	0.453253	0.044822	5	500	{'max_dep 5, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn_vanilla,y_train_churn)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=1,
              min_child_weight=1, missing=None, n_estimators=300, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=12.614703880190605,
              seed=None, silent=None, subsample=1, verbosity=1)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn_vanilla)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn_vanilla)[:,1]
```

In [ ]:

```
gbdt_train_auc_score_churn_vanilla = roc_auc_score(y_train_churn, y_train_churn_pred)
gbdt_test_auc_score_churn_vanilla = roc_auc_score(y_test_churn, y_test_churn_pred)
```

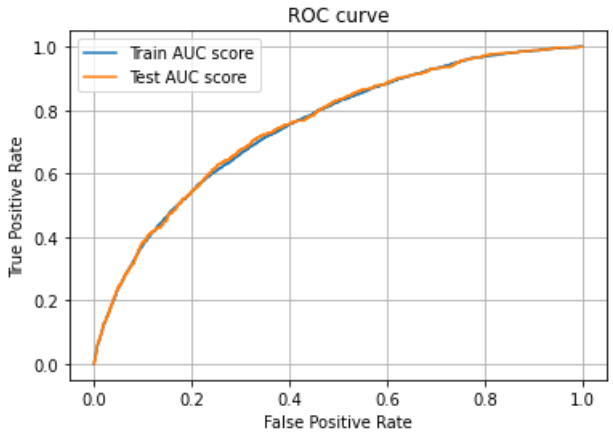
In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
```

```
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve')
plt.grid()
plt.show()
```



In [ ]:

```
gbdt_churn_vanilla_score = ['GBDT (Churn Vanilla)', gbdt_train_auc_score_churn_vanilla, gbdt_test_auc_s
core_churn_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_churn_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+---------------------+--------------------+------------------+
|        Model        |     Train AUC      |     Test AUC     |
+---------------------+--------------------+------------------+
| GBDT (Churn Vanilla) | 0.7486085630837572 | 0.75076905048547 |
+---------------------+--------------------+------------------+
```

In [ ]:

```
lr_churn_vanilla_score = ['Logistic Regression (Churn Vanilla)', '0.6842109623015381','0.69675474075036
85']
rf_churn_vanilla_score = ['Random Forest (Churn Vanilla)', '0.7946363330829114', '0.7351618321589403']
gbdt_churn_vanilla_score = ['GBDT (Churn Vanilla)', '0.7486085630837572',  '0.75076905048547']
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_vanilla_score)
score_table.add_row(rf_churn_vanilla_score)
score_table.add_row(gbdt_churn_vanilla_score)

print(score_table)
```

```
+-----------------------------------+--------------------+--------------------+
|                Model              |     Train AUC      |      Test AUC      |
+-----------------------------------+--------------------+--------------------+
| Logistic Regression (Churn Vanilla) | 0.6842109623015381 | 0.6967547407503685 |
|    Random Forest (Churn Vanilla)    | 0.7946363330829114 | 0.7351618321589403 |
|         GBDT (Churn Vanilla)        | 0.7486085630837572 |  0.75076905048547  |
+-----------------------------------+--------------------+--------------------+
```

Observation:

- **GBDT is performing best on Churn Vanilla Dataset**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.1, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 500, max_depth= 7, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_churn, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 300, max_depth= 1, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.001,0.01,0.1,1,10]}
```

In [ ]:

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_churn_vanilla,y_train_churn)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 20.9min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
                                                                       eta0=0.0,
                                                                       fit_intercept=True,
```

```
                                                    l1_ratio=0.15,
                                                    learning_rate='optimal',
                                                    loss='log',
                                                    max_iter=1000,
                                                    n_iter_no_change=5,
                                                    n_jobs=-1,
                                                    penalty='l2',
                                                    power_t=0.5,
                                                    random_state=None,
                                                    shuffle=True,
                                                    t...
                                                       validation_fraction=0.1,
                                                       verbose=0,
                                                       warm_start=False),
                                    shuffle=True,
                                    store_train_meta_features=False,
                                    stratify=True, use_clones=True,
                                    use_features_in_secondary=False,
                                    use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	81.726316	0.516360	0.680670	0.009544	0.001	{'meta_classifier__al 0.001}
1	82.915531	0.927563	0.738008	0.042789	0.01	{'meta_classifier__al 0.01}
2	81.400961	0.177815	0.672055	0.005134	0.1	{'meta_classifier__al 0.1}
3	81.104166	0.169014	0.673726	0.001384	1	{'meta_classifier__al 1}
4	81.389621	0.012089	0.674111	0.004087	10	{'meta_classifier__al 10}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn_vanilla,y_train_churn)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1, average=False,
                                                class_weight='balanced',
                                                early_stopping=False,
                                                epsilon=0.1, eta0=0.0,
                                                fit_intercept=True,
                                                l1_ratio=0.15,
                                                learning_rate='optimal',
                                                loss='log', max_iter=1000,
```

```
                                        n_iter_no_change=5, n_jobs=-1,
                                        penalty='l2', power_t=0.5,
                                        random_state=None, shuffle=True,
                                        tol=0.001,
                                        validation_fraction=0.1,
                                        verbose=0, w...
                                           fit_intercept=True,
                                           l1_ratio=0.15,
                                           learning_rate='optimal',
                                           loss='log', max_iter=1000,
                                           n_iter_no_change=5,
                                           n_jobs=-1, penalty='l2',
                                           power_t=0.5,
                                           random_state=None,
                                           shuffle=True, tol=0.001,
                                           validation_fraction=0.1,
                                           verbose=0,
                                           warm_start=False),
                    shuffle=True, store_train_meta_features=False,
                    stratify=True, use_clones=True,
                    use_features_in_secondary=False, use_probas=True,
                    verbose=0)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn_vanilla)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn_vanilla)[:,1]
```

In [ ]:

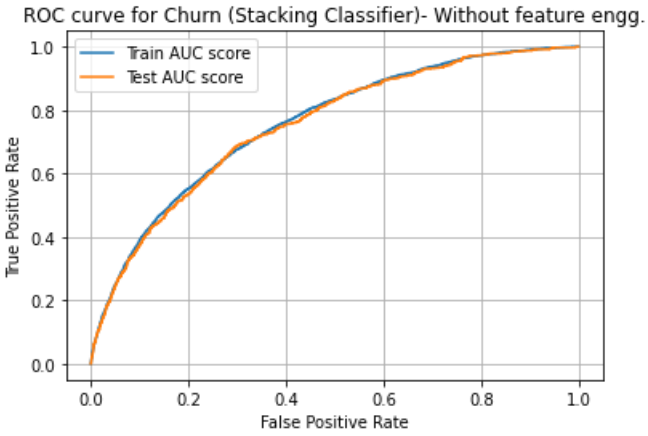
```
stack_train_auc_score_churn_vanilla = roc_auc_score(y_train_churn, y_train_churn_pred)
stack_test_auc_score_churn_vanilla = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Stacking Classifier)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_churn_vanilla_score = ['Stacking Classifier (churn Vanilla)', stack_train_auc_score_churn_vanilla
```

```
, stack_test_auc_score_churn_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(stack_churn_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------+--------------------+--------------------+
|               Model               |     Train AUC      |      Test AUC      |
+-----------------------------------+--------------------+--------------------+
| Stacking Classifier (churn Vanilla) | 0.7569440059210848 | 0.7516441109891073 |
+-----------------------------------+--------------------+--------------------+
```

Upselling

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_upselling_vanilla, y_train_upselling)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  15 out of  15 | elapsed:    3.1s finished
```

```
-- Epoch 1
Norm: 1.62, NNZs: 72, Bias: -0.744418, T: 40000, Avg. loss: 1.975277
Total training time: 0.03 seconds.
-- Epoch 2
Norm: 1.54, NNZs: 72, Bias: -0.673673, T: 80000, Avg. loss: 0.559542
Total training time: 0.05 seconds.
-- Epoch 3
Norm: 1.49, NNZs: 72, Bias: -0.712260, T: 120000, Avg. loss: 0.553406
Total training time: 0.09 seconds.
-- Epoch 4
Norm: 1.50, NNZs: 72, Bias: -0.650362, T: 160000, Avg. loss: 0.542605
Total training time: 0.11 seconds.
-- Epoch 5
Norm: 1.49, NNZs: 72, Bias: -0.697259, T: 200000, Avg. loss: 0.548246
Total training time: 0.14 seconds.
-- Epoch 6
Norm: 1.49, NNZs: 72, Bias: -0.683812, T: 240000, Avg. loss: 0.541815
Total training time: 0.16 seconds.
-- Epoch 7
```

```
Norm: 1.47, NNZs: 72, Bias: -0.674632, T: 280000, Avg. loss: 0.540752
Total training time: 0.19 seconds.
-- Epoch 8
Norm: 1.46, NNZs: 72, Bias: -0.692469, T: 320000, Avg. loss: 0.543860
Total training time: 0.22 seconds.
-- Epoch 9
Norm: 1.48, NNZs: 72, Bias: -0.670301, T: 360000, Avg. loss: 0.537956
Total training time: 0.24 seconds.
-- Epoch 10
Norm: 1.48, NNZs: 72, Bias: -0.661715, T: 400000, Avg. loss: 0.538397
Total training time: 0.27 seconds.
-- Epoch 11
Norm: 1.47, NNZs: 72, Bias: -0.667997, T: 440000, Avg. loss: 0.540646
Total training time: 0.30 seconds.
-- Epoch 12
Norm: 1.47, NNZs: 72, Bias: -0.676296, T: 480000, Avg. loss: 0.541530
Total training time: 0.33 seconds.
-- Epoch 13
Norm: 1.47, NNZs: 72, Bias: -0.671202, T: 520000, Avg. loss: 0.538997
Total training time: 0.35 seconds.
-- Epoch 14
Norm: 1.47, NNZs: 72, Bias: -0.678008, T: 560000, Avg. loss: 0.540901
Total training time: 0.38 seconds.
Convergence after 14 epochs took 0.38 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=None,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
0	0.781771	0.123422	0.013597	0.003639	0.001	{'alpha': 0.001}	0.786653	0.798252
1	0.407265	0.027677	0.012708	0.002858	0.01	{'alpha': 0.01}	0.799428	0.795469
2	0.219520	0.012390	0.010671	0.000339	0.1	{'alpha': 0.1}	0.795893	0.786419
3	0.191356	0.010884	0.010669	0.000247	1	{'alpha': 1}	0.755372	0.739598
4	0.164769	0.007879	0.009447	0.001883	10	{'alpha': 10}	0.710804	0.698827

In [ ]:

```
clf = classifier.best_estimator_
```

In [ ]:

```
# clf = SGDClassifier(loss = 'log',alpha = 0.001, class_weight= 'balanced', n_jobs = -1)
```

In [ ]:

```
clf.fit(X_train_upselling_vanilla,y_train_upselling)
```

```
-- Epoch 1
Norm: 1.62, NNZs: 72, Bias: -0.744418, T: 40000, Avg. loss: 1.975277
Total training time: 0.03 seconds.
-- Epoch 2
Norm: 1.54, NNZs: 72, Bias: -0.673673, T: 80000, Avg. loss: 0.559542
Total training time: 0.06 seconds.
-- Epoch 3
Norm: 1.49, NNZs: 72, Bias: -0.712260, T: 120000, Avg. loss: 0.553406
Total training time: 0.08 seconds.
-- Epoch 4
Norm: 1.50, NNZs: 72, Bias: -0.650362, T: 160000, Avg. loss: 0.542605
Total training time: 0.11 seconds.
-- Epoch 5
Norm: 1.49, NNZs: 72, Bias: -0.697259, T: 200000, Avg. loss: 0.548246
Total training time: 0.13 seconds.
-- Epoch 6
Norm: 1.49, NNZs: 72, Bias: -0.683812, T: 240000, Avg. loss: 0.541815
Total training time: 0.16 seconds.
-- Epoch 7
Norm: 1.47, NNZs: 72, Bias: -0.674632, T: 280000, Avg. loss: 0.540752
Total training time: 0.19 seconds.
-- Epoch 8
Norm: 1.46, NNZs: 72, Bias: -0.692469, T: 320000, Avg. loss: 0.543860
Total training time: 0.23 seconds.
-- Epoch 9
Norm: 1.48, NNZs: 72, Bias: -0.670301, T: 360000, Avg. loss: 0.537956
Total training time: 0.26 seconds.
-- Epoch 10
Norm: 1.48, NNZs: 72, Bias: -0.661715, T: 400000, Avg. loss: 0.538397
Total training time: 0.29 seconds.
-- Epoch 11
Norm: 1.47, NNZs: 72, Bias: -0.667997, T: 440000, Avg. loss: 0.540646
Total training time: 0.32 seconds.
-- Epoch 12
Norm: 1.47, NNZs: 72, Bias: -0.676296, T: 480000, Avg. loss: 0.541530
Total training time: 0.35 seconds.
-- Epoch 13
Norm: 1.47, NNZs: 72, Bias: -0.671202, T: 520000, Avg. loss: 0.538997
Total training time: 0.38 seconds.
-- Epoch 14
Norm: 1.47, NNZs: 72, Bias: -0.678008, T: 560000, Avg. loss: 0.540901
Total training time: 0.41 seconds.
Convergence after 14 epochs took 0.41 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.01, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=None, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling_vanilla)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling_vanilla)[:,1]
```

In [ ]:

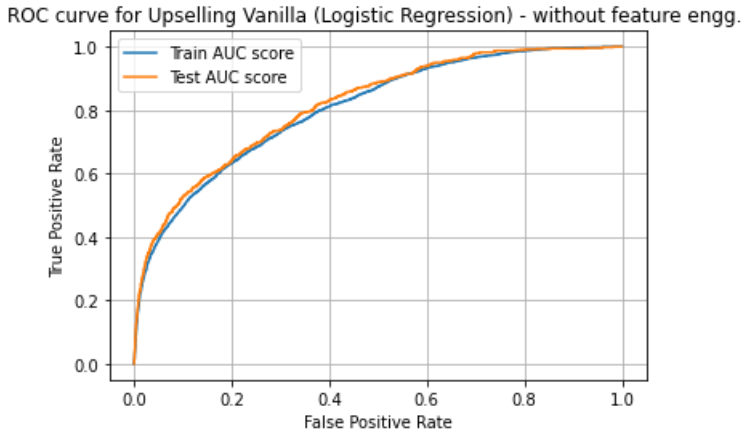
```
lr_train_auc_score_upselling_vanilla = roc_auc_score(y_train_upselling, y_train_upselling_pred)
lr_test_auc_score_upselling_vanilla = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling Vanilla (Logistic Regression) - without feature engg. ')
plt.grid()
plt.show()
```



In [ ]:

```
lr_upselling_vanilla_score = ['Logistic Regression (Upselling Vanilla)', lr_train_auc_score_upselling_v
anilla, lr_test_auc_score_upselling_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+----------------------------------------+--------------------+--------------------+
|                 Model                  |     Train AUC      |      Test AUC      |
+----------------------------------------+--------------------+--------------------+
| Logistic Regression (Upselling Vanilla) | 0.8049619771089853 | 0.8160040972065781 |
+----------------------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced',n_jobs = -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_upselling_vanilla,y_train_upselling)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:  2.8min
[Parallel(n_jobs=-1)]: Done  90 out of  90 | elapsed:  9.4min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	0.495963	0.037263	0.120906	0.009820	3	10	{'max_dep 3, 'n_estima 10}
1	0.798819	0.035540	0.121094	0.006240	3	20	{'max_dep 3, 'n_estima 20}
2	1.995262	0.038069	0.124629	0.009300	3	50	{'max_dep 3, 'n_estima

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	50} pa
3	3.818902	0.033190	0.272564	0.046657	3	100	{'max_dep 3, 'n_estima 100}
4	7.886909	0.181493	0.441068	0.002986	3	200	{'max_dep 3, 'n_estima 200}
5	19.373818	1.113196	1.013573	0.048290	3	500	{'max_dep 3, 'n_estima 500}
6	0.722123	0.067484	0.136581	0.006197	5	10	{'max_dep 5, 'n_estima 10}
7	1.313881	0.013615	0.135076	0.003358	5	20	{'max_dep 5, 'n_estima 20}
8	2.942913	0.179190	0.197201	0.055355	5	50	{'max_dep 5, 'n_estima 50}
9	5.804356	0.093853	0.270202	0.047316	5	100	{'max_dep 5, 'n_estima 100}
10	11.536702	0.202056	0.570745	0.052032	5	200	{'max_dep 5, 'n_estima 200}
11	28.571529	0.794921	1.069774	0.051285	5	500	{'max_dep 5, 'n_estima 500}
12	0.929942	0.030831	0.138526	0.002250	7	10	{'max_dep 7, 'n_estima 10}
13	1.690416	0.056286	0.137614	0.001774	7	20	{'max_dep 7, 'n_estima 20}
14	3.878578	0.133886	0.227780	0.010124	7	50	{'max_dep 7, 'n_estima 50}
15	7.415753	0.098200	0.340079	0.005405	7	100	{'max_dep 7, 'n_estima 100}
16	15.105763	0.220238	0.542877	0.005703	7	200	{'max_dep 7, 'n_estima 200}
							{'max_dep 7,

17	36.574232	1.300453	1.245832	0.083243	7	500	'n_estima
	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
							500}
18	1.144152	0.045375	0.136883	0.003314	10	10	{'max_dep 10, 'n_estima 10}
19	2.149566	0.064297	0.136513	0.000932	10	20	{'max_dep 10, 'n_estima 20}
20	4.895023	0.304559	0.235376	0.004013	10	50	{'max_dep 10, 'n_estima 50}
21	9.563269	0.038552	0.372946	0.046166	10	100	{'max_dep 10, 'n_estima 100}
22	18.918013	0.056506	0.641323	0.006785	10	200	{'max_dep 10, 'n_estima 200}
23	47.050353	1.354340	1.372628	0.048051	10	500	{'max_dep 10, 'n_estima 500}
24	1.433366	0.052592	0.133216	0.006049	15	10	{'max_dep 15, 'n_estima 10}
25	2.568910	0.115933	0.136369	0.002304	15	20	{'max_dep 15, 'n_estima 20}
26	6.122596	0.224198	0.232859	0.006030	15	50	{'max_dep 15, 'n_estima 50}
27	11.532823	0.067432	0.443077	0.006235	15	100	{'max_dep 15, 'n_estima 100}
28	23.152334	0.193008	0.667630	0.051755	15	200	{'max_dep 15, 'n_estima 200}
29	50.584068	10.585120	1.438424	0.442735	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train,y_train)
```

```
clf.fit(X_train_upselling_vanilla,y_train_upselling)
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=7, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=500,
                       n_jobs=-1, oob_score=False, random_state=None, verbose=0,
                       warm_start=False)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling_vanilla)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling_vanilla)[:,1]
```

In [ ]:

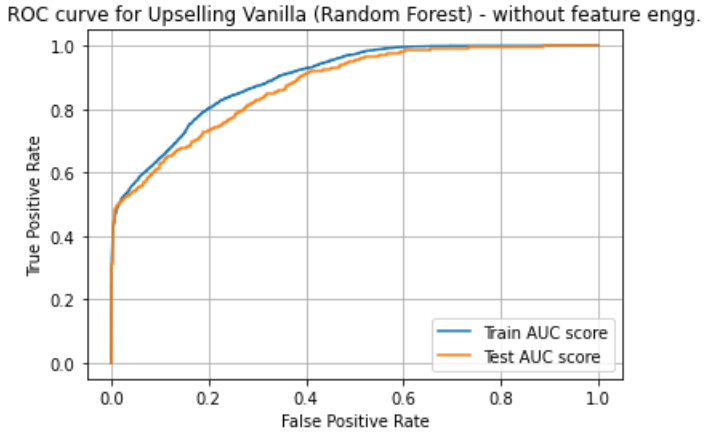
```
rf_train_auc_score_upselling_vanilla = roc_auc_score(y_train_upselling, y_train_upselling_pred)
rf_test_auc_score_upselling_vanilla = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling Vanilla (Random Forest) - without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_upselling_vanilla_score = ['Random Forest (Upselling Vanilla)', rf_train_auc_score_upselling_vanilla
, rf_test_auc_score_upselling_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_upselling_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------+--------------------+--------------------+
|               Model               |   Avg Train AUC    |    Avg Test AUC    |
+-----------------------------------+--------------------+--------------------+
| Random Forest (Upselling Vanilla) | 0.8963542355442349 | 0.8739862581662536 |
+-----------------------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_upselling, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,300,500], 'max_depth' : [1,2,3,4,5]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_upselling_vanilla,y_train_upselling)
```

```
Fitting 3 folds for each of 35 candidates, totalling 105 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:  3.2min
[Parallel(n_jobs=-1)]: Done 105 out of 105 | elapsed: 13.9min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=12.577732518669382,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=-1,
             param_grid={'max_depth': [1, 2, 3, 4, 5],
                         'n_estimators': [10, 20, 50, 100, 200, 300, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	0.799208	0.042533	0.058376	0.015238	1	10	{'max_dep 1, 'n_estima 10}
1	1.307191	0.183696	0.046451	0.019374	1	20	{'max_dep 1, 'n_estima 20}
2	2.762892	0.237335	0.035934	0.003019	1	50	{'max_dep 1, 'n_estima 50}
3	4.226590	0.183571	0.044686	0.000751	1	100	{'max_dep 1, 'n_estima 100}
4	7.934646	0.060116	0.066977	0.000338	1	200	{'max_dep 1, 'n_estima 200}
5	11.814854	0.031798	0.094211	0.004894	1	300	{'max_dep 1, 'n_estima 300}
6	19.416563	0.040227	0.139026	0.001726	1	500	{'max_dep 1, 'n_estima 500}
7	0.882797	0.002894	0.031229	0.008485	2	10	{'max_dep 2, 'n_estima 10}
8	1.500194	0.006253	0.028290	0.000292	2	20	{'max_dep 2, 'n_estima 20}
9	3.370905	0.005139	0.040924	0.001165	2	50	{'max_dep 2, 'n_estima 50}
10	6.580183	0.054963	0.063147	0.000546	2	100	{'max_dep 2, 'n_estima 100}
11	12.639682	0.116795	0.106541	0.001806	2	200	{'max_dep 2, 'n_estima 200}
12	18.928178	0.084349	0.148693	0.002740	2	300	{'max_dep 2, 'n_estima 300}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
13	31.419397	0.185423	0.233487	0.004636	2	500	{'max_dep 2, 'n_estima 500}
14	1.130068	0.006136	0.026775	0.000276	3	10	{'max_dep 3, 'n_estima 10}
15	2.011468	0.007860	0.030655	0.000269	3	20	{'max_dep 3, 'n_estima 20}
16	4.635302	0.024668	0.048593	0.000264	3	50	{'max_dep 3, 'n_estima 50}
17	9.067476	0.041803	0.080608	0.000742	3	100	{'max_dep 3, 'n_estima 100}
18	17.751084	0.066845	0.138562	0.002036	3	200	{'max_dep 3, 'n_estima 200}
19	26.374044	0.022132	0.198028	0.004814	3	300	{'max_dep 3, 'n_estima 300}
20	43.860117	0.120963	0.315005	0.004037	3	500	{'max_dep 3, 'n_estima 500}
21	1.419754	0.010218	0.028174	0.000200	4	10	{'max_dep 4, 'n_estima 10}
22	2.584914	0.021348	0.034662	0.000799	4	20	{'max_dep 4, 'n_estima 20}
23	6.059435	0.026196	0.056071	0.000740	4	50	{'max_dep 4, 'n_estima 50}
24	11.779335	0.037044	0.096947	0.002396	4	100	{'max_dep 4, 'n_estima 100}
25	23.210216	0.026957	0.174977	0.006743	4	200	{'max_dep 4, 'n_estima 200}
26	34.494946	0.069606	0.252948	0.008183	4	300	{'max_dep 4, 'n_estima 300}
27	57.276719	0.058984	0.406219	0.007105	4	500	{'max_dep 4,

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	'n_estimat pa 500}
28	1.752034	0.005632	0.029566	0.000408	5	10	{'max_dep 5, 'n_estima 10}
29	3.232225	0.022953	0.036707	0.000788	5	20	{'max_dep 5, 'n_estima 20}
30	7.676363	0.017654	0.063418	0.001630	5	50	{'max_dep 5, 'n_estima 50}
31	14.847343	0.083636	0.109798	0.001183	5	100	{'max_dep 5, 'n_estima 100}
32	29.324392	0.277661	0.208075	0.004512	5	200	{'max_dep 5, 'n_estima 200}
33	44.007388	0.075388	0.324122	0.012156	5	300	{'max_dep 5, 'n_estima 300}
34	66.845533	8.314101	0.460960	0.047911	5	500	{'max_dep 5, 'n_estima 500}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling_vanilla,y_train_upselling)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=2,
              min_child_weight=1, missing=None, n_estimators=100, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=12.577732518669382,
              seed=None, silent=None, subsample=1, verbosity=1)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling_vanilla)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling_vanilla)[:,1]
```

In [ ]:

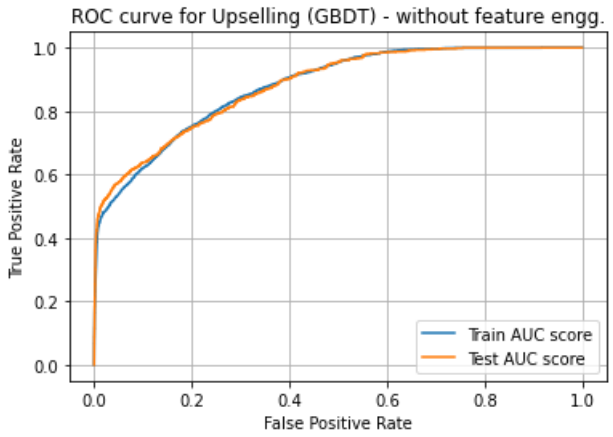
```
gbdt_train_auc_score_upselling_vanilla = roc_auc_score(y_train_upselling, y_train_upselling_pred)
gbdt_test_auc_score_upselling_vanilla = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling (GBDT) - without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
gbdt_upselling_vanilla_score = ['GBDT (Upselling Vanilla)', gbdt_train_auc_score_upselling_vanilla, gbd
t_test_auc_score_upselling_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_upselling_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------------------+--------------------+--------------------+
|          Model           |     Train AUC      |      Test AUC      |
+--------------------------+--------------------+--------------------+
| GBDT (Upselling Vanilla) | 0.8779311865389684 | 0.8797164954804011 |
+--------------------------+--------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_vanilla_score)
score_table.add_row(rf_upselling_vanilla_score)
score_table.add_row(gbdt_upselling_vanilla_score)

print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
| Logistic Regression (Upselling Vanilla) | 0.8049619771089853 | 0.8160040972065781 |
|    Random Forest (Upselling Vanilla)    | 0.8049619771089853 | 0.8160040972065781 |
```

```
|          GBDT (Upselling Vanilla)         | 0.8779311865389684 | 0.8797164954804011 |
+-------------------------------------------+--------------------+--------------------+
```

Observation:

- **GBDT is performing well on Upselling Vanilla Dataset**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.01, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 500, max_depth= 7, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_upselling, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 100, max_depth= 2, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.001,0.01,0.1,1,10]}
```

In [ ]:

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_upselling_vanilla,y_train_upselling)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits

[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 18.6min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.01,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
```

```
                                                        eta0=0.0,
                                                        fit_intercept=True,
                                                        l1_ratio=0.15,
                                                        learning_rate='optimal',
                                                        loss='log',
                                                        max_iter=1000,
                                                        n_iter_no_change=5,
                                                        n_jobs=-1,
                                                        penalty='l2',
                                                        power_t=0.5,
                                                        random_state=None,
                                                        shuffle=True,...
                                                           validation_fraction=0.1,
                                                           verbose=0,
                                                           warm_start=False),
                                        shuffle=True,
                                        store_train_meta_features=False,
                                        stratify=True, use_clones=True,
                                        use_features_in_secondary=False,
                                        use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	72.717597	0.346755	0.688122	0.052380	0.001	{'meta_classifier__al 0.001}
1	72.624091	0.230355	0.654626	0.003786	0.01	{'meta_classifier__al 0.01}
2	72.681169	0.554895	0.685843	0.044604	0.1	{'meta_classifier__al 0.1}
3	72.868510	0.090381	0.659026	0.002591	1	{'meta_classifier__al 1}
4	72.225790	0.327302	0.654480	0.003487	10	{'meta_classifier__al 10}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling_vanilla,y_train_upselling)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.01, average=False,
                                                class_weight='balanced',
                                                early_stopping=False,
                                                epsilon=0.1, eta0=0.0,
                                                fit_intercept=True,
                                                l1_ratio=0.15,
                                                learning_rate='optimal',
```

```
                                        loss='log', max_iter=1000,
                                        n_iter_no_change=5, n_jobs=-1,
                                        penalty='l2', power_t=0.5,
                                        random_state=None, shuffle=True,
                                        tol=0.001,
                                        validation_fraction=0.1,
                                        verbose=0,...
                                           fit_intercept=True,
                                           l1_ratio=0.15,
                                           learning_rate='optimal',
                                           loss='log', max_iter=1000,
                                           n_iter_no_change=5,
                                           n_jobs=-1, penalty='l2',
                                           power_t=0.5,
                                           random_state=None,
                                           shuffle=True, tol=0.001,
                                           validation_fraction=0.1,
                                           verbose=0,
                                           warm_start=False),
                    shuffle=True, store_train_meta_features=False,
                    stratify=True, use_clones=True,
                    use_features_in_secondary=False, use_probas=True,
                    verbose=0)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling_vanilla)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling_vanilla)[:,1]
```

In [ ]:

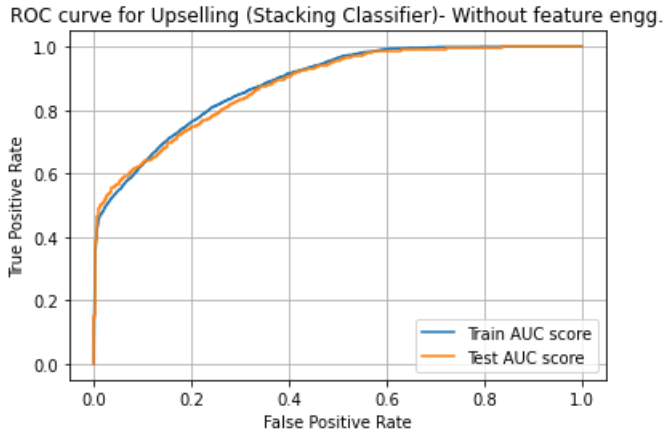
```
stack_train_auc_score_upselling_vanilla = roc_auc_score(y_train_upselling, y_train_upselling_pred)
stack_test_auc_score_upselling_vanilla = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling (Stacking Classifier)- Without feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_upselling_vanilla_score = ['Stacking Classifier (upselling Vanilla)', stack_train_auc_score_upsel
ling_vanilla, stack_test_auc_score_upselling_vanilla]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(stack_upselling_vanilla_score)
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
| Stacking Classifier (upselling Vanilla) | 0.8842637363329957 | 0.8795341187485919 |
+-----------------------------------------+--------------------+--------------------+
```

Dataset with feature engg.

Appetency

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13, n_jobs= -1)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= Tr
ue)
```

In [ ]:

```
classifier.fit(X_train_appetency, y_train_appetency)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
```

```
-- Epoch 1
Norm: 885.74, NNZs: 2297, Bias: -186.260486, T: 26666, Avg. loss: 1147.066881
Total training time: 0.20 seconds.
-- Epoch 2
Norm: 421.77, NNZs: 2297, Bias: -198.576212, T: 53332, Avg. loss: 171.520287
Total training time: 0.47 seconds.
-- Epoch 3
Norm: 281.69, NNZs: 2297, Bias: -191.218814, T: 79998, Avg. loss: 64.886984
Total training time: 0.72 seconds.
-- Epoch 4
Norm: 215.67, NNZs: 2297, Bias: -185.962289, T: 106664, Avg. loss: 38.048479
Total training time: 0.98 seconds.
-- Epoch 5
Norm: 178.53, NNZs: 2297, Bias: -179.065158, T: 133330, Avg. loss: 27.191797
Total training time: 1.22 seconds.
-- Epoch 6
```

```
-- Epoch 6
Norm: 154.34, NNZs: 2297, Bias: -173.733669, T: 159996, Avg. loss: 20.973034
Total training time: 1.46 seconds.
-- Epoch 7
Norm: 137.85, NNZs: 2297, Bias: -168.690458, T: 186662, Avg. loss: 14.866735
Total training time: 1.69 seconds.
-- Epoch 8
Norm: 124.84, NNZs: 2297, Bias: -165.224409, T: 213328, Avg. loss: 13.325569
Total training time: 1.93 seconds.
-- Epoch 9
Norm: 115.39, NNZs: 2297, Bias: -161.552035, T: 239994, Avg. loss: 11.092587
Total training time: 2.16 seconds.
-- Epoch 10
Norm: 109.26, NNZs: 2297, Bias: -157.832093, T: 266660, Avg. loss: 10.782811
Total training time: 2.40 seconds.
-- Epoch 11
Norm: 103.34, NNZs: 2297, Bias: -154.905693, T: 293326, Avg. loss: 8.571021
Total training time: 2.64 seconds.
-- Epoch 12
Norm: 98.28, NNZs: 2297, Bias: -152.519100, T: 319992, Avg. loss: 7.687741
Total training time: 2.88 seconds.
-- Epoch 13
Norm: 94.98, NNZs: 2297, Bias: -150.031540, T: 346658, Avg. loss: 7.642568
Total training time: 3.12 seconds.
-- Epoch 14
Norm: 92.38, NNZs: 2297, Bias: -147.378566, T: 373324, Avg. loss: 7.045796
Total training time: 3.36 seconds.
-- Epoch 15
Norm: 89.44, NNZs: 2297, Bias: -145.153379, T: 399990, Avg. loss: 7.002969
Total training time: 3.60 seconds.
-- Epoch 16
Norm: 86.80, NNZs: 2297, Bias: -143.194546, T: 426656, Avg. loss: 6.115847
Total training time: 3.84 seconds.
-- Epoch 17
Norm: 84.57, NNZs: 2297, Bias: -141.288631, T: 453322, Avg. loss: 5.754008
Total training time: 4.08 seconds.
-- Epoch 18
Norm: 82.90, NNZs: 2297, Bias: -139.388652, T: 479988, Avg. loss: 5.952241
Total training time: 4.33 seconds.
-- Epoch 19
Norm: 80.49, NNZs: 2297, Bias: -138.058488, T: 506654, Avg. loss: 5.022612
Total training time: 4.57 seconds.
-- Epoch 20
Norm: 78.76, NNZs: 2297, Bias: -136.588327, T: 533320, Avg. loss: 5.424740
Total training time: 4.81 seconds.
-- Epoch 21
Norm: 77.79, NNZs: 2297, Bias: -134.930955, T: 559986, Avg. loss: 4.885519
Total training time: 5.05 seconds.
-- Epoch 22
Norm: 76.60, NNZs: 2297, Bias: -133.428283, T: 586652, Avg. loss: 4.628902
Total training time: 5.30 seconds.
-- Epoch 23
Norm: 75.40, NNZs: 2297, Bias: -132.095736, T: 613318, Avg. loss: 4.301231
Total training time: 5.54 seconds.
-- Epoch 24
Norm: 73.99, NNZs: 2297, Bias: -130.886093, T: 639984, Avg. loss: 4.194104
Total training time: 5.78 seconds.
-- Epoch 25
Norm: 72.52, NNZs: 2297, Bias: -129.898584, T: 666650, Avg. loss: 3.733776
Total training time: 6.02 seconds.
-- Epoch 26
Norm: 72.01, NNZs: 2297, Bias: -128.534024, T: 693316, Avg. loss: 3.619723
Total training time: 6.25 seconds.
-- Epoch 27
Norm: 71.30, NNZs: 2297, Bias: -127.292584, T: 719982, Avg. loss: 3.632244
Total training time: 6.49 seconds.
-- Epoch 28
Norm: 70.95, NNZs: 2297, Bias: -126.048910, T: 746648, Avg. loss: 3.650890
Total training time: 6.72 seconds.
-- Epoch 29
Norm: 70.38, NNZs: 2297, Bias: -124.883704, T: 773314, Avg. loss: 3.432910
Total training time: 6.96 seconds.
-- Epoch 30
Norm: 69.24, NNZs: 2297, Bias: -124.091176, T: 799980, Avg. loss: 3.421070
Total training time: 7.19 seconds.
-- Epoch 31
Norm: 68.59, NNZs: 2297, Bias: -123.084980, T: 826646, Avg. loss: 3.189466
Total training time: 7.43 seconds.
```

```
Total training time: 7.43 seconds.
-- Epoch 32
Norm: 67.30, NNZs: 2297, Bias: -122.518703, T: 853312, Avg. loss: 3.082487
Total training time: 7.65 seconds.
-- Epoch 33
Norm: 66.91, NNZs: 2297, Bias: -121.588683, T: 879978, Avg. loss: 2.809412
Total training time: 7.89 seconds.
-- Epoch 34
Norm: 66.64, NNZs: 2297, Bias: -120.559762, T: 906644, Avg. loss: 2.849923
Total training time: 8.12 seconds.
-- Epoch 35
Norm: 66.12, NNZs: 2297, Bias: -119.719518, T: 933310, Avg. loss: 3.025893
Total training time: 8.35 seconds.
-- Epoch 36
Norm: 65.80, NNZs: 2297, Bias: -118.914239, T: 959976, Avg. loss: 2.824527
Total training time: 8.58 seconds.
-- Epoch 37
Norm: 65.13, NNZs: 2297, Bias: -118.192816, T: 986642, Avg. loss: 2.770913
Total training time: 8.81 seconds.
-- Epoch 38
Norm: 64.66, NNZs: 2297, Bias: -117.425986, T: 1013308, Avg. loss: 2.565359
Total training time: 9.04 seconds.
-- Epoch 39
Norm: 64.98, NNZs: 2297, Bias: -116.336415, T: 1039974, Avg. loss: 2.579260
Total training time: 9.27 seconds.
-- Epoch 40
Norm: 64.17, NNZs: 2297, Bias: -115.758107, T: 1066640, Avg. loss: 2.672514
Total training time: 9.50 seconds.
-- Epoch 41
Norm: 63.21, NNZs: 2297, Bias: -115.320653, T: 1093306, Avg. loss: 2.446951
Total training time: 9.73 seconds.
-- Epoch 42
Norm: 62.76, NNZs: 2297, Bias: -114.657682, T: 1119972, Avg. loss: 2.521005
Total training time: 9.97 seconds.
-- Epoch 43
Norm: 63.15, NNZs: 2297, Bias: -113.807325, T: 1146638, Avg. loss: 2.376549
Total training time: 10.19 seconds.
-- Epoch 44
Norm: 62.33, NNZs: 2297, Bias: -113.349084, T: 1173304, Avg. loss: 2.876498
Total training time: 10.42 seconds.
-- Epoch 45
Norm: 61.87, NNZs: 2297, Bias: -112.770800, T: 1199970, Avg. loss: 2.499887
Total training time: 10.65 seconds.
-- Epoch 46
Norm: 61.54, NNZs: 2297, Bias: -112.151329, T: 1226636, Avg. loss: 2.337654
Total training time: 10.88 seconds.
-- Epoch 47
Norm: 61.18, NNZs: 2297, Bias: -111.536046, T: 1253302, Avg. loss: 2.331595
Total training time: 11.10 seconds.
-- Epoch 48
Norm: 60.54, NNZs: 2297, Bias: -111.108468, T: 1279968, Avg. loss: 2.018838
Total training time: 11.33 seconds.
-- Epoch 49
Norm: 60.28, NNZs: 2297, Bias: -110.506997, T: 1306634, Avg. loss: 2.172510
Total training time: 11.56 seconds.
-- Epoch 50
Norm: 60.04, NNZs: 2297, Bias: -109.908518, T: 1333300, Avg. loss: 2.060846
Total training time: 11.78 seconds.
-- Epoch 51
Norm: 59.68, NNZs: 2297, Bias: -109.410622, T: 1359966, Avg. loss: 2.003822
Total training time: 12.01 seconds.
-- Epoch 52
Norm: 59.40, NNZs: 2297, Bias: -108.872132, T: 1386632, Avg. loss: 2.157914
Total training time: 12.23 seconds.
-- Epoch 53
Norm: 59.21, NNZs: 2297, Bias: -108.294089, T: 1413298, Avg. loss: 1.853835
Total training time: 12.46 seconds.
-- Epoch 54
Norm: 58.92, NNZs: 2297, Bias: -107.805186, T: 1439964, Avg. loss: 1.989796
Total training time: 12.69 seconds.
-- Epoch 55
Norm: 59.66, NNZs: 2297, Bias: -107.059932, T: 1466630, Avg. loss: 2.167928
Total training time: 12.92 seconds.
-- Epoch 56
Norm: 59.37, NNZs: 2297, Bias: -106.571236, T: 1493296, Avg. loss: 1.926271
Total training time: 13.14 seconds.
-- Epoch 57
```

```
Norm: 58.89, NNZs: 2297, Bias: -106.162725, T: 1519962, Avg. loss: 1.985603
Total training time: 13.36 seconds.
-- Epoch 58
Norm: 58.60, NNZs: 2297, Bias: -105.698303, T: 1546628, Avg. loss: 1.737021
Total training time: 13.59 seconds.
-- Epoch 59
Norm: 58.08, NNZs: 2297, Bias: -105.393137, T: 1573294, Avg. loss: 1.750621
Total training time: 13.82 seconds.
-- Epoch 60
Norm: 57.88, NNZs: 2297, Bias: -104.919246, T: 1599960, Avg. loss: 1.865353
Total training time: 14.04 seconds.
-- Epoch 61
Norm: 57.37, NNZs: 2297, Bias: -104.604323, T: 1626626, Avg. loss: 1.674946
Total training time: 14.26 seconds.
-- Epoch 62
Norm: 57.37, NNZs: 2297, Bias: -104.058616, T: 1653292, Avg. loss: 1.648201
Total training time: 14.50 seconds.
-- Epoch 63
Norm: 56.87, NNZs: 2297, Bias: -103.764123, T: 1679958, Avg. loss: 1.773796
Total training time: 14.72 seconds.
-- Epoch 64
Norm: 56.86, NNZs: 2297, Bias: -103.242615, T: 1706624, Avg. loss: 1.603282
Total training time: 14.94 seconds.
-- Epoch 65
Norm: 56.65, NNZs: 2297, Bias: -102.823598, T: 1733290, Avg. loss: 1.638542
Total training time: 15.16 seconds.
-- Epoch 66
Norm: 56.22, NNZs: 2297, Bias: -102.525253, T: 1759956, Avg. loss: 1.584920
Total training time: 15.39 seconds.
-- Epoch 67
Norm: 55.98, NNZs: 2297, Bias: -102.149727, T: 1786622, Avg. loss: 1.568192
Total training time: 15.62 seconds.
-- Epoch 68
Norm: 55.81, NNZs: 2297, Bias: -101.736171, T: 1813288, Avg. loss: 1.501834
Total training time: 15.84 seconds.
-- Epoch 69
Norm: 55.67, NNZs: 2297, Bias: -101.336498, T: 1839954, Avg. loss: 1.604779
Total training time: 16.07 seconds.
-- Epoch 70
Norm: 55.35, NNZs: 2297, Bias: -101.020401, T: 1866620, Avg. loss: 1.580018
Total training time: 16.29 seconds.
-- Epoch 71
Norm: 55.38, NNZs: 2297, Bias: -100.533165, T: 1893286, Avg. loss: 1.575150
Total training time: 16.52 seconds.
-- Epoch 72
Norm: 54.92, NNZs: 2297, Bias: -100.309086, T: 1919952, Avg. loss: 1.483635
Total training time: 16.75 seconds.
-- Epoch 73
Norm: 54.82, NNZs: 2297, Bias: -99.906627, T: 1946618, Avg. loss: 1.443803
Total training time: 16.97 seconds.
-- Epoch 74
Norm: 54.52, NNZs: 2297, Bias: -99.624594, T: 1973284, Avg. loss: 1.489043
Total training time: 17.19 seconds.
-- Epoch 75
Norm: 54.37, NNZs: 2297, Bias: -99.267422, T: 1999950, Avg. loss: 1.404626
Total training time: 17.43 seconds.
-- Epoch 76
Norm: 54.17, NNZs: 2297, Bias: -98.941934, T: 2026616, Avg. loss: 1.486633
Total training time: 17.65 seconds.
-- Epoch 77
Norm: 53.97, NNZs: 2297, Bias: -98.642650, T: 2053282, Avg. loss: 1.413570
Total training time: 17.88 seconds.
-- Epoch 78
Norm: 53.86, NNZs: 2297, Bias: -98.289986, T: 2079948, Avg. loss: 1.414060
Total training time: 18.10 seconds.
-- Epoch 79
Norm: 53.80, NNZs: 2297, Bias: -97.911559, T: 2106614, Avg. loss: 1.372963
Total training time: 18.32 seconds.
-- Epoch 80
Norm: 53.56, NNZs: 2297, Bias: -97.633984, T: 2133280, Avg. loss: 1.406809
Total training time: 18.55 seconds.
-- Epoch 81
Norm: 53.49, NNZs: 2297, Bias: -97.268121, T: 2159946, Avg. loss: 1.342754
Total training time: 18.77 seconds.
-- Epoch 82
Norm: 53.32, NNZs: 2297, Bias: -96.959876, T: 2186612, Avg. loss: 1.294346
Total training time: 18.99 seconds.
```

```
-- Epoch 83
Norm: 53.06, NNZs: 2297, Bias: -96.720899, T: 2213278, Avg. loss: 1.295240
Total training time: 19.22 seconds.
-- Epoch 84
Norm: 52.96, NNZs: 2297, Bias: -96.391693, T: 2239944, Avg. loss: 1.364508
Total training time: 19.44 seconds.
-- Epoch 85
Norm: 52.62, NNZs: 2297, Bias: -96.190973, T: 2266610, Avg. loss: 1.277521
Total training time: 19.67 seconds.
-- Epoch 86
Norm: 52.54, NNZs: 2297, Bias: -95.866905, T: 2293276, Avg. loss: 1.305253
Total training time: 19.89 seconds.
-- Epoch 87
Norm: 52.46, NNZs: 2297, Bias: -95.545872, T: 2319942, Avg. loss: 1.226121
Total training time: 20.12 seconds.
-- Epoch 88
Norm: 52.31, NNZs: 2297, Bias: -95.259929, T: 2346608, Avg. loss: 1.312943
Total training time: 20.35 seconds.
-- Epoch 89
Norm: 52.24, NNZs: 2297, Bias: -94.944412, T: 2373274, Avg. loss: 1.217994
Total training time: 20.58 seconds.
-- Epoch 90
Norm: 52.01, NNZs: 2297, Bias: -94.714809, T: 2399940, Avg. loss: 1.190015
Total training time: 20.80 seconds.
-- Epoch 91
Norm: 51.79, NNZs: 2297, Bias: -94.489234, T: 2426606, Avg. loss: 1.224681
Total training time: 21.02 seconds.
-- Epoch 92
Norm: 51.66, NNZs: 2297, Bias: -94.219757, T: 2453272, Avg. loss: 1.193917
Total training time: 21.25 seconds.
-- Epoch 93
Norm: 51.52, NNZs: 2297, Bias: -93.956476, T: 2479938, Avg. loss: 1.144430
Total training time: 21.47 seconds.
-- Epoch 94
Norm: 51.42, NNZs: 2297, Bias: -93.704526, T: 2506604, Avg. loss: 1.218194
Total training time: 21.70 seconds.
-- Epoch 95
Norm: 51.31, NNZs: 2297, Bias: -93.435860, T: 2533270, Avg. loss: 1.177889
Total training time: 21.92 seconds.
-- Epoch 96
Norm: 51.23, NNZs: 2297, Bias: -93.154120, T: 2559936, Avg. loss: 1.164151
Total training time: 22.14 seconds.
-- Epoch 97
Norm: 50.94, NNZs: 2297, Bias: -92.985613, T: 2586602, Avg. loss: 1.147860
Total training time: 22.36 seconds.
-- Epoch 98
Norm: 50.84, NNZs: 2297, Bias: -92.727713, T: 2613268, Avg. loss: 1.201395
Total training time: 22.58 seconds.
Convergence after 98 epochs took 22.58 seconds
-- Epoch 1
Norm: 274.51, NNZs: 2298, Bias: -145.044768, T: 26667, Avg. loss: 535.079461
Total training time: 0.23 seconds.
-- Epoch 2
Norm: 261.49, NNZs: 2298, Bias: -138.216271, T: 53334, Avg. loss: 128.910422
Total training time: 0.49 seconds.
-- Epoch 3
Norm: 223.13, NNZs: 2298, Bias: -134.121591, T: 80001, Avg. loss: 63.630134
Total training time: 0.73 seconds.
-- Epoch 4
Norm: 164.68, NNZs: 2298, Bias: -132.562049, T: 106668, Avg. loss: 41.953433
Total training time: 0.96 seconds.
-- Epoch 5
Norm: 133.80, NNZs: 2298, Bias: -131.017572, T: 133335, Avg. loss: 24.157069
Total training time: 1.19 seconds.
-- Epoch 6
Norm: 117.94, NNZs: 2298, Bias: -126.362899, T: 160002, Avg. loss: 19.390755
Total training time: 1.42 seconds.
-- Epoch 7
Norm: 107.46, NNZs: 2298, Bias: -122.079139, T: 186669, Avg. loss: 15.338036
Total training time: 1.65 seconds.
-- Epoch 8
Norm: 98.71, NNZs: 2298, Bias: -119.300115, T: 213336, Avg. loss: 12.923650
Total training time: 1.87 seconds.
-- Epoch 9
Norm: 92.52, NNZs: 2298, Bias: -116.931173, T: 240003, Avg. loss: 11.798927
Total training time: 2.10 seconds.
-- Epoch 10
```

```
Norm: 86.36, NNZs: 2298, Bias: -114.655113, T: 266670, Avg. loss: 9.088363
Total training time: 2.33 seconds.
-- Epoch 11
Norm: 82.45, NNZs: 2298, Bias: -112.422039, T: 293337, Avg. loss: 8.389786
Total training time: 2.56 seconds.
-- Epoch 12
Norm: 79.52, NNZs: 2298, Bias: -110.455922, T: 320004, Avg. loss: 7.977381
Total training time: 2.78 seconds.
-- Epoch 13
Norm: 77.19, NNZs: 2298, Bias: -108.677983, T: 346671, Avg. loss: 7.209605
Total training time: 3.00 seconds.
-- Epoch 14
Norm: 74.76, NNZs: 2298, Bias: -106.982833, T: 373338, Avg. loss: 7.415671
Total training time: 3.23 seconds.
-- Epoch 15
Norm: 72.17, NNZs: 2298, Bias: -105.402713, T: 400005, Avg. loss: 5.678267
Total training time: 3.46 seconds.
-- Epoch 16
Norm: 70.44, NNZs: 2298, Bias: -103.696071, T: 426672, Avg. loss: 5.433717
Total training time: 3.68 seconds.
-- Epoch 17
Norm: 68.79, NNZs: 2298, Bias: -102.227030, T: 453339, Avg. loss: 5.201780
Total training time: 3.91 seconds.
-- Epoch 18
Norm: 67.24, NNZs: 2298, Bias: -100.924633, T: 480006, Avg. loss: 5.025373
Total training time: 4.13 seconds.
-- Epoch 19
Norm: 65.79, NNZs: 2298, Bias: -99.775148, T: 506673, Avg. loss: 4.986434
Total training time: 4.35 seconds.
-- Epoch 20
Norm: 63.91, NNZs: 2298, Bias: -98.899633, T: 533340, Avg. loss: 4.350824
Total training time: 4.57 seconds.
-- Epoch 21
Norm: 63.12, NNZs: 2298, Bias: -97.497335, T: 560007, Avg. loss: 4.113335
Total training time: 4.79 seconds.
-- Epoch 22
Norm: 61.79, NNZs: 2298, Bias: -96.578231, T: 586674, Avg. loss: 3.882319
Total training time: 5.01 seconds.
-- Epoch 23
Norm: 60.59, NNZs: 2298, Bias: -95.598605, T: 613341, Avg. loss: 3.923576
Total training time: 5.24 seconds.
-- Epoch 24
Norm: 59.21, NNZs: 2298, Bias: -94.815164, T: 640008, Avg. loss: 3.342339
Total training time: 5.47 seconds.
-- Epoch 25
Norm: 58.05, NNZs: 2298, Bias: -94.098509, T: 666675, Avg. loss: 2.996577
Total training time: 5.69 seconds.
-- Epoch 26
Norm: 57.69, NNZs: 2298, Bias: -92.940348, T: 693342, Avg. loss: 3.263143
Total training time: 5.92 seconds.
-- Epoch 27
Norm: 59.11, NNZs: 2298, Bias: -92.049238, T: 720009, Avg. loss: 4.401987
Total training time: 6.14 seconds.
-- Epoch 28
Norm: 58.75, NNZs: 2298, Bias: -91.012419, T: 746676, Avg. loss: 3.564561
Total training time: 6.36 seconds.
-- Epoch 29
Norm: 57.36, NNZs: 2298, Bias: -90.447634, T: 773343, Avg. loss: 3.105558
Total training time: 6.59 seconds.
-- Epoch 30
Norm: 56.35, NNZs: 2298, Bias: -89.834659, T: 800010, Avg. loss: 2.906695
Total training time: 6.81 seconds.
-- Epoch 31
Norm: 55.89, NNZs: 2298, Bias: -88.978550, T: 826677, Avg. loss: 2.886995
Total training time: 7.04 seconds.
-- Epoch 32
Norm: 54.93, NNZs: 2298, Bias: -88.446802, T: 853344, Avg. loss: 2.566249
Total training time: 7.27 seconds.
-- Epoch 33
Norm: 54.03, NNZs: 2298, Bias: -87.929082, T: 880011, Avg. loss: 2.513563
Total training time: 7.50 seconds.
-- Epoch 34
Norm: 53.54, NNZs: 2298, Bias: -87.288003, T: 906678, Avg. loss: 2.347704
Total training time: 7.73 seconds.
-- Epoch 35
Norm: 53.03, NNZs: 2298, Bias: -86.620467, T: 933345, Avg. loss: 2.323120
Total training time: 7.95 seconds.
```

```
-- Epoch 36
Norm: 52.58, NNZs: 2298, Bias: -85.975434, T: 960012, Avg. loss: 2.352985
Total training time: 8.18 seconds.
-- Epoch 37
Norm: 52.39, NNZs: 2298, Bias: -85.260583, T: 986679, Avg. loss: 2.192825
Total training time: 8.40 seconds.
-- Epoch 38
Norm: 51.86, NNZs: 2298, Bias: -84.815654, T: 1013346, Avg. loss: 2.417977
Total training time: 8.63 seconds.
-- Epoch 39
Norm: 51.67, NNZs: 2298, Bias: -84.144163, T: 1040013, Avg. loss: 2.238391
Total training time: 8.85 seconds.
-- Epoch 40
Norm: 51.30, NNZs: 2298, Bias: -83.596016, T: 1066680, Avg. loss: 2.114802
Total training time: 9.08 seconds.
-- Epoch 41
Norm: 50.70, NNZs: 2298, Bias: -83.133465, T: 1093347, Avg. loss: 2.093421
Total training time: 9.30 seconds.
-- Epoch 42
Norm: 50.25, NNZs: 2298, Bias: -82.635091, T: 1120014, Avg. loss: 1.928454
Total training time: 9.53 seconds.
-- Epoch 43
Norm: 50.17, NNZs: 2298, Bias: -81.970905, T: 1146681, Avg. loss: 1.985341
Total training time: 9.75 seconds.
-- Epoch 44
Norm: 49.88, NNZs: 2298, Bias: -81.464171, T: 1173348, Avg. loss: 2.050347
Total training time: 9.98 seconds.
-- Epoch 45
Norm: 49.36, NNZs: 2298, Bias: -81.058994, T: 1200015, Avg. loss: 1.873039
Total training time: 10.21 seconds.
-- Epoch 46
Norm: 48.86, NNZs: 2298, Bias: -80.725646, T: 1226682, Avg. loss: 1.933512
Total training time: 10.43 seconds.
-- Epoch 47
Norm: 48.56, NNZs: 2298, Bias: -80.242961, T: 1253349, Avg. loss: 1.833787
Total training time: 10.66 seconds.
-- Epoch 48
Norm: 48.04, NNZs: 2298, Bias: -79.906143, T: 1280016, Avg. loss: 1.703311
Total training time: 10.89 seconds.
-- Epoch 49
Norm: 47.38, NNZs: 2298, Bias: -79.676374, T: 1306683, Avg. loss: 1.606546
Total training time: 11.11 seconds.
-- Epoch 50
Norm: 47.38, NNZs: 2298, Bias: -79.101739, T: 1333350, Avg. loss: 1.645374
Total training time: 11.34 seconds.
-- Epoch 51
Norm: 47.05, NNZs: 2298, Bias: -78.717979, T: 1360017, Avg. loss: 1.535257
Total training time: 11.57 seconds.
-- Epoch 52
Norm: 46.79, NNZs: 2298, Bias: -78.302250, T: 1386684, Avg. loss: 1.510209
Total training time: 11.79 seconds.
-- Epoch 53
Norm: 46.30, NNZs: 2298, Bias: -78.033621, T: 1413351, Avg. loss: 1.522067
Total training time: 12.01 seconds.
-- Epoch 54
Norm: 46.07, NNZs: 2298, Bias: -77.654364, T: 1440018, Avg. loss: 1.548980
Total training time: 12.25 seconds.
-- Epoch 55
Norm: 45.83, NNZs: 2298, Bias: -77.279887, T: 1466685, Avg. loss: 1.433994
Total training time: 12.47 seconds.
-- Epoch 56
Norm: 45.75, NNZs: 2298, Bias: -76.857269, T: 1493352, Avg. loss: 1.635983
Total training time: 12.70 seconds.
-- Epoch 57
Norm: 45.32, NNZs: 2298, Bias: -76.588108, T: 1520019, Avg. loss: 1.438139
Total training time: 12.92 seconds.
-- Epoch 58
Norm: 45.06, NNZs: 2298, Bias: -76.276905, T: 1546686, Avg. loss: 1.397127
Total training time: 13.15 seconds.
-- Epoch 59
Norm: 45.11, NNZs: 2298, Bias: -75.789491, T: 1573353, Avg. loss: 1.325396
Total training time: 13.38 seconds.
-- Epoch 60
Norm: 44.83, NNZs: 2298, Bias: -75.481397, T: 1600020, Avg. loss: 1.389144
Total training time: 13.61 seconds.
-- Epoch 61
Norm: 44.61, NNZs: 2298, Bias: -75.173155, T: 1626687, Avg. loss: 1.383206
```

```
Total training time: 13.83 seconds.
-- Epoch 62
Norm: 44.28, NNZs: 2298, Bias: -74.922941, T: 1653354, Avg. loss: 1.314269
Total training time: 14.05 seconds.
-- Epoch 63
Norm: 44.07, NNZs: 2298, Bias: -74.612798, T: 1680021, Avg. loss: 1.245326
Total training time: 14.28 seconds.
-- Epoch 64
Norm: 43.71, NNZs: 2298, Bias: -74.392609, T: 1706688, Avg. loss: 1.262947
Total training time: 14.51 seconds.
-- Epoch 65
Norm: 43.55, NNZs: 2298, Bias: -74.071122, T: 1733355, Avg. loss: 1.248135
Total training time: 14.73 seconds.
-- Epoch 66
Norm: 43.40, NNZs: 2298, Bias: -73.756798, T: 1760022, Avg. loss: 1.196758
Total training time: 14.95 seconds.
-- Epoch 67
Norm: 43.41, NNZs: 2298, Bias: -73.376730, T: 1786689, Avg. loss: 1.252863
Total training time: 15.17 seconds.
-- Epoch 68
Norm: 43.00, NNZs: 2298, Bias: -73.203231, T: 1813356, Avg. loss: 1.220346
Total training time: 15.40 seconds.
-- Epoch 69
Norm: 42.93, NNZs: 2298, Bias: -72.870578, T: 1840023, Avg. loss: 1.211318
Total training time: 15.62 seconds.
-- Epoch 70
Norm: 42.62, NNZs: 2298, Bias: -72.668901, T: 1866690, Avg. loss: 1.158295
Total training time: 15.84 seconds.
-- Epoch 71
Norm: 42.28, NNZs: 2298, Bias: -72.481730, T: 1893357, Avg. loss: 1.105938
Total training time: 16.07 seconds.
-- Epoch 72
Norm: 42.36, NNZs: 2298, Bias: -72.081063, T: 1920024, Avg. loss: 1.092740
Total training time: 16.29 seconds.
-- Epoch 73
Norm: 42.27, NNZs: 2298, Bias: -71.796904, T: 1946691, Avg. loss: 1.110088
Total training time: 16.51 seconds.
-- Epoch 74
Norm: 41.87, NNZs: 2298, Bias: -71.666546, T: 1973358, Avg. loss: 1.091902
Total training time: 16.74 seconds.
-- Epoch 75
Norm: 41.88, NNZs: 2298, Bias: -71.321318, T: 2000025, Avg. loss: 1.103695
Total training time: 16.96 seconds.
-- Epoch 76
Norm: 41.60, NNZs: 2298, Bias: -71.131108, T: 2026692, Avg. loss: 1.053667
Total training time: 17.20 seconds.
-- Epoch 77
Norm: 41.32, NNZs: 2298, Bias: -70.956831, T: 2053359, Avg. loss: 1.071363
Total training time: 17.42 seconds.
-- Epoch 78
Norm: 41.22, NNZs: 2298, Bias: -70.703886, T: 2080026, Avg. loss: 1.044512
Total training time: 17.65 seconds.
-- Epoch 79
Norm: 41.71, NNZs: 2298, Bias: -70.314338, T: 2106693, Avg. loss: 1.308562
Total training time: 17.87 seconds.
-- Epoch 80
Norm: 41.59, NNZs: 2298, Bias: -70.060376, T: 2133360, Avg. loss: 1.229879
Total training time: 18.11 seconds.
-- Epoch 81
Norm: 41.35, NNZs: 2298, Bias: -69.867517, T: 2160027, Avg. loss: 1.115379
Total training time: 18.33 seconds.
-- Epoch 82
Norm: 41.22, NNZs: 2298, Bias: -69.625931, T: 2186694, Avg. loss: 1.076499
Total training time: 18.55 seconds.
-- Epoch 83
Norm: 41.04, NNZs: 2298, Bias: -69.425781, T: 2213361, Avg. loss: 1.078251
Total training time: 18.77 seconds.
Convergence after 83 epochs took 18.77 seconds
-- Epoch 1
Norm: 485.86, NNZs: 2298, Bias: -187.384250, T: 26667, Avg. loss: 755.540296
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 266.74, NNZs: 2298, Bias: -187.826792, T: 53334, Avg. loss: 119.374447
Total training time: 0.46 seconds.
-- Epoch 3
Norm: 191.35, NNZs: 2298, Bias: -179.399269, T: 80001, Avg. loss: 47.462182
Total training time: 0.71 seconds.
```

```
-- Epoch 4
Norm: 239.51, NNZs: 2298, Bias: -176.348813, T: 106668, Avg. loss: 79.088338
Total training time: 0.95 seconds.
-- Epoch 5
Norm: 186.40, NNZs: 2298, Bias: -170.992968, T: 133335, Avg. loss: 40.807251
Total training time: 1.18 seconds.
-- Epoch 6
Norm: 158.51, NNZs: 2298, Bias: -165.699507, T: 160002, Avg. loss: 23.963472
Total training time: 1.42 seconds.
-- Epoch 7
Norm: 140.98, NNZs: 2298, Bias: -160.676883, T: 186669, Avg. loss: 16.031790
Total training time: 1.64 seconds.
-- Epoch 8
Norm: 126.81, NNZs: 2298, Bias: -156.524090, T: 213336, Avg. loss: 14.128189
Total training time: 1.87 seconds.
-- Epoch 9
Norm: 118.22, NNZs: 2298, Bias: -152.311141, T: 240003, Avg. loss: 11.591269
Total training time: 2.11 seconds.
-- Epoch 10
Norm: 111.55, NNZs: 2298, Bias: -148.380580, T: 266670, Avg. loss: 11.893633
Total training time: 2.34 seconds.
-- Epoch 11
Norm: 104.44, NNZs: 2298, Bias: -145.810493, T: 293337, Avg. loss: 9.452046
Total training time: 2.58 seconds.
-- Epoch 12
Norm: 99.28, NNZs: 2298, Bias: -143.201312, T: 320004, Avg. loss: 8.537355
Total training time: 2.81 seconds.
-- Epoch 13
Norm: 93.98, NNZs: 2298, Bias: -141.229110, T: 346671, Avg. loss: 7.716702
Total training time: 3.04 seconds.
-- Epoch 14
Norm: 92.71, NNZs: 2298, Bias: -138.098633, T: 373338, Avg. loss: 6.919069
Total training time: 3.29 seconds.
-- Epoch 15
Norm: 88.71, NNZs: 2298, Bias: -136.240531, T: 400005, Avg. loss: 6.758827
Total training time: 3.52 seconds.
-- Epoch 16
Norm: 85.66, NNZs: 2298, Bias: -134.334755, T: 426672, Avg. loss: 6.026631
Total training time: 3.76 seconds.
-- Epoch 17
Norm: 83.12, NNZs: 2298, Bias: -132.622536, T: 453339, Avg. loss: 5.285451
Total training time: 3.99 seconds.
-- Epoch 18
Norm: 82.34, NNZs: 2298, Bias: -130.436256, T: 480006, Avg. loss: 5.772021
Total training time: 4.23 seconds.
-- Epoch 19
Norm: 80.71, NNZs: 2298, Bias: -128.879877, T: 506673, Avg. loss: 6.193950
Total training time: 4.47 seconds.
-- Epoch 20
Norm: 78.67, NNZs: 2298, Bias: -127.314578, T: 533340, Avg. loss: 4.991367
Total training time: 4.70 seconds.
-- Epoch 21
Norm: 76.83, NNZs: 2298, Bias: -125.981271, T: 560007, Avg. loss: 4.929378
Total training time: 4.94 seconds.
-- Epoch 22
Norm: 75.41, NNZs: 2298, Bias: -124.685535, T: 586674, Avg. loss: 4.516634
Total training time: 5.17 seconds.
-- Epoch 23
Norm: 74.40, NNZs: 2298, Bias: -123.169803, T: 613341, Avg. loss: 4.559798
Total training time: 5.41 seconds.
-- Epoch 24
Norm: 73.06, NNZs: 2298, Bias: -121.988710, T: 640008, Avg. loss: 4.350640
Total training time: 5.64 seconds.
-- Epoch 25
Norm: 71.68, NNZs: 2298, Bias: -120.915564, T: 666675, Avg. loss: 4.043247
Total training time: 5.88 seconds.
-- Epoch 26
Norm: 70.70, NNZs: 2298, Bias: -119.774996, T: 693342, Avg. loss: 3.783923
Total training time: 6.11 seconds.
-- Epoch 27
Norm: 69.67, NNZs: 2298, Bias: -118.701668, T: 720009, Avg. loss: 3.596770
Total training time: 6.34 seconds.
-- Epoch 28
Norm: 68.44, NNZs: 2298, Bias: -117.798035, T: 746676, Avg. loss: 3.559352
Total training time: 6.57 seconds.
-- Epoch 29
Norm: 67.88, NNZs: 2298, Bias: -116.670256, T: 773343, Avg. loss: 3.905212
```

```
Total training time: 6.80 seconds.
-- Epoch 30
Norm: 67.60, NNZs: 2298, Bias: -115.544329, T: 800010, Avg. loss: 3.340712
Total training time: 7.03 seconds.
-- Epoch 31
Norm: 66.90, NNZs: 2298, Bias: -114.571376, T: 826677, Avg. loss: 3.183339
Total training time: 7.27 seconds.
-- Epoch 32
Norm: 67.25, NNZs: 2298, Bias: -113.285083, T: 853344, Avg. loss: 3.292368
Total training time: 7.50 seconds.
-- Epoch 33
Norm: 66.01, NNZs: 2298, Bias: -112.682646, T: 880011, Avg. loss: 3.029776
Total training time: 7.73 seconds.
-- Epoch 34
Norm: 65.60, NNZs: 2298, Bias: -111.633844, T: 906678, Avg. loss: 3.073258
Total training time: 7.96 seconds.
-- Epoch 35
Norm: 64.77, NNZs: 2298, Bias: -110.912613, T: 933345, Avg. loss: 2.793143
Total training time: 8.20 seconds.
-- Epoch 36
Norm: 64.30, NNZs: 2298, Bias: -110.115615, T: 960012, Avg. loss: 2.844974
Total training time: 8.43 seconds.
-- Epoch 37
Norm: 63.89, NNZs: 2298, Bias: -109.293710, T: 986679, Avg. loss: 2.940548
Total training time: 8.66 seconds.
-- Epoch 38
Norm: 62.82, NNZs: 2298, Bias: -108.773946, T: 1013346, Avg. loss: 2.612208
Total training time: 8.89 seconds.
-- Epoch 39
Norm: 62.43, NNZs: 2298, Bias: -107.966884, T: 1040013, Avg. loss: 2.495321
Total training time: 9.12 seconds.
-- Epoch 40
Norm: 61.77, NNZs: 2298, Bias: -107.361630, T: 1066680, Avg. loss: 2.582807
Total training time: 9.35 seconds.
-- Epoch 41
Norm: 61.24, NNZs: 2298, Bias: -106.722239, T: 1093347, Avg. loss: 2.324954
Total training time: 9.58 seconds.
-- Epoch 42
Norm: 60.82, NNZs: 2298, Bias: -106.048029, T: 1120014, Avg. loss: 2.465254
Total training time: 9.80 seconds.
-- Epoch 43
Norm: 60.77, NNZs: 2298, Bias: -105.204797, T: 1146681, Avg. loss: 2.297762
Total training time: 10.02 seconds.
-- Epoch 44
Norm: 60.25, NNZs: 2298, Bias: -104.645898, T: 1173348, Avg. loss: 2.262435
Total training time: 10.25 seconds.
-- Epoch 45
Norm: 59.91, NNZs: 2298, Bias: -104.007902, T: 1200015, Avg. loss: 2.299720
Total training time: 10.48 seconds.
-- Epoch 46
Norm: 59.27, NNZs: 2298, Bias: -103.539229, T: 1226682, Avg. loss: 2.311615
Total training time: 10.70 seconds.
-- Epoch 47
Norm: 59.06, NNZs: 2298, Bias: -102.867249, T: 1253349, Avg. loss: 2.258992
Total training time: 10.93 seconds.
-- Epoch 48
Norm: 58.59, NNZs: 2298, Bias: -102.400150, T: 1280016, Avg. loss: 2.235786
Total training time: 11.15 seconds.
-- Epoch 49
Norm: 58.48, NNZs: 2298, Bias: -101.716819, T: 1306683, Avg. loss: 2.068447
Total training time: 11.38 seconds.
-- Epoch 50
Norm: 58.06, NNZs: 2298, Bias: -101.229213, T: 1333350, Avg. loss: 2.123350
Total training time: 11.60 seconds.
-- Epoch 51
Norm: 57.46, NNZs: 2298, Bias: -100.821561, T: 1360017, Avg. loss: 1.981628
Total training time: 11.83 seconds.
-- Epoch 52
Norm: 57.18, NNZs: 2298, Bias: -100.287603, T: 1386684, Avg. loss: 2.001950
Total training time: 12.06 seconds.
-- Epoch 53
Norm: 56.95, NNZs: 2298, Bias: -99.742087, T: 1413351, Avg. loss: 1.938240
Total training time: 12.28 seconds.
-- Epoch 54
Norm: 56.87, NNZs: 2298, Bias: -99.154383, T: 1440018, Avg. loss: 1.896842
Total training time: 12.50 seconds.
-- Epoch 55
```

```
Norm: 56.36, NNZs: 2298, Bias: -98.789188, T: 1466685, Avg. loss: 1.871890
Total training time: 12.73 seconds.
-- Epoch 56
Norm: 56.13, NNZs: 2298, Bias: -98.286595, T: 1493352, Avg. loss: 1.837604
Total training time: 12.95 seconds.
-- Epoch 57
Norm: 56.14, NNZs: 2298, Bias: -97.671389, T: 1520019, Avg. loss: 1.723669
Total training time: 13.18 seconds.
-- Epoch 58
Norm: 55.48, NNZs: 2298, Bias: -97.423051, T: 1546686, Avg. loss: 1.778431
Total training time: 13.40 seconds.
-- Epoch 59
Norm: 55.57, NNZs: 2298, Bias: -96.783017, T: 1573353, Avg. loss: 1.694274
Total training time: 13.63 seconds.
-- Epoch 60
Norm: 55.35, NNZs: 2298, Bias: -96.312706, T: 1600020, Avg. loss: 1.730917
Total training time: 13.85 seconds.
-- Epoch 61
Norm: 54.88, NNZs: 2298, Bias: -95.990819, T: 1626687, Avg. loss: 1.678342
Total training time: 14.07 seconds.
-- Epoch 62
Norm: 54.51, NNZs: 2298, Bias: -95.639093, T: 1653354, Avg. loss: 1.588548
Total training time: 14.30 seconds.
-- Epoch 63
Norm: 54.10, NNZs: 2298, Bias: -95.330448, T: 1680021, Avg. loss: 1.474003
Total training time: 14.52 seconds.
-- Epoch 64
Norm: 53.85, NNZs: 2298, Bias: -94.952049, T: 1706688, Avg. loss: 1.570010
Total training time: 14.75 seconds.
-- Epoch 65
Norm: 54.61, NNZs: 2298, Bias: -94.238418, T: 1733355, Avg. loss: 1.946738
Total training time: 14.97 seconds.
-- Epoch 66
Norm: 54.33, NNZs: 2298, Bias: -93.853734, T: 1760022, Avg. loss: 1.709193
Total training time: 15.20 seconds.
-- Epoch 67
Norm: 53.90, NNZs: 2298, Bias: -93.577178, T: 1786689, Avg. loss: 1.675389
Total training time: 15.42 seconds.
-- Epoch 68
Norm: 53.72, NNZs: 2298, Bias: -93.171163, T: 1813356, Avg. loss: 1.537486
Total training time: 15.64 seconds.
Convergence after 68 epochs took 15.64 seconds
-- Epoch 1
Norm: 88.68, NNZs: 2297, Bias: -8.689834, T: 26666, Avg. loss: 125.818640
Total training time: 0.23 seconds.
-- Epoch 2
Norm: 41.76, NNZs: 2297, Bias: -11.084262, T: 53332, Avg. loss: 17.988145
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 27.06, NNZs: 2297, Bias: -11.156199, T: 79998, Avg. loss: 6.282410
Total training time: 0.68 seconds.
-- Epoch 4
Norm: 20.73, NNZs: 2297, Bias: -10.981878, T: 106664, Avg. loss: 3.582898
Total training time: 0.90 seconds.
-- Epoch 5
Norm: 16.95, NNZs: 2297, Bias: -10.583729, T: 133330, Avg. loss: 2.675325
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 14.13, NNZs: 2297, Bias: -10.460542, T: 159996, Avg. loss: 1.833229
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 12.53, NNZs: 2297, Bias: -10.075524, T: 186662, Avg. loss: 1.476967
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 11.08, NNZs: 2297, Bias: -9.914690, T: 213328, Avg. loss: 1.110729
Total training time: 1.78 seconds.
-- Epoch 9
Norm: 10.08, NNZs: 2297, Bias: -9.711357, T: 239994, Avg. loss: 1.080080
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 9.32, NNZs: 2297, Bias: -9.514499, T: 266660, Avg. loss: 0.908337
Total training time: 2.24 seconds.
-- Epoch 11
Norm: 8.75, NNZs: 2297, Bias: -9.298854, T: 293326, Avg. loss: 0.861901
Total training time: 2.47 seconds.
-- Epoch 12
Norm: 8.29, NNZs: 2297, Bias: -9.148993, T: 319992, Avg. loss: 0.718047
```

```
Total training time: 2.69 seconds.
-- Epoch 13
Norm: 7.72, NNZs: 2297, Bias: -9.061514, T: 346658, Avg. loss: 0.698231
Total training time: 2.92 seconds.
-- Epoch 14
Norm: 7.42, NNZs: 2297, Bias: -8.933326, T: 373324, Avg. loss: 0.621068
Total training time: 3.14 seconds.
-- Epoch 15
Norm: 7.12, NNZs: 2297, Bias: -8.785762, T: 399990, Avg. loss: 0.636444
Total training time: 3.37 seconds.
-- Epoch 16
Norm: 6.84, NNZs: 2297, Bias: -8.684282, T: 426656, Avg. loss: 0.576303
Total training time: 3.59 seconds.
-- Epoch 17
Norm: 6.62, NNZs: 2297, Bias: -8.582133, T: 453322, Avg. loss: 0.541998
Total training time: 3.81 seconds.
-- Epoch 18
Norm: 6.40, NNZs: 2297, Bias: -8.498289, T: 479988, Avg. loss: 0.529638
Total training time: 4.03 seconds.
-- Epoch 19
Norm: 6.19, NNZs: 2297, Bias: -8.418787, T: 506654, Avg. loss: 0.485687
Total training time: 4.25 seconds.
-- Epoch 20
Norm: 6.04, NNZs: 2297, Bias: -8.328299, T: 533320, Avg. loss: 0.518389
Total training time: 4.48 seconds.
-- Epoch 21
Norm: 5.91, NNZs: 2297, Bias: -8.243792, T: 559986, Avg. loss: 0.471454
Total training time: 4.70 seconds.
-- Epoch 22
Norm: 5.77, NNZs: 2297, Bias: -8.170955, T: 586652, Avg. loss: 0.452414
Total training time: 4.93 seconds.
-- Epoch 23
Norm: 5.70, NNZs: 2297, Bias: -8.076428, T: 613318, Avg. loss: 0.440290
Total training time: 5.15 seconds.
-- Epoch 24
Norm: 5.55, NNZs: 2297, Bias: -8.026504, T: 639984, Avg. loss: 0.458448
Total training time: 5.38 seconds.
-- Epoch 25
Norm: 5.44, NNZs: 2297, Bias: -7.971405, T: 666650, Avg. loss: 0.424060
Total training time: 5.60 seconds.
-- Epoch 26
Norm: 5.39, NNZs: 2297, Bias: -7.895707, T: 693316, Avg. loss: 0.424777
Total training time: 5.83 seconds.
-- Epoch 27
Norm: 5.29, NNZs: 2297, Bias: -7.842854, T: 719982, Avg. loss: 0.429772
Total training time: 6.05 seconds.
-- Epoch 28
Norm: 5.24, NNZs: 2297, Bias: -7.775695, T: 746648, Avg. loss: 0.417200
Total training time: 6.27 seconds.
-- Epoch 29
Norm: 5.17, NNZs: 2297, Bias: -7.716634, T: 773314, Avg. loss: 0.412685
Total training time: 6.49 seconds.
-- Epoch 30
Norm: 5.09, NNZs: 2297, Bias: -7.677583, T: 799980, Avg. loss: 0.415503
Total training time: 6.72 seconds.
-- Epoch 31
Norm: 5.03, NNZs: 2297, Bias: -7.628027, T: 826646, Avg. loss: 0.393914
Total training time: 6.95 seconds.
-- Epoch 32
Norm: 4.95, NNZs: 2297, Bias: -7.595016, T: 853312, Avg. loss: 0.411459
Total training time: 7.17 seconds.
-- Epoch 33
Norm: 4.91, NNZs: 2297, Bias: -7.546385, T: 879978, Avg. loss: 0.381791
Total training time: 7.39 seconds.
-- Epoch 34
Norm: 4.87, NNZs: 2297, Bias: -7.497507, T: 906644, Avg. loss: 0.378584
Total training time: 7.62 seconds.
-- Epoch 35
Norm: 4.82, NNZs: 2297, Bias: -7.458573, T: 933310, Avg. loss: 0.396117
Total training time: 7.84 seconds.
-- Epoch 36
Norm: 4.78, NNZs: 2297, Bias: -7.415083, T: 959976, Avg. loss: 0.380070
Total training time: 8.06 seconds.
-- Epoch 37
Norm: 4.73, NNZs: 2297, Bias: -7.382058, T: 986642, Avg. loss: 0.384982
Total training time: 8.28 seconds.
-- Epoch 38
```

```
Norm: 4.70, NNZs: 2297, Bias: -7.340582, T: 1013308, Avg. loss: 0.369407
Total training time: 8.51 seconds.
-- Epoch 39
Norm: 4.67, NNZs: 2297, Bias: -7.300581, T: 1039974, Avg. loss: 0.378306
Total training time: 8.73 seconds.
-- Epoch 40
Norm: 4.63, NNZs: 2297, Bias: -7.271690, T: 1066640, Avg. loss: 0.377871
Total training time: 8.95 seconds.
-- Epoch 41
Norm: 4.58, NNZs: 2297, Bias: -7.244908, T: 1093306, Avg. loss: 0.377805
Total training time: 9.17 seconds.
-- Epoch 42
Norm: 4.55, NNZs: 2297, Bias: -7.214956, T: 1119972, Avg. loss: 0.372375
Total training time: 9.40 seconds.
-- Epoch 43
Norm: 4.54, NNZs: 2297, Bias: -7.174122, T: 1146638, Avg. loss: 0.371094
Total training time: 9.62 seconds.
Convergence after 43 epochs took 9.62 seconds
-- Epoch 1
Norm: 26.84, NNZs: 2298, Bias: -2.858463, T: 26667, Avg. loss: 71.576683
Total training time: 0.23 seconds.
-- Epoch 2
Norm: 25.47, NNZs: 2298, Bias: -3.925198, T: 53334, Avg. loss: 13.464303
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 21.50, NNZs: 2298, Bias: -4.380512, T: 80001, Avg. loss: 7.020807
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 15.55, NNZs: 2298, Bias: -4.788150, T: 106668, Avg. loss: 3.946338
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 12.72, NNZs: 2298, Bias: -4.871353, T: 133335, Avg. loss: 2.429735
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 10.88, NNZs: 2298, Bias: -4.815068, T: 160002, Avg. loss: 1.860971
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 9.55, NNZs: 2298, Bias: -4.797497, T: 186669, Avg. loss: 1.403363
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 8.50, NNZs: 2298, Bias: -4.758582, T: 213336, Avg. loss: 1.154656
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 7.72, NNZs: 2298, Bias: -4.752561, T: 240003, Avg. loss: 1.013204
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 7.16, NNZs: 2298, Bias: -4.684991, T: 266670, Avg. loss: 0.859506
Total training time: 2.24 seconds.
-- Epoch 11
Norm: 6.69, NNZs: 2298, Bias: -4.604328, T: 293337, Avg. loss: 0.789921
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 6.34, NNZs: 2298, Bias: -4.547923, T: 320004, Avg. loss: 0.738123
Total training time: 2.68 seconds.
-- Epoch 13
Norm: 6.01, NNZs: 2298, Bias: -4.506521, T: 346671, Avg. loss: 0.696306
Total training time: 2.90 seconds.
-- Epoch 14
Norm: 5.78, NNZs: 2298, Bias: -4.436604, T: 373338, Avg. loss: 0.612642
Total training time: 3.12 seconds.
-- Epoch 15
Norm: 5.50, NNZs: 2298, Bias: -4.411259, T: 400005, Avg. loss: 0.610908
Total training time: 3.35 seconds.
-- Epoch 16
Norm: 5.29, NNZs: 2298, Bias: -4.377479, T: 426672, Avg. loss: 0.568999
Total training time: 3.58 seconds.
-- Epoch 17
Norm: 5.12, NNZs: 2298, Bias: -4.328009, T: 453339, Avg. loss: 0.541539
Total training time: 3.81 seconds.
-- Epoch 18
Norm: 4.98, NNZs: 2298, Bias: -4.290422, T: 480006, Avg. loss: 0.532779
Total training time: 4.03 seconds.
-- Epoch 19
Norm: 4.84, NNZs: 2298, Bias: -4.262939, T: 506673, Avg. loss: 0.508493
Total training time: 4.26 seconds.
-- Epoch 20
Norm: 4.67, NNZs: 2298, Bias: -4.249462, T: 533340, Avg. loss: 0.475933
```

```
Total training time: 4.48 seconds.
-- Epoch 21
Norm: 4.59, NNZs: 2298, Bias: -4.196252, T: 560007, Avg. loss: 0.467488
Total training time: 4.70 seconds.
-- Epoch 22
Norm: 4.49, NNZs: 2298, Bias: -4.177392, T: 586674, Avg. loss: 0.464193
Total training time: 4.91 seconds.
-- Epoch 23
Norm: 4.39, NNZs: 2298, Bias: -4.154541, T: 613341, Avg. loss: 0.453380
Total training time: 5.13 seconds.
-- Epoch 24
Norm: 4.31, NNZs: 2298, Bias: -4.134670, T: 640008, Avg. loss: 0.437144
Total training time: 5.36 seconds.
-- Epoch 25
Norm: 4.23, NNZs: 2298, Bias: -4.121361, T: 666675, Avg. loss: 0.433196
Total training time: 5.58 seconds.
-- Epoch 26
Norm: 4.17, NNZs: 2298, Bias: -4.097338, T: 693342, Avg. loss: 0.438371
Total training time: 5.80 seconds.
-- Epoch 27
Norm: 4.09, NNZs: 2298, Bias: -4.077878, T: 720009, Avg. loss: 0.431679
Total training time: 6.02 seconds.
-- Epoch 28
Norm: 4.04, NNZs: 2298, Bias: -4.054009, T: 746676, Avg. loss: 0.416638
Total training time: 6.25 seconds.
-- Epoch 29
Norm: 3.97, NNZs: 2298, Bias: -4.041287, T: 773343, Avg. loss: 0.411721
Total training time: 6.47 seconds.
-- Epoch 30
Norm: 3.92, NNZs: 2298, Bias: -4.028665, T: 800010, Avg. loss: 0.406935
Total training time: 6.69 seconds.
-- Epoch 31
Norm: 3.88, NNZs: 2298, Bias: -4.003654, T: 826677, Avg. loss: 0.409086
Total training time: 6.92 seconds.
-- Epoch 32
Norm: 3.84, NNZs: 2298, Bias: -3.984305, T: 853344, Avg. loss: 0.395894
Total training time: 7.15 seconds.
-- Epoch 33
Norm: 3.78, NNZs: 2298, Bias: -3.983553, T: 880011, Avg. loss: 0.405542
Total training time: 7.37 seconds.
-- Epoch 34
Norm: 3.77, NNZs: 2298, Bias: -3.957683, T: 906678, Avg. loss: 0.393793
Total training time: 7.60 seconds.
-- Epoch 35
Norm: 3.73, NNZs: 2298, Bias: -3.941056, T: 933345, Avg. loss: 0.399455
Total training time: 7.82 seconds.
-- Epoch 36
Norm: 3.70, NNZs: 2298, Bias: -3.925116, T: 960012, Avg. loss: 0.396258
Total training time: 8.04 seconds.
-- Epoch 37
Norm: 3.68, NNZs: 2298, Bias: -3.907521, T: 986679, Avg. loss: 0.390415
Total training time: 8.26 seconds.
-- Epoch 38
Norm: 3.64, NNZs: 2298, Bias: -3.900634, T: 1013346, Avg. loss: 0.390723
Total training time: 8.48 seconds.
-- Epoch 39
Norm: 3.61, NNZs: 2298, Bias: -3.890639, T: 1040013, Avg. loss: 0.390952
Total training time: 8.70 seconds.
-- Epoch 40
Norm: 3.59, NNZs: 2298, Bias: -3.879042, T: 1066680, Avg. loss: 0.385563
Total training time: 8.92 seconds.
-- Epoch 41
Norm: 3.57, NNZs: 2298, Bias: -3.869614, T: 1093347, Avg. loss: 0.386787
Total training time: 9.14 seconds.
-- Epoch 42
Norm: 3.55, NNZs: 2298, Bias: -3.855576, T: 1120014, Avg. loss: 0.381066
Total training time: 9.37 seconds.
-- Epoch 43
Norm: 3.54, NNZs: 2298, Bias: -3.833073, T: 1146681, Avg. loss: 0.380096
Total training time: 9.59 seconds.
-- Epoch 44
Norm: 3.52, NNZs: 2298, Bias: -3.826492, T: 1173348, Avg. loss: 0.386751
Total training time: 9.81 seconds.
-- Epoch 45
Norm: 3.50, NNZs: 2298, Bias: -3.816749, T: 1200015, Avg. loss: 0.382062
Total training time: 10.03 seconds.
-- Epoch 46
```

```
-- Epoch 46
Norm: 3.48, NNZs: 2298, Bias: -3.809313, T: 1226682, Avg. loss: 0.383858
Total training time: 10.26 seconds.
-- Epoch 47
Norm: 3.47, NNZs: 2298, Bias: -3.794651, T: 1253349, Avg. loss: 0.378539
Total training time: 10.49 seconds.
-- Epoch 48
Norm: 3.45, NNZs: 2298, Bias: -3.789573, T: 1280016, Avg. loss: 0.376314
Total training time: 10.71 seconds.
-- Epoch 49
Norm: 3.43, NNZs: 2298, Bias: -3.780385, T: 1306683, Avg. loss: 0.377379
Total training time: 10.94 seconds.
-- Epoch 50
Norm: 3.43, NNZs: 2298, Bias: -3.766590, T: 1333350, Avg. loss: 0.376579
Total training time: 11.17 seconds.
-- Epoch 51
Norm: 3.41, NNZs: 2298, Bias: -3.758764, T: 1360017, Avg. loss: 0.375485
Total training time: 11.39 seconds.
-- Epoch 52
Norm: 3.40, NNZs: 2298, Bias: -3.751713, T: 1386684, Avg. loss: 0.371151
Total training time: 11.62 seconds.
-- Epoch 53
Norm: 3.38, NNZs: 2298, Bias: -3.746375, T: 1413351, Avg. loss: 0.378721
Total training time: 11.84 seconds.
-- Epoch 54
Norm: 3.36, NNZs: 2298, Bias: -3.741597, T: 1440018, Avg. loss: 0.370236
Total training time: 12.07 seconds.
-- Epoch 55
Norm: 3.36, NNZs: 2298, Bias: -3.728850, T: 1466685, Avg. loss: 0.366576
Total training time: 12.30 seconds.
-- Epoch 56
Norm: 3.35, NNZs: 2298, Bias: -3.721086, T: 1493352, Avg. loss: 0.372320
Total training time: 12.54 seconds.
-- Epoch 57
Norm: 3.33, NNZs: 2298, Bias: -3.715356, T: 1520019, Avg. loss: 0.373263
Total training time: 12.76 seconds.
-- Epoch 58
Norm: 3.32, NNZs: 2298, Bias: -3.709024, T: 1546686, Avg. loss: 0.371765
Total training time: 12.98 seconds.
-- Epoch 59
Norm: 3.32, NNZs: 2298, Bias: -3.696964, T: 1573353, Avg. loss: 0.363681
Total training time: 13.20 seconds.
-- Epoch 60
Norm: 3.30, NNZs: 2298, Bias: -3.694615, T: 1600020, Avg. loss: 0.373470
Total training time: 13.42 seconds.
-- Epoch 61
Norm: 3.30, NNZs: 2298, Bias: -3.685850, T: 1626687, Avg. loss: 0.367301
Total training time: 13.65 seconds.
-- Epoch 62
Norm: 3.29, NNZs: 2298, Bias: -3.680852, T: 1653354, Avg. loss: 0.368254
Total training time: 13.87 seconds.
-- Epoch 63
Norm: 3.28, NNZs: 2298, Bias: -3.676421, T: 1680021, Avg. loss: 0.364959
Total training time: 14.10 seconds.
-- Epoch 64
Norm: 3.26, NNZs: 2298, Bias: -3.674888, T: 1706688, Avg. loss: 0.369647
Total training time: 14.32 seconds.
Convergence after 64 epochs took 14.32 seconds
-- Epoch 1
Norm: 48.47, NNZs: 2298, Bias: -16.446306, T: 26667, Avg. loss: 92.578510
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 25.43, NNZs: 2298, Bias: -16.459845, T: 53334, Avg. loss: 10.244275
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 17.99, NNZs: 2298, Bias: -15.694862, T: 80001, Avg. loss: 4.874868
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 23.58, NNZs: 2298, Bias: -15.527670, T: 106668, Avg. loss: 7.576441
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 18.74, NNZs: 2298, Bias: -15.095505, T: 133335, Avg. loss: 4.539552
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 15.57, NNZs: 2298, Bias: -14.691133, T: 160002, Avg. loss: 2.527538
Total training time: 1.33 seconds.
-- Epoch 7
Norm: 13.68, NNZs: 2298, Bias: -14.288818, T: 186669, Avg. loss: 1.541289
```

```
Norm: 13.00, NNZs: 2298, Bias: -14.288618, T: 186669, Avg. loss: 1.541269
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 12.30, NNZs: 2298, Bias: -13.909188, T: 213336, Avg. loss: 1.412794
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 11.33, NNZs: 2298, Bias: -13.556494, T: 240003, Avg. loss: 1.143538
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 10.51, NNZs: 2298, Bias: -13.252040, T: 266670, Avg. loss: 0.988105
Total training time: 2.23 seconds.
-- Epoch 11
Norm: 9.93, NNZs: 2298, Bias: -12.954588, T: 293337, Avg. loss: 0.957392
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 9.38, NNZs: 2298, Bias: -12.748033, T: 320004, Avg. loss: 0.815841
Total training time: 2.68 seconds.
-- Epoch 13
Norm: 8.87, NNZs: 2298, Bias: -12.573621, T: 346671, Avg. loss: 0.762634
Total training time: 2.91 seconds.
-- Epoch 14
Norm: 8.63, NNZs: 2298, Bias: -12.328868, T: 373338, Avg. loss: 0.683384
Total training time: 3.13 seconds.
-- Epoch 15
Norm: 8.34, NNZs: 2298, Bias: -12.127913, T: 400005, Avg. loss: 0.693543
Total training time: 3.35 seconds.
-- Epoch 16
Norm: 8.07, NNZs: 2298, Bias: -11.962274, T: 426672, Avg. loss: 0.668388
Total training time: 3.58 seconds.
-- Epoch 17
Norm: 7.85, NNZs: 2298, Bias: -11.800713, T: 453339, Avg. loss: 0.624303
Total training time: 3.80 seconds.
-- Epoch 18
Norm: 7.61, NNZs: 2298, Bias: -11.664494, T: 480006, Avg. loss: 0.589080
Total training time: 4.02 seconds.
-- Epoch 19
Norm: 7.43, NNZs: 2298, Bias: -11.524568, T: 506673, Avg. loss: 0.549414
Total training time: 4.24 seconds.
-- Epoch 20
Norm: 7.26, NNZs: 2298, Bias: -11.387703, T: 533340, Avg. loss: 0.523922
Total training time: 4.47 seconds.
-- Epoch 21
Norm: 7.07, NNZs: 2298, Bias: -11.285017, T: 560007, Avg. loss: 0.508519
Total training time: 4.69 seconds.
-- Epoch 22
Norm: 6.98, NNZs: 2298, Bias: -11.158259, T: 586674, Avg. loss: 0.507829
Total training time: 4.91 seconds.
-- Epoch 23
Norm: 6.84, NNZs: 2298, Bias: -11.052037, T: 613341, Avg. loss: 0.502731
Total training time: 5.14 seconds.
-- Epoch 24
Norm: 6.73, NNZs: 2298, Bias: -10.948910, T: 640008, Avg. loss: 0.471653
Total training time: 5.36 seconds.
-- Epoch 25
Norm: 6.61, NNZs: 2298, Bias: -10.856320, T: 666675, Avg. loss: 0.468068
Total training time: 5.59 seconds.
-- Epoch 26
Norm: 6.50, NNZs: 2298, Bias: -10.771485, T: 693342, Avg. loss: 0.447582
Total training time: 5.81 seconds.
-- Epoch 27
Norm: 6.45, NNZs: 2298, Bias: -10.666459, T: 720009, Avg. loss: 0.444910
Total training time: 6.03 seconds.
-- Epoch 28
Norm: 6.35, NNZs: 2298, Bias: -10.588849, T: 746676, Avg. loss: 0.454362
Total training time: 6.26 seconds.
-- Epoch 29
Norm: 6.26, NNZs: 2298, Bias: -10.512327, T: 773343, Avg. loss: 0.442503
Total training time: 6.48 seconds.
-- Epoch 30
Norm: 6.19, NNZs: 2298, Bias: -10.433482, T: 800010, Avg. loss: 0.422577
Total training time: 6.71 seconds.
-- Epoch 31
Norm: 6.14, NNZs: 2298, Bias: -10.352280, T: 826677, Avg. loss: 0.418118
Total training time: 6.93 seconds.
-- Epoch 32
Norm: 6.11, NNZs: 2298, Bias: -10.264118, T: 853344, Avg. loss: 0.398506
Total training time: 7.15 seconds.
-- Epoch 33
```

```
-- Epoch 33
Norm: 6.02, NNZs: 2298, Bias: -10.211200, T: 880011, Avg. loss: 0.417448
Total training time: 7.37 seconds.
-- Epoch 34
Norm: 5.98, NNZs: 2298, Bias: -10.136125, T: 906678, Avg. loss: 0.404744
Total training time: 7.60 seconds.
-- Epoch 35
Norm: 5.93, NNZs: 2298, Bias: -10.068485, T: 933345, Avg. loss: 0.410400
Total training time: 7.83 seconds.
-- Epoch 36
Norm: 5.89, NNZs: 2298, Bias: -10.004750, T: 960012, Avg. loss: 0.395344
Total training time: 8.06 seconds.
-- Epoch 37
Norm: 5.82, NNZs: 2298, Bias: -9.954182, T: 986679, Avg. loss: 0.395808
Total training time: 8.28 seconds.
-- Epoch 38
Norm: 5.78, NNZs: 2298, Bias: -9.896042, T: 1013346, Avg. loss: 0.392977
Total training time: 8.50 seconds.
-- Epoch 39
Norm: 5.80, NNZs: 2298, Bias: -9.832128, T: 1040013, Avg. loss: 0.412602
Total training time: 8.73 seconds.
-- Epoch 40
Norm: 5.77, NNZs: 2298, Bias: -9.770480, T: 1066680, Avg. loss: 0.391019
Total training time: 8.95 seconds.
-- Epoch 41
Norm: 5.72, NNZs: 2298, Bias: -9.721230, T: 1093347, Avg. loss: 0.396955
Total training time: 9.17 seconds.
-- Epoch 42
Norm: 5.67, NNZs: 2298, Bias: -9.672539, T: 1120014, Avg. loss: 0.387350
Total training time: 9.40 seconds.
-- Epoch 43
Norm: 5.65, NNZs: 2298, Bias: -9.614550, T: 1146681, Avg. loss: 0.381432
Total training time: 9.62 seconds.
-- Epoch 44
Norm: 5.61, NNZs: 2298, Bias: -9.570468, T: 1173348, Avg. loss: 0.383831
Total training time: 9.85 seconds.
-- Epoch 45
Norm: 5.57, NNZs: 2298, Bias: -9.524447, T: 1200015, Avg. loss: 0.379120
Total training time: 10.07 seconds.
-- Epoch 46
Norm: 5.52, NNZs: 2298, Bias: -9.487187, T: 1226682, Avg. loss: 0.382880
Total training time: 10.29 seconds.
-- Epoch 47
Norm: 5.50, NNZs: 2298, Bias: -9.436096, T: 1253349, Avg. loss: 0.369654
Total training time: 10.52 seconds.
-- Epoch 48
Norm: 5.46, NNZs: 2298, Bias: -9.399064, T: 1280016, Avg. loss: 0.374910
Total training time: 10.75 seconds.
-- Epoch 49
Norm: 5.43, NNZs: 2298, Bias: -9.357004, T: 1306683, Avg. loss: 0.371976
Total training time: 10.98 seconds.
-- Epoch 50
Norm: 5.41, NNZs: 2298, Bias: -9.312838, T: 1333350, Avg. loss: 0.369820
Total training time: 11.20 seconds.
-- Epoch 51
Norm: 5.38, NNZs: 2298, Bias: -9.274885, T: 1360017, Avg. loss: 0.369161
Total training time: 11.42 seconds.
-- Epoch 52
Norm: 5.35, NNZs: 2298, Bias: -9.239383, T: 1386684, Avg. loss: 0.372592
Total training time: 11.65 seconds.
Convergence after 52 epochs took 11.65 seconds
-- Epoch 1
Norm: 5.63, NNZs: 2297, Bias: 7.445706, T: 26666, Avg. loss: 7.961795
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 5.02, NNZs: 2297, Bias: 6.129656, T: 53332, Avg. loss: 3.203413
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 3.79, NNZs: 2297, Bias: 5.237892, T: 79998, Avg. loss: 2.051165
Total training time: 0.68 seconds.
-- Epoch 4
Norm: 3.35, NNZs: 2297, Bias: 4.635640, T: 106664, Avg. loss: 1.334895
Total training time: 0.90 seconds.
-- Epoch 5
Norm: 3.07, NNZs: 2297, Bias: 4.242795, T: 133330, Avg. loss: 1.065958
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 2.89, NNZs: 2297, Bias: 3.944582, T: 159996, Avg. loss: 0.960460
```

```
Norm: 2.89, NNZs: 2297, Bias: 3.944582, T: 159996, Avg. loss: 0.960460
Total training time: 1.35 seconds.
-- Epoch 7
Norm: 2.74, NNZs: 2297, Bias: 3.716428, T: 186662, Avg. loss: 0.882345
Total training time: 1.57 seconds.
-- Epoch 8
Norm: 2.62, NNZs: 2297, Bias: 3.525654, T: 213328, Avg. loss: 0.847256
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 2.52, NNZs: 2297, Bias: 3.364472, T: 239994, Avg. loss: 0.820305
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 2.44, NNZs: 2297, Bias: 3.225814, T: 266660, Avg. loss: 0.795771
Total training time: 2.23 seconds.
-- Epoch 11
Norm: 2.37, NNZs: 2297, Bias: 3.105761, T: 293326, Avg. loss: 0.773474
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 2.31, NNZs: 2297, Bias: 2.998849, T: 319992, Avg. loss: 0.758064
Total training time: 2.68 seconds.
-- Epoch 13
Norm: 2.26, NNZs: 2297, Bias: 2.902110, T: 346658, Avg. loss: 0.748873
Total training time: 2.91 seconds.
-- Epoch 14
Norm: 2.21, NNZs: 2297, Bias: 2.815161, T: 373324, Avg. loss: 0.737237
Total training time: 3.13 seconds.
-- Epoch 15
Norm: 2.16, NNZs: 2297, Bias: 2.737290, T: 399990, Avg. loss: 0.721629
Total training time: 3.35 seconds.
-- Epoch 16
Norm: 2.13, NNZs: 2297, Bias: 2.663935, T: 426656, Avg. loss: 0.719116
Total training time: 3.58 seconds.
-- Epoch 17
Norm: 2.09, NNZs: 2297, Bias: 2.595779, T: 453322, Avg. loss: 0.712650
Total training time: 3.80 seconds.
-- Epoch 18
Norm: 2.06, NNZs: 2297, Bias: 2.533752, T: 479988, Avg. loss: 0.703283
Total training time: 4.01 seconds.
-- Epoch 19
Norm: 2.03, NNZs: 2297, Bias: 2.474743, T: 506654, Avg. loss: 0.699264
Total training time: 4.24 seconds.
-- Epoch 20
Norm: 2.00, NNZs: 2297, Bias: 2.420246, T: 533320, Avg. loss: 0.692351
Total training time: 4.46 seconds.
-- Epoch 21
Norm: 1.98, NNZs: 2297, Bias: 2.369561, T: 559986, Avg. loss: 0.683840
Total training time: 4.69 seconds.
-- Epoch 22
Norm: 1.95, NNZs: 2297, Bias: 2.322229, T: 586652, Avg. loss: 0.678080
Total training time: 4.91 seconds.
-- Epoch 23
Norm: 1.93, NNZs: 2297, Bias: 2.278483, T: 613318, Avg. loss: 0.668967
Total training time: 5.13 seconds.
-- Epoch 24
Norm: 1.91, NNZs: 2297, Bias: 2.235437, T: 639984, Avg. loss: 0.673537
Total training time: 5.36 seconds.
-- Epoch 25
Norm: 1.89, NNZs: 2297, Bias: 2.194864, T: 666650, Avg. loss: 0.667718
Total training time: 5.58 seconds.
-- Epoch 26
Norm: 1.87, NNZs: 2297, Bias: 2.157774, T: 693316, Avg. loss: 0.655636
Total training time: 5.82 seconds.
-- Epoch 27
Norm: 1.85, NNZs: 2297, Bias: 2.121087, T: 719982, Avg. loss: 0.659299
Total training time: 6.04 seconds.
-- Epoch 28
Norm: 1.83, NNZs: 2297, Bias: 2.086346, T: 746648, Avg. loss: 0.655610
Total training time: 6.26 seconds.
-- Epoch 29
Norm: 1.82, NNZs: 2297, Bias: 2.053790, T: 773314, Avg. loss: 0.648064
Total training time: 6.49 seconds.
-- Epoch 30
Norm: 1.80, NNZs: 2297, Bias: 2.021844, T: 799980, Avg. loss: 0.650169
Total training time: 6.70 seconds.
-- Epoch 31
Norm: 1.78, NNZs: 2297, Bias: 1.991119, T: 826646, Avg. loss: 0.647875
Total training time: 6.93 seconds.
-- Epoch 32
```

```
-- Epoch 32
Norm: 1.77, NNZs: 2297, Bias: 1.962183, T: 853312, Avg. loss: 0.641273
Total training time: 7.15 seconds.
-- Epoch 33
Norm: 1.76, NNZs: 2297, Bias: 1.934444, T: 879978, Avg. loss: 0.637376
Total training time: 7.38 seconds.
-- Epoch 34
Norm: 1.74, NNZs: 2297, Bias: 1.907936, T: 906644, Avg. loss: 0.633958
Total training time: 7.61 seconds.
-- Epoch 35
Norm: 1.73, NNZs: 2297, Bias: 1.882000, T: 933310, Avg. loss: 0.635560
Total training time: 7.84 seconds.
-- Epoch 36
Norm: 1.72, NNZs: 2297, Bias: 1.857032, T: 959976, Avg. loss: 0.632702
Total training time: 8.06 seconds.
-- Epoch 37
Norm: 1.71, NNZs: 2297, Bias: 1.832811, T: 986642, Avg. loss: 0.631361
Total training time: 8.28 seconds.
-- Epoch 38
Norm: 1.70, NNZs: 2297, Bias: 1.810073, T: 1013308, Avg. loss: 0.623506
Total training time: 8.51 seconds.
-- Epoch 39
Norm: 1.69, NNZs: 2297, Bias: 1.787433, T: 1039974, Avg. loss: 0.627651
Total training time: 8.73 seconds.
-- Epoch 40
Norm: 1.68, NNZs: 2297, Bias: 1.765649, T: 1066640, Avg. loss: 0.624089
Total training time: 8.96 seconds.
-- Epoch 41
Norm: 1.67, NNZs: 2297, Bias: 1.744492, T: 1093306, Avg. loss: 0.623342
Total training time: 9.18 seconds.
-- Epoch 42
Norm: 1.66, NNZs: 2297, Bias: 1.724075, T: 1119972, Avg. loss: 0.620629
Total training time: 9.40 seconds.
-- Epoch 43
Norm: 1.65, NNZs: 2297, Bias: 1.704396, T: 1146638, Avg. loss: 0.617109
Total training time: 9.62 seconds.
-- Epoch 44
Norm: 1.64, NNZs: 2297, Bias: 1.685266, T: 1173304, Avg. loss: 0.615344
Total training time: 9.85 seconds.
-- Epoch 45
Norm: 1.63, NNZs: 2297, Bias: 1.667008, T: 1199970, Avg. loss: 0.610662
Total training time: 10.07 seconds.
-- Epoch 46
Norm: 1.62, NNZs: 2297, Bias: 1.648791, T: 1226636, Avg. loss: 0.614273
Total training time: 10.29 seconds.
-- Epoch 47
Norm: 1.61, NNZs: 2297, Bias: 1.631299, T: 1253302, Avg. loss: 0.610667
Total training time: 10.52 seconds.
-- Epoch 48
Norm: 1.60, NNZs: 2297, Bias: 1.614246, T: 1279968, Avg. loss: 0.608980
Total training time: 10.74 seconds.
-- Epoch 49
Norm: 1.60, NNZs: 2297, Bias: 1.597585, T: 1306634, Avg. loss: 0.608125
Total training time: 10.96 seconds.
-- Epoch 50
Norm: 1.59, NNZs: 2297, Bias: 1.581403, T: 1333300, Avg. loss: 0.606351
Total training time: 11.18 seconds.
-- Epoch 51
Norm: 1.58, NNZs: 2297, Bias: 1.565523, T: 1359966, Avg. loss: 0.606425
Total training time: 11.40 seconds.
-- Epoch 52
Norm: 1.57, NNZs: 2297, Bias: 1.550212, T: 1386632, Avg. loss: 0.603078
Total training time: 11.63 seconds.
-- Epoch 53
Norm: 1.57, NNZs: 2297, Bias: 1.535174, T: 1413298, Avg. loss: 0.602844
Total training time: 11.85 seconds.
-- Epoch 54
Norm: 1.56, NNZs: 2297, Bias: 1.520797, T: 1439964, Avg. loss: 0.597805
Total training time: 12.07 seconds.
-- Epoch 55
Norm: 1.55, NNZs: 2297, Bias: 1.505931, T: 1466630, Avg. loss: 0.606456
Total training time: 12.30 seconds.
-- Epoch 56
Norm: 1.55, NNZs: 2297, Bias: 1.492224, T: 1493296, Avg. loss: 0.595530
Total training time: 12.52 seconds.
-- Epoch 57
Norm: 1.54, NNZs: 2297, Bias: 1.478416, T: 1519962, Avg. loss: 0.599478
Total training time: 12.74 seconds.
```

```
Total training time: 12.74 seconds.
-- Epoch 58
Norm: 1.54, NNZs: 2297, Bias: 1.465024, T: 1546628, Avg. loss: 0.597017
Total training time: 12.96 seconds.
-- Epoch 59
Norm: 1.53, NNZs: 2297, Bias: 1.451740, T: 1573294, Avg. loss: 0.598019
Total training time: 13.19 seconds.
-- Epoch 60
Norm: 1.52, NNZs: 2297, Bias: 1.439153, T: 1599960, Avg. loss: 0.592096
Total training time: 13.41 seconds.
-- Epoch 61
Norm: 1.52, NNZs: 2297, Bias: 1.426379, T: 1626626, Avg. loss: 0.596672
Total training time: 13.63 seconds.
-- Epoch 62
Norm: 1.51, NNZs: 2297, Bias: 1.414122, T: 1653292, Avg. loss: 0.592299
Total training time: 13.85 seconds.
-- Epoch 63
Norm: 1.51, NNZs: 2297, Bias: 1.402055, T: 1679958, Avg. loss: 0.592430
Total training time: 14.08 seconds.
-- Epoch 64
Norm: 1.50, NNZs: 2297, Bias: 1.390269, T: 1706624, Avg. loss: 0.590749
Total training time: 14.30 seconds.
-- Epoch 65
Norm: 1.50, NNZs: 2297, Bias: 1.378740, T: 1733290, Avg. loss: 0.589011
Total training time: 14.52 seconds.
-- Epoch 66
Norm: 1.49, NNZs: 2297, Bias: 1.367539, T: 1759956, Avg. loss: 0.587291
Total training time: 14.74 seconds.
-- Epoch 67
Norm: 1.49, NNZs: 2297, Bias: 1.356454, T: 1786622, Avg. loss: 0.587443
Total training time: 14.96 seconds.
-- Epoch 68
Norm: 1.48, NNZs: 2297, Bias: 1.345419, T: 1813288, Avg. loss: 0.589086
Total training time: 15.18 seconds.
-- Epoch 69
Norm: 1.48, NNZs: 2297, Bias: 1.334829, T: 1839954, Avg. loss: 0.584824
Total training time: 15.40 seconds.
-- Epoch 70
Norm: 1.47, NNZs: 2297, Bias: 1.324343, T: 1866620, Avg. loss: 0.584764
Total training time: 15.63 seconds.
-- Epoch 71
Norm: 1.47, NNZs: 2297, Bias: 1.313874, T: 1893286, Avg. loss: 0.586739
Total training time: 15.85 seconds.
-- Epoch 72
Norm: 1.46, NNZs: 2297, Bias: 1.303845, T: 1919952, Avg. loss: 0.582008
Total training time: 16.08 seconds.
-- Epoch 73
Norm: 1.46, NNZs: 2297, Bias: 1.294175, T: 1946618, Avg. loss: 0.578141
Total training time: 16.31 seconds.
-- Epoch 74
Norm: 1.46, NNZs: 2297, Bias: 1.284344, T: 1973284, Avg. loss: 0.582733
Total training time: 16.53 seconds.
-- Epoch 75
Norm: 1.45, NNZs: 2297, Bias: 1.274693, T: 1999950, Avg. loss: 0.581657
Total training time: 16.75 seconds.
-- Epoch 76
Norm: 1.45, NNZs: 2297, Bias: 1.265329, T: 2026616, Avg. loss: 0.578965
Total training time: 16.98 seconds.
-- Epoch 77
Norm: 1.44, NNZs: 2297, Bias: 1.255821, T: 2053282, Avg. loss: 0.582918
Total training time: 17.21 seconds.
-- Epoch 78
Norm: 1.44, NNZs: 2297, Bias: 1.246802, T: 2079948, Avg. loss: 0.577095
Total training time: 17.43 seconds.
-- Epoch 79
Norm: 1.43, NNZs: 2297, Bias: 1.237843, T: 2106614, Avg. loss: 0.577322
Total training time: 17.65 seconds.
-- Epoch 80
Norm: 1.43, NNZs: 2297, Bias: 1.228994, T: 2133280, Avg. loss: 0.577651
Total training time: 17.88 seconds.
-- Epoch 81
Norm: 1.43, NNZs: 2297, Bias: 1.220435, T: 2159946, Avg. loss: 0.574187
Total training time: 18.10 seconds.
-- Epoch 82
Norm: 1.42, NNZs: 2297, Bias: 1.211644, T: 2186612, Avg. loss: 0.579921
Total training time: 18.32 seconds.
-- Epoch 83
```

```
Norm: 1.42, NNZs: 2297, Bias: 1.203215, T: 2213278, Avg. loss: 0.575354
Total training time: 18.56 seconds.
-- Epoch 84
Norm: 1.42, NNZs: 2297, Bias: 1.195074, T: 2239944, Avg. loss: 0.571879
Total training time: 18.78 seconds.
-- Epoch 85
Norm: 1.41, NNZs: 2297, Bias: 1.186830, T: 2266610, Avg. loss: 0.575194
Total training time: 19.00 seconds.
-- Epoch 86
Norm: 1.41, NNZs: 2297, Bias: 1.178757, T: 2293276, Avg. loss: 0.573919
Total training time: 19.22 seconds.
-- Epoch 87
Norm: 1.41, NNZs: 2297, Bias: 1.170879, T: 2319942, Avg. loss: 0.571494
Total training time: 19.44 seconds.
-- Epoch 88
Norm: 1.40, NNZs: 2297, Bias: 1.163057, T: 2346608, Avg. loss: 0.572095
Total training time: 19.67 seconds.
-- Epoch 89
Norm: 1.40, NNZs: 2297, Bias: 1.155468, T: 2373274, Avg. loss: 0.569174
Total training time: 19.89 seconds.
-- Epoch 90
Norm: 1.40, NNZs: 2297, Bias: 1.147870, T: 2399940, Avg. loss: 0.570803
Total training time: 20.11 seconds.
-- Epoch 91
Norm: 1.39, NNZs: 2297, Bias: 1.140362, T: 2426606, Avg. loss: 0.570678
Total training time: 20.33 seconds.
-- Epoch 92
Norm: 1.39, NNZs: 2297, Bias: 1.132914, T: 2453272, Avg. loss: 0.570870
Total training time: 20.56 seconds.
-- Epoch 93
Norm: 1.39, NNZs: 2297, Bias: 1.125717, T: 2479938, Avg. loss: 0.567185
Total training time: 20.79 seconds.
-- Epoch 94
Norm: 1.38, NNZs: 2297, Bias: 1.118503, T: 2506604, Avg. loss: 0.569343
Total training time: 21.01 seconds.
-- Epoch 95
Norm: 1.38, NNZs: 2297, Bias: 1.111489, T: 2533270, Avg. loss: 0.566937
Total training time: 21.23 seconds.
-- Epoch 96
Norm: 1.38, NNZs: 2297, Bias: 1.104505, T: 2559936, Avg. loss: 0.567394
Total training time: 21.45 seconds.
-- Epoch 97
Norm: 1.37, NNZs: 2297, Bias: 1.097598, T: 2586602, Avg. loss: 0.566959
Total training time: 21.68 seconds.
-- Epoch 98
Norm: 1.37, NNZs: 2297, Bias: 1.090878, T: 2613268, Avg. loss: 0.564744
Total training time: 21.90 seconds.
-- Epoch 99
Norm: 1.37, NNZs: 2297, Bias: 1.084147, T: 2639934, Avg. loss: 0.565970
Total training time: 22.12 seconds.
-- Epoch 100
Norm: 1.37, NNZs: 2297, Bias: 1.077410, T: 2666600, Avg. loss: 0.567121
Total training time: 22.34 seconds.
-- Epoch 101
Norm: 1.36, NNZs: 2297, Bias: 1.070884, T: 2693266, Avg. loss: 0.564510
Total training time: 22.56 seconds.
-- Epoch 102
Norm: 1.36, NNZs: 2297, Bias: 1.064451, T: 2719932, Avg. loss: 0.563514
Total training time: 22.78 seconds.
-- Epoch 103
Norm: 1.36, NNZs: 2297, Bias: 1.058159, T: 2746598, Avg. loss: 0.561748
Total training time: 23.01 seconds.
-- Epoch 104
Norm: 1.36, NNZs: 2297, Bias: 1.051804, T: 2773264, Avg. loss: 0.564480
Total training time: 23.24 seconds.
-- Epoch 105
Norm: 1.35, NNZs: 2297, Bias: 1.045604, T: 2799930, Avg. loss: 0.562041
Total training time: 23.46 seconds.
-- Epoch 106
Norm: 1.35, NNZs: 2297, Bias: 1.039337, T: 2826596, Avg. loss: 0.564809
Total training time: 23.69 seconds.
-- Epoch 107
Norm: 1.35, NNZs: 2297, Bias: 1.033368, T: 2853262, Avg. loss: 0.559609
Total training time: 23.91 seconds.
-- Epoch 108
Norm: 1.34, NNZs: 2297, Bias: 1.027251, T: 2879928, Avg. loss: 0.563818
Total training time: 24.13 seconds.
```

```
-- Epoch 109
Norm: 1.34, NNZs: 2297, Bias: 1.021400, T: 2906594, Avg. loss: 0.558841
Total training time: 24.35 seconds.
-- Epoch 110
Norm: 1.34, NNZs: 2297, Bias: 1.015446, T: 2933260, Avg. loss: 0.562515
Total training time: 24.57 seconds.
-- Epoch 111
Norm: 1.34, NNZs: 2297, Bias: 1.009740, T: 2959926, Avg. loss: 0.558134
Total training time: 24.80 seconds.
-- Epoch 112
Norm: 1.34, NNZs: 2297, Bias: 1.003933, T: 2986592, Avg. loss: 0.561351
Total training time: 25.03 seconds.
Convergence after 112 epochs took 25.03 seconds
-- Epoch 1
Norm: 2.76, NNZs: 2298, Bias: 1.077676, T: 26667, Avg. loss: 8.714622
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 2.51, NNZs: 2298, Bias: 0.720623, T: 53334, Avg. loss: 1.547192
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 1.81, NNZs: 2298, Bias: 0.490475, T: 80001, Avg. loss: 0.847551
Total training time: 0.68 seconds.
-- Epoch 4
Norm: 1.50, NNZs: 2298, Bias: 0.348805, T: 106668, Avg. loss: 0.647837
Total training time: 0.90 seconds.
-- Epoch 5
Norm: 1.34, NNZs: 2298, Bias: 0.247687, T: 133335, Avg. loss: 0.587712
Total training time: 1.13 seconds.
-- Epoch 6
Norm: 1.25, NNZs: 2298, Bias: 0.178027, T: 160002, Avg. loss: 0.561440
Total training time: 1.35 seconds.
-- Epoch 7
Norm: 1.19, NNZs: 2298, Bias: 0.121242, T: 186669, Avg. loss: 0.543872
Total training time: 1.57 seconds.
-- Epoch 8
Norm: 1.15, NNZs: 2298, Bias: 0.072534, T: 213336, Avg. loss: 0.530187
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 1.11, NNZs: 2298, Bias: 0.027427, T: 240003, Avg. loss: 0.530945
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 1.08, NNZs: 2298, Bias: -0.008880, T: 266670, Avg. loss: 0.521158
Total training time: 2.24 seconds.
-- Epoch 11
Norm: 1.07, NNZs: 2298, Bias: -0.037541, T: 293337, Avg. loss: 0.509970
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 1.05, NNZs: 2298, Bias: -0.065882, T: 320004, Avg. loss: 0.509004
Total training time: 2.69 seconds.
-- Epoch 13
Norm: 1.04, NNZs: 2298, Bias: -0.092888, T: 346671, Avg. loss: 0.509054
Total training time: 2.91 seconds.
-- Epoch 14
Norm: 1.03, NNZs: 2298, Bias: -0.115400, T: 373338, Avg. loss: 0.502733
Total training time: 3.13 seconds.
-- Epoch 15
Norm: 1.02, NNZs: 2298, Bias: -0.135723, T: 400005, Avg. loss: 0.499937
Total training time: 3.36 seconds.
-- Epoch 16
Norm: 1.01, NNZs: 2298, Bias: -0.155363, T: 426672, Avg. loss: 0.500190
Total training time: 3.58 seconds.
-- Epoch 17
Norm: 1.00, NNZs: 2298, Bias: -0.174093, T: 453339, Avg. loss: 0.500031
Total training time: 3.81 seconds.
-- Epoch 18
Norm: 0.99, NNZs: 2298, Bias: -0.190357, T: 480006, Avg. loss: 0.492803
Total training time: 4.04 seconds.
-- Epoch 19
Norm: 0.99, NNZs: 2298, Bias: -0.205113, T: 506673, Avg. loss: 0.490384
Total training time: 4.26 seconds.
-- Epoch 20
Norm: 0.98, NNZs: 2298, Bias: -0.220251, T: 533340, Avg. loss: 0.494933
Total training time: 4.48 seconds.
-- Epoch 21
Norm: 0.98, NNZs: 2298, Bias: -0.233443, T: 560007, Avg. loss: 0.489475
Total training time: 4.70 seconds.
-- Epoch 22
```

```
Norm: 0.98, NNZs: 2298, Bias: -0.246085, T: 586674, Avg. loss: 0.489533
Total training time: 4.92 seconds.
-- Epoch 23
Norm: 0.97, NNZs: 2298, Bias: -0.258321, T: 613341, Avg. loss: 0.489533
Total training time: 5.15 seconds.
-- Epoch 24
Norm: 0.97, NNZs: 2298, Bias: -0.270029, T: 640008, Avg. loss: 0.485947
Total training time: 5.37 seconds.
-- Epoch 25
Norm: 0.96, NNZs: 2298, Bias: -0.281311, T: 666675, Avg. loss: 0.486771
Total training time: 5.59 seconds.
-- Epoch 26
Norm: 0.96, NNZs: 2298, Bias: -0.291690, T: 693342, Avg. loss: 0.484761
Total training time: 5.81 seconds.
-- Epoch 27
Norm: 0.96, NNZs: 2298, Bias: -0.301269, T: 720009, Avg. loss: 0.482652
Total training time: 6.03 seconds.
-- Epoch 28
Norm: 0.96, NNZs: 2298, Bias: -0.310825, T: 746676, Avg. loss: 0.483972
Total training time: 6.26 seconds.
-- Epoch 29
Norm: 0.95, NNZs: 2298, Bias: -0.319982, T: 773343, Avg. loss: 0.482831
Total training time: 6.47 seconds.
-- Epoch 30
Norm: 0.95, NNZs: 2298, Bias: -0.328277, T: 800010, Avg. loss: 0.479857
Total training time: 6.70 seconds.
-- Epoch 31
Norm: 0.95, NNZs: 2298, Bias: -0.336179, T: 826677, Avg. loss: 0.478945
Total training time: 6.92 seconds.
-- Epoch 32
Norm: 0.95, NNZs: 2298, Bias: -0.344640, T: 853344, Avg. loss: 0.482381
Total training time: 7.14 seconds.
-- Epoch 33
Norm: 0.94, NNZs: 2298, Bias: -0.352329, T: 880011, Avg. loss: 0.479502
Total training time: 7.36 seconds.
-- Epoch 34
Norm: 0.94, NNZs: 2298, Bias: -0.360028, T: 906678, Avg. loss: 0.480957
Total training time: 7.59 seconds.
-- Epoch 35
Norm: 0.94, NNZs: 2298, Bias: -0.366683, T: 933345, Avg. loss: 0.475961
Total training time: 7.81 seconds.
-- Epoch 36
Norm: 0.94, NNZs: 2298, Bias: -0.373368, T: 960012, Avg. loss: 0.476993
Total training time: 8.03 seconds.
-- Epoch 37
Norm: 0.94, NNZs: 2298, Bias: -0.379578, T: 986679, Avg. loss: 0.474236
Total training time: 8.25 seconds.
-- Epoch 38
Norm: 0.94, NNZs: 2298, Bias: -0.385927, T: 1013346, Avg. loss: 0.476420
Total training time: 8.48 seconds.
-- Epoch 39
Norm: 0.93, NNZs: 2298, Bias: -0.392152, T: 1040013, Avg. loss: 0.476541
Total training time: 8.71 seconds.
-- Epoch 40
Norm: 0.93, NNZs: 2298, Bias: -0.397906, T: 1066680, Avg. loss: 0.473899
Total training time: 8.93 seconds.
-- Epoch 41
Norm: 0.93, NNZs: 2298, Bias: -0.403860, T: 1093347, Avg. loss: 0.475419
Total training time: 9.15 seconds.
-- Epoch 42
Norm: 0.93, NNZs: 2298, Bias: -0.409367, T: 1120014, Avg. loss: 0.474118
Total training time: 9.38 seconds.
Convergence after 42 epochs took 9.38 seconds
-- Epoch 1
Norm: 4.82, NNZs: 2298, Bias: -1.421417, T: 26667, Avg. loss: 10.389381
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 2.36, NNZs: 2298, Bias: -1.508059, T: 53334, Avg. loss: 1.112665
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 1.75, NNZs: 2298, Bias: -1.490210, T: 80001, Avg. loss: 0.630995
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 1.44, NNZs: 2298, Bias: -1.501553, T: 106668, Avg. loss: 0.535074
Total training time: 0.91 seconds.
-- Epoch 5
Norm: 1.29, NNZs: 2298, Bias: -1.486769, T: 133335, Avg. loss: 0.488437
```

```
Total training time: 1.13 seconds.
-- Epoch 6
Norm: 1.20, NNZs: 2298, Bias: -1.477750, T: 160002, Avg. loss: 0.466609
Total training time: 1.36 seconds.
-- Epoch 7
Norm: 1.20, NNZs: 2298, Bias: -1.475379, T: 186669, Avg. loss: 0.470330
Total training time: 1.58 seconds.
-- Epoch 8
Norm: 1.12, NNZs: 2298, Bias: -1.473206, T: 213336, Avg. loss: 0.475545
Total training time: 1.80 seconds.
-- Epoch 9
Norm: 1.07, NNZs: 2298, Bias: -1.471150, T: 240003, Avg. loss: 0.450042
Total training time: 2.03 seconds.
-- Epoch 10
Norm: 1.04, NNZs: 2298, Bias: -1.468029, T: 266670, Avg. loss: 0.442161
Total training time: 2.25 seconds.
-- Epoch 11
Norm: 1.03, NNZs: 2298, Bias: -1.463927, T: 293337, Avg. loss: 0.433328
Total training time: 2.48 seconds.
-- Epoch 12
Norm: 1.01, NNZs: 2298, Bias: -1.462197, T: 320004, Avg. loss: 0.435027
Total training time: 2.70 seconds.
-- Epoch 13
Norm: 0.99, NNZs: 2298, Bias: -1.462804, T: 346671, Avg. loss: 0.436701
Total training time: 2.92 seconds.
-- Epoch 14
Norm: 0.98, NNZs: 2298, Bias: -1.459651, T: 373338, Avg. loss: 0.431360
Total training time: 3.15 seconds.
-- Epoch 15
Norm: 0.97, NNZs: 2298, Bias: -1.459559, T: 400005, Avg. loss: 0.431594
Total training time: 3.37 seconds.
-- Epoch 16
Norm: 0.97, NNZs: 2298, Bias: -1.457697, T: 426672, Avg. loss: 0.427609
Total training time: 3.60 seconds.
-- Epoch 17
Norm: 0.96, NNZs: 2298, Bias: -1.456492, T: 453339, Avg. loss: 0.427743
Total training time: 3.83 seconds.
-- Epoch 18
Norm: 0.95, NNZs: 2298, Bias: -1.456237, T: 480006, Avg. loss: 0.430006
Total training time: 4.05 seconds.
-- Epoch 19
Norm: 0.95, NNZs: 2298, Bias: -1.453658, T: 506673, Avg. loss: 0.421099
Total training time: 4.27 seconds.
-- Epoch 20
Norm: 0.95, NNZs: 2298, Bias: -1.453282, T: 533340, Avg. loss: 0.427892
Total training time: 4.50 seconds.
-- Epoch 21
Norm: 0.94, NNZs: 2298, Bias: -1.454157, T: 560007, Avg. loss: 0.432201
Total training time: 4.72 seconds.
-- Epoch 22
Norm: 0.94, NNZs: 2298, Bias: -1.452442, T: 586674, Avg. loss: 0.423216
Total training time: 4.95 seconds.
-- Epoch 23
Norm: 0.93, NNZs: 2298, Bias: -1.452266, T: 613341, Avg. loss: 0.427072
Total training time: 5.17 seconds.
-- Epoch 24
Norm: 0.93, NNZs: 2298, Bias: -1.452405, T: 640008, Avg. loss: 0.426380
Total training time: 5.40 seconds.
Convergence after 24 epochs took 5.40 seconds
-- Epoch 1
Norm: 0.77, NNZs: 2297, Bias: 10.447278, T: 26666, Avg. loss: 9.748296
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.78, NNZs: 2297, Bias: 10.100819, T: 53332, Avg. loss: 8.901375
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 0.78, NNZs: 2297, Bias: 9.898201, T: 79998, Avg. loss: 8.640400
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 0.78, NNZs: 2297, Bias: 9.754465, T: 106664, Avg. loss: 8.469931
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.78, NNZs: 2297, Bias: 9.642978, T: 133330, Avg. loss: 8.343757
Total training time: 1.11 seconds.
-- Epoch 6
Norm: 0.78, NNZs: 2297, Bias: 9.551877, T: 159996, Avg. loss: 8.244948
Total training time: 1.33 seconds.
```

```
-- Epoch 7
Norm: 0.78, NNZs: 2297, Bias: 9.474865, T: 186662, Avg. loss: 8.161865
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 0.78, NNZs: 2297, Bias: 9.408155, T: 213328, Avg. loss: 8.091406
Total training time: 1.78 seconds.
-- Epoch 9
Norm: 0.78, NNZs: 2297, Bias: 9.349317, T: 239994, Avg. loss: 8.029432
Total training time: 2.00 seconds.
-- Epoch 10
Norm: 0.78, NNZs: 2297, Bias: 9.296688, T: 266660, Avg. loss: 7.974416
Total training time: 2.22 seconds.
-- Epoch 11
Norm: 0.78, NNZs: 2297, Bias: 9.249080, T: 293326, Avg. loss: 7.925208
Total training time: 2.44 seconds.
-- Epoch 12
Norm: 0.78, NNZs: 2297, Bias: 9.205619, T: 319992, Avg. loss: 7.880342
Total training time: 2.67 seconds.
-- Epoch 13
Norm: 0.78, NNZs: 2297, Bias: 9.165643, T: 346658, Avg. loss: 7.839352
Total training time: 2.89 seconds.
-- Epoch 14
Norm: 0.78, NNZs: 2297, Bias: 9.128632, T: 373324, Avg. loss: 7.801485
Total training time: 3.12 seconds.
-- Epoch 15
Norm: 0.78, NNZs: 2297, Bias: 9.094173, T: 399990, Avg. loss: 7.766183
Total training time: 3.34 seconds.
-- Epoch 16
Norm: 0.78, NNZs: 2297, Bias: 9.061943, T: 426656, Avg. loss: 7.733540
Total training time: 3.57 seconds.
-- Epoch 17
Norm: 0.78, NNZs: 2297, Bias: 9.031668, T: 453322, Avg. loss: 7.702762
Total training time: 3.79 seconds.
-- Epoch 18
Norm: 0.78, NNZs: 2297, Bias: 9.003125, T: 479988, Avg. loss: 7.674163
Total training time: 4.01 seconds.
-- Epoch 19
Norm: 0.78, NNZs: 2297, Bias: 8.976127, T: 506654, Avg. loss: 7.646849
Total training time: 4.24 seconds.
-- Epoch 20
Norm: 0.78, NNZs: 2297, Bias: 8.950515, T: 533320, Avg. loss: 7.620972
Total training time: 4.46 seconds.
-- Epoch 21
Norm: 0.78, NNZs: 2297, Bias: 8.926154, T: 559986, Avg. loss: 7.596483
Total training time: 4.68 seconds.
-- Epoch 22
Norm: 0.78, NNZs: 2297, Bias: 8.902926, T: 586652, Avg. loss: 7.573203
Total training time: 4.90 seconds.
-- Epoch 23
Norm: 0.78, NNZs: 2297, Bias: 8.880731, T: 613318, Avg. loss: 7.550981
Total training time: 5.14 seconds.
-- Epoch 24
Norm: 0.78, NNZs: 2297, Bias: 8.859482, T: 639984, Avg. loss: 7.529599
Total training time: 5.36 seconds.
-- Epoch 25
Norm: 0.78, NNZs: 2297, Bias: 8.839102, T: 666650, Avg. loss: 7.509041
Total training time: 5.58 seconds.
-- Epoch 26
Norm: 0.78, NNZs: 2297, Bias: 8.819521, T: 693316, Avg. loss: 7.489526
Total training time: 5.80 seconds.
-- Epoch 27
Norm: 0.78, NNZs: 2297, Bias: 8.800679, T: 719982, Avg. loss: 7.470616
Total training time: 6.03 seconds.
-- Epoch 28
Norm: 0.78, NNZs: 2297, Bias: 8.782524, T: 746648, Avg. loss: 7.452569
Total training time: 6.25 seconds.
-- Epoch 29
Norm: 0.78, NNZs: 2297, Bias: 8.765006, T: 773314, Avg. loss: 7.434976
Total training time: 6.48 seconds.
-- Epoch 30
Norm: 0.78, NNZs: 2297, Bias: 8.748083, T: 799980, Avg. loss: 7.418052
Total training time: 6.70 seconds.
-- Epoch 31
Norm: 0.78, NNZs: 2297, Bias: 8.731714, T: 826646, Avg. loss: 7.401780
Total training time: 6.92 seconds.
-- Epoch 32
Norm: 0.78, NNZs: 2297, Bias: 8.715866, T: 853312, Avg. loss: 7.386023
```

```
Total training time: 7.14 seconds.
-- Epoch 33
Norm: 0.78, NNZs: 2297, Bias: 8.700506, T: 879978, Avg. loss: 7.370661
Total training time: 7.36 seconds.
-- Epoch 34
Norm: 0.78, NNZs: 2297, Bias: 8.685605, T: 906644, Avg. loss: 7.355876
Total training time: 7.59 seconds.
-- Epoch 35
Norm: 0.78, NNZs: 2297, Bias: 8.671136, T: 933310, Avg. loss: 7.341441
Total training time: 7.81 seconds.
-- Epoch 36
Norm: 0.78, NNZs: 2297, Bias: 8.657076, T: 959976, Avg. loss: 7.327408
Total training time: 8.03 seconds.
-- Epoch 37
Norm: 0.78, NNZs: 2297, Bias: 8.643400, T: 986642, Avg. loss: 7.314005
Total training time: 8.26 seconds.
-- Epoch 38
Norm: 0.78, NNZs: 2297, Bias: 8.630090, T: 1013308, Avg. loss: 7.300715
Total training time: 8.49 seconds.
-- Epoch 39
Norm: 0.78, NNZs: 2297, Bias: 8.617126, T: 1039974, Avg. loss: 7.287877
Total training time: 8.71 seconds.
-- Epoch 40
Norm: 0.78, NNZs: 2297, Bias: 8.604490, T: 1066640, Avg. loss: 7.275325
Total training time: 8.93 seconds.
-- Epoch 41
Norm: 0.78, NNZs: 2297, Bias: 8.592167, T: 1093306, Avg. loss: 7.263120
Total training time: 9.15 seconds.
-- Epoch 42
Norm: 0.78, NNZs: 2297, Bias: 8.580141, T: 1119972, Avg. loss: 7.251198
Total training time: 9.37 seconds.
-- Epoch 43
Norm: 0.78, NNZs: 2297, Bias: 8.568397, T: 1146638, Avg. loss: 7.239563
Total training time: 9.60 seconds.
-- Epoch 44
Norm: 0.78, NNZs: 2297, Bias: 8.556925, T: 1173304, Avg. loss: 7.228100
Total training time: 9.82 seconds.
-- Epoch 45
Norm: 0.78, NNZs: 2297, Bias: 8.545710, T: 1199970, Avg. loss: 7.217081
Total training time: 10.04 seconds.
-- Epoch 46
Norm: 0.78, NNZs: 2297, Bias: 8.534742, T: 1226636, Avg. loss: 7.206139
Total training time: 10.26 seconds.
-- Epoch 47
Norm: 0.78, NNZs: 2297, Bias: 8.524010, T: 1253302, Avg. loss: 7.195538
Total training time: 10.49 seconds.
-- Epoch 48
Norm: 0.78, NNZs: 2297, Bias: 8.513504, T: 1279968, Avg. loss: 7.185114
Total training time: 10.71 seconds.
-- Epoch 49
Norm: 0.78, NNZs: 2297, Bias: 8.503215, T: 1306634, Avg. loss: 7.174985
Total training time: 10.94 seconds.
-- Epoch 50
Norm: 0.78, NNZs: 2297, Bias: 8.493134, T: 1333300, Avg. loss: 7.165049
Total training time: 11.16 seconds.
-- Epoch 51
Norm: 0.77, NNZs: 2297, Bias: 8.483253, T: 1359966, Avg. loss: 7.155258
Total training time: 11.38 seconds.
-- Epoch 52
Norm: 0.77, NNZs: 2297, Bias: 8.473565, T: 1386632, Avg. loss: 7.145610
Total training time: 11.60 seconds.
-- Epoch 53
Norm: 0.77, NNZs: 2297, Bias: 8.464060, T: 1413298, Avg. loss: 7.136276
Total training time: 11.82 seconds.
-- Epoch 54
Norm: 0.77, NNZs: 2297, Bias: 8.454734, T: 1439964, Avg. loss: 7.126978
Total training time: 12.05 seconds.
-- Epoch 55
Norm: 0.77, NNZs: 2297, Bias: 8.445579, T: 1466630, Avg. loss: 7.117978
Total training time: 12.27 seconds.
-- Epoch 56
Norm: 0.77, NNZs: 2297, Bias: 8.436588, T: 1493296, Avg. loss: 7.109121
Total training time: 12.49 seconds.
-- Epoch 57
Norm: 0.77, NNZs: 2297, Bias: 8.427758, T: 1519962, Avg. loss: 7.100401
Total training time: 12.72 seconds.
-- Epoch 58
```

```
Norm: 0.77, NNZs: 2297, Bias: 8.419081, T: 1546628, Avg. loss: 7.091818
Total training time: 12.94 seconds.
-- Epoch 59
Norm: 0.77, NNZs: 2297, Bias: 8.410552, T: 1573294, Avg. loss: 7.083411
Total training time: 13.16 seconds.
-- Epoch 60
Norm: 0.77, NNZs: 2297, Bias: 8.402167, T: 1599960, Avg. loss: 7.075145
Total training time: 13.38 seconds.
-- Epoch 61
Norm: 0.77, NNZs: 2297, Bias: 8.393921, T: 1626626, Avg. loss: 7.067032
Total training time: 13.60 seconds.
-- Epoch 62
Norm: 0.77, NNZs: 2297, Bias: 8.385809, T: 1653292, Avg. loss: 7.059045
Total training time: 13.83 seconds.
-- Epoch 63
Norm: 0.77, NNZs: 2297, Bias: 8.377827, T: 1679958, Avg. loss: 7.051139
Total training time: 14.05 seconds.
-- Epoch 64
Norm: 0.77, NNZs: 2297, Bias: 8.369970, T: 1706624, Avg. loss: 7.043424
Total training time: 14.28 seconds.
-- Epoch 65
Norm: 0.77, NNZs: 2297, Bias: 8.362236, T: 1733290, Avg. loss: 7.035709
Total training time: 14.50 seconds.
-- Epoch 66
Norm: 0.77, NNZs: 2297, Bias: 8.354620, T: 1759956, Avg. loss: 7.028256
Total training time: 14.73 seconds.
-- Epoch 67
Norm: 0.77, NNZs: 2297, Bias: 8.347118, T: 1786622, Avg. loss: 7.020836
Total training time: 14.95 seconds.
-- Epoch 68
Norm: 0.77, NNZs: 2297, Bias: 8.339728, T: 1813288, Avg. loss: 7.013514
Total training time: 15.17 seconds.
-- Epoch 69
Norm: 0.77, NNZs: 2297, Bias: 8.332446, T: 1839954, Avg. loss: 7.006309
Total training time: 15.41 seconds.
-- Epoch 70
Norm: 0.77, NNZs: 2297, Bias: 8.325268, T: 1866620, Avg. loss: 6.999342
Total training time: 15.64 seconds.
-- Epoch 71
Norm: 0.77, NNZs: 2297, Bias: 8.318193, T: 1893286, Avg. loss: 6.992331
Total training time: 15.86 seconds.
-- Epoch 72
Norm: 0.77, NNZs: 2297, Bias: 8.311217, T: 1919952, Avg. loss: 6.985498
Total training time: 16.08 seconds.
-- Epoch 73
Norm: 0.77, NNZs: 2297, Bias: 8.304337, T: 1946618, Avg. loss: 6.978664
Total training time: 16.30 seconds.
-- Epoch 74
Norm: 0.77, NNZs: 2297, Bias: 8.297550, T: 1973284, Avg. loss: 6.972000
Total training time: 16.53 seconds.
-- Epoch 75
Norm: 0.77, NNZs: 2297, Bias: 8.290855, T: 1999950, Avg. loss: 6.965443
Total training time: 16.76 seconds.
-- Epoch 76
Norm: 0.77, NNZs: 2297, Bias: 8.284249, T: 2026616, Avg. loss: 6.958884
Total training time: 16.98 seconds.
-- Epoch 77
Norm: 0.77, NNZs: 2297, Bias: 8.277729, T: 2053282, Avg. loss: 6.952473
Total training time: 17.21 seconds.
-- Epoch 78
Norm: 0.77, NNZs: 2297, Bias: 8.271293, T: 2079948, Avg. loss: 6.946151
Total training time: 17.43 seconds.
-- Epoch 79
Norm: 0.77, NNZs: 2297, Bias: 8.264939, T: 2106614, Avg. loss: 6.939883
Total training time: 17.65 seconds.
-- Epoch 80
Norm: 0.77, NNZs: 2297, Bias: 8.258666, T: 2133280, Avg. loss: 6.933778
Total training time: 17.88 seconds.
-- Epoch 81
Norm: 0.77, NNZs: 2297, Bias: 8.252470, T: 2159946, Avg. loss: 6.927656
Total training time: 18.10 seconds.
-- Epoch 82
Norm: 0.77, NNZs: 2297, Bias: 8.246351, T: 2186612, Avg. loss: 6.921596
Total training time: 18.32 seconds.
-- Epoch 83
Norm: 0.77, NNZs: 2297, Bias: 8.240305, T: 2213278, Avg. loss: 6.915651
Total training time: 18.54 seconds.
```

```
-- Epoch 84
Norm: 0.77, NNZs: 2297, Bias: 8.234333, T: 2239944, Avg. loss: 6.909806
Total training time: 18.78 seconds.
-- Epoch 85
Norm: 0.77, NNZs: 2297, Bias: 8.228431, T: 2266610, Avg. loss: 6.903972
Total training time: 19.01 seconds.
-- Epoch 86
Norm: 0.77, NNZs: 2297, Bias: 8.222598, T: 2293276, Avg. loss: 6.898260
Total training time: 19.23 seconds.
-- Epoch 87
Norm: 0.77, NNZs: 2297, Bias: 8.216833, T: 2319942, Avg. loss: 6.892586
Total training time: 19.45 seconds.
-- Epoch 88
Norm: 0.77, NNZs: 2297, Bias: 8.211133, T: 2346608, Avg. loss: 6.886992
Total training time: 19.67 seconds.
-- Epoch 89
Norm: 0.77, NNZs: 2297, Bias: 8.205498, T: 2373274, Avg. loss: 6.881453
Total training time: 19.89 seconds.
-- Epoch 90
Norm: 0.77, NNZs: 2297, Bias: 8.199926, T: 2399940, Avg. loss: 6.875979
Total training time: 20.11 seconds.
-- Epoch 91
Norm: 0.77, NNZs: 2297, Bias: 8.194416, T: 2426606, Avg. loss: 6.870579
Total training time: 20.33 seconds.
-- Epoch 92
Norm: 0.77, NNZs: 2297, Bias: 8.188966, T: 2453272, Avg. loss: 6.865205
Total training time: 20.55 seconds.
-- Epoch 93
Norm: 0.77, NNZs: 2297, Bias: 8.183575, T: 2479938, Avg. loss: 6.859846
Total training time: 20.78 seconds.
-- Epoch 94
Norm: 0.77, NNZs: 2297, Bias: 8.178242, T: 2506604, Avg. loss: 6.854650
Total training time: 21.00 seconds.
-- Epoch 95
Norm: 0.77, NNZs: 2297, Bias: 8.172965, T: 2533270, Avg. loss: 6.849451
Total training time: 21.23 seconds.
-- Epoch 96
Norm: 0.77, NNZs: 2297, Bias: 8.167744, T: 2559936, Avg. loss: 6.844362
Total training time: 21.45 seconds.
-- Epoch 97
Norm: 0.77, NNZs: 2297, Bias: 8.162577, T: 2586602, Avg. loss: 6.839273
Total training time: 21.67 seconds.
-- Epoch 98
Norm: 0.77, NNZs: 2297, Bias: 8.157463, T: 2613268, Avg. loss: 6.834268
Total training time: 21.89 seconds.
-- Epoch 99
Norm: 0.77, NNZs: 2297, Bias: 8.152400, T: 2639934, Avg. loss: 6.829261
Total training time: 22.11 seconds.
-- Epoch 100
Norm: 0.77, NNZs: 2297, Bias: 8.147389, T: 2666600, Avg. loss: 6.824348
Total training time: 22.34 seconds.
-- Epoch 101
Norm: 0.77, NNZs: 2297, Bias: 8.142428, T: 2693266, Avg. loss: 6.819479
Total training time: 22.56 seconds.
-- Epoch 102
Norm: 0.77, NNZs: 2297, Bias: 8.137516, T: 2719932, Avg. loss: 6.814673
Total training time: 22.78 seconds.
-- Epoch 103
Norm: 0.77, NNZs: 2297, Bias: 8.132651, T: 2746598, Avg. loss: 6.809878
Total training time: 23.01 seconds.
-- Epoch 104
Norm: 0.77, NNZs: 2297, Bias: 8.127834, T: 2773264, Avg. loss: 6.805150
Total training time: 23.23 seconds.
-- Epoch 105
Norm: 0.77, NNZs: 2297, Bias: 8.123063, T: 2799930, Avg. loss: 6.800456
Total training time: 23.45 seconds.
-- Epoch 106
Norm: 0.77, NNZs: 2297, Bias: 8.118337, T: 2826596, Avg. loss: 6.795840
Total training time: 23.68 seconds.
-- Epoch 107
Norm: 0.77, NNZs: 2297, Bias: 8.113655, T: 2853262, Avg. loss: 6.791246
Total training time: 23.90 seconds.
-- Epoch 108
Norm: 0.77, NNZs: 2297, Bias: 8.109018, T: 2879928, Avg. loss: 6.786691
Total training time: 24.12 seconds.
-- Epoch 109
Norm: 0.77, NNZs: 2297, Bias: 8.104423, T: 2906594, Avg. loss: 6.782173
```

```
Total training time: 24.34 seconds.
-- Epoch 110
Norm: 0.77, NNZs: 2297, Bias: 8.099870, T: 2933260, Avg. loss: 6.777705
Total training time: 24.56 seconds.
-- Epoch 111
Norm: 0.77, NNZs: 2297, Bias: 8.095358, T: 2959926, Avg. loss: 6.773267
Total training time: 24.79 seconds.
-- Epoch 112
Norm: 0.77, NNZs: 2297, Bias: 8.090887, T: 2986592, Avg. loss: 6.768889
Total training time: 25.01 seconds.
-- Epoch 113
Norm: 0.77, NNZs: 2297, Bias: 8.086455, T: 3013258, Avg. loss: 6.764563
Total training time: 25.23 seconds.
-- Epoch 114
Norm: 0.77, NNZs: 2297, Bias: 8.082063, T: 3039924, Avg. loss: 6.760224
Total training time: 25.46 seconds.
-- Epoch 115
Norm: 0.77, NNZs: 2297, Bias: 8.077709, T: 3066590, Avg. loss: 6.755982
Total training time: 25.70 seconds.
-- Epoch 116
Norm: 0.77, NNZs: 2297, Bias: 8.073393, T: 3093256, Avg. loss: 6.751714
Total training time: 25.93 seconds.
-- Epoch 117
Norm: 0.77, NNZs: 2297, Bias: 8.069114, T: 3119922, Avg. loss: 6.747560
Total training time: 26.15 seconds.
-- Epoch 118
Norm: 0.77, NNZs: 2297, Bias: 8.064871, T: 3146588, Avg. loss: 6.743371
Total training time: 26.37 seconds.
-- Epoch 119
Norm: 0.77, NNZs: 2297, Bias: 8.060664, T: 3173254, Avg. loss: 6.739227
Total training time: 26.60 seconds.
-- Epoch 120
Norm: 0.77, NNZs: 2297, Bias: 8.056493, T: 3199920, Avg. loss: 6.735135
Total training time: 26.82 seconds.
-- Epoch 121
Norm: 0.77, NNZs: 2297, Bias: 8.052356, T: 3226586, Avg. loss: 6.731082
Total training time: 27.04 seconds.
-- Epoch 122
Norm: 0.77, NNZs: 2297, Bias: 8.048253, T: 3253252, Avg. loss: 6.727075
Total training time: 27.26 seconds.
-- Epoch 123
Norm: 0.77, NNZs: 2297, Bias: 8.044184, T: 3279918, Avg. loss: 6.723082
Total training time: 27.48 seconds.
-- Epoch 124
Norm: 0.77, NNZs: 2297, Bias: 8.040148, T: 3306584, Avg. loss: 6.719112
Total training time: 27.71 seconds.
-- Epoch 125
Norm: 0.77, NNZs: 2297, Bias: 8.036144, T: 3333250, Avg. loss: 6.715189
Total training time: 27.93 seconds.
-- Epoch 126
Norm: 0.77, NNZs: 2297, Bias: 8.032172, T: 3359916, Avg. loss: 6.711319
Total training time: 28.15 seconds.
-- Epoch 127
Norm: 0.77, NNZs: 2297, Bias: 8.028232, T: 3386582, Avg. loss: 6.707433
Total training time: 28.37 seconds.
-- Epoch 128
Norm: 0.77, NNZs: 2297, Bias: 8.024322, T: 3413248, Avg. loss: 6.703621
Total training time: 28.59 seconds.
-- Epoch 129
Norm: 0.77, NNZs: 2297, Bias: 8.020443, T: 3439914, Avg. loss: 6.699832
Total training time: 28.82 seconds.
-- Epoch 130
Norm: 0.77, NNZs: 2297, Bias: 8.016594, T: 3466580, Avg. loss: 6.696056
Total training time: 29.05 seconds.
-- Epoch 131
Norm: 0.77, NNZs: 2297, Bias: 8.012775, T: 3493246, Avg. loss: 6.692281
Total training time: 29.27 seconds.
-- Epoch 132
Norm: 0.77, NNZs: 2297, Bias: 8.008985, T: 3519912, Avg. loss: 6.688590
Total training time: 29.49 seconds.
-- Epoch 133
Norm: 0.77, NNZs: 2297, Bias: 8.005223, T: 3546578, Avg. loss: 6.684901
Total training time: 29.71 seconds.
-- Epoch 134
Norm: 0.77, NNZs: 2297, Bias: 8.001490, T: 3573244, Avg. loss: 6.681233
Total training time: 29.94 seconds.
-- Epoch 135
```

```
Norm: 0.77, NNZs: 2297, Bias: 7.997784, T: 3599910, Avg. loss: 6.677604
Total training time: 30.16 seconds.
-- Epoch 136
Norm: 0.77, NNZs: 2297, Bias: 7.994106, T: 3626576, Avg. loss: 6.674010
Total training time: 30.38 seconds.
-- Epoch 137
Norm: 0.77, NNZs: 2297, Bias: 7.990454, T: 3653242, Avg. loss: 6.670416
Total training time: 30.60 seconds.
-- Epoch 138
Norm: 0.77, NNZs: 2297, Bias: 7.986830, T: 3679908, Avg. loss: 6.666876
Total training time: 30.83 seconds.
-- Epoch 139
Norm: 0.77, NNZs: 2297, Bias: 7.983231, T: 3706574, Avg. loss: 6.663348
Total training time: 31.05 seconds.
-- Epoch 140
Norm: 0.77, NNZs: 2297, Bias: 7.979658, T: 3733240, Avg. loss: 6.659844
Total training time: 31.27 seconds.
-- Epoch 141
Norm: 0.77, NNZs: 2297, Bias: 7.976111, T: 3759906, Avg. loss: 6.656345
Total training time: 31.49 seconds.
-- Epoch 142
Norm: 0.77, NNZs: 2297, Bias: 7.972589, T: 3786572, Avg. loss: 6.652909
Total training time: 31.71 seconds.
-- Epoch 143
Norm: 0.77, NNZs: 2297, Bias: 7.969092, T: 3813238, Avg. loss: 6.649509
Total training time: 31.93 seconds.
-- Epoch 144
Norm: 0.77, NNZs: 2297, Bias: 7.965619, T: 3839904, Avg. loss: 6.646122
Total training time: 32.16 seconds.
-- Epoch 145
Norm: 0.77, NNZs: 2297, Bias: 7.962170, T: 3866570, Avg. loss: 6.642703
Total training time: 32.38 seconds.
-- Epoch 146
Norm: 0.77, NNZs: 2297, Bias: 7.958745, T: 3893236, Avg. loss: 6.639385
Total training time: 32.60 seconds.
-- Epoch 147
Norm: 0.77, NNZs: 2297, Bias: 7.955343, T: 3919902, Avg. loss: 6.636028
Total training time: 32.83 seconds.
-- Epoch 148
Norm: 0.77, NNZs: 2297, Bias: 7.951965, T: 3946568, Avg. loss: 6.632717
Total training time: 33.05 seconds.
-- Epoch 149
Norm: 0.77, NNZs: 2297, Bias: 7.948609, T: 3973234, Avg. loss: 6.629442
Total training time: 33.27 seconds.
-- Epoch 150
Norm: 0.77, NNZs: 2297, Bias: 7.945275, T: 3999900, Avg. loss: 6.626179
Total training time: 33.49 seconds.
-- Epoch 151
Norm: 0.77, NNZs: 2297, Bias: 7.941964, T: 4026566, Avg. loss: 6.622922
Total training time: 33.71 seconds.
-- Epoch 152
Norm: 0.77, NNZs: 2297, Bias: 7.938675, T: 4053232, Avg. loss: 6.619727
Total training time: 33.94 seconds.
-- Epoch 153
Norm: 0.77, NNZs: 2297, Bias: 7.935407, T: 4079898, Avg. loss: 6.616517
Total training time: 34.16 seconds.
-- Epoch 154
Norm: 0.77, NNZs: 2297, Bias: 7.932161, T: 4106564, Avg. loss: 6.613335
Total training time: 34.38 seconds.
-- Epoch 155
Norm: 0.77, NNZs: 2297, Bias: 7.928936, T: 4133230, Avg. loss: 6.610175
Total training time: 34.60 seconds.
-- Epoch 156
Norm: 0.77, NNZs: 2297, Bias: 7.925731, T: 4159896, Avg. loss: 6.607021
Total training time: 34.83 seconds.
-- Epoch 157
Norm: 0.77, NNZs: 2297, Bias: 7.922547, T: 4186562, Avg. loss: 6.603912
Total training time: 35.05 seconds.
-- Epoch 158
Norm: 0.77, NNZs: 2297, Bias: 7.919383, T: 4213228, Avg. loss: 6.600813
Total training time: 35.27 seconds.
-- Epoch 159
Norm: 0.77, NNZs: 2297, Bias: 7.916239, T: 4239894, Avg. loss: 6.597760
Total training time: 35.50 seconds.
-- Epoch 160
Norm: 0.77, NNZs: 2297, Bias: 7.913115, T: 4266560, Avg. loss: 6.594710
Total training time: 35.72 seconds.
```

```
-- Epoch 161
Norm: 0.77, NNZs: 2297, Bias: 7.910011, T: 4293226, Avg. loss: 6.591643
Total training time: 35.96 seconds.
-- Epoch 162
Norm: 0.77, NNZs: 2297, Bias: 7.906926, T: 4319892, Avg. loss: 6.588596
Total training time: 36.18 seconds.
-- Epoch 163
Norm: 0.77, NNZs: 2297, Bias: 7.903859, T: 4346558, Avg. loss: 6.585634
Total training time: 36.40 seconds.
-- Epoch 164
Norm: 0.77, NNZs: 2297, Bias: 7.900812, T: 4373224, Avg. loss: 6.582644
Total training time: 36.63 seconds.
-- Epoch 165
Norm: 0.77, NNZs: 2297, Bias: 7.897783, T: 4399890, Avg. loss: 6.579705
Total training time: 36.85 seconds.
-- Epoch 166
Norm: 0.77, NNZs: 2297, Bias: 7.894772, T: 4426556, Avg. loss: 6.576749
Total training time: 37.08 seconds.
-- Epoch 167
Norm: 0.77, NNZs: 2297, Bias: 7.891780, T: 4453222, Avg. loss: 6.573802
Total training time: 37.30 seconds.
-- Epoch 168
Norm: 0.77, NNZs: 2297, Bias: 7.888805, T: 4479888, Avg. loss: 6.570889
Total training time: 37.52 seconds.
-- Epoch 169
Norm: 0.77, NNZs: 2297, Bias: 7.885848, T: 4506554, Avg. loss: 6.567980
Total training time: 37.75 seconds.
-- Epoch 170
Norm: 0.77, NNZs: 2297, Bias: 7.882909, T: 4533220, Avg. loss: 6.565139
Total training time: 37.97 seconds.
-- Epoch 171
Norm: 0.77, NNZs: 2297, Bias: 7.879986, T: 4559886, Avg. loss: 6.562257
Total training time: 38.19 seconds.
-- Epoch 172
Norm: 0.77, NNZs: 2297, Bias: 7.877081, T: 4586552, Avg. loss: 6.559447
Total training time: 38.41 seconds.
-- Epoch 173
Norm: 0.77, NNZs: 2297, Bias: 7.874193, T: 4613218, Avg. loss: 6.556599
Total training time: 38.63 seconds.
-- Epoch 174
Norm: 0.77, NNZs: 2297, Bias: 7.871321, T: 4639884, Avg. loss: 6.553792
Total training time: 38.86 seconds.
-- Epoch 175
Norm: 0.77, NNZs: 2297, Bias: 7.868466, T: 4666550, Avg. loss: 6.551011
Total training time: 39.08 seconds.
-- Epoch 176
Norm: 0.77, NNZs: 2297, Bias: 7.865628, T: 4693216, Avg. loss: 6.548231
Total training time: 39.31 seconds.
-- Epoch 177
Norm: 0.77, NNZs: 2297, Bias: 7.862805, T: 4719882, Avg. loss: 6.545477
Total training time: 39.54 seconds.
-- Epoch 178
Norm: 0.77, NNZs: 2297, Bias: 7.859998, T: 4746548, Avg. loss: 6.542721
Total training time: 39.76 seconds.
-- Epoch 179
Norm: 0.77, NNZs: 2297, Bias: 7.857207, T: 4773214, Avg. loss: 6.540007
Total training time: 39.98 seconds.
-- Epoch 180
Norm: 0.77, NNZs: 2297, Bias: 7.854432, T: 4799880, Avg. loss: 6.537294
Total training time: 40.20 seconds.
-- Epoch 181
Norm: 0.77, NNZs: 2297, Bias: 7.851672, T: 4826546, Avg. loss: 6.534585
Total training time: 40.42 seconds.
-- Epoch 182
Norm: 0.77, NNZs: 2297, Bias: 7.848927, T: 4853212, Avg. loss: 6.531893
Total training time: 40.64 seconds.
-- Epoch 183
Norm: 0.77, NNZs: 2297, Bias: 7.846197, T: 4879878, Avg. loss: 6.529218
Total training time: 40.87 seconds.
-- Epoch 184
Norm: 0.77, NNZs: 2297, Bias: 7.843482, T: 4906544, Avg. loss: 6.526581
Total training time: 41.09 seconds.
-- Epoch 185
Norm: 0.77, NNZs: 2297, Bias: 7.840782, T: 4933210, Avg. loss: 6.523924
Total training time: 41.31 seconds.
-- Epoch 186
Norm: 0.77, NNZs: 2297, Bias: 7.838097, T: 4959876, Avg. loss: 6.521312
```

```
Norm: 0.77, NNZs: 2297, Bias: [illegible], T: [illegible], Avg. loss: [illegible]
Total training time: 41.53 seconds.
-- Epoch 187
Norm: 0.77, NNZs: 2297, Bias: 7.835426, T: 4986542, Avg. loss: 6.518690
Total training time: 41.76 seconds.
-- Epoch 188
Norm: 0.77, NNZs: 2297, Bias: 7.832769, T: 5013208, Avg. loss: 6.516091
Total training time: 41.98 seconds.
-- Epoch 189
Norm: 0.77, NNZs: 2297, Bias: 7.830126, T: 5039874, Avg. loss: 6.513507
Total training time: 42.20 seconds.
-- Epoch 190
Norm: 0.77, NNZs: 2297, Bias: 7.827497, T: 5066540, Avg. loss: 6.510934
Total training time: 42.42 seconds.
-- Epoch 191
Norm: 0.77, NNZs: 2297, Bias: 7.824882, T: 5093206, Avg. loss: 6.508396
Total training time: 42.64 seconds.
-- Epoch 192
Norm: 0.77, NNZs: 2297, Bias: 7.822281, T: 5119872, Avg. loss: 6.505851
Total training time: 42.87 seconds.
-- Epoch 193
Norm: 0.77, NNZs: 2297, Bias: 7.819693, T: 5146538, Avg. loss: 6.503302
Total training time: 43.10 seconds.
-- Epoch 194
Norm: 0.77, NNZs: 2297, Bias: 7.817119, T: 5173204, Avg. loss: 6.500794
Total training time: 43.32 seconds.
-- Epoch 195
Norm: 0.77, NNZs: 2297, Bias: 7.814558, T: 5199870, Avg. loss: 6.498301
Total training time: 43.55 seconds.
-- Epoch 196
Norm: 0.77, NNZs: 2297, Bias: 7.812010, T: 5226536, Avg. loss: 6.495789
Total training time: 43.77 seconds.
-- Epoch 197
Norm: 0.77, NNZs: 2297, Bias: 7.809475, T: 5253202, Avg. loss: 6.493311
Total training time: 43.99 seconds.
-- Epoch 198
Norm: 0.77, NNZs: 2297, Bias: 7.806953, T: 5279868, Avg. loss: 6.490844
Total training time: 44.22 seconds.
-- Epoch 199
Norm: 0.77, NNZs: 2297, Bias: 7.804444, T: 5306534, Avg. loss: 6.488403
Total training time: 44.43 seconds.
-- Epoch 200
Norm: 0.77, NNZs: 2297, Bias: 7.801947, T: 5333200, Avg. loss: 6.485958
Total training time: 44.66 seconds.
-- Epoch 201
Norm: 0.77, NNZs: 2297, Bias: 7.799463, T: 5359866, Avg. loss: 6.483543
Total training time: 44.88 seconds.
-- Epoch 202
Norm: 0.77, NNZs: 2297, Bias: 7.796991, T: 5386532, Avg. loss: 6.481126
Total training time: 45.10 seconds.
-- Epoch 203
Norm: 0.77, NNZs: 2297, Bias: 7.794531, T: 5413198, Avg. loss: 6.478719
Total training time: 45.32 seconds.
-- Epoch 204
Norm: 0.77, NNZs: 2297, Bias: 7.792083, T: 5439864, Avg. loss: 6.476329
Total training time: 45.54 seconds.
-- Epoch 205
Norm: 0.77, NNZs: 2297, Bias: 7.789648, T: 5466530, Avg. loss: 6.473948
Total training time: 45.77 seconds.
-- Epoch 206
Norm: 0.77, NNZs: 2297, Bias: 7.787224, T: 5493196, Avg. loss: 6.471593
Total training time: 46.00 seconds.
-- Epoch 207
Norm: 0.77, NNZs: 2297, Bias: 7.784812, T: 5519862, Avg. loss: 6.469215
Total training time: 46.23 seconds.
-- Epoch 208
Norm: 0.77, NNZs: 2297, Bias: 7.782412, T: 5546528, Avg. loss: 6.466874
Total training time: 46.45 seconds.
-- Epoch 209
Norm: 0.77, NNZs: 2297, Bias: 7.780023, T: 5573194, Avg. loss: 6.464528
Total training time: 46.68 seconds.
-- Epoch 210
Norm: 0.77, NNZs: 2297, Bias: 7.777645, T: 5599860, Avg. loss: 6.462208
Total training time: 46.90 seconds.
-- Epoch 211
Norm: 0.77, NNZs: 2297, Bias: 7.775279, T: 5626526, Avg. loss: 6.459897
Total training time: 47.13 seconds.
-- Epoch 212
```

-- Epoch 212
Norm: 0.77, NNZs: 2297, Bias: 7.772924, T: 5653192, Avg. loss: 6.457605
Total training time: 47.35 seconds.
-- Epoch 213
Norm: 0.77, NNZs: 2297, Bias: 7.770581, T: 5679858, Avg. loss: 6.455322
Total training time: 47.57 seconds.
-- Epoch 214
Norm: 0.77, NNZs: 2297, Bias: 7.768248, T: 5706524, Avg. loss: 6.453039
Total training time: 47.79 seconds.
-- Epoch 215
Norm: 0.77, NNZs: 2297, Bias: 7.765926, T: 5733190, Avg. loss: 6.450764
Total training time: 48.02 seconds.
-- Epoch 216
Norm: 0.77, NNZs: 2297, Bias: 7.763615, T: 5759856, Avg. loss: 6.448504
Total training time: 48.24 seconds.
-- Epoch 217
Norm: 0.77, NNZs: 2297, Bias: 7.761315, T: 5786522, Avg. loss: 6.446259
Total training time: 48.46 seconds.
-- Epoch 218
Norm: 0.77, NNZs: 2297, Bias: 7.759025, T: 5813188, Avg. loss: 6.444035
Total training time: 48.68 seconds.
-- Epoch 219
Norm: 0.77, NNZs: 2297, Bias: 7.756745, T: 5839854, Avg. loss: 6.441802
Total training time: 48.90 seconds.
-- Epoch 220
Norm: 0.77, NNZs: 2297, Bias: 7.754477, T: 5866520, Avg. loss: 6.439601
Total training time: 49.13 seconds.
-- Epoch 221
Norm: 0.77, NNZs: 2297, Bias: 7.752218, T: 5893186, Avg. loss: 6.437372
Total training time: 49.35 seconds.
-- Epoch 222
Norm: 0.77, NNZs: 2297, Bias: 7.749970, T: 5919852, Avg. loss: 6.435171
Total training time: 49.58 seconds.
-- Epoch 223
Norm: 0.77, NNZs: 2297, Bias: 7.747731, T: 5946518, Avg. loss: 6.432997
Total training time: 49.81 seconds.
-- Epoch 224
Norm: 0.77, NNZs: 2297, Bias: 7.745503, T: 5973184, Avg. loss: 6.430812
Total training time: 50.03 seconds.
-- Epoch 225
Norm: 0.77, NNZs: 2297, Bias: 7.743285, T: 5999850, Avg. loss: 6.428645
Total training time: 50.25 seconds.
-- Epoch 226
Norm: 0.77, NNZs: 2297, Bias: 7.741076, T: 6026516, Avg. loss: 6.426504
Total training time: 50.47 seconds.
-- Epoch 227
Norm: 0.77, NNZs: 2297, Bias: 7.738878, T: 6053182, Avg. loss: 6.424337
Total training time: 50.69 seconds.
-- Epoch 228
Norm: 0.77, NNZs: 2297, Bias: 7.736689, T: 6079848, Avg. loss: 6.422216
Total training time: 50.92 seconds.
-- Epoch 229
Norm: 0.77, NNZs: 2297, Bias: 7.734509, T: 6106514, Avg. loss: 6.420087
Total training time: 51.15 seconds.
-- Epoch 230
Norm: 0.77, NNZs: 2297, Bias: 7.732340, T: 6133180, Avg. loss: 6.417945
Total training time: 51.37 seconds.
-- Epoch 231
Norm: 0.77, NNZs: 2297, Bias: 7.730179, T: 6159846, Avg. loss: 6.415837
Total training time: 51.59 seconds.
-- Epoch 232
Norm: 0.77, NNZs: 2297, Bias: 7.728028, T: 6186512, Avg. loss: 6.413751
Total training time: 51.81 seconds.
-- Epoch 233
Norm: 0.77, NNZs: 2297, Bias: 7.725886, T: 6213178, Avg. loss: 6.411655
Total training time: 52.03 seconds.
-- Epoch 234
Norm: 0.77, NNZs: 2297, Bias: 7.723754, T: 6239844, Avg. loss: 6.409586
Total training time: 52.25 seconds.
-- Epoch 235
Norm: 0.77, NNZs: 2297, Bias: 7.721630, T: 6266510, Avg. loss: 6.407504
Total training time: 52.47 seconds.
-- Epoch 236
Norm: 0.77, NNZs: 2297, Bias: 7.719516, T: 6293176, Avg. loss: 6.405431
Total training time: 52.70 seconds.
-- Epoch 237
Norm: 0.77, NNZs: 2297, Bias: 7.717410, T: 6319842, Avg. loss: 6.403381
Total training time: 52.92 seconds.

```
Total training time: 52.92 seconds.
-- Epoch 238
Norm: 0.77, NNZs: 2297, Bias: 7.715314, T: 6346508, Avg. loss: 6.401328
Total training time: 53.15 seconds.
-- Epoch 239
Norm: 0.77, NNZs: 2297, Bias: 7.713226, T: 6373174, Avg. loss: 6.399304
Total training time: 53.37 seconds.
-- Epoch 240
Norm: 0.77, NNZs: 2297, Bias: 7.711147, T: 6399840, Avg. loss: 6.397277
Total training time: 53.60 seconds.
-- Epoch 241
Norm: 0.77, NNZs: 2297, Bias: 7.709076, T: 6426506, Avg. loss: 6.395259
Total training time: 53.82 seconds.
-- Epoch 242
Norm: 0.77, NNZs: 2297, Bias: 7.707014, T: 6453172, Avg. loss: 6.393232
Total training time: 54.05 seconds.
-- Epoch 243
Norm: 0.77, NNZs: 2297, Bias: 7.704961, T: 6479838, Avg. loss: 6.391216
Total training time: 54.27 seconds.
-- Epoch 244
Norm: 0.77, NNZs: 2297, Bias: 7.702916, T: 6506504, Avg. loss: 6.389219
Total training time: 54.50 seconds.
-- Epoch 245
Norm: 0.77, NNZs: 2297, Bias: 7.700880, T: 6533170, Avg. loss: 6.387239
Total training time: 54.72 seconds.
-- Epoch 246
Norm: 0.77, NNZs: 2297, Bias: 7.698851, T: 6559836, Avg. loss: 6.385251
Total training time: 54.95 seconds.
-- Epoch 247
Norm: 0.77, NNZs: 2297, Bias: 7.696832, T: 6586502, Avg. loss: 6.383290
Total training time: 55.17 seconds.
-- Epoch 248
Norm: 0.77, NNZs: 2297, Bias: 7.694820, T: 6613168, Avg. loss: 6.381321
Total training time: 55.39 seconds.
-- Epoch 249
Norm: 0.77, NNZs: 2297, Bias: 7.692816, T: 6639834, Avg. loss: 6.379376
Total training time: 55.62 seconds.
-- Epoch 250
Norm: 0.77, NNZs: 2297, Bias: 7.690820, T: 6666500, Avg. loss: 6.377427
Total training time: 55.85 seconds.
-- Epoch 251
Norm: 0.77, NNZs: 2297, Bias: 7.688833, T: 6693166, Avg. loss: 6.375475
Total training time: 56.07 seconds.
-- Epoch 252
Norm: 0.77, NNZs: 2297, Bias: 7.686853, T: 6719832, Avg. loss: 6.373547
Total training time: 56.29 seconds.
-- Epoch 253
Norm: 0.77, NNZs: 2297, Bias: 7.684881, T: 6746498, Avg. loss: 6.371619
Total training time: 56.53 seconds.
-- Epoch 254
Norm: 0.77, NNZs: 2297, Bias: 7.682917, T: 6773164, Avg. loss: 6.369701
Total training time: 56.75 seconds.
-- Epoch 255
Norm: 0.77, NNZs: 2297, Bias: 7.680961, T: 6799830, Avg. loss: 6.367802
Total training time: 56.97 seconds.
-- Epoch 256
Norm: 0.77, NNZs: 2297, Bias: 7.679012, T: 6826496, Avg. loss: 6.365907
Total training time: 57.20 seconds.
-- Epoch 257
Norm: 0.77, NNZs: 2297, Bias: 7.677071, T: 6853162, Avg. loss: 6.364005
Total training time: 57.42 seconds.
-- Epoch 258
Norm: 0.77, NNZs: 2297, Bias: 7.675138, T: 6879828, Avg. loss: 6.362108
Total training time: 57.64 seconds.
-- Epoch 259
Norm: 0.77, NNZs: 2297, Bias: 7.673212, T: 6906494, Avg. loss: 6.360240
Total training time: 57.86 seconds.
-- Epoch 260
Norm: 0.77, NNZs: 2297, Bias: 7.671293, T: 6933160, Avg. loss: 6.358353
Total training time: 58.08 seconds.
-- Epoch 261
Norm: 0.77, NNZs: 2297, Bias: 7.669382, T: 6959826, Avg. loss: 6.356502
Total training time: 58.30 seconds.
-- Epoch 262
Norm: 0.77, NNZs: 2297, Bias: 7.667478, T: 6986492, Avg. loss: 6.354653
Total training time: 58.53 seconds.
-- Epoch 263
Norm: 0.77, NNZs: 2297, Bias: 7.665581, T: 7013158, Avg. loss: 6.352788
```

```
Norm: 0.77, NNZs: 2297, Bias: 7.665581, T: 7013158, Avg. loss: 6.352788
Total training time: 58.75 seconds.
-- Epoch 264
Norm: 0.77, NNZs: 2297, Bias: 7.663692, T: 7039824, Avg. loss: 6.350955
Total training time: 58.97 seconds.
-- Epoch 265
Norm: 0.77, NNZs: 2297, Bias: 7.661810, T: 7066490, Avg. loss: 6.349111
Total training time: 59.21 seconds.
-- Epoch 266
Norm: 0.77, NNZs: 2297, Bias: 7.659935, T: 7093156, Avg. loss: 6.347276
Total training time: 59.43 seconds.
-- Epoch 267
Norm: 0.77, NNZs: 2297, Bias: 7.658067, T: 7119822, Avg. loss: 6.345449
Total training time: 59.66 seconds.
-- Epoch 268
Norm: 0.77, NNZs: 2297, Bias: 7.656205, T: 7146488, Avg. loss: 6.343640
Total training time: 59.88 seconds.
-- Epoch 269
Norm: 0.77, NNZs: 2297, Bias: 7.654351, T: 7173154, Avg. loss: 6.341817
Total training time: 60.10 seconds.
-- Epoch 270
Norm: 0.77, NNZs: 2297, Bias: 7.652504, T: 7199820, Avg. loss: 6.340030
Total training time: 60.33 seconds.
-- Epoch 271
Norm: 0.77, NNZs: 2297, Bias: 7.650664, T: 7226486, Avg. loss: 6.338231
Total training time: 60.55 seconds.
-- Epoch 272
Norm: 0.77, NNZs: 2297, Bias: 7.648830, T: 7253152, Avg. loss: 6.336447
Total training time: 60.78 seconds.
-- Epoch 273
Norm: 0.77, NNZs: 2297, Bias: 7.647003, T: 7279818, Avg. loss: 6.334665
Total training time: 61.00 seconds.
-- Epoch 274
Norm: 0.77, NNZs: 2297, Bias: 7.645183, T: 7306484, Avg. loss: 6.332890
Total training time: 61.22 seconds.
-- Epoch 275
Norm: 0.77, NNZs: 2297, Bias: 7.643370, T: 7333150, Avg. loss: 6.331109
Total training time: 61.44 seconds.
-- Epoch 276
Norm: 0.77, NNZs: 2297, Bias: 7.641563, T: 7359816, Avg. loss: 6.329362
Total training time: 61.67 seconds.
-- Epoch 277
Norm: 0.77, NNZs: 2297, Bias: 7.639762, T: 7386482, Avg. loss: 6.327598
Total training time: 61.89 seconds.
-- Epoch 278
Norm: 0.77, NNZs: 2297, Bias: 7.637968, T: 7413148, Avg. loss: 6.325844
Total training time: 62.11 seconds.
-- Epoch 279
Norm: 0.77, NNZs: 2297, Bias: 7.636181, T: 7439814, Avg. loss: 6.324100
Total training time: 62.34 seconds.
-- Epoch 280
Norm: 0.77, NNZs: 2297, Bias: 7.634400, T: 7466480, Avg. loss: 6.322368
Total training time: 62.56 seconds.
-- Epoch 281
Norm: 0.77, NNZs: 2297, Bias: 7.632625, T: 7493146, Avg. loss: 6.320640
Total training time: 62.78 seconds.
-- Epoch 282
Norm: 0.77, NNZs: 2297, Bias: 7.630857, T: 7519812, Avg. loss: 6.318907
Total training time: 63.01 seconds.
-- Epoch 283
Norm: 0.77, NNZs: 2297, Bias: 7.629095, T: 7546478, Avg. loss: 6.317194
Total training time: 63.23 seconds.
-- Epoch 284
Norm: 0.77, NNZs: 2297, Bias: 7.627339, T: 7573144, Avg. loss: 6.315487
Total training time: 63.46 seconds.
-- Epoch 285
Norm: 0.77, NNZs: 2297, Bias: 7.625589, T: 7599810, Avg. loss: 6.313767
Total training time: 63.68 seconds.
-- Epoch 286
Norm: 0.77, NNZs: 2297, Bias: 7.623846, T: 7626476, Avg. loss: 6.312085
Total training time: 63.90 seconds.
-- Epoch 287
Norm: 0.77, NNZs: 2297, Bias: 7.622108, T: 7653142, Avg. loss: 6.310371
Total training time: 64.12 seconds.
-- Epoch 288
Norm: 0.77, NNZs: 2297, Bias: 7.620377, T: 7679808, Avg. loss: 6.308694
Total training time: 64.35 seconds.
-- Epoch 289
```

```
-- Epoch 289
Norm: 0.77, NNZs: 2297, Bias: 7.618652, T: 7706474, Avg. loss: 6.307010
Total training time: 64.58 seconds.
-- Epoch 290
Norm: 0.77, NNZs: 2297, Bias: 7.616932, T: 7733140, Avg. loss: 6.305331
Total training time: 64.80 seconds.
-- Epoch 291
Norm: 0.77, NNZs: 2297, Bias: 7.615219, T: 7759806, Avg. loss: 6.303660
Total training time: 65.02 seconds.
-- Epoch 292
Norm: 0.77, NNZs: 2297, Bias: 7.613511, T: 7786472, Avg. loss: 6.301988
Total training time: 65.24 seconds.
-- Epoch 293
Norm: 0.77, NNZs: 2297, Bias: 7.611809, T: 7813138, Avg. loss: 6.300326
Total training time: 65.46 seconds.
-- Epoch 294
Norm: 0.77, NNZs: 2297, Bias: 7.610114, T: 7839804, Avg. loss: 6.298683
Total training time: 65.69 seconds.
-- Epoch 295
Norm: 0.77, NNZs: 2297, Bias: 7.608423, T: 7866470, Avg. loss: 6.297033
Total training time: 65.90 seconds.
-- Epoch 296
Norm: 0.77, NNZs: 2297, Bias: 7.606739, T: 7893136, Avg. loss: 6.295395
Total training time: 66.13 seconds.
-- Epoch 297
Norm: 0.77, NNZs: 2297, Bias: 7.605060, T: 7919802, Avg. loss: 6.293738
Total training time: 66.36 seconds.
-- Epoch 298
Norm: 0.77, NNZs: 2297, Bias: 7.603387, T: 7946468, Avg. loss: 6.292113
Total training time: 66.59 seconds.
-- Epoch 299
Norm: 0.77, NNZs: 2297, Bias: 7.601720, T: 7973134, Avg. loss: 6.290491
Total training time: 66.82 seconds.
-- Epoch 300
Norm: 0.77, NNZs: 2297, Bias: 7.600058, T: 7999800, Avg. loss: 6.288867
Total training time: 67.04 seconds.
-- Epoch 301
Norm: 0.77, NNZs: 2297, Bias: 7.598402, T: 8026466, Avg. loss: 6.287254
Total training time: 67.27 seconds.
-- Epoch 302
Norm: 0.77, NNZs: 2297, Bias: 7.596751, T: 8053132, Avg. loss: 6.285660
Total training time: 67.49 seconds.
-- Epoch 303
Norm: 0.77, NNZs: 2297, Bias: 7.595105, T: 8079798, Avg. loss: 6.284044
Total training time: 67.71 seconds.
-- Epoch 304
Norm: 0.77, NNZs: 2297, Bias: 7.593466, T: 8106464, Avg. loss: 6.282452
Total training time: 67.93 seconds.
-- Epoch 305
Norm: 0.77, NNZs: 2297, Bias: 7.591831, T: 8133130, Avg. loss: 6.280853
Total training time: 68.15 seconds.
-- Epoch 306
Norm: 0.77, NNZs: 2297, Bias: 7.590202, T: 8159796, Avg. loss: 6.279266
Total training time: 68.38 seconds.
-- Epoch 307
Norm: 0.77, NNZs: 2297, Bias: 7.588578, T: 8186462, Avg. loss: 6.277675
Total training time: 68.61 seconds.
-- Epoch 308
Norm: 0.77, NNZs: 2297, Bias: 7.586959, T: 8213128, Avg. loss: 6.276100
Total training time: 68.83 seconds.
-- Epoch 309
Norm: 0.77, NNZs: 2297, Bias: 7.585346, T: 8239794, Avg. loss: 6.274539
Total training time: 69.05 seconds.
-- Epoch 310
Norm: 0.77, NNZs: 2297, Bias: 7.583738, T: 8266460, Avg. loss: 6.272967
Total training time: 69.28 seconds.
-- Epoch 311
Norm: 0.77, NNZs: 2297, Bias: 7.582135, T: 8293126, Avg. loss: 6.271397
Total training time: 69.50 seconds.
-- Epoch 312
Norm: 0.77, NNZs: 2297, Bias: 7.580538, T: 8319792, Avg. loss: 6.269842
Total training time: 69.72 seconds.
-- Epoch 313
Norm: 0.77, NNZs: 2297, Bias: 7.578945, T: 8346458, Avg. loss: 6.268300
Total training time: 69.94 seconds.
-- Epoch 314
Norm: 0.77, NNZs: 2297, Bias: 7.577357, T: 8373124, Avg. loss: 6.266738
Total training time: 70.17 seconds.
```

```
Total training time: 70.17 seconds.
-- Epoch 315
Norm: 0.77, NNZs: 2297, Bias: 7.575775, T: 8399790, Avg. loss: 6.265205
Total training time: 70.39 seconds.
-- Epoch 316
Norm: 0.77, NNZs: 2297, Bias: 7.574198, T: 8426456, Avg. loss: 6.263668
Total training time: 70.61 seconds.
-- Epoch 317
Norm: 0.77, NNZs: 2297, Bias: 7.572625, T: 8453122, Avg. loss: 6.262140
Total training time: 70.84 seconds.
-- Epoch 318
Norm: 0.77, NNZs: 2297, Bias: 7.571058, T: 8479788, Avg. loss: 6.260602
Total training time: 71.07 seconds.
-- Epoch 319
Norm: 0.77, NNZs: 2297, Bias: 7.569495, T: 8506454, Avg. loss: 6.259081
Total training time: 71.30 seconds.
-- Epoch 320
Norm: 0.77, NNZs: 2297, Bias: 7.567938, T: 8533120, Avg. loss: 6.257557
Total training time: 71.52 seconds.
-- Epoch 321
Norm: 0.77, NNZs: 2297, Bias: 7.566385, T: 8559786, Avg. loss: 6.256040
Total training time: 71.75 seconds.
-- Epoch 322
Norm: 0.77, NNZs: 2297, Bias: 7.564837, T: 8586452, Avg. loss: 6.254540
Total training time: 71.96 seconds.
-- Epoch 323
Norm: 0.77, NNZs: 2297, Bias: 7.563294, T: 8613118, Avg. loss: 6.253041
Total training time: 72.19 seconds.
-- Epoch 324
Norm: 0.77, NNZs: 2297, Bias: 7.561755, T: 8639784, Avg. loss: 6.251537
Total training time: 72.41 seconds.
-- Epoch 325
Norm: 0.77, NNZs: 2297, Bias: 7.560222, T: 8666450, Avg. loss: 6.250040
Total training time: 72.63 seconds.
-- Epoch 326
Norm: 0.77, NNZs: 2297, Bias: 7.558693, T: 8693116, Avg. loss: 6.248548
Total training time: 72.86 seconds.
-- Epoch 327
Norm: 0.77, NNZs: 2297, Bias: 7.557169, T: 8719782, Avg. loss: 6.247065
Total training time: 73.09 seconds.
-- Epoch 328
Norm: 0.77, NNZs: 2297, Bias: 7.555649, T: 8746448, Avg. loss: 6.245585
Total training time: 73.31 seconds.
-- Epoch 329
Norm: 0.77, NNZs: 2297, Bias: 7.554134, T: 8773114, Avg. loss: 6.244112
Total training time: 73.53 seconds.
-- Epoch 330
Norm: 0.77, NNZs: 2297, Bias: 7.552624, T: 8799780, Avg. loss: 6.242636
Total training time: 73.76 seconds.
-- Epoch 331
Norm: 0.77, NNZs: 2297, Bias: 7.551119, T: 8826446, Avg. loss: 6.241163
Total training time: 73.98 seconds.
-- Epoch 332
Norm: 0.77, NNZs: 2297, Bias: 7.549617, T: 8853112, Avg. loss: 6.239705
Total training time: 74.20 seconds.
-- Epoch 333
Norm: 0.77, NNZs: 2297, Bias: 7.548121, T: 8879778, Avg. loss: 6.238235
Total training time: 74.42 seconds.
-- Epoch 334
Norm: 0.77, NNZs: 2297, Bias: 7.546629, T: 8906444, Avg. loss: 6.236791
Total training time: 74.64 seconds.
-- Epoch 335
Norm: 0.77, NNZs: 2297, Bias: 7.545141, T: 8933110, Avg. loss: 6.235349
Total training time: 74.86 seconds.
-- Epoch 336
Norm: 0.77, NNZs: 2297, Bias: 7.543658, T: 8959776, Avg. loss: 6.233893
Total training time: 75.09 seconds.
-- Epoch 337
Norm: 0.77, NNZs: 2297, Bias: 7.542179, T: 8986442, Avg. loss: 6.232454
Total training time: 75.31 seconds.
-- Epoch 338
Norm: 0.77, NNZs: 2297, Bias: 7.540705, T: 9013108, Avg. loss: 6.231037
Total training time: 75.54 seconds.
-- Epoch 339
Norm: 0.77, NNZs: 2297, Bias: 7.539235, T: 9039774, Avg. loss: 6.229581
Total training time: 75.76 seconds.
-- Epoch 340
```

```
Norm: 0.77, NNZs: 2297, Bias: 7.537769, T: 9066440, Avg. loss: 6.228163
Total training time: 75.98 seconds.
-- Epoch 341
Norm: 0.77, NNZs: 2297, Bias: 7.536308, T: 9093106, Avg. loss: 6.226737
Total training time: 76.21 seconds.
-- Epoch 342
Norm: 0.77, NNZs: 2297, Bias: 7.534851, T: 9119772, Avg. loss: 6.225314
Total training time: 76.43 seconds.
-- Epoch 343
Norm: 0.77, NNZs: 2297, Bias: 7.533398, T: 9146438, Avg. loss: 6.223900
Total training time: 76.65 seconds.
-- Epoch 344
Norm: 0.77, NNZs: 2297, Bias: 7.531949, T: 9173104, Avg. loss: 6.222486
Total training time: 76.88 seconds.
-- Epoch 345
Norm: 0.77, NNZs: 2297, Bias: 7.530505, T: 9199770, Avg. loss: 6.221084
Total training time: 77.10 seconds.
-- Epoch 346
Norm: 0.77, NNZs: 2297, Bias: 7.529065, T: 9226436, Avg. loss: 6.219668
Total training time: 77.33 seconds.
-- Epoch 347
Norm: 0.77, NNZs: 2297, Bias: 7.527629, T: 9253102, Avg. loss: 6.218285
Total training time: 77.55 seconds.
-- Epoch 348
Norm: 0.77, NNZs: 2297, Bias: 7.526197, T: 9279768, Avg. loss: 6.216888
Total training time: 77.78 seconds.
-- Epoch 349
Norm: 0.77, NNZs: 2297, Bias: 7.524769, T: 9306434, Avg. loss: 6.215490
Total training time: 78.00 seconds.
-- Epoch 350
Norm: 0.77, NNZs: 2297, Bias: 7.523345, T: 9333100, Avg. loss: 6.214104
Total training time: 78.22 seconds.
-- Epoch 351
Norm: 0.77, NNZs: 2297, Bias: 7.521926, T: 9359766, Avg. loss: 6.212733
Total training time: 78.45 seconds.
-- Epoch 352
Norm: 0.77, NNZs: 2297, Bias: 7.520510, T: 9386432, Avg. loss: 6.211349
Total training time: 78.66 seconds.
-- Epoch 353
Norm: 0.77, NNZs: 2297, Bias: 7.519099, T: 9413098, Avg. loss: 6.209964
Total training time: 78.89 seconds.
-- Epoch 354
Norm: 0.77, NNZs: 2297, Bias: 7.517691, T: 9439764, Avg. loss: 6.208597
Total training time: 79.11 seconds.
-- Epoch 355
Norm: 0.77, NNZs: 2297, Bias: 7.516288, T: 9466430, Avg. loss: 6.207227
Total training time: 79.33 seconds.
-- Epoch 356
Norm: 0.77, NNZs: 2297, Bias: 7.514888, T: 9493096, Avg. loss: 6.205863
Total training time: 79.56 seconds.
-- Epoch 357
Norm: 0.77, NNZs: 2297, Bias: 7.513493, T: 9519762, Avg. loss: 6.204499
Total training time: 79.78 seconds.
-- Epoch 358
Norm: 0.77, NNZs: 2297, Bias: 7.512101, T: 9546428, Avg. loss: 6.203154
Total training time: 80.00 seconds.
-- Epoch 359
Norm: 0.77, NNZs: 2297, Bias: 7.510713, T: 9573094, Avg. loss: 6.201803
Total training time: 80.22 seconds.
-- Epoch 360
Norm: 0.77, NNZs: 2297, Bias: 7.509329, T: 9599760, Avg. loss: 6.200450
Total training time: 80.45 seconds.
-- Epoch 361
Norm: 0.77, NNZs: 2297, Bias: 7.507949, T: 9626426, Avg. loss: 6.199106
Total training time: 80.67 seconds.
-- Epoch 362
Norm: 0.77, NNZs: 2297, Bias: 7.506573, T: 9653092, Avg. loss: 6.197764
Total training time: 80.89 seconds.
-- Epoch 363
Norm: 0.77, NNZs: 2297, Bias: 7.505200, T: 9679758, Avg. loss: 6.196428
Total training time: 81.11 seconds.
-- Epoch 364
Norm: 0.77, NNZs: 2297, Bias: 7.503831, T: 9706424, Avg. loss: 6.195096
Total training time: 81.35 seconds.
-- Epoch 365
Norm: 0.77, NNZs: 2297, Bias: 7.502466, T: 9733090, Avg. loss: 6.193763
Total training time: 81.57 seconds.
```

```
-- Epoch 366
Norm: 0.77, NNZs: 2297, Bias: 7.501105, T: 9759756, Avg. loss: 6.192441
Total training time: 81.79 seconds.
-- Epoch 367
Norm: 0.77, NNZs: 2297, Bias: 7.499748, T: 9786422, Avg. loss: 6.191109
Total training time: 82.01 seconds.
-- Epoch 368
Norm: 0.77, NNZs: 2297, Bias: 7.498394, T: 9813088, Avg. loss: 6.189803
Total training time: 82.24 seconds.
-- Epoch 369
Norm: 0.77, NNZs: 2297, Bias: 7.497044, T: 9839754, Avg. loss: 6.188477
Total training time: 82.46 seconds.
-- Epoch 370
Norm: 0.77, NNZs: 2297, Bias: 7.495697, T: 9866420, Avg. loss: 6.187165
Total training time: 82.68 seconds.
-- Epoch 371
Norm: 0.77, NNZs: 2297, Bias: 7.494355, T: 9893086, Avg. loss: 6.185867
Total training time: 82.90 seconds.
-- Epoch 372
Norm: 0.77, NNZs: 2297, Bias: 7.493015, T: 9919752, Avg. loss: 6.184568
Total training time: 83.13 seconds.
-- Epoch 373
Norm: 0.77, NNZs: 2297, Bias: 7.491680, T: 9946418, Avg. loss: 6.183260
Total training time: 83.35 seconds.
-- Epoch 374
Norm: 0.77, NNZs: 2297, Bias: 7.490348, T: 9973084, Avg. loss: 6.181972
Total training time: 83.57 seconds.
-- Epoch 375
Norm: 0.77, NNZs: 2297, Bias: 7.489020, T: 9999750, Avg. loss: 6.180669
Total training time: 83.79 seconds.
-- Epoch 376
Norm: 0.77, NNZs: 2297, Bias: 7.487695, T: 10026416, Avg. loss: 6.179371
Total training time: 84.02 seconds.
-- Epoch 377
Norm: 0.77, NNZs: 2297, Bias: 7.486373, T: 10053082, Avg. loss: 6.178094
Total training time: 84.24 seconds.
-- Epoch 378
Norm: 0.77, NNZs: 2297, Bias: 7.485055, T: 10079748, Avg. loss: 6.176807
Total training time: 84.47 seconds.
-- Epoch 379
Norm: 0.77, NNZs: 2297, Bias: 7.483741, T: 10106414, Avg. loss: 6.175526
Total training time: 84.69 seconds.
-- Epoch 380
Norm: 0.77, NNZs: 2297, Bias: 7.482430, T: 10133080, Avg. loss: 6.174256
Total training time: 84.91 seconds.
-- Epoch 381
Norm: 0.77, NNZs: 2297, Bias: 7.481123, T: 10159746, Avg. loss: 6.172977
Total training time: 85.13 seconds.
-- Epoch 382
Norm: 0.77, NNZs: 2297, Bias: 7.479819, T: 10186412, Avg. loss: 6.171711
Total training time: 85.35 seconds.
-- Epoch 383
Norm: 0.77, NNZs: 2297, Bias: 7.478518, T: 10213078, Avg. loss: 6.170445
Total training time: 85.58 seconds.
-- Epoch 384
Norm: 0.77, NNZs: 2297, Bias: 7.477221, T: 10239744, Avg. loss: 6.169185
Total training time: 85.81 seconds.
-- Epoch 385
Norm: 0.77, NNZs: 2297, Bias: 7.475927, T: 10266410, Avg. loss: 6.167916
Total training time: 86.03 seconds.
-- Epoch 386
Norm: 0.77, NNZs: 2297, Bias: 7.474637, T: 10293076, Avg. loss: 6.166660
Total training time: 86.25 seconds.
-- Epoch 387
Norm: 0.77, NNZs: 2297, Bias: 7.473350, T: 10319742, Avg. loss: 6.165409
Total training time: 86.47 seconds.
-- Epoch 388
Norm: 0.77, NNZs: 2297, Bias: 7.472066, T: 10346408, Avg. loss: 6.164151
Total training time: 86.70 seconds.
-- Epoch 389
Norm: 0.77, NNZs: 2297, Bias: 7.470786, T: 10373074, Avg. loss: 6.162915
Total training time: 86.92 seconds.
-- Epoch 390
Norm: 0.77, NNZs: 2297, Bias: 7.469509, T: 10399740, Avg. loss: 6.161664
Total training time: 87.15 seconds.
-- Epoch 391
Norm: 0.77, NNZs: 2297, Bias: 7.468235, T: 10426406, Avg. loss: 6.160431
```

```
Total training time: 87.37 seconds.
-- Epoch 392
Norm: 0.77, NNZs: 2297, Bias: 7.466964, T: 10453072, Avg. loss: 6.159190
Total training time: 87.60 seconds.
-- Epoch 393
Norm: 0.77, NNZs: 2297, Bias: 7.465697, T: 10479738, Avg. loss: 6.157957
Total training time: 87.82 seconds.
-- Epoch 394
Norm: 0.77, NNZs: 2297, Bias: 7.464433, T: 10506404, Avg. loss: 6.156727
Total training time: 88.05 seconds.
-- Epoch 395
Norm: 0.77, NNZs: 2297, Bias: 7.463172, T: 10533070, Avg. loss: 6.155501
Total training time: 88.27 seconds.
-- Epoch 396
Norm: 0.77, NNZs: 2297, Bias: 7.461914, T: 10559736, Avg. loss: 6.154276
Total training time: 88.50 seconds.
-- Epoch 397
Norm: 0.77, NNZs: 2297, Bias: 7.460659, T: 10586402, Avg. loss: 6.153053
Total training time: 88.71 seconds.
-- Epoch 398
Norm: 0.77, NNZs: 2297, Bias: 7.459408, T: 10613068, Avg. loss: 6.151840
Total training time: 88.94 seconds.
-- Epoch 399
Norm: 0.77, NNZs: 2297, Bias: 7.458160, T: 10639734, Avg. loss: 6.150618
Total training time: 89.16 seconds.
-- Epoch 400
Norm: 0.77, NNZs: 2297, Bias: 7.456915, T: 10666400, Avg. loss: 6.149409
Total training time: 89.38 seconds.
-- Epoch 401
Norm: 0.77, NNZs: 2297, Bias: 7.455673, T: 10693066, Avg. loss: 6.148208
Total training time: 89.61 seconds.
-- Epoch 402
Norm: 0.77, NNZs: 2297, Bias: 7.454434, T: 10719732, Avg. loss: 6.146994
Total training time: 89.83 seconds.
-- Epoch 403
Norm: 0.77, NNZs: 2297, Bias: 7.453198, T: 10746398, Avg. loss: 6.145788
Total training time: 90.05 seconds.
-- Epoch 404
Norm: 0.77, NNZs: 2297, Bias: 7.451965, T: 10773064, Avg. loss: 6.144586
Total training time: 90.28 seconds.
-- Epoch 405
Norm: 0.77, NNZs: 2297, Bias: 7.450736, T: 10799730, Avg. loss: 6.143385
Total training time: 90.50 seconds.
-- Epoch 406
Norm: 0.77, NNZs: 2297, Bias: 7.449509, T: 10826396, Avg. loss: 6.142201
Total training time: 90.73 seconds.
-- Epoch 407
Norm: 0.77, NNZs: 2297, Bias: 7.448285, T: 10853062, Avg. loss: 6.141007
Total training time: 90.95 seconds.
-- Epoch 408
Norm: 0.77, NNZs: 2297, Bias: 7.447065, T: 10879728, Avg. loss: 6.139818
Total training time: 91.18 seconds.
-- Epoch 409
Norm: 0.77, NNZs: 2297, Bias: 7.445847, T: 10906394, Avg. loss: 6.138633
Total training time: 91.40 seconds.
-- Epoch 410
Norm: 0.77, NNZs: 2297, Bias: 7.444632, T: 10933060, Avg. loss: 6.137448
Total training time: 91.64 seconds.
-- Epoch 411
Norm: 0.77, NNZs: 2297, Bias: 7.443421, T: 10959726, Avg. loss: 6.136267
Total training time: 91.87 seconds.
-- Epoch 412
Norm: 0.77, NNZs: 2297, Bias: 7.442212, T: 10986392, Avg. loss: 6.135092
Total training time: 92.09 seconds.
-- Epoch 413
Norm: 0.77, NNZs: 2297, Bias: 7.441006, T: 11013058, Avg. loss: 6.133916
Total training time: 92.31 seconds.
-- Epoch 414
Norm: 0.77, NNZs: 2297, Bias: 7.439803, T: 11039724, Avg. loss: 6.132754
Total training time: 92.53 seconds.
-- Epoch 415
Norm: 0.77, NNZs: 2297, Bias: 7.438604, T: 11066390, Avg. loss: 6.131582
Total training time: 92.76 seconds.
-- Epoch 416
Norm: 0.77, NNZs: 2297, Bias: 7.437406, T: 11093056, Avg. loss: 6.130420
Total training time: 92.98 seconds.
-- Epoch 417
```

```
Norm: 0.77, NNZs: 2297, Bias: 7.436212, T: 11119722, Avg. loss: 6.129258
Total training time: 93.21 seconds.
-- Epoch 418
Norm: 0.77, NNZs: 2297, Bias: 7.435021, T: 11146388, Avg. loss: 6.128099
Total training time: 93.43 seconds.
-- Epoch 419
Norm: 0.77, NNZs: 2297, Bias: 7.433833, T: 11173054, Avg. loss: 6.126943
Total training time: 93.67 seconds.
-- Epoch 420
Norm: 0.77, NNZs: 2297, Bias: 7.432647, T: 11199720, Avg. loss: 6.125791
Total training time: 93.88 seconds.
-- Epoch 421
Norm: 0.77, NNZs: 2297, Bias: 7.431464, T: 11226386, Avg. loss: 6.124629
Total training time: 94.11 seconds.
-- Epoch 422
Norm: 0.77, NNZs: 2297, Bias: 7.430284, T: 11253052, Avg. loss: 6.123490
Total training time: 94.34 seconds.
-- Epoch 423
Norm: 0.77, NNZs: 2297, Bias: 7.429107, T: 11279718, Avg. loss: 6.122344
Total training time: 94.55 seconds.
-- Epoch 424
Norm: 0.77, NNZs: 2297, Bias: 7.427933, T: 11306384, Avg. loss: 6.121199
Total training time: 94.78 seconds.
-- Epoch 425
Norm: 0.77, NNZs: 2297, Bias: 7.426761, T: 11333050, Avg. loss: 6.120058
Total training time: 95.01 seconds.
-- Epoch 426
Norm: 0.77, NNZs: 2297, Bias: 7.425592, T: 11359716, Avg. loss: 6.118920
Total training time: 95.23 seconds.
-- Epoch 427
Norm: 0.77, NNZs: 2297, Bias: 7.424426, T: 11386382, Avg. loss: 6.117777
Total training time: 95.45 seconds.
-- Epoch 428
Norm: 0.77, NNZs: 2297, Bias: 7.423263, T: 11413048, Avg. loss: 6.116659
Total training time: 95.67 seconds.
-- Epoch 429
Norm: 0.77, NNZs: 2297, Bias: 7.422102, T: 11439714, Avg. loss: 6.115526
Total training time: 95.90 seconds.
-- Epoch 430
Norm: 0.77, NNZs: 2297, Bias: 7.420944, T: 11466380, Avg. loss: 6.114405
Total training time: 96.12 seconds.
-- Epoch 431
Norm: 0.77, NNZs: 2297, Bias: 7.419789, T: 11493046, Avg. loss: 6.113279
Total training time: 96.34 seconds.
-- Epoch 432
Norm: 0.77, NNZs: 2297, Bias: 7.418636, T: 11519712, Avg. loss: 6.112148
Total training time: 96.57 seconds.
-- Epoch 433
Norm: 0.77, NNZs: 2297, Bias: 7.417486, T: 11546378, Avg. loss: 6.111030
Total training time: 96.80 seconds.
-- Epoch 434
Norm: 0.77, NNZs: 2297, Bias: 7.416339, T: 11573044, Avg. loss: 6.109908
Total training time: 97.02 seconds.
-- Epoch 435
Norm: 0.77, NNZs: 2297, Bias: 7.415195, T: 11599710, Avg. loss: 6.108803
Total training time: 97.24 seconds.
-- Epoch 436
Norm: 0.77, NNZs: 2297, Bias: 7.414053, T: 11626376, Avg. loss: 6.107689
Total training time: 97.47 seconds.
-- Epoch 437
Norm: 0.77, NNZs: 2297, Bias: 7.412913, T: 11653042, Avg. loss: 6.106585
Total training time: 97.69 seconds.
-- Epoch 438
Norm: 0.77, NNZs: 2297, Bias: 7.411777, T: 11679708, Avg. loss: 6.105483
Total training time: 97.92 seconds.
-- Epoch 439
Norm: 0.77, NNZs: 2297, Bias: 7.410643, T: 11706374, Avg. loss: 6.104373
Total training time: 98.14 seconds.
-- Epoch 440
Norm: 0.77, NNZs: 2297, Bias: 7.409511, T: 11733040, Avg. loss: 6.103275
Total training time: 98.37 seconds.
-- Epoch 441
Norm: 0.77, NNZs: 2297, Bias: 7.408382, T: 11759706, Avg. loss: 6.102173
Total training time: 98.59 seconds.
-- Epoch 442
Norm: 0.77, NNZs: 2297, Bias: 7.407256, T: 11786372, Avg. loss: 6.101073
Total training time: 98.82 seconds.
```

```
-- Epoch 443
Norm: 0.77, NNZs: 2297, Bias: 7.406132, T: 11813038, Avg. loss: 6.099985
Total training time: 99.04 seconds.
-- Epoch 444
Norm: 0.77, NNZs: 2297, Bias: 7.405011, T: 11839704, Avg. loss: 6.098896
Total training time: 99.27 seconds.
-- Epoch 445
Norm: 0.77, NNZs: 2297, Bias: 7.403892, T: 11866370, Avg. loss: 6.097797
Total training time: 99.49 seconds.
-- Epoch 446
Norm: 0.77, NNZs: 2297, Bias: 7.402775, T: 11893036, Avg. loss: 6.096717
Total training time: 99.72 seconds.
-- Epoch 447
Norm: 0.77, NNZs: 2297, Bias: 7.401662, T: 11919702, Avg. loss: 6.095636
Total training time: 99.94 seconds.
-- Epoch 448
Norm: 0.77, NNZs: 2297, Bias: 7.400550, T: 11946368, Avg. loss: 6.094555
Total training time: 100.16 seconds.
-- Epoch 449
Norm: 0.77, NNZs: 2297, Bias: 7.399442, T: 11973034, Avg. loss: 6.093472
Total training time: 100.38 seconds.
-- Epoch 450
Norm: 0.77, NNZs: 2297, Bias: 7.398335, T: 11999700, Avg. loss: 6.092392
Total training time: 100.61 seconds.
-- Epoch 451
Norm: 0.77, NNZs: 2297, Bias: 7.397232, T: 12026366, Avg. loss: 6.091323
Total training time: 100.84 seconds.
-- Epoch 452
Norm: 0.77, NNZs: 2297, Bias: 7.396130, T: 12053032, Avg. loss: 6.090257
Total training time: 101.06 seconds.
-- Epoch 453
Norm: 0.77, NNZs: 2297, Bias: 7.395031, T: 12079698, Avg. loss: 6.089184
Total training time: 101.29 seconds.
-- Epoch 454
Norm: 0.77, NNZs: 2297, Bias: 7.393935, T: 12106364, Avg. loss: 6.088119
Total training time: 101.51 seconds.
-- Epoch 455
Norm: 0.77, NNZs: 2297, Bias: 7.392841, T: 12133030, Avg. loss: 6.087059
Total training time: 101.73 seconds.
-- Epoch 456
Norm: 0.77, NNZs: 2297, Bias: 7.391749, T: 12159696, Avg. loss: 6.085993
Total training time: 101.97 seconds.
-- Epoch 457
Norm: 0.77, NNZs: 2297, Bias: 7.390660, T: 12186362, Avg. loss: 6.084934
Total training time: 102.19 seconds.
-- Epoch 458
Norm: 0.77, NNZs: 2297, Bias: 7.389573, T: 12213028, Avg. loss: 6.083879
Total training time: 102.42 seconds.
-- Epoch 459
Norm: 0.77, NNZs: 2297, Bias: 7.388488, T: 12239694, Avg. loss: 6.082819
Total training time: 102.64 seconds.
-- Epoch 460
Norm: 0.77, NNZs: 2297, Bias: 7.387406, T: 12266360, Avg. loss: 6.081766
Total training time: 102.86 seconds.
-- Epoch 461
Norm: 0.77, NNZs: 2297, Bias: 7.386326, T: 12293026, Avg. loss: 6.080713
Total training time: 103.09 seconds.
-- Epoch 462
Norm: 0.77, NNZs: 2297, Bias: 7.385249, T: 12319692, Avg. loss: 6.079666
Total training time: 103.32 seconds.
-- Epoch 463
Norm: 0.77, NNZs: 2297, Bias: 7.384174, T: 12346358, Avg. loss: 6.078622
Total training time: 103.54 seconds.
-- Epoch 464
Norm: 0.77, NNZs: 2297, Bias: 7.383101, T: 12373024, Avg. loss: 6.077586
Total training time: 103.77 seconds.
-- Epoch 465
Norm: 0.77, NNZs: 2297, Bias: 7.382030, T: 12399690, Avg. loss: 6.076540
Total training time: 103.99 seconds.
-- Epoch 466
Norm: 0.77, NNZs: 2297, Bias: 7.380962, T: 12426356, Avg. loss: 6.075502
Total training time: 104.21 seconds.
-- Epoch 467
Norm: 0.77, NNZs: 2297, Bias: 7.379896, T: 12453022, Avg. loss: 6.074466
Total training time: 104.43 seconds.
-- Epoch 468
Norm: 0.77, NNZs: 2297, Bias: 7.378833, T: 12479688, Avg. loss: 6.073429
```

```
Total training time: 104.66 seconds.
-- Epoch 469
Norm: 0.77, NNZs: 2297, Bias: 7.377771, T: 12506354, Avg. loss: 6.072394
Total training time: 104.89 seconds.
-- Epoch 470
Norm: 0.77, NNZs: 2297, Bias: 7.376712, T: 12533020, Avg. loss: 6.071368
Total training time: 105.11 seconds.
-- Epoch 471
Norm: 0.77, NNZs: 2297, Bias: 7.375656, T: 12559686, Avg. loss: 6.070344
Total training time: 105.33 seconds.
-- Epoch 472
Norm: 0.77, NNZs: 2297, Bias: 7.374601, T: 12586352, Avg. loss: 6.069315
Total training time: 105.56 seconds.
-- Epoch 473
Norm: 0.77, NNZs: 2297, Bias: 7.373549, T: 12613018, Avg. loss: 6.068284
Total training time: 105.78 seconds.
-- Epoch 474
Norm: 0.77, NNZs: 2297, Bias: 7.372499, T: 12639684, Avg. loss: 6.067267
Total training time: 106.01 seconds.
-- Epoch 475
Norm: 0.77, NNZs: 2297, Bias: 7.371451, T: 12666350, Avg. loss: 6.066244
Total training time: 106.23 seconds.
-- Epoch 476
Norm: 0.77, NNZs: 2297, Bias: 7.370405, T: 12693016, Avg. loss: 6.065225
Total training time: 106.45 seconds.
-- Epoch 477
Norm: 0.77, NNZs: 2297, Bias: 7.369362, T: 12719682, Avg. loss: 6.064217
Total training time: 106.67 seconds.
-- Epoch 478
Norm: 0.77, NNZs: 2297, Bias: 7.368320, T: 12746348, Avg. loss: 6.063207
Total training time: 106.90 seconds.
-- Epoch 479
Norm: 0.77, NNZs: 2297, Bias: 7.367281, T: 12773014, Avg. loss: 6.062190
Total training time: 107.12 seconds.
-- Epoch 480
Norm: 0.77, NNZs: 2297, Bias: 7.366244, T: 12799680, Avg. loss: 6.061185
Total training time: 107.34 seconds.
-- Epoch 481
Norm: 0.77, NNZs: 2297, Bias: 7.365210, T: 12826346, Avg. loss: 6.060176
Total training time: 107.57 seconds.
-- Epoch 482
Norm: 0.77, NNZs: 2297, Bias: 7.364177, T: 12853012, Avg. loss: 6.059170
Total training time: 107.80 seconds.
-- Epoch 483
Norm: 0.77, NNZs: 2297, Bias: 7.363146, T: 12879678, Avg. loss: 6.058174
Total training time: 108.02 seconds.
-- Epoch 484
Norm: 0.77, NNZs: 2297, Bias: 7.362118, T: 12906344, Avg. loss: 6.057179
Total training time: 108.25 seconds.
-- Epoch 485
Norm: 0.77, NNZs: 2297, Bias: 7.361092, T: 12933010, Avg. loss: 6.056174
Total training time: 108.47 seconds.
-- Epoch 486
Norm: 0.77, NNZs: 2297, Bias: 7.360068, T: 12959676, Avg. loss: 6.055174
Total training time: 108.69 seconds.
-- Epoch 487
Norm: 0.77, NNZs: 2297, Bias: 7.359046, T: 12986342, Avg. loss: 6.054183
Total training time: 108.92 seconds.
-- Epoch 488
Norm: 0.77, NNZs: 2297, Bias: 7.358026, T: 13013008, Avg. loss: 6.053192
Total training time: 109.14 seconds.
-- Epoch 489
Norm: 0.77, NNZs: 2297, Bias: 7.357008, T: 13039674, Avg. loss: 6.052196
Total training time: 109.37 seconds.
-- Epoch 490
Norm: 0.77, NNZs: 2297, Bias: 7.355993, T: 13066340, Avg. loss: 6.051214
Total training time: 109.59 seconds.
-- Epoch 491
Norm: 0.77, NNZs: 2297, Bias: 7.354979, T: 13093006, Avg. loss: 6.050235
Total training time: 109.82 seconds.
Convergence after 491 epochs took 109.82 seconds
-- Epoch 1
Norm: 0.34, NNZs: 2298, Bias: -0.170533, T: 26667, Avg. loss: 1.291390
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.32, NNZs: 2298, Bias: -0.205848, T: 53334, Avg. loss: 0.657748
Total training time: 0.45 seconds.
```

```
-- Epoch 3
Norm: 0.28, NNZs: 2298, Bias: -0.227777, T: 80001, Avg. loss: 0.634393
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 0.27, NNZs: 2298, Bias: -0.242422, T: 106668, Avg. loss: 0.619774
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.26, NNZs: 2298, Bias: -0.253522, T: 133335, Avg. loss: 0.615251
Total training time: 1.11 seconds.
-- Epoch 6
Norm: 0.26, NNZs: 2298, Bias: -0.262210, T: 160002, Avg. loss: 0.609610
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 0.26, NNZs: 2298, Bias: -0.269486, T: 186669, Avg. loss: 0.606911
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 0.26, NNZs: 2298, Bias: -0.275555, T: 213336, Avg. loss: 0.601192
Total training time: 1.78 seconds.
-- Epoch 9
Norm: 0.26, NNZs: 2298, Bias: -0.281136, T: 240003, Avg. loss: 0.605124
Total training time: 2.00 seconds.
-- Epoch 10
Norm: 0.26, NNZs: 2298, Bias: -0.286077, T: 266670, Avg. loss: 0.604102
Total training time: 2.23 seconds.
-- Epoch 11
Norm: 0.26, NNZs: 2298, Bias: -0.290404, T: 293337, Avg. loss: 0.599683
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 0.26, NNZs: 2298, Bias: -0.294253, T: 320004, Avg. loss: 0.596740
Total training time: 2.68 seconds.
-- Epoch 13
Norm: 0.26, NNZs: 2298, Bias: -0.297916, T: 346671, Avg. loss: 0.599658
Total training time: 2.90 seconds.
-- Epoch 14
Norm: 0.26, NNZs: 2298, Bias: -0.301214, T: 373338, Avg. loss: 0.596665
Total training time: 3.12 seconds.
-- Epoch 15
Norm: 0.26, NNZs: 2298, Bias: -0.304258, T: 400005, Avg. loss: 0.595725
Total training time: 3.34 seconds.
-- Epoch 16
Norm: 0.26, NNZs: 2298, Bias: -0.307009, T: 426672, Avg. loss: 0.593332
Total training time: 3.57 seconds.
-- Epoch 17
Norm: 0.26, NNZs: 2298, Bias: -0.309673, T: 453339, Avg. loss: 0.595421
Total training time: 3.79 seconds.
-- Epoch 18
Norm: 0.26, NNZs: 2298, Bias: -0.312175, T: 480006, Avg. loss: 0.594771
Total training time: 4.01 seconds.
-- Epoch 19
Norm: 0.26, NNZs: 2298, Bias: -0.314439, T: 506673, Avg. loss: 0.590665
Total training time: 4.23 seconds.
-- Epoch 20
Norm: 0.26, NNZs: 2298, Bias: -0.316645, T: 533340, Avg. loss: 0.593338
Total training time: 4.46 seconds.
-- Epoch 21
Norm: 0.26, NNZs: 2298, Bias: -0.318747, T: 560007, Avg. loss: 0.592768
Total training time: 4.68 seconds.
-- Epoch 22
Norm: 0.26, NNZs: 2298, Bias: -0.320705, T: 586674, Avg. loss: 0.591199
Total training time: 4.90 seconds.
-- Epoch 23
Norm: 0.26, NNZs: 2298, Bias: -0.322590, T: 613341, Avg. loss: 0.591467
Total training time: 5.13 seconds.
-- Epoch 24
Norm: 0.26, NNZs: 2298, Bias: -0.324366, T: 640008, Avg. loss: 0.589677
Total training time: 5.35 seconds.
Convergence after 24 epochs took 5.35 seconds
-- Epoch 1
Norm: 0.49, NNZs: 2298, Bias: -0.510189, T: 26667, Avg. loss: 1.365324
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.32, NNZs: 2298, Bias: -0.528091, T: 53334, Avg. loss: 0.586579
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 0.29, NNZs: 2298, Bias: -0.537593, T: 80001, Avg. loss: 0.558096
Total training time: 0.67 seconds.
-- Epoch 4
```

```
Norm: 0.29, NNZs: 2298, Bias: -0.544605, T: 106668, Avg. loss: 0.554664
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.28, NNZs: 2298, Bias: -0.549603, T: 133335, Avg. loss: 0.549874
Total training time: 1.11 seconds.
-- Epoch 6
Norm: 0.27, NNZs: 2298, Bias: -0.553686, T: 160002, Avg. loss: 0.546223
Total training time: 1.33 seconds.
-- Epoch 7
Norm: 0.27, NNZs: 2298, Bias: -0.556829, T: 186669, Avg. loss: 0.542117
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 0.27, NNZs: 2298, Bias: -0.559726, T: 213336, Avg. loss: 0.542782
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 0.27, NNZs: 2298, Bias: -0.562289, T: 240003, Avg. loss: 0.541910
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 0.27, NNZs: 2298, Bias: -0.564438, T: 266670, Avg. loss: 0.539264
Total training time: 2.23 seconds.
-- Epoch 11
Norm: 0.27, NNZs: 2298, Bias: -0.566546, T: 293337, Avg. loss: 0.540826
Total training time: 2.45 seconds.
-- Epoch 12
Norm: 0.27, NNZs: 2298, Bias: -0.568228, T: 320004, Avg. loss: 0.535535
Total training time: 2.68 seconds.
-- Epoch 13
Norm: 0.27, NNZs: 2298, Bias: -0.570017, T: 346671, Avg. loss: 0.540899
Total training time: 2.90 seconds.
-- Epoch 14
Norm: 0.27, NNZs: 2298, Bias: -0.571554, T: 373338, Avg. loss: 0.538049
Total training time: 3.12 seconds.
-- Epoch 15
Norm: 0.27, NNZs: 2298, Bias: -0.573016, T: 400005, Avg. loss: 0.537840
Total training time: 3.34 seconds.
-- Epoch 16
Norm: 0.27, NNZs: 2298, Bias: -0.574255, T: 426672, Avg. loss: 0.535016
Total training time: 3.57 seconds.
-- Epoch 17
Norm: 0.27, NNZs: 2298, Bias: -0.575487, T: 453339, Avg. loss: 0.536131
Total training time: 3.80 seconds.
Convergence after 17 epochs took 3.80 seconds
-- Epoch 1
Norm: 0.08, NNZs: 2297, Bias: 7.489190, T: 26666, Avg. loss: 7.314097
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.08, NNZs: 2297, Bias: 7.454574, T: 53332, Avg. loss: 7.220275
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 0.08, NNZs: 2297, Bias: 7.434327, T: 79998, Avg. loss: 7.194308
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 0.08, NNZs: 2297, Bias: 7.419962, T: 106664, Avg. loss: 7.177252
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.08, NNZs: 2297, Bias: 7.408820, T: 133330, Avg. loss: 7.164663
Total training time: 1.11 seconds.
-- Epoch 6
Norm: 0.08, NNZs: 2297, Bias: 7.399716, T: 159996, Avg. loss: 7.154769
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 0.08, NNZs: 2297, Bias: 7.392018, T: 186662, Avg. loss: 7.146472
Total training time: 1.56 seconds.
-- Epoch 8
Norm: 0.08, NNZs: 2297, Bias: 7.385351, T: 213328, Avg. loss: 7.139432
Total training time: 1.78 seconds.
-- Epoch 9
Norm: 0.08, NNZs: 2297, Bias: 7.379469, T: 239994, Avg. loss: 7.133238
Total training time: 2.01 seconds.
-- Epoch 10
Norm: 0.08, NNZs: 2297, Bias: 7.374209, T: 266660, Avg. loss: 7.127729
Total training time: 2.25 seconds.
-- Epoch 11
Norm: 0.08, NNZs: 2297, Bias: 7.369450, T: 293326, Avg. loss: 7.122807
Total training time: 2.47 seconds.
-- Epoch 12
Norm: 0.08, NNZs: 2297, Bias: 7.365105, T: 319992, Avg. loss: 7.118318
```

```
Total training time: 2.69 seconds.
-- Epoch 13
Norm: 0.08, NNZs: 2297, Bias: 7.361108, T: 346658, Avg. loss: 7.114212
Total training time: 2.92 seconds.
-- Epoch 14
Norm: 0.08, NNZs: 2297, Bias: 7.357408, T: 373324, Avg. loss: 7.110410
Total training time: 3.14 seconds.
-- Epoch 15
Norm: 0.08, NNZs: 2297, Bias: 7.353963, T: 399990, Avg. loss: 7.106887
Total training time: 3.37 seconds.
-- Epoch 16
Norm: 0.08, NNZs: 2297, Bias: 7.350741, T: 426656, Avg. loss: 7.103611
Total training time: 3.59 seconds.
-- Epoch 17
Norm: 0.08, NNZs: 2297, Bias: 7.347714, T: 453322, Avg. loss: 7.100531
Total training time: 3.81 seconds.
-- Epoch 18
Norm: 0.08, NNZs: 2297, Bias: 7.344860, T: 479988, Avg. loss: 7.097679
Total training time: 4.04 seconds.
-- Epoch 19
Norm: 0.08, NNZs: 2297, Bias: 7.342160, T: 506654, Avg. loss: 7.094938
Total training time: 4.26 seconds.
-- Epoch 20
Norm: 0.08, NNZs: 2297, Bias: 7.339599, T: 533320, Avg. loss: 7.092350
Total training time: 4.48 seconds.
-- Epoch 21
Norm: 0.08, NNZs: 2297, Bias: 7.337163, T: 559986, Avg. loss: 7.089896
Total training time: 4.71 seconds.
-- Epoch 22
Norm: 0.08, NNZs: 2297, Bias: 7.334841, T: 586652, Avg. loss: 7.087560
Total training time: 4.93 seconds.
-- Epoch 23
Norm: 0.08, NNZs: 2297, Bias: 7.332621, T: 613318, Avg. loss: 7.085343
Total training time: 5.15 seconds.
-- Epoch 24
Norm: 0.08, NNZs: 2297, Bias: 7.330497, T: 639984, Avg. loss: 7.083196
Total training time: 5.38 seconds.
-- Epoch 25
Norm: 0.08, NNZs: 2297, Bias: 7.328458, T: 666650, Avg. loss: 7.081137
Total training time: 5.60 seconds.
-- Epoch 26
Norm: 0.08, NNZs: 2297, Bias: 7.326500, T: 693316, Avg. loss: 7.079175
Total training time: 5.82 seconds.
-- Epoch 27
Norm: 0.08, NNZs: 2297, Bias: 7.324616, T: 719982, Avg. loss: 7.077283
Total training time: 6.05 seconds.
-- Epoch 28
Norm: 0.08, NNZs: 2297, Bias: 7.322800, T: 746648, Avg. loss: 7.075474
Total training time: 6.27 seconds.
-- Epoch 29
Norm: 0.08, NNZs: 2297, Bias: 7.321048, T: 773314, Avg. loss: 7.073717
Total training time: 6.50 seconds.
-- Epoch 30
Norm: 0.08, NNZs: 2297, Bias: 7.319355, T: 799980, Avg. loss: 7.072010
Total training time: 6.72 seconds.
-- Epoch 31
Norm: 0.08, NNZs: 2297, Bias: 7.317718, T: 826646, Avg. loss: 7.070382
Total training time: 6.95 seconds.
-- Epoch 32
Norm: 0.08, NNZs: 2297, Bias: 7.316133, T: 853312, Avg. loss: 7.068803
Total training time: 7.17 seconds.
-- Epoch 33
Norm: 0.08, NNZs: 2297, Bias: 7.314597, T: 879978, Avg. loss: 7.067260
Total training time: 7.39 seconds.
-- Epoch 34
Norm: 0.08, NNZs: 2297, Bias: 7.313106, T: 906644, Avg. loss: 7.065782
Total training time: 7.62 seconds.
-- Epoch 35
Norm: 0.08, NNZs: 2297, Bias: 7.311659, T: 933310, Avg. loss: 7.064334
Total training time: 7.84 seconds.
-- Epoch 36
Norm: 0.08, NNZs: 2297, Bias: 7.310253, T: 959976, Avg. loss: 7.062921
Total training time: 8.06 seconds.
-- Epoch 37
Norm: 0.08, NNZs: 2297, Bias: 7.308885, T: 986642, Avg. loss: 7.061579
Total training time: 8.29 seconds.
-- Epoch 38
```

```
Norm: 0.08, NNZs: 2297, Bias: 7.307553, T: 1013308, Avg. loss: 7.060248
Total training time: 8.51 seconds.
-- Epoch 39
Norm: 0.08, NNZs: 2297, Bias: 7.306256, T: 1039974, Avg. loss: 7.058955
Total training time: 8.73 seconds.
-- Epoch 40
Norm: 0.08, NNZs: 2297, Bias: 7.304992, T: 1066640, Avg. loss: 7.057697
Total training time: 8.95 seconds.
-- Epoch 41
Norm: 0.08, NNZs: 2297, Bias: 7.303760, T: 1093306, Avg. loss: 7.056475
Total training time: 9.18 seconds.
-- Epoch 42
Norm: 0.08, NNZs: 2297, Bias: 7.302557, T: 1119972, Avg. loss: 7.055278
Total training time: 9.40 seconds.
-- Epoch 43
Norm: 0.08, NNZs: 2297, Bias: 7.301382, T: 1146638, Avg. loss: 7.054113
Total training time: 9.63 seconds.
-- Epoch 44
Norm: 0.08, NNZs: 2297, Bias: 7.300234, T: 1173304, Avg. loss: 7.052964
Total training time: 9.85 seconds.
-- Epoch 45
Norm: 0.08, NNZs: 2297, Bias: 7.299112, T: 1199970, Avg. loss: 7.051859
Total training time: 10.08 seconds.
-- Epoch 46
Norm: 0.08, NNZs: 2297, Bias: 7.298015, T: 1226636, Avg. loss: 7.050759
Total training time: 10.30 seconds.
-- Epoch 47
Norm: 0.08, NNZs: 2297, Bias: 7.296941, T: 1253302, Avg. loss: 7.049694
Total training time: 10.53 seconds.
-- Epoch 48
Norm: 0.08, NNZs: 2297, Bias: 7.295890, T: 1279968, Avg. loss: 7.048647
Total training time: 10.75 seconds.
-- Epoch 49
Norm: 0.08, NNZs: 2297, Bias: 7.294861, T: 1306634, Avg. loss: 7.047633
Total training time: 10.98 seconds.
-- Epoch 50
Norm: 0.08, NNZs: 2297, Bias: 7.293852, T: 1333300, Avg. loss: 7.046632
Total training time: 11.20 seconds.
-- Epoch 51
Norm: 0.08, NNZs: 2297, Bias: 7.292863, T: 1359966, Avg. loss: 7.045650
Total training time: 11.42 seconds.
-- Epoch 52
Norm: 0.08, NNZs: 2297, Bias: 7.291894, T: 1386632, Avg. loss: 7.044680
Total training time: 11.64 seconds.
-- Epoch 53
Norm: 0.08, NNZs: 2297, Bias: 7.290943, T: 1413298, Avg. loss: 7.043743
Total training time: 11.86 seconds.
-- Epoch 54
Norm: 0.08, NNZs: 2297, Bias: 7.290010, T: 1439964, Avg. loss: 7.042811
Total training time: 12.09 seconds.
-- Epoch 55
Norm: 0.08, NNZs: 2297, Bias: 7.289094, T: 1466630, Avg. loss: 7.041910
Total training time: 12.31 seconds.
Convergence after 55 epochs took 12.31 seconds
-- Epoch 1
Norm: 0.09, NNZs: 2298, Bias: -0.189922, T: 26667, Avg. loss: 0.656016
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.08, NNZs: 2298, Bias: -0.191152, T: 53334, Avg. loss: 0.628680
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 0.08, NNZs: 2298, Bias: -0.191902, T: 80001, Avg. loss: 0.627735
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 0.08, NNZs: 2298, Bias: -0.192452, T: 106668, Avg. loss: 0.626795
Total training time: 0.90 seconds.
-- Epoch 5
Norm: 0.08, NNZs: 2298, Bias: -0.192893, T: 133335, Avg. loss: 0.627333
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 0.08, NNZs: 2298, Bias: -0.193245, T: 160002, Avg. loss: 0.626259
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 0.08, NNZs: 2298, Bias: -0.193546, T: 186669, Avg. loss: 0.626065
Total training time: 1.58 seconds.
Convergence after 7 epochs took 1.58 seconds
-- Epoch 1
```

```
-- Epoch 1
Norm: 0.10, NNZs: 2298, Bias: -0.227003, T: 26667, Avg. loss: 0.650882
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.09, NNZs: 2298, Bias: -0.228089, T: 53334, Avg. loss: 0.615936
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 0.09, NNZs: 2298, Bias: -0.228785, T: 80001, Avg. loss: 0.614705
Total training time: 0.67 seconds.
-- Epoch 4
Norm: 0.09, NNZs: 2298, Bias: -0.229284, T: 106668, Avg. loss: 0.613883
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.09, NNZs: 2298, Bias: -0.229671, T: 133335, Avg. loss: 0.614138
Total training time: 1.12 seconds.
-- Epoch 6
Norm: 0.09, NNZs: 2298, Bias: -0.229988, T: 160002, Avg. loss: 0.613864
Total training time: 1.34 seconds.
-- Epoch 7
Norm: 0.09, NNZs: 2298, Bias: -0.230244, T: 186669, Avg. loss: 0.612833
Total training time: 1.57 seconds.
-- Epoch 8
Norm: 0.09, NNZs: 2298, Bias: -0.230469, T: 213336, Avg. loss: 0.612755
Total training time: 1.79 seconds.
-- Epoch 9
Norm: 0.09, NNZs: 2298, Bias: -0.230670, T: 240003, Avg. loss: 0.612886
Total training time: 2.02 seconds.
-- Epoch 10
Norm: 0.09, NNZs: 2298, Bias: -0.230849, T: 266670, Avg. loss: 0.612248
Total training time: 2.24 seconds.
-- Epoch 11
Norm: 0.09, NNZs: 2298, Bias: -0.231018, T: 293337, Avg. loss: 0.612846
Total training time: 2.46 seconds.
-- Epoch 12
Norm: 0.09, NNZs: 2298, Bias: -0.231167, T: 320004, Avg. loss: 0.612042
Total training time: 2.69 seconds.
Convergence after 12 epochs took 2.69 seconds

[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed:  4.6min finished

-- Epoch 1
Norm: 5.96, NNZs: 2298, Bias: -5.465162, T: 40000, Avg. loss: 11.097174
Total training time: 0.33 seconds.
-- Epoch 2
Norm: 3.32, NNZs: 2298, Bias: -5.068262, T: 80000, Avg. loss: 1.314259
Total training time: 0.66 seconds.
-- Epoch 3
Norm: 2.44, NNZs: 2298, Bias: -4.797968, T: 120000, Avg. loss: 0.561795
Total training time: 0.99 seconds.
-- Epoch 4
Norm: 2.10, NNZs: 2298, Bias: -4.602059, T: 160000, Avg. loss: 0.445532
Total training time: 1.31 seconds.
-- Epoch 5
Norm: 1.91, NNZs: 2298, Bias: -4.455615, T: 200000, Avg. loss: 0.409067
Total training time: 1.65 seconds.
-- Epoch 6
Norm: 1.80, NNZs: 2298, Bias: -4.341571, T: 240000, Avg. loss: 0.392574
Total training time: 1.97 seconds.
-- Epoch 7
Norm: 1.71, NNZs: 2298, Bias: -4.252221, T: 280000, Avg. loss: 0.385269
Total training time: 2.30 seconds.
-- Epoch 8
Norm: 1.66, NNZs: 2298, Bias: -4.171628, T: 320000, Avg. loss: 0.380347
Total training time: 2.62 seconds.
-- Epoch 9
Norm: 1.61, NNZs: 2298, Bias: -4.104134, T: 360000, Avg. loss: 0.380509
Total training time: 2.95 seconds.
-- Epoch 10
Norm: 1.58, NNZs: 2298, Bias: -4.043437, T: 400000, Avg. loss: 0.372749
Total training time: 3.27 seconds.
-- Epoch 11
Norm: 1.55, NNZs: 2298, Bias: -3.988347, T: 440000, Avg. loss: 0.370648
Total training time: 3.60 seconds.
-- Epoch 12
Norm: 1.53, NNZs: 2298, Bias: -3.940540, T: 480000, Avg. loss: 0.374970
Total training time: 3.92 seconds.
```

```
-- Epoch 13
Norm: 1.51, NNZs: 2298, Bias: -3.896748, T: 520000, Avg. loss: 0.371117
Total training time: 4.24 seconds.
-- Epoch 14
Norm: 1.49, NNZs: 2298, Bias: -3.856574, T: 560000, Avg. loss: 0.367475
Total training time: 4.58 seconds.
-- Epoch 15
Norm: 1.48, NNZs: 2298, Bias: -3.819748, T: 600000, Avg. loss: 0.368201
Total training time: 4.91 seconds.
-- Epoch 16
Norm: 1.46, NNZs: 2298, Bias: -3.787973, T: 640000, Avg. loss: 0.375359
Total training time: 5.24 seconds.
-- Epoch 17
Norm: 1.45, NNZs: 2298, Bias: -3.755563, T: 680000, Avg. loss: 0.365896
Total training time: 5.56 seconds.
-- Epoch 18
Norm: 1.44, NNZs: 2298, Bias: -3.726300, T: 720000, Avg. loss: 0.368921
Total training time: 5.89 seconds.
-- Epoch 19
Norm: 1.43, NNZs: 2298, Bias: -3.698373, T: 760000, Avg. loss: 0.369372
Total training time: 6.21 seconds.
-- Epoch 20
Norm: 1.42, NNZs: 2298, Bias: -3.673379, T: 800000, Avg. loss: 0.372054
Total training time: 6.54 seconds.
-- Epoch 21
Norm: 1.42, NNZs: 2298, Bias: -3.647958, T: 840000, Avg. loss: 0.364078
Total training time: 6.86 seconds.
-- Epoch 22
Norm: 1.40, NNZs: 2298, Bias: -3.626343, T: 880000, Avg. loss: 0.373744
Total training time: 7.19 seconds.
-- Epoch 23
Norm: 1.40, NNZs: 2298, Bias: -3.604150, T: 920000, Avg. loss: 0.368032
Total training time: 7.52 seconds.
-- Epoch 24
Norm: 1.38, NNZs: 2298, Bias: -3.584873, T: 960000, Avg. loss: 0.374524
Total training time: 7.84 seconds.
-- Epoch 25
Norm: 1.38, NNZs: 2298, Bias: -3.564784, T: 1000000, Avg. loss: 0.367819
Total training time: 8.17 seconds.
-- Epoch 26
Norm: 1.37, NNZs: 2298, Bias: -3.547208, T: 1040000, Avg. loss: 0.374486
Total training time: 8.50 seconds.
Convergence after 26 epochs took 8.50 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=-1,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
0	19.433624	2.867074	0.060713	0.009647	0.001	{'alpha': 0.001}	0.719689	0.711379
1	12.237974	1.909266	0.057043	0.007077	0.01	{'alpha': 0.01}	0.752528	0.760771
2	13.657126	8.523234	0.058858	0.008870	0.1	{'alpha': 0.1}	0.789858	0.807171
3	40.012982	49.638022	0.063406	0.007919	1	{'alpha': 1}	0.673183	0.827017
4	5.891976	4.828762	0.057103	0.008190	10	{'alpha': 10}	0.659822	0.806946

In [ ]:

```
# clf = SGDClassifier(loss = 'log',alpha = 0.001, class_weight= 'balanced', n_jobs = -1)
sgd = classifier.best_estimator_
```

In [ ]:

```
sgd.fit(X_train_appetency,y_train_appetency)
```

```
-- Epoch 1
Norm: 5.96, NNZs: 2298, Bias: -5.465162, T: 40000, Avg. loss: 11.097174
Total training time: 0.32 seconds.
-- Epoch 2
Norm: 3.32, NNZs: 2298, Bias: -5.068262, T: 80000, Avg. loss: 1.314259
Total training time: 0.65 seconds.
-- Epoch 3
Norm: 2.44, NNZs: 2298, Bias: -4.797968, T: 120000, Avg. loss: 0.561795
Total training time: 0.98 seconds.
-- Epoch 4
Norm: 2.10, NNZs: 2298, Bias: -4.602059, T: 160000, Avg. loss: 0.445532
Total training time: 1.31 seconds.
-- Epoch 5
Norm: 1.91, NNZs: 2298, Bias: -4.455615, T: 200000, Avg. loss: 0.409067
Total training time: 1.63 seconds.
-- Epoch 6
Norm: 1.80, NNZs: 2298, Bias: -4.341571, T: 240000, Avg. loss: 0.392574
Total training time: 1.96 seconds.
-- Epoch 7
Norm: 1.71, NNZs: 2298, Bias: -4.252221, T: 280000, Avg. loss: 0.385269
Total training time: 2.29 seconds.
-- Epoch 8
Norm: 1.66, NNZs: 2298, Bias: -4.171628, T: 320000, Avg. loss: 0.380347
Total training time: 2.61 seconds.
-- Epoch 9
Norm: 1.61, NNZs: 2298, Bias: -4.104134, T: 360000, Avg. loss: 0.380509
Total training time: 2.94 seconds.
-- Epoch 10
Norm: 1.58, NNZs: 2298, Bias: -4.043437, T: 400000, Avg. loss: 0.372749
Total training time: 3.27 seconds.
-- Epoch 11
Norm: 1.55, NNZs: 2298, Bias: -3.988347, T: 440000, Avg. loss: 0.370648
Total training time: 3.59 seconds.
-- Epoch 12
Norm: 1.53, NNZs: 2298, Bias: -3.940540, T: 480000, Avg. loss: 0.374970
Total training time: 3.93 seconds.
-- Epoch 13
Norm: 1.51, NNZs: 2298, Bias: -3.896748, T: 520000, Avg. loss: 0.371117
Total training time: 4.25 seconds.
-- Epoch 14
Norm: 1.49, NNZs: 2298, Bias: -3.856574, T: 560000, Avg. loss: 0.367475
Total training time: 4.58 seconds.
-- Epoch 15
Norm: 1.48, NNZs: 2298, Bias: -3.819748, T: 600000, Avg. loss: 0.368201
Total training time: 4.91 seconds.
-- Epoch 16
Norm: 1.46, NNZs: 2298, Bias: -3.787973, T: 640000, Avg. loss: 0.375359
```

```
Total training time: 5.23 seconds.
-- Epoch 17
Norm: 1.45, NNZs: 2298, Bias: -3.755563, T: 680000, Avg. loss: 0.365896
Total training time: 5.55 seconds.
-- Epoch 18
Norm: 1.44, NNZs: 2298, Bias: -3.726300, T: 720000, Avg. loss: 0.368921
Total training time: 5.88 seconds.
-- Epoch 19
Norm: 1.43, NNZs: 2298, Bias: -3.698373, T: 760000, Avg. loss: 0.369372
Total training time: 6.21 seconds.
-- Epoch 20
Norm: 1.42, NNZs: 2298, Bias: -3.673379, T: 800000, Avg. loss: 0.372054
Total training time: 6.53 seconds.
-- Epoch 21
Norm: 1.42, NNZs: 2298, Bias: -3.647958, T: 840000, Avg. loss: 0.364078
Total training time: 6.85 seconds.
-- Epoch 22
Norm: 1.40, NNZs: 2298, Bias: -3.626343, T: 880000, Avg. loss: 0.373744
Total training time: 7.18 seconds.
-- Epoch 23
Norm: 1.40, NNZs: 2298, Bias: -3.604150, T: 920000, Avg. loss: 0.368032
Total training time: 7.51 seconds.
-- Epoch 24
Norm: 1.38, NNZs: 2298, Bias: -3.584873, T: 960000, Avg. loss: 0.374524
Total training time: 7.84 seconds.
-- Epoch 25
Norm: 1.38, NNZs: 2298, Bias: -3.564784, T: 1000000, Avg. loss: 0.367819
Total training time: 8.17 seconds.
-- Epoch 26
Norm: 1.37, NNZs: 2298, Bias: -3.547208, T: 1040000, Avg. loss: 0.374486
Total training time: 8.49 seconds.
Convergence after 26 epochs took 8.49 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.1, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=-1, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_appetency_pred = sgd.predict_proba(X_train_appetency)[:,1]
y_test_appetency_pred = sgd.predict_proba(X_test_appetency)[:,1]
```

In [ ]:

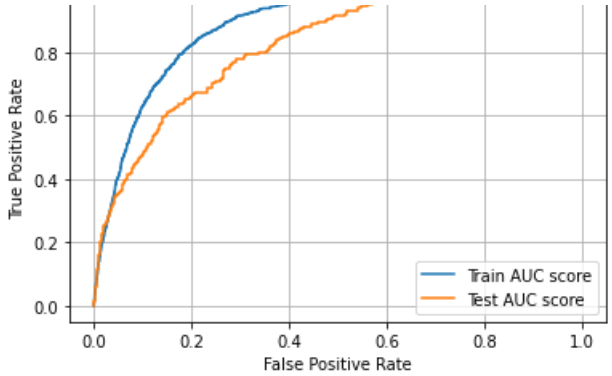
```
lr_train_auc_score_appetency = roc_auc_score(y_train_appetency, y_train_appetency_pred)
lr_test_auc_score_appetency = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Logisitic Regression) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
lr_appetency_score = ['Logistic Regression (Appetency)', lr_train_auc_score_appetency, lr_test_auc_scor
e_appetency]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_score)
```

In [ ]:

```
print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency) | 0.8866793102619893 | 0.8190018280448158 |
+---------------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced',n_jobs = -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= Tr
ue)
```

In [ ]:

```
classifier.fit(X_train_appetency,y_train_appetency)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  90 out of  90 | elapsed: 60.0min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
```

```
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	2.404836	0.672921	0.195657	0.014780	3	10	{'max_dep 3, 'n_estima 10}
1	3.199440	0.163946	0.227598	0.066346	3	20	{'max_dep 3, 'n_estima 20}
2	7.715899	0.194233	0.188699	0.004464	3	50	{'max_dep 3, 'n_estima 50}
3	14.522333	0.077993	0.285643	0.001073	3	100	{'max_dep 3, 'n_estima 100}
4	28.247815	0.046415	0.385333	0.000907	3	200	{'max_dep 3, 'n_estima 200}
5	69.868657	0.123985	0.689737	0.009073	3	500	{'max_dep 3, 'n_estima 500}
6	2.796354	0.037016	0.182027	0.004830	5	10	{'max_dep 5, 'n_estima 10}
7	4.960933	0.028561	0.182110	0.004774	5	20	{'max_dep 5, 'n_estima

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	20} pa
8	11.342389	0.060556	0.181496	0.004161	5	50	{'max_dep 5, 'n_estima 50}
9	22.010760	0.049281	0.284619	0.003715	5	100	{'max_dep 5, 'n_estima 100}
10	43.718124	0.708521	0.492453	0.006016	5	200	{'max_dep 5, 'n_estima 200}
11	105.760339	0.471855	0.892544	0.006845	5	500	{'max_dep 5, 'n_estima 500}
12	3.322253	0.054223	0.179559	0.002621	7	10	{'max_dep 7, 'n_estima 10}
13	5.995891	0.043548	0.178607	0.001250	7	20	{'max_dep 7, 'n_estima 20}
14	14.155534	0.087581	0.247436	0.045932	7	50	{'max_dep 7, 'n_estima 50}
15	27.739626	0.047551	0.282702	0.002809	7	100	{'max_dep 7, 'n_estima 100}
16	55.521966	0.692381	0.490392	0.006718	7	200	{'max_dep 7, 'n_estima 200}
17	139.063323	0.474064	1.090393	0.006166	7	500	{'max_dep 7, 'n_estima 500}
18	3.893305	0.077066	0.179612	0.002275	10	10	{'max_dep 10, 'n_estima 10}
19	7.175249	0.058715	0.180779	0.002154	10	20	{'max_dep 10, 'n_estima 20}
20	16.907720	0.082601	0.280775	0.000641	10	50	{'max_dep 10, 'n_estima 50}
21	33.874245	0.639884	0.397504	0.021898	10	100	{'max_dep 10, 'n_estima 100}
							{'max_dep 10

22	65.598499	0.359091	0.586026	0.006177	10	200	pa 'n_estima 200}
	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	
23	163.407354	0.921737	1.306817	0.095925	10	500	{'max_dep 10, 'n_estima 500}
24	4.165524	0.019994	0.180654	0.001030	15	10	{'max_dep 15, 'n_estima 10}
25	7.531246	0.023514	0.179766	0.000995	15	20	{'max_dep 15, 'n_estima 20}
26	18.047251	0.151678	0.282876	0.001336	15	50	{'max_dep 15, 'n_estima 50}
27	35.460121	0.596604	0.386842	0.003537	15	100	{'max_dep 15, 'n_estima 100}
28	70.105403	0.892473	0.652016	0.043635	15	200	{'max_dep 15, 'n_estima 200}
29	174.632221	1.599515	1.391962	0.004348	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
clf = RandomForestClassifier(n_estimators= 500,max_depth= 5, n_jobs= -1, verbose=1, class_weight= 'balanced')
# best_rf = classifier_rf.best_estimator_
```

In [ ]:

```
clf.fit(X_train_appetency,y_train_appetency)
```

```
[Parallel(n_jobs=-1)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:   11.4s
[Parallel(n_jobs=-1)]: Done 196 tasks      | elapsed:   48.5s
[Parallel(n_jobs=-1)]: Done 446 tasks      | elapsed:  1.8min
[Parallel(n_jobs=-1)]: Done 500 out of 500 | elapsed:  2.0min finished
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=5, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=500,
                       n_jobs=-1, oob_score=False, random_state=None, verbose=1,
                       warm_start=False)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency)[:,1]
```

```
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.3s
[Parallel(n_jobs=2)]: Done 196 tasks      | elapsed:    1.2s
[Parallel(n_jobs=2)]: Done 446 tasks      | elapsed:    2.7s
[Parallel(n_jobs=2)]: Done 500 out of 500 | elapsed:    3.0s finished
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.1s
[Parallel(n_jobs=2)]: Done 196 tasks      | elapsed:    0.3s
[Parallel(n_jobs=2)]: Done 446 tasks      | elapsed:    0.6s
[Parallel(n_jobs=2)]: Done 500 out of 500 | elapsed:    0.7s finished
```

In [ ]:

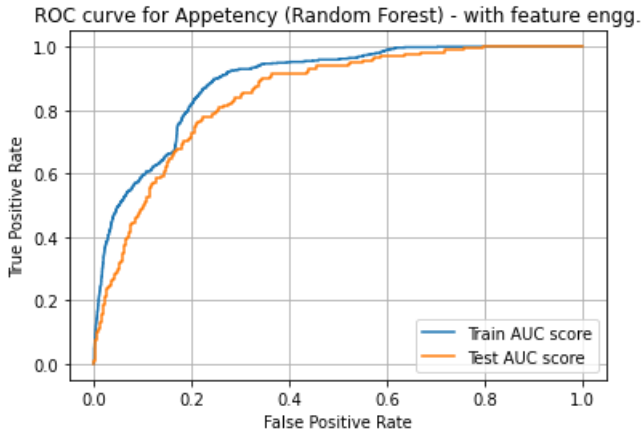
```
rf_train_auc_score_appetency = roc_auc_score(y_train_appetency, y_train_appetency_pred)
rf_test_auc_score_appetency = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Random Forest) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_appetency_score = ['Random Forest (Appetency)', rf_train_auc_score_appetency, rf_test_auc_score_appe
tency]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_appetency_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------------------+--------------------+--------------------+
```

```
|           Model           |     Train AUC      |      Test AUC      |
+---------------------------+--------------------+--------------------+
| Random Forest (Appetency) | 0.8883534212350627 | 0.8459632011604309 |
+---------------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_appetency, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs = -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,300], 'max_depth' : [1,2,3,4]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= Tr
ue)
```

In [ ]:

```
classifier.fit(X_train_appetency,y_train_appetency)
```

```
Fitting 3 folds for each of 20 candidates, totalling 60 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  60 out of  60 | elapsed: 108.6min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=55.17977528089887,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [1, 2, 3, 4],
                         'n_estimators': [10, 20, 50, 100, 300]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	10.896650	1.006601	0.256308	0.012611	1	10	{'max_dep 1, 'n_estima 10}
1	15.024864	0.044993	0.252607	0.009390	1	20	{'max_dep 1, 'n_estima 20}
2	30.008708	0.069633	0.264239	0.006869	1	50	{'max_dep 1, 'n_estima 50}
3	55.690285	1.036762	0.434746	0.234567	1	100	{'max_dep 1, 'n_estima 100}
4	154.877685	0.459197	0.320326	0.003197	1	300	{'max_dep 1, 'n_estima 300}
5	13.850022	0.055023	0.271097	0.001952	2	10	{'max_dep 2, 'n_estima 10}
6	22.537746	0.026305	0.274178	0.007134	2	20	{'max_dep 2, 'n_estima 20}
7	49.662612	0.924375	0.404167	0.161926	2	50	{'max_dep 2, 'n_estima 50}
8	92.289390	0.276860	0.291416	0.009416	2	100	{'max_dep 2, 'n_estima 100}
9	265.418526	0.786601	0.510139	0.242667	2	300	{'max_dep 2, 'n_estima 300}
10	18.819978	1.409670	0.421020	0.214798	3	10	{'max_dep 3, 'n_estima 10}
11	30.660357	0.101253	0.266253	0.003326	3	20	{'max_dep 3, 'n_estima 20}
12	68.550236	0.102595	0.283168	0.010766	3	50	{'max_dep 3, 'n_estima 50}
13	131.494644	0.613038	0.305041	0.004697	3	100	{'max_dep 3, 'n_estima 100}
							{'max_dep

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
14	378.184749	0.604861	0.354010	0.020870	3	300	3, 'n_estima 300}
15	21.814586	0.046694	0.259070	0.016768	4	10	{'max_dep 4, 'n_estima 10}
16	38.651040	0.031177	0.260928	0.012717	4	20	{'max_dep 4, 'n_estima 20}
17	87.742131	0.113951	0.280561	0.015706	4	50	{'max_dep 4, 'n_estima 50}
18	169.248799	0.221922	0.301061	0.013599	4	100	{'max_dep 4, 'n_estima 100}
19	496.849267	1.474120	0.381283	0.013197	4	300	{'max_dep 4, 'n_estima 300}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_appetency,y_train_appetency)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=1,
              min_child_weight=1, missing=None, n_estimators=100, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=55.17977528089887,
              seed=None, silent=None, subsample=1, verbosity=1)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency)[:,1]
```

In [ ]:

```
gbdt_train_auc_score_appetency = roc_auc_score(y_train_appetency, y_train_appetency_pred)
gbdt_test_auc_score_appetency = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

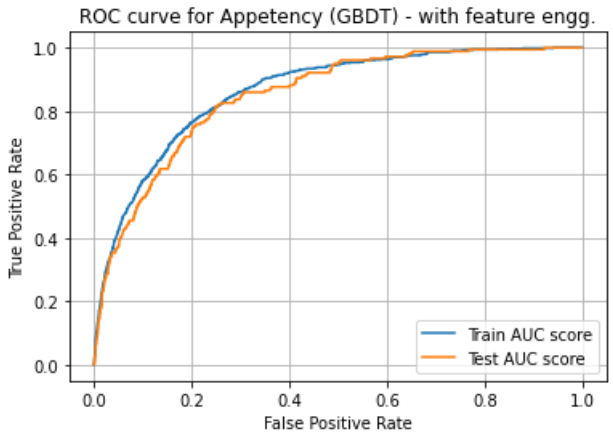
In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
```

```
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (GBDT) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
gbdt_appetency_score = ['GBDT (Appetency)', gbdt_train_auc_score_appetency, gbdt_test_auc_score_appeten
cy]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_appetency_score)
```

In [ ]:

```
print(score_table)
```

```
+------------------+-------------------+--------------------+
|      Model       |     Train AUC     |      Test AUC      |
+------------------+-------------------+--------------------+
| GBDT (Appetency) | 0.863205203607357 | 0.8496999398278114 |
+------------------+-------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_score)
score_table.add_row(rf_appetency_score)
score_table.add_row(gbdt_appetency_score)

print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency) | 0.8866793102619893 | 0.8190018280448158 |
|    Random Forest (Appetency)    | 0.8883534212350627 | 0.8459632011604309 |
|         GBDT (Appetency)        | 0.863205203607357  | 0.8496999398278114 |
+---------------------------------+--------------------+--------------------+
```

Observation:

- **Both RF and GBDT have near equal Test AUC score in Appetency Dataset.**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.1, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 500, max_depth= 5, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_appetency, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 100, max_depth= 1, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.001,0.01,0.1,1,10]}
```

In [ ]:

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_appetency,y_train_appetency)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 120.1min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
                                                                       eta0=0.0,
                                                                       fit_intercept=True,
                                                                       l1_ratio=0.15,
                                                                       learning_rate='optimal',
                                                                       loss='log',
                                                                       max_iter=1000,
                                                                       n_iter_no_change=5,
                                                                       n_jobs=-1,
                                                                       penalty='l2',
```

```
                                                        power_t=0.5,
                                                        random_state=None,
                                                        shuffle=True,
                                                        t...
                                                           validation_fraction=0.1,
                                                           verbose=0,
                                                           warm_start=False),
                                        shuffle=True,
                                        store_train_meta_features=False,
                                        stratify=True, use_clones=True,
                                        use_features_in_secondary=False,
                                        use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	452.164535	14.372543	1.170939	0.059586	0.001	{'meta_classifier__al 0.001}
1	459.108813	2.877518	1.088599	0.036774	0.01	{'meta_classifier__al 0.01}
2	459.782142	14.620391	1.199291	0.090281	0.1	{'meta_classifier__al 0.1}
3	493.833181	44.660702	1.287732	0.083239	1	{'meta_classifier__al 1}
4	517.381957	9.461368	1.772353	0.435529	10	{'meta_classifier__al 10}

In [ ]:

```
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_appetency,y_train_appetency)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1, average=False,
                                                 class_weight='balanced',
                                                 early_stopping=False,
                                                 epsilon=0.1, eta0=0.0,
                                                 fit_intercept=True,
                                                 l1_ratio=0.15,
                                                 learning_rate='optimal',
                                                 loss='log', max_iter=1000,
                                                 n_iter_no_change=5, n_jobs=-1,
                                                 penalty='l2', power_t=0.5,
                                                 random_state=None, shuffle=True,
                                                 tol=0.001,
                                                 validation_fraction=0.1,
                                                 verbose=0, w...
                                                    fit_intercept=True,
                                                    l1_ratio=0.15,
```

```
                                        learning_rate='optimal',
                                        loss='log', max_iter=1000,
                                        n_iter_no_change=5,
                                        n_jobs=-1, penalty='l2',
                                        power_t=0.5,
                                        random_state=None,
                                        shuffle=True, tol=0.001,
                                        validation_fraction=0.1,
                                        verbose=0,
                                        warm_start=False),
                    shuffle=True, store_train_meta_features=False,
                    stratify=True, use_clones=True,
                    use_features_in_secondary=False, use_probas=True,
                    verbose=0)
```

In [ ]:

```
y_train_appetency_pred = clf.predict_proba(X_train_appetency)[:,1]
y_test_appetency_pred = clf.predict_proba(X_test_appetency)[:,1]
```

In [ ]:

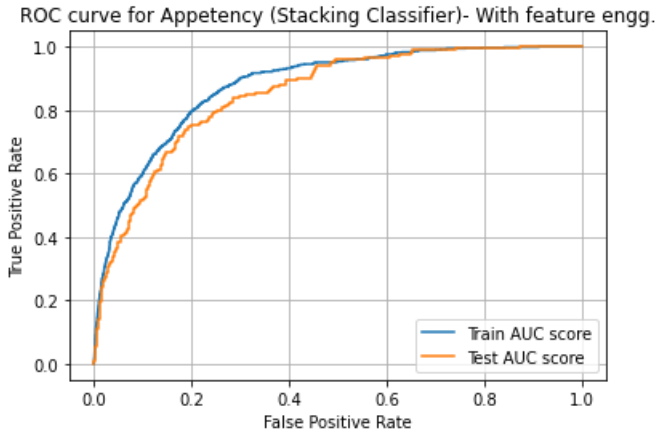
```
stack_train_auc_score_appetency = roc_auc_score(y_train_appetency, y_train_appetency_pred)
stack_test_auc_score_appetency = roc_auc_score(y_test_appetency, y_test_appetency_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_appetency,y_train_appetency_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_appetency,y_test_appetency_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Appetency (Stacking Classifier)- With feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_appetency_score = ['Stacking Classifier (apptency)', stack_train_auc_score_appetency, stack_test_
auc_score_appetency]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
```

```
score_table.add_row(stack_appetency_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------------------------+--------------------+--------------------+
|             Model              |     Train AUC      |      Test AUC      |
+--------------------------------+--------------------+--------------------+
| Stacking Classifier (apptency) | 0.8777009204857703 | 0.8510881328089429 |
+--------------------------------+--------------------+--------------------+
```

Churn

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13, n_jobs= -1)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= True)
```

In [ ]:

```
classifier.fit(X_train_churn, y_train_churn)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
```

```
-- Epoch 1
Norm: 146.75, NNZs: 2298, Bias: -52.560917, T: 26666, Avg. loss: 208.833706
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 91.83, NNZs: 2298, Bias: -44.833715, T: 53332, Avg. loss: 33.889825
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 62.59, NNZs: 2298, Bias: -40.650446, T: 79998, Avg. loss: 18.394564
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 52.80, NNZs: 2298, Bias: -36.375323, T: 106664, Avg. loss: 12.985991
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 45.90, NNZs: 2298, Bias: -33.047807, T: 133330, Avg. loss: 9.580014
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 41.21, NNZs: 2298, Bias: -31.111095, T: 159996, Avg. loss: 7.795679
Total training time: 1.31 seconds.
-- Epoch 7
Norm: 36.53, NNZs: 2298, Bias: -29.468897, T: 186662, Avg. loss: 6.591227
Total training time: 1.53 seconds.
-- Epoch 8
Norm: 33.60, NNZs: 2298, Bias: -27.863739, T: 213328, Avg. loss: 5.645612
Total training time: 1.75 seconds.
-- Epoch 9
Norm: 32.58, NNZs: 2298, Bias: -26.210960, T: 239994, Avg. loss: 4.887295
Total training time: 1.98 seconds.
-- Epoch 10
```

```
Norm: 29.88, NNZs: 2298, Bias: -25.151800, T: 266660, Avg. loss: 4.407773
Total training time: 2.20 seconds.
-- Epoch 11
Norm: 28.40, NNZs: 2298, Bias: -24.429031, T: 293326, Avg. loss: 4.024331
Total training time: 2.41 seconds.
-- Epoch 12
Norm: 26.26, NNZs: 2298, Bias: -23.788074, T: 319992, Avg. loss: 3.638492
Total training time: 2.63 seconds.
-- Epoch 13
Norm: 26.00, NNZs: 2298, Bias: -22.535193, T: 346658, Avg. loss: 3.402198
Total training time: 2.84 seconds.
-- Epoch 14
Norm: 24.71, NNZs: 2298, Bias: -21.956816, T: 373324, Avg. loss: 3.006286
Total training time: 3.06 seconds.
-- Epoch 15
Norm: 23.56, NNZs: 2298, Bias: -21.424114, T: 399990, Avg. loss: 2.914827
Total training time: 3.28 seconds.
-- Epoch 16
Norm: 22.89, NNZs: 2298, Bias: -20.758483, T: 426656, Avg. loss: 2.607199
Total training time: 3.50 seconds.
-- Epoch 17
Norm: 22.10, NNZs: 2298, Bias: -20.320769, T: 453322, Avg. loss: 2.521963
Total training time: 3.71 seconds.
-- Epoch 18
Norm: 21.50, NNZs: 2298, Bias: -19.816035, T: 479988, Avg. loss: 2.366383
Total training time: 3.93 seconds.
-- Epoch 19
Norm: 20.72, NNZs: 2298, Bias: -19.425383, T: 506654, Avg. loss: 2.208455
Total training time: 4.15 seconds.
-- Epoch 20
Norm: 20.16, NNZs: 2298, Bias: -18.980684, T: 533320, Avg. loss: 2.092447
Total training time: 4.36 seconds.
-- Epoch 21
Norm: 19.87, NNZs: 2298, Bias: -18.543434, T: 559986, Avg. loss: 2.059824
Total training time: 4.57 seconds.
-- Epoch 22
Norm: 19.24, NNZs: 2298, Bias: -18.220421, T: 586652, Avg. loss: 1.818143
Total training time: 4.79 seconds.
-- Epoch 23
Norm: 19.05, NNZs: 2298, Bias: -17.751216, T: 613318, Avg. loss: 1.894374
Total training time: 5.01 seconds.
-- Epoch 24
Norm: 18.57, NNZs: 2298, Bias: -17.468222, T: 639984, Avg. loss: 1.776827
Total training time: 5.23 seconds.
-- Epoch 25
Norm: 17.99, NNZs: 2298, Bias: -17.282250, T: 666650, Avg. loss: 1.656797
Total training time: 5.44 seconds.
-- Epoch 26
Norm: 17.89, NNZs: 2298, Bias: -16.809779, T: 693316, Avg. loss: 1.541662
Total training time: 5.66 seconds.
-- Epoch 27
Norm: 17.55, NNZs: 2298, Bias: -16.553572, T: 719982, Avg. loss: 1.584923
Total training time: 5.87 seconds.
-- Epoch 28
Norm: 17.12, NNZs: 2298, Bias: -16.409782, T: 746648, Avg. loss: 1.582100
Total training time: 6.10 seconds.
-- Epoch 29
Norm: 16.89, NNZs: 2298, Bias: -16.078383, T: 773314, Avg. loss: 1.475682
Total training time: 6.31 seconds.
-- Epoch 30
Norm: 16.69, NNZs: 2298, Bias: -15.902674, T: 799980, Avg. loss: 1.499792
Total training time: 6.53 seconds.
-- Epoch 31
Norm: 16.37, NNZs: 2298, Bias: -15.682564, T: 826646, Avg. loss: 1.349246
Total training time: 6.74 seconds.
-- Epoch 32
Norm: 16.10, NNZs: 2298, Bias: -15.489238, T: 853312, Avg. loss: 1.356520
Total training time: 6.96 seconds.
-- Epoch 33
Norm: 15.82, NNZs: 2298, Bias: -15.272798, T: 879978, Avg. loss: 1.247882
Total training time: 7.18 seconds.
-- Epoch 34
Norm: 15.69, NNZs: 2298, Bias: -15.049969, T: 906644, Avg. loss: 1.236129
Total training time: 7.39 seconds.
-- Epoch 35
Norm: 15.40, NNZs: 2298, Bias: -14.841065, T: 933310, Avg. loss: 1.262491
Total training time: 7.61 seconds.
```

```
-- Epoch 36
Norm: 15.20, NNZs: 2298, Bias: -14.754756, T: 959976, Avg. loss: 1.215508
Total training time: 7.83 seconds.
-- Epoch 37
Norm: 15.07, NNZs: 2298, Bias: -14.543944, T: 986642, Avg. loss: 1.204542
Total training time: 8.05 seconds.
-- Epoch 38
Norm: 14.84, NNZs: 2298, Bias: -14.366440, T: 1013308, Avg. loss: 1.175419
Total training time: 8.26 seconds.
-- Epoch 39
Norm: 14.59, NNZs: 2298, Bias: -14.255010, T: 1039974, Avg. loss: 1.112006
Total training time: 8.48 seconds.
-- Epoch 40
Norm: 14.37, NNZs: 2298, Bias: -14.137707, T: 1066640, Avg. loss: 1.145285
Total training time: 8.69 seconds.
-- Epoch 41
Norm: 14.33, NNZs: 2298, Bias: -13.852994, T: 1093306, Avg. loss: 1.040097
Total training time: 8.91 seconds.
-- Epoch 42
Norm: 14.13, NNZs: 2298, Bias: -13.786485, T: 1119972, Avg. loss: 1.039693
Total training time: 9.13 seconds.
-- Epoch 43
Norm: 13.90, NNZs: 2298, Bias: -13.673049, T: 1146638, Avg. loss: 1.047516
Total training time: 9.34 seconds.
-- Epoch 44
Norm: 13.76, NNZs: 2298, Bias: -13.535623, T: 1173304, Avg. loss: 1.029011
Total training time: 9.55 seconds.
-- Epoch 45
Norm: 13.62, NNZs: 2298, Bias: -13.402313, T: 1199970, Avg. loss: 0.996853
Total training time: 9.77 seconds.
-- Epoch 46
Norm: 13.52, NNZs: 2298, Bias: -13.242754, T: 1226636, Avg. loss: 1.048759
Total training time: 9.99 seconds.
-- Epoch 47
Norm: 13.36, NNZs: 2298, Bias: -13.133301, T: 1253302, Avg. loss: 1.006262
Total training time: 10.22 seconds.
-- Epoch 48
Norm: 13.15, NNZs: 2298, Bias: -13.085816, T: 1279968, Avg. loss: 0.939719
Total training time: 10.44 seconds.
-- Epoch 49
Norm: 13.07, NNZs: 2298, Bias: -12.915718, T: 1306634, Avg. loss: 0.940974
Total training time: 10.65 seconds.
-- Epoch 50
Norm: 12.99, NNZs: 2298, Bias: -12.780686, T: 1333300, Avg. loss: 0.938711
Total training time: 10.87 seconds.
-- Epoch 51
Norm: 12.87, NNZs: 2298, Bias: -12.669299, T: 1359966, Avg. loss: 0.935268
Total training time: 11.09 seconds.
-- Epoch 52
Norm: 12.70, NNZs: 2298, Bias: -12.632724, T: 1386632, Avg. loss: 0.915533
Total training time: 11.31 seconds.
-- Epoch 53
Norm: 12.62, NNZs: 2298, Bias: -12.505936, T: 1413298, Avg. loss: 0.915592
Total training time: 11.52 seconds.
-- Epoch 54
Norm: 12.44, NNZs: 2298, Bias: -12.466150, T: 1439964, Avg. loss: 0.884460
Total training time: 11.73 seconds.
-- Epoch 55
Norm: 12.40, NNZs: 2298, Bias: -12.327101, T: 1466630, Avg. loss: 0.870550
Total training time: 11.95 seconds.
-- Epoch 56
Norm: 12.33, NNZs: 2298, Bias: -12.218199, T: 1493296, Avg. loss: 0.898800
Total training time: 12.18 seconds.
-- Epoch 57
Norm: 12.24, NNZs: 2298, Bias: -12.098331, T: 1519962, Avg. loss: 0.881674
Total training time: 12.40 seconds.
-- Epoch 58
Norm: 12.12, NNZs: 2298, Bias: -12.070452, T: 1546628, Avg. loss: 0.889839
Total training time: 12.61 seconds.
-- Epoch 59
Norm: 12.06, NNZs: 2298, Bias: -11.945323, T: 1573294, Avg. loss: 0.858921
Total training time: 12.82 seconds.
-- Epoch 60
Norm: 11.93, NNZs: 2298, Bias: -11.891466, T: 1599960, Avg. loss: 0.852548
Total training time: 13.04 seconds.
-- Epoch 61
Norm: 11.87, NNZs: 2298, Bias: -11.779828, T: 1626626, Avg. loss: 0.822454
```

```
Total training time: 13.26 seconds.
-- Epoch 62
Norm: 11.79, NNZs: 2298, Bias: -11.673207, T: 1653292, Avg. loss: 0.826016
Total training time: 13.48 seconds.
-- Epoch 63
Norm: 11.65, NNZs: 2298, Bias: -11.652931, T: 1679958, Avg. loss: 0.817054
Total training time: 13.69 seconds.
-- Epoch 64
Norm: 11.53, NNZs: 2298, Bias: -11.598421, T: 1706624, Avg. loss: 0.825319
Total training time: 13.91 seconds.
-- Epoch 65
Norm: 11.49, NNZs: 2298, Bias: -11.477080, T: 1733290, Avg. loss: 0.788022
Total training time: 14.13 seconds.
-- Epoch 66
Norm: 11.41, NNZs: 2298, Bias: -11.436815, T: 1759956, Avg. loss: 0.784673
Total training time: 14.34 seconds.
-- Epoch 67
Norm: 11.35, NNZs: 2298, Bias: -11.344135, T: 1786622, Avg. loss: 0.797408
Total training time: 14.56 seconds.
-- Epoch 68
Norm: 11.26, NNZs: 2298, Bias: -11.296411, T: 1813288, Avg. loss: 0.790967
Total training time: 14.78 seconds.
-- Epoch 69
Norm: 11.20, NNZs: 2298, Bias: -11.220407, T: 1839954, Avg. loss: 0.780886
Total training time: 14.99 seconds.
-- Epoch 70
Norm: 11.16, NNZs: 2298, Bias: -11.132756, T: 1866620, Avg. loss: 0.788912
Total training time: 15.22 seconds.
-- Epoch 71
Norm: 11.12, NNZs: 2298, Bias: -11.062089, T: 1893286, Avg. loss: 0.785674
Total training time: 15.43 seconds.
-- Epoch 72
Norm: 11.01, NNZs: 2298, Bias: -11.025619, T: 1919952, Avg. loss: 0.786827
Total training time: 15.64 seconds.
-- Epoch 73
Norm: 10.97, NNZs: 2298, Bias: -10.955925, T: 1946618, Avg. loss: 0.746310
Total training time: 15.86 seconds.
-- Epoch 74
Norm: 10.88, NNZs: 2298, Bias: -10.899443, T: 1973284, Avg. loss: 0.763372
Total training time: 16.08 seconds.
-- Epoch 75
Norm: 10.83, NNZs: 2298, Bias: -10.837433, T: 1999950, Avg. loss: 0.754965
Total training time: 16.30 seconds.
-- Epoch 76
Norm: 10.79, NNZs: 2298, Bias: -10.774655, T: 2026616, Avg. loss: 0.751089
Total training time: 16.51 seconds.
-- Epoch 77
Norm: 10.67, NNZs: 2298, Bias: -10.773266, T: 2053282, Avg. loss: 0.744405
Total training time: 16.73 seconds.
-- Epoch 78
Norm: 10.59, NNZs: 2298, Bias: -10.733922, T: 2079948, Avg. loss: 0.740615
Total training time: 16.95 seconds.
-- Epoch 79
Norm: 10.61, NNZs: 2298, Bias: -10.618054, T: 2106614, Avg. loss: 0.731538
Total training time: 17.17 seconds.
-- Epoch 80
Norm: 10.48, NNZs: 2298, Bias: -10.618035, T: 2133280, Avg. loss: 0.726742
Total training time: 17.38 seconds.
-- Epoch 81
Norm: 10.42, NNZs: 2298, Bias: -10.571299, T: 2159946, Avg. loss: 0.728540
Total training time: 17.60 seconds.
-- Epoch 82
Norm: 10.38, NNZs: 2298, Bias: -10.509214, T: 2186612, Avg. loss: 0.728681
Total training time: 17.81 seconds.
-- Epoch 83
Norm: 10.40, NNZs: 2298, Bias: -10.428037, T: 2213278, Avg. loss: 0.743882
Total training time: 18.03 seconds.
-- Epoch 84
Norm: 10.34, NNZs: 2298, Bias: -10.372799, T: 2239944, Avg. loss: 0.720800
Total training time: 18.25 seconds.
-- Epoch 85
Norm: 10.32, NNZs: 2298, Bias: -10.304476, T: 2266610, Avg. loss: 0.709410
Total training time: 18.47 seconds.
-- Epoch 86
Norm: 10.22, NNZs: 2298, Bias: -10.283758, T: 2293276, Avg. loss: 0.715182
Total training time: 18.68 seconds.
-- Epoch 87
```

```
Norm: 10.12, NNZs: 2298, Bias: -10.275978, T: 2319942, Avg. loss: 0.699101
Total training time: 18.90 seconds.
-- Epoch 88
Norm: 10.08, NNZs: 2298, Bias: -10.228152, T: 2346608, Avg. loss: 0.702047
Total training time: 19.12 seconds.
-- Epoch 89
Norm: 10.04, NNZs: 2298, Bias: -10.173500, T: 2373274, Avg. loss: 0.698982
Total training time: 19.34 seconds.
-- Epoch 90
Norm: 9.98, NNZs: 2298, Bias: -10.133184, T: 2399940, Avg. loss: 0.697063
Total training time: 19.56 seconds.
-- Epoch 91
Norm: 9.96, NNZs: 2298, Bias: -10.061769, T: 2426606, Avg. loss: 0.695899
Total training time: 19.77 seconds.
-- Epoch 92
Norm: 9.89, NNZs: 2298, Bias: -10.038060, T: 2453272, Avg. loss: 0.684554
Total training time: 19.99 seconds.
-- Epoch 93
Norm: 9.89, NNZs: 2298, Bias: -9.967454, T: 2479938, Avg. loss: 0.690796
Total training time: 20.21 seconds.
-- Epoch 94
Norm: 9.82, NNZs: 2298, Bias: -9.954043, T: 2506604, Avg. loss: 0.681584
Total training time: 20.43 seconds.
-- Epoch 95
Norm: 9.79, NNZs: 2298, Bias: -9.905260, T: 2533270, Avg. loss: 0.691276
Total training time: 20.65 seconds.
-- Epoch 96
Norm: 9.74, NNZs: 2298, Bias: -9.863895, T: 2559936, Avg. loss: 0.687448
Total training time: 20.87 seconds.
-- Epoch 97
Norm: 9.69, NNZs: 2298, Bias: -9.837094, T: 2586602, Avg. loss: 0.691729
Total training time: 21.09 seconds.
-- Epoch 98
Norm: 9.67, NNZs: 2298, Bias: -9.792899, T: 2613268, Avg. loss: 0.673571
Total training time: 21.31 seconds.
-- Epoch 99
Norm: 9.63, NNZs: 2298, Bias: -9.745238, T: 2639934, Avg. loss: 0.688729
Total training time: 21.52 seconds.
-- Epoch 100
Norm: 9.61, NNZs: 2298, Bias: -9.696071, T: 2666600, Avg. loss: 0.661784
Total training time: 21.74 seconds.
-- Epoch 101
Norm: 9.56, NNZs: 2298, Bias: -9.662312, T: 2693266, Avg. loss: 0.688323
Total training time: 21.95 seconds.
-- Epoch 102
Norm: 9.55, NNZs: 2298, Bias: -9.606092, T: 2719932, Avg. loss: 0.661766
Total training time: 22.17 seconds.
-- Epoch 103
Norm: 9.50, NNZs: 2298, Bias: -9.574581, T: 2746598, Avg. loss: 0.671170
Total training time: 22.38 seconds.
-- Epoch 104
Norm: 9.44, NNZs: 2298, Bias: -9.562287, T: 2773264, Avg. loss: 0.671826
Total training time: 22.60 seconds.
-- Epoch 105
Norm: 9.40, NNZs: 2298, Bias: -9.531207, T: 2799930, Avg. loss: 0.660227
Total training time: 22.82 seconds.
-- Epoch 106
Norm: 9.40, NNZs: 2298, Bias: -9.485877, T: 2826596, Avg. loss: 0.669929
Total training time: 23.03 seconds.
-- Epoch 107
Norm: 9.36, NNZs: 2298, Bias: -9.449613, T: 2853262, Avg. loss: 0.677203
Total training time: 23.25 seconds.
-- Epoch 108
Norm: 9.33, NNZs: 2298, Bias: -9.405991, T: 2879928, Avg. loss: 0.657289
Total training time: 23.47 seconds.
-- Epoch 109
Norm: 9.30, NNZs: 2298, Bias: -9.363079, T: 2906594, Avg. loss: 0.656306
Total training time: 23.68 seconds.
-- Epoch 110
Norm: 9.26, NNZs: 2298, Bias: -9.340739, T: 2933260, Avg. loss: 0.659615
Total training time: 23.90 seconds.
-- Epoch 111
Norm: 9.22, NNZs: 2298, Bias: -9.314977, T: 2959926, Avg. loss: 0.659692
Total training time: 24.11 seconds.
-- Epoch 112
Norm: 9.20, NNZs: 2298, Bias: -9.271555, T: 2986592, Avg. loss: 0.663348
Total training time: 24.33 seconds.
```

```
-- Epoch 113
Norm: 9.15, NNZs: 2298, Bias: -9.255689, T: 3013258, Avg. loss: 0.655445
Total training time: 24.55 seconds.
Convergence after 113 epochs took 24.55 seconds
-- Epoch 1
Norm: 159.38, NNZs: 2297, Bias: -58.483978, T: 26667, Avg. loss: 212.756219
Total training time: 0.23 seconds.
-- Epoch 2
Norm: 109.87, NNZs: 2297, Bias: -46.375844, T: 53334, Avg. loss: 32.807901
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 67.87, NNZs: 2297, Bias: -43.226448, T: 80001, Avg. loss: 23.621297
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 54.99, NNZs: 2297, Bias: -38.863267, T: 106668, Avg. loss: 12.726874
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 52.09, NNZs: 2297, Bias: -35.298440, T: 133335, Avg. loss: 12.752421
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 43.46, NNZs: 2297, Bias: -33.513618, T: 160002, Avg. loss: 8.346450
Total training time: 1.31 seconds.
-- Epoch 7
Norm: 40.14, NNZs: 2297, Bias: -31.213863, T: 186669, Avg. loss: 6.747427
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 36.59, NNZs: 2297, Bias: -29.504714, T: 213336, Avg. loss: 5.994054
Total training time: 1.73 seconds.
-- Epoch 9
Norm: 34.10, NNZs: 2297, Bias: -28.056046, T: 240003, Avg. loss: 5.170506
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 31.59, NNZs: 2297, Bias: -26.734382, T: 266670, Avg. loss: 4.460100
Total training time: 2.25 seconds.
-- Epoch 11
Norm: 29.75, NNZs: 2297, Bias: -25.722296, T: 293337, Avg. loss: 3.962375
Total training time: 2.60 seconds.
-- Epoch 12
Norm: 27.75, NNZs: 2297, Bias: -24.771927, T: 320004, Avg. loss: 3.484632
Total training time: 2.93 seconds.
-- Epoch 13
Norm: 29.03, NNZs: 2297, Bias: -23.351619, T: 346671, Avg. loss: 3.463138
Total training time: 3.14 seconds.
-- Epoch 14
Norm: 26.42, NNZs: 2297, Bias: -23.107572, T: 373338, Avg. loss: 3.640075
Total training time: 3.36 seconds.
-- Epoch 15
Norm: 25.30, NNZs: 2297, Bias: -22.479621, T: 400005, Avg. loss: 2.913975
Total training time: 3.57 seconds.
-- Epoch 16
Norm: 24.49, NNZs: 2297, Bias: -21.773507, T: 426672, Avg. loss: 2.796107
Total training time: 3.79 seconds.
-- Epoch 17
Norm: 23.24, NNZs: 2297, Bias: -21.440874, T: 453339, Avg. loss: 2.446166
Total training time: 4.01 seconds.
-- Epoch 18
Norm: 22.43, NNZs: 2297, Bias: -21.007525, T: 480006, Avg. loss: 2.204111
Total training time: 4.22 seconds.
-- Epoch 19
Norm: 22.05, NNZs: 2297, Bias: -20.264789, T: 506673, Avg. loss: 2.251653
Total training time: 4.44 seconds.
-- Epoch 20
Norm: 21.21, NNZs: 2297, Bias: -20.082990, T: 533340, Avg. loss: 2.110078
Total training time: 4.66 seconds.
-- Epoch 21
Norm: 20.64, NNZs: 2297, Bias: -19.597403, T: 560007, Avg. loss: 2.036991
Total training time: 4.88 seconds.
-- Epoch 22
Norm: 20.34, NNZs: 2297, Bias: -19.004567, T: 586674, Avg. loss: 1.882769
Total training time: 5.09 seconds.
-- Epoch 23
Norm: 19.60, NNZs: 2297, Bias: -18.869921, T: 613341, Avg. loss: 1.854589
Total training time: 5.31 seconds.
-- Epoch 24
Norm: 19.25, NNZs: 2297, Bias: -18.422757, T: 640008, Avg. loss: 1.759489
Total training time: 5.53 seconds.
-- Epoch 25
```

```
Norm: 18.77, NNZs: 2297, Bias: -18.114030, T: 666675, Avg. loss: 1.738995
Total training time: 5.74 seconds.
-- Epoch 26
Norm: 19.31, NNZs: 2297, Bias: -17.555740, T: 693342, Avg. loss: 2.011626
Total training time: 5.96 seconds.
-- Epoch 27
Norm: 18.50, NNZs: 2297, Bias: -17.516204, T: 720009, Avg. loss: 1.700681
Total training time: 6.18 seconds.
-- Epoch 28
Norm: 17.95, NNZs: 2297, Bias: -17.309761, T: 746676, Avg. loss: 1.495746
Total training time: 6.41 seconds.
-- Epoch 29
Norm: 17.70, NNZs: 2297, Bias: -16.946697, T: 773343, Avg. loss: 1.464090
Total training time: 6.63 seconds.
-- Epoch 30
Norm: 17.39, NNZs: 2297, Bias: -16.734682, T: 800010, Avg. loss: 1.496033
Total training time: 6.84 seconds.
-- Epoch 31
Norm: 16.90, NNZs: 2297, Bias: -16.567790, T: 826677, Avg. loss: 1.380137
Total training time: 7.06 seconds.
-- Epoch 32
Norm: 16.70, NNZs: 2297, Bias: -16.271767, T: 853344, Avg. loss: 1.297341
Total training time: 7.28 seconds.
-- Epoch 33
Norm: 16.49, NNZs: 2297, Bias: -16.011116, T: 880011, Avg. loss: 1.264388
Total training time: 7.50 seconds.
-- Epoch 34
Norm: 16.03, NNZs: 2297, Bias: -15.954701, T: 906678, Avg. loss: 1.290558
Total training time: 7.71 seconds.
-- Epoch 35
Norm: 15.90, NNZs: 2297, Bias: -15.701006, T: 933345, Avg. loss: 1.235575
Total training time: 7.93 seconds.
-- Epoch 36
Norm: 15.64, NNZs: 2297, Bias: -15.510423, T: 960012, Avg. loss: 1.195073
Total training time: 8.14 seconds.
-- Epoch 37
Norm: 16.00, NNZs: 2297, Bias: -15.142006, T: 986679, Avg. loss: 1.340233
Total training time: 8.36 seconds.
-- Epoch 38
Norm: 15.42, NNZs: 2297, Bias: -15.163559, T: 1013346, Avg. loss: 1.254770
Total training time: 8.57 seconds.
-- Epoch 39
Norm: 15.36, NNZs: 2297, Bias: -14.890291, T: 1040013, Avg. loss: 1.133338
Total training time: 8.79 seconds.
-- Epoch 40
Norm: 15.14, NNZs: 2297, Bias: -14.727769, T: 1066680, Avg. loss: 1.130897
Total training time: 9.00 seconds.
-- Epoch 41
Norm: 14.87, NNZs: 2297, Bias: -14.654151, T: 1093347, Avg. loss: 1.086348
Total training time: 9.22 seconds.
-- Epoch 42
Norm: 14.70, NNZs: 2297, Bias: -14.484366, T: 1120014, Avg. loss: 1.064307
Total training time: 9.44 seconds.
-- Epoch 43
Norm: 14.47, NNZs: 2297, Bias: -14.382089, T: 1146681, Avg. loss: 1.044123
Total training time: 9.65 seconds.
-- Epoch 44
Norm: 14.30, NNZs: 2297, Bias: -14.260981, T: 1173348, Avg. loss: 1.049950
Total training time: 9.87 seconds.
-- Epoch 45
Norm: 14.06, NNZs: 2297, Bias: -14.166352, T: 1200015, Avg. loss: 1.039313
Total training time: 10.08 seconds.
-- Epoch 46
Norm: 14.01, NNZs: 2297, Bias: -13.978680, T: 1226682, Avg. loss: 0.986397
Total training time: 10.30 seconds.
-- Epoch 47
Norm: 13.89, NNZs: 2297, Bias: -13.791624, T: 1253349, Avg. loss: 1.005366
Total training time: 10.52 seconds.
-- Epoch 48
Norm: 13.76, NNZs: 2297, Bias: -13.708284, T: 1280016, Avg. loss: 1.001180
Total training time: 10.73 seconds.
-- Epoch 49
Norm: 13.56, NNZs: 2297, Bias: -13.622633, T: 1306683, Avg. loss: 0.981268
Total training time: 10.95 seconds.
-- Epoch 50
Norm: 13.41, NNZs: 2297, Bias: -13.501161, T: 1333350, Avg. loss: 0.915341
Total training time: 11.16 seconds.
```

```
-- Epoch 51
Norm: 13.28, NNZs: 2297, Bias: -13.394857, T: 1360017, Avg. loss: 0.953859
Total training time: 11.38 seconds.
-- Epoch 52
Norm: 13.19, NNZs: 2297, Bias: -13.253266, T: 1386684, Avg. loss: 0.897536
Total training time: 11.59 seconds.
-- Epoch 53
Norm: 13.06, NNZs: 2297, Bias: -13.170868, T: 1413351, Avg. loss: 0.917642
Total training time: 11.80 seconds.
-- Epoch 54
Norm: 12.93, NNZs: 2297, Bias: -13.064066, T: 1440018, Avg. loss: 0.898711
Total training time: 12.02 seconds.
-- Epoch 55
Norm: 12.89, NNZs: 2297, Bias: -12.928280, T: 1466685, Avg. loss: 0.876679
Total training time: 12.24 seconds.
-- Epoch 56
Norm: 12.64, NNZs: 2297, Bias: -12.939481, T: 1493352, Avg. loss: 0.880039
Total training time: 12.45 seconds.
-- Epoch 57
Norm: 12.54, NNZs: 2297, Bias: -12.835160, T: 1520019, Avg. loss: 0.851270
Total training time: 12.67 seconds.
-- Epoch 58
Norm: 12.54, NNZs: 2297, Bias: -12.661798, T: 1546686, Avg. loss: 0.872435
Total training time: 12.89 seconds.
-- Epoch 59
Norm: 12.47, NNZs: 2297, Bias: -12.590623, T: 1573353, Avg. loss: 0.902467
Total training time: 13.10 seconds.
-- Epoch 60
Norm: 12.38, NNZs: 2297, Bias: -12.486959, T: 1600020, Avg. loss: 0.827540
Total training time: 13.32 seconds.
-- Epoch 61
Norm: 12.26, NNZs: 2297, Bias: -12.442491, T: 1626687, Avg. loss: 0.839962
Total training time: 13.54 seconds.
-- Epoch 62
Norm: 12.11, NNZs: 2297, Bias: -12.397884, T: 1653354, Avg. loss: 0.819760
Total training time: 13.75 seconds.
-- Epoch 63
Norm: 12.06, NNZs: 2297, Bias: -12.279017, T: 1680021, Avg. loss: 0.835206
Total training time: 13.97 seconds.
-- Epoch 64
Norm: 12.10, NNZs: 2297, Bias: -12.203347, T: 1706688, Avg. loss: 0.962988
Total training time: 14.19 seconds.
-- Epoch 65
Norm: 12.03, NNZs: 2297, Bias: -12.109871, T: 1733355, Avg. loss: 0.825572
Total training time: 14.41 seconds.
-- Epoch 66
Norm: 11.92, NNZs: 2297, Bias: -12.042757, T: 1760022, Avg. loss: 0.835442
Total training time: 14.63 seconds.
-- Epoch 67
Norm: 11.87, NNZs: 2297, Bias: -11.947862, T: 1786689, Avg. loss: 0.788086
Total training time: 14.84 seconds.
-- Epoch 68
Norm: 11.71, NNZs: 2297, Bias: -11.945774, T: 1813356, Avg. loss: 0.813665
Total training time: 15.05 seconds.
-- Epoch 69
Norm: 11.68, NNZs: 2297, Bias: -11.823920, T: 1840023, Avg. loss: 0.770654
Total training time: 15.27 seconds.
-- Epoch 70
Norm: 11.62, NNZs: 2297, Bias: -11.734041, T: 1866690, Avg. loss: 0.790256
Total training time: 15.49 seconds.
-- Epoch 71
Norm: 11.48, NNZs: 2297, Bias: -11.722157, T: 1893357, Avg. loss: 0.767091
Total training time: 15.70 seconds.
-- Epoch 72
Norm: 11.44, NNZs: 2297, Bias: -11.617091, T: 1920024, Avg. loss: 0.771206
Total training time: 15.92 seconds.
-- Epoch 73
Norm: 11.35, NNZs: 2297, Bias: -11.561887, T: 1946691, Avg. loss: 0.769334
Total training time: 16.14 seconds.
-- Epoch 74
Norm: 11.28, NNZs: 2297, Bias: -11.494219, T: 1973358, Avg. loss: 0.765642
Total training time: 16.36 seconds.
-- Epoch 75
Norm: 11.20, NNZs: 2297, Bias: -11.448873, T: 2000025, Avg. loss: 0.748332
Total training time: 16.58 seconds.
-- Epoch 76
Norm: 11.13, NNZs: 2297, Bias: -11.392190, T: 2026692, Avg. loss: 0.745394
```

```
Norm: 11.15, NNZs: 2297, Bias: -11.392190, T: 2026692, Avg. loss: 0.745394
Total training time: 16.80 seconds.
-- Epoch 77
Norm: 11.11, NNZs: 2297, Bias: -11.274068, T: 2053359, Avg. loss: 0.746619
Total training time: 17.02 seconds.
-- Epoch 78
Norm: 11.01, NNZs: 2297, Bias: -11.256751, T: 2080026, Avg. loss: 0.756471
Total training time: 17.23 seconds.
-- Epoch 79
Norm: 11.03, NNZs: 2297, Bias: -11.165308, T: 2106693, Avg. loss: 0.766146
Total training time: 17.45 seconds.
-- Epoch 80
Norm: 10.94, NNZs: 2297, Bias: -11.124906, T: 2133360, Avg. loss: 0.750649
Total training time: 17.66 seconds.
-- Epoch 81
Norm: 10.86, NNZs: 2297, Bias: -11.087007, T: 2160027, Avg. loss: 0.732728
Total training time: 17.88 seconds.
-- Epoch 82
Norm: 10.82, NNZs: 2297, Bias: -11.018712, T: 2186694, Avg. loss: 0.734655
Total training time: 18.09 seconds.
-- Epoch 83
Norm: 10.73, NNZs: 2297, Bias: -10.987804, T: 2213361, Avg. loss: 0.727085
Total training time: 18.31 seconds.
-- Epoch 84
Norm: 10.70, NNZs: 2297, Bias: -10.902543, T: 2240028, Avg. loss: 0.718197
Total training time: 18.53 seconds.
-- Epoch 85
Norm: 10.61, NNZs: 2297, Bias: -10.894941, T: 2266695, Avg. loss: 0.712390
Total training time: 18.74 seconds.
-- Epoch 86
Norm: 10.59, NNZs: 2297, Bias: -10.808824, T: 2293362, Avg. loss: 0.711597
Total training time: 18.96 seconds.
-- Epoch 87
Norm: 10.53, NNZs: 2297, Bias: -10.779062, T: 2320029, Avg. loss: 0.723243
Total training time: 19.18 seconds.
-- Epoch 88
Norm: 10.48, NNZs: 2297, Bias: -10.730276, T: 2346696, Avg. loss: 0.720160
Total training time: 19.39 seconds.
-- Epoch 89
Norm: 10.44, NNZs: 2297, Bias: -10.669725, T: 2373363, Avg. loss: 0.708540
Total training time: 19.61 seconds.
-- Epoch 90
Norm: 10.39, NNZs: 2297, Bias: -10.630489, T: 2400030, Avg. loss: 0.712688
Total training time: 19.82 seconds.
-- Epoch 91
Norm: 10.33, NNZs: 2297, Bias: -10.598923, T: 2426697, Avg. loss: 0.696730
Total training time: 20.04 seconds.
-- Epoch 92
Norm: 10.31, NNZs: 2297, Bias: -10.508929, T: 2453364, Avg. loss: 0.700086
Total training time: 20.25 seconds.
-- Epoch 93
Norm: 10.23, NNZs: 2297, Bias: -10.492732, T: 2480031, Avg. loss: 0.693830
Total training time: 20.48 seconds.
-- Epoch 94
Norm: 10.17, NNZs: 2297, Bias: -10.465032, T: 2506698, Avg. loss: 0.693407
Total training time: 20.69 seconds.
-- Epoch 95
Norm: 10.16, NNZs: 2297, Bias: -10.390662, T: 2533365, Avg. loss: 0.686398
Total training time: 20.91 seconds.
-- Epoch 96
Norm: 10.11, NNZs: 2297, Bias: -10.341535, T: 2560032, Avg. loss: 0.688456
Total training time: 21.12 seconds.
-- Epoch 97
Norm: 10.16, NNZs: 2297, Bias: -10.269760, T: 2586699, Avg. loss: 0.779993
Total training time: 21.34 seconds.
-- Epoch 98
Norm: 10.05, NNZs: 2297, Bias: -10.283903, T: 2613366, Avg. loss: 0.699919
Total training time: 21.55 seconds.
-- Epoch 99
Norm: 10.02, NNZs: 2297, Bias: -10.231378, T: 2640033, Avg. loss: 0.688938
Total training time: 21.76 seconds.
-- Epoch 100
Norm: 9.99, NNZs: 2297, Bias: -10.177293, T: 2666700, Avg. loss: 0.673602
Total training time: 21.98 seconds.
-- Epoch 101
Norm: 9.94, NNZs: 2297, Bias: -10.144143, T: 2693367, Avg. loss: 0.687533
Total training time: 22.20 seconds.
-- Epoch 102
```

```
-- Epoch 102
Norm: 9.87, NNZs: 2297, Bias: -10.128879, T: 2720034, Avg. loss: 0.676150
Total training time: 22.41 seconds.
-- Epoch 103
Norm: 9.84, NNZs: 2297, Bias: -10.080916, T: 2746701, Avg. loss: 0.670401
Total training time: 22.63 seconds.
-- Epoch 104
Norm: 9.81, NNZs: 2297, Bias: -10.027900, T: 2773368, Avg. loss: 0.670150
Total training time: 22.85 seconds.
-- Epoch 105
Norm: 9.77, NNZs: 2297, Bias: -9.997218, T: 2800035, Avg. loss: 0.661040
Total training time: 23.06 seconds.
-- Epoch 106
Norm: 9.75, NNZs: 2297, Bias: -9.936403, T: 2826702, Avg. loss: 0.678851
Total training time: 23.28 seconds.
-- Epoch 107
Norm: 9.71, NNZs: 2297, Bias: -9.910634, T: 2853369, Avg. loss: 0.666695
Total training time: 23.50 seconds.
-- Epoch 108
Norm: 9.63, NNZs: 2297, Bias: -9.901498, T: 2880036, Avg. loss: 0.667757
Total training time: 23.71 seconds.
-- Epoch 109
Norm: 9.62, NNZs: 2297, Bias: -9.848706, T: 2906703, Avg. loss: 0.656667
Total training time: 23.93 seconds.
-- Epoch 110
Norm: 9.59, NNZs: 2297, Bias: -9.804263, T: 2933370, Avg. loss: 0.654255
Total training time: 24.15 seconds.
-- Epoch 111
Norm: 9.56, NNZs: 2297, Bias: -9.771703, T: 2960037, Avg. loss: 0.657540
Total training time: 24.36 seconds.
-- Epoch 112
Norm: 9.51, NNZs: 2297, Bias: -9.747239, T: 2986704, Avg. loss: 0.654689
Total training time: 24.58 seconds.
-- Epoch 113
Norm: 9.46, NNZs: 2297, Bias: -9.735893, T: 3013371, Avg. loss: 0.661652
Total training time: 24.80 seconds.
-- Epoch 114
Norm: 9.42, NNZs: 2297, Bias: -9.703166, T: 3040038, Avg. loss: 0.650946
Total training time: 25.02 seconds.
-- Epoch 115
Norm: 9.42, NNZs: 2297, Bias: -9.643617, T: 3066705, Avg. loss: 0.639036
Total training time: 25.23 seconds.
-- Epoch 116
Norm: 9.39, NNZs: 2297, Bias: -9.609692, T: 3093372, Avg. loss: 0.658552
Total training time: 25.45 seconds.
-- Epoch 117
Norm: 9.38, NNZs: 2297, Bias: -9.558836, T: 3120039, Avg. loss: 0.644947
Total training time: 25.67 seconds.
-- Epoch 118
Norm: 9.32, NNZs: 2297, Bias: -9.549034, T: 3146706, Avg. loss: 0.658027
Total training time: 25.88 seconds.
-- Epoch 119
Norm: 9.29, NNZs: 2297, Bias: -9.528299, T: 3173373, Avg. loss: 0.641932
Total training time: 26.10 seconds.
-- Epoch 120
Norm: 9.25, NNZs: 2297, Bias: -9.498683, T: 3200040, Avg. loss: 0.653032
Total training time: 26.31 seconds.
Convergence after 120 epochs took 26.31 seconds
-- Epoch 1
Norm: 200.38, NNZs: 2298, Bias: -77.658473, T: 26667, Avg. loss: 196.316794
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 87.85, NNZs: 2298, Bias: -63.585883, T: 53334, Avg. loss: 41.443432
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 65.63, NNZs: 2298, Bias: -55.090533, T: 80001, Avg. loss: 17.504425
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 56.09, NNZs: 2298, Bias: -50.338782, T: 106668, Avg. loss: 12.981759
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 49.27, NNZs: 2298, Bias: -45.694872, T: 133335, Avg. loss: 9.473457
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 45.97, NNZs: 2298, Bias: -42.760941, T: 160002, Avg. loss: 9.475700
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 41.90, NNZs: 2298, Bias: -39.913432, T: 186669, Avg. loss: 6.563982
```

```
Norm: 41.90, NNZs: 2298, Bias: -39.913432, T: 186669, Avg. loss: 6.563982
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 39.61, NNZs: 2298, Bias: -37.195296, T: 213336, Avg. loss: 5.655255
Total training time: 1.73 seconds.
-- Epoch 9
Norm: 36.76, NNZs: 2298, Bias: -35.798443, T: 240003, Avg. loss: 4.872967
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 35.11, NNZs: 2298, Bias: -34.133639, T: 266670, Avg. loss: 4.644644
Total training time: 2.17 seconds.
-- Epoch 11
Norm: 34.09, NNZs: 2298, Bias: -32.957965, T: 293337, Avg. loss: 4.951942
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 31.92, NNZs: 2298, Bias: -31.753401, T: 320004, Avg. loss: 3.511837
Total training time: 2.60 seconds.
-- Epoch 13
Norm: 31.04, NNZs: 2298, Bias: -30.547774, T: 346671, Avg. loss: 3.282955
Total training time: 2.82 seconds.
-- Epoch 14
Norm: 29.69, NNZs: 2298, Bias: -29.451420, T: 373338, Avg. loss: 2.970317
Total training time: 3.03 seconds.
-- Epoch 15
Norm: 28.66, NNZs: 2298, Bias: -28.716677, T: 400005, Avg. loss: 2.897136
Total training time: 3.25 seconds.
-- Epoch 16
Norm: 27.73, NNZs: 2298, Bias: -27.907477, T: 426672, Avg. loss: 2.535354
Total training time: 3.47 seconds.
-- Epoch 17
Norm: 26.52, NNZs: 2298, Bias: -27.389486, T: 453339, Avg. loss: 2.449556
Total training time: 3.68 seconds.
-- Epoch 18
Norm: 25.78, NNZs: 2298, Bias: -26.684008, T: 480006, Avg. loss: 2.310076
Total training time: 3.90 seconds.
-- Epoch 19
Norm: 25.14, NNZs: 2298, Bias: -26.048730, T: 506673, Avg. loss: 2.196970
Total training time: 4.12 seconds.
-- Epoch 20
Norm: 24.43, NNZs: 2298, Bias: -25.477038, T: 533340, Avg. loss: 2.068030
Total training time: 4.33 seconds.
-- Epoch 21
Norm: 24.18, NNZs: 2298, Bias: -24.733471, T: 560007, Avg. loss: 1.945530
Total training time: 4.54 seconds.
-- Epoch 22
Norm: 23.45, NNZs: 2298, Bias: -24.331890, T: 586674, Avg. loss: 1.850322
Total training time: 4.76 seconds.
-- Epoch 23
Norm: 22.98, NNZs: 2298, Bias: -23.884880, T: 613341, Avg. loss: 1.842879
Total training time: 4.98 seconds.
-- Epoch 24
Norm: 22.54, NNZs: 2298, Bias: -23.394209, T: 640008, Avg. loss: 1.663678
Total training time: 5.20 seconds.
-- Epoch 25
Norm: 22.12, NNZs: 2298, Bias: -23.012297, T: 666675, Avg. loss: 1.663965
Total training time: 5.41 seconds.
-- Epoch 26
Norm: 22.17, NNZs: 2298, Bias: -22.482058, T: 693342, Avg. loss: 1.917207
Total training time: 5.63 seconds.
-- Epoch 27
Norm: 21.44, NNZs: 2298, Bias: -22.283558, T: 720009, Avg. loss: 1.620144
Total training time: 5.84 seconds.
-- Epoch 28
Norm: 20.93, NNZs: 2298, Bias: -21.970931, T: 746676, Avg. loss: 1.457130
Total training time: 6.06 seconds.
-- Epoch 29
Norm: 20.68, NNZs: 2298, Bias: -21.515740, T: 773343, Avg. loss: 1.359922
Total training time: 6.27 seconds.
-- Epoch 30
Norm: 20.22, NNZs: 2298, Bias: -21.286241, T: 800010, Avg. loss: 1.377281
Total training time: 6.49 seconds.
-- Epoch 31
Norm: 20.00, NNZs: 2298, Bias: -20.946347, T: 826677, Avg. loss: 1.348384
Total training time: 6.70 seconds.
-- Epoch 32
Norm: 19.80, NNZs: 2298, Bias: -20.555153, T: 853344, Avg. loss: 1.350403
Total training time: 6.92 seconds.
-- Epoch 33
```

```
-- Epoch 33
Norm: 19.51, NNZs: 2298, Bias: -20.257218, T: 880011, Avg. loss: 1.221747
Total training time: 7.15 seconds.
-- Epoch 34
Norm: 19.12, NNZs: 2298, Bias: -20.108815, T: 906678, Avg. loss: 1.255283
Total training time: 7.37 seconds.
-- Epoch 35
Norm: 18.92, NNZs: 2298, Bias: -19.780982, T: 933345, Avg. loss: 1.200727
Total training time: 7.58 seconds.
-- Epoch 36
Norm: 18.65, NNZs: 2298, Bias: -19.550931, T: 960012, Avg. loss: 1.139345
Total training time: 7.80 seconds.
-- Epoch 37
Norm: 18.51, NNZs: 2298, Bias: -19.263217, T: 986679, Avg. loss: 1.150775
Total training time: 8.02 seconds.
-- Epoch 38
Norm: 18.14, NNZs: 2298, Bias: -19.148081, T: 1013346, Avg. loss: 1.155831
Total training time: 8.23 seconds.
-- Epoch 39
Norm: 17.87, NNZs: 2298, Bias: -18.955784, T: 1040013, Avg. loss: 1.089551
Total training time: 8.45 seconds.
-- Epoch 40
Norm: 17.83, NNZs: 2298, Bias: -18.770245, T: 1066680, Avg. loss: 1.301633
Total training time: 8.66 seconds.
-- Epoch 41
Norm: 17.56, NNZs: 2298, Bias: -18.601816, T: 1093347, Avg. loss: 1.106091
Total training time: 8.89 seconds.
-- Epoch 42
Norm: 17.32, NNZs: 2298, Bias: -18.432296, T: 1120014, Avg. loss: 1.022080
Total training time: 9.10 seconds.
-- Epoch 43
Norm: 17.17, NNZs: 2298, Bias: -18.214833, T: 1146681, Avg. loss: 1.043308
Total training time: 9.32 seconds.
-- Epoch 44
Norm: 17.00, NNZs: 2298, Bias: -18.022205, T: 1173348, Avg. loss: 1.013419
Total training time: 9.53 seconds.
-- Epoch 45
Norm: 16.82, NNZs: 2298, Bias: -17.856001, T: 1200015, Avg. loss: 0.971126
Total training time: 9.75 seconds.
-- Epoch 46
Norm: 16.71, NNZs: 2298, Bias: -17.652962, T: 1226682, Avg. loss: 1.001384
Total training time: 9.97 seconds.
-- Epoch 47
Norm: 16.54, NNZs: 2298, Bias: -17.485639, T: 1253349, Avg. loss: 0.950320
Total training time: 10.19 seconds.
-- Epoch 48
Norm: 16.33, NNZs: 2298, Bias: -17.356206, T: 1280016, Avg. loss: 0.953212
Total training time: 10.40 seconds.
-- Epoch 49
Norm: 16.16, NNZs: 2298, Bias: -17.214350, T: 1306683, Avg. loss: 0.922588
Total training time: 10.62 seconds.
-- Epoch 50
Norm: 16.08, NNZs: 2298, Bias: -17.029744, T: 1333350, Avg. loss: 0.919440
Total training time: 10.83 seconds.
-- Epoch 51
Norm: 15.94, NNZs: 2298, Bias: -16.888352, T: 1360017, Avg. loss: 0.912062
Total training time: 11.06 seconds.
-- Epoch 52
Norm: 15.73, NNZs: 2298, Bias: -16.793086, T: 1386684, Avg. loss: 0.867755
Total training time: 11.27 seconds.
-- Epoch 53
Norm: 15.65, NNZs: 2298, Bias: -16.635754, T: 1413351, Avg. loss: 0.895073
Total training time: 11.49 seconds.
-- Epoch 54
Norm: 15.55, NNZs: 2298, Bias: -16.476678, T: 1440018, Avg. loss: 0.879344
Total training time: 11.70 seconds.
-- Epoch 55
Norm: 15.35, NNZs: 2298, Bias: -16.395891, T: 1466685, Avg. loss: 0.867537
Total training time: 11.92 seconds.
-- Epoch 56
Norm: 15.23, NNZs: 2298, Bias: -16.281989, T: 1493352, Avg. loss: 0.859309
Total training time: 12.13 seconds.
-- Epoch 57
Norm: 15.15, NNZs: 2298, Bias: -16.123372, T: 1520019, Avg. loss: 0.838616
Total training time: 12.35 seconds.
-- Epoch 58
Norm: 14.94, NNZs: 2298, Bias: -16.072392, T: 1546686, Avg. loss: 0.822276
Total training time: 12.56 seconds.
```

```
Total training time: 12.56 seconds.
-- Epoch 59
Norm: 14.95, NNZs: 2298, Bias: -15.865175, T: 1573353, Avg. loss: 0.841978
Total training time: 12.78 seconds.
-- Epoch 60
Norm: 14.78, NNZs: 2298, Bias: -15.800881, T: 1600020, Avg. loss: 0.823486
Total training time: 13.00 seconds.
-- Epoch 61
Norm: 14.70, NNZs: 2298, Bias: -15.670662, T: 1626687, Avg. loss: 0.804534
Total training time: 13.22 seconds.
-- Epoch 62
Norm: 14.57, NNZs: 2298, Bias: -15.598757, T: 1653354, Avg. loss: 0.830316
Total training time: 13.43 seconds.
-- Epoch 63
Norm: 14.47, NNZs: 2298, Bias: -15.487935, T: 1680021, Avg. loss: 0.797620
Total training time: 13.65 seconds.
-- Epoch 64
Norm: 14.38, NNZs: 2298, Bias: -15.368867, T: 1706688, Avg. loss: 0.792855
Total training time: 13.86 seconds.
-- Epoch 65
Norm: 14.29, NNZs: 2298, Bias: -15.260571, T: 1733355, Avg. loss: 0.794642
Total training time: 14.08 seconds.
-- Epoch 66
Norm: 14.26, NNZs: 2298, Bias: -15.107728, T: 1760022, Avg. loss: 0.787965
Total training time: 14.30 seconds.
-- Epoch 67
Norm: 14.07, NNZs: 2298, Bias: -15.090778, T: 1786689, Avg. loss: 0.789684
Total training time: 14.51 seconds.
-- Epoch 68
Norm: 13.98, NNZs: 2298, Bias: -14.996558, T: 1813356, Avg. loss: 0.769157
Total training time: 14.73 seconds.
-- Epoch 69
Norm: 13.91, NNZs: 2298, Bias: -14.886337, T: 1840023, Avg. loss: 0.766471
Total training time: 14.95 seconds.
-- Epoch 70
Norm: 13.85, NNZs: 2298, Bias: -14.786235, T: 1866690, Avg. loss: 0.771260
Total training time: 15.17 seconds.
-- Epoch 71
Norm: 13.71, NNZs: 2298, Bias: -14.735627, T: 1893357, Avg. loss: 0.765311
Total training time: 15.38 seconds.
-- Epoch 72
Norm: 13.64, NNZs: 2298, Bias: -14.635892, T: 1920024, Avg. loss: 0.745897
Total training time: 15.60 seconds.
-- Epoch 73
Norm: 13.57, NNZs: 2298, Bias: -14.542786, T: 1946691, Avg. loss: 0.759144
Total training time: 15.82 seconds.
-- Epoch 74
Norm: 13.48, NNZs: 2298, Bias: -14.473945, T: 1973358, Avg. loss: 0.743710
Total training time: 16.04 seconds.
-- Epoch 75
Norm: 13.41, NNZs: 2298, Bias: -14.377947, T: 2000025, Avg. loss: 0.744894
Total training time: 16.26 seconds.
-- Epoch 76
Norm: 13.33, NNZs: 2298, Bias: -14.295736, T: 2026692, Avg. loss: 0.735425
Total training time: 16.48 seconds.
-- Epoch 77
Norm: 13.28, NNZs: 2298, Bias: -14.194893, T: 2053359, Avg. loss: 0.727542
Total training time: 16.69 seconds.
-- Epoch 78
Norm: 13.21, NNZs: 2298, Bias: -14.168723, T: 2080026, Avg. loss: 0.861123
Total training time: 16.90 seconds.
-- Epoch 79
Norm: 13.15, NNZs: 2298, Bias: -14.085828, T: 2106693, Avg. loss: 0.731778
Total training time: 17.13 seconds.
-- Epoch 80
Norm: 13.10, NNZs: 2298, Bias: -13.989414, T: 2133360, Avg. loss: 0.709333
Total training time: 17.35 seconds.
-- Epoch 81
Norm: 12.97, NNZs: 2298, Bias: -13.975559, T: 2160027, Avg. loss: 0.722258
Total training time: 17.56 seconds.
-- Epoch 82
Norm: 12.99, NNZs: 2298, Bias: -13.830183, T: 2186694, Avg. loss: 0.703479
Total training time: 17.77 seconds.
-- Epoch 83
Norm: 12.89, NNZs: 2298, Bias: -13.796399, T: 2213361, Avg. loss: 0.719235
Total training time: 18.00 seconds.
-- Epoch 84
```

```
Norm: 12.85, NNZs: 2298, Bias: -13.701324, T: 2240028, Avg. loss: 0.702202
Total training time: 18.22 seconds.
-- Epoch 85
Norm: 12.72, NNZs: 2298, Bias: -13.695358, T: 2266695, Avg. loss: 0.710122
Total training time: 18.44 seconds.
-- Epoch 86
Norm: 12.65, NNZs: 2298, Bias: -13.631836, T: 2293362, Avg. loss: 0.706301
Total training time: 18.65 seconds.
-- Epoch 87
Norm: 12.63, NNZs: 2298, Bias: -13.526653, T: 2320029, Avg. loss: 0.683446
Total training time: 18.87 seconds.
-- Epoch 88
Norm: 12.59, NNZs: 2298, Bias: -13.447239, T: 2346696, Avg. loss: 0.698255
Total training time: 19.09 seconds.
-- Epoch 89
Norm: 12.50, NNZs: 2298, Bias: -13.410966, T: 2373363, Avg. loss: 0.702028
Total training time: 19.31 seconds.
-- Epoch 90
Norm: 12.42, NNZs: 2298, Bias: -13.360002, T: 2400030, Avg. loss: 0.687975
Total training time: 19.53 seconds.
-- Epoch 91
Norm: 12.36, NNZs: 2298, Bias: -13.308380, T: 2426697, Avg. loss: 0.686571
Total training time: 19.74 seconds.
-- Epoch 92
Norm: 12.34, NNZs: 2298, Bias: -13.209173, T: 2453364, Avg. loss: 0.669258
Total training time: 19.96 seconds.
-- Epoch 93
Norm: 12.29, NNZs: 2298, Bias: -13.175705, T: 2480031, Avg. loss: 0.716657
Total training time: 20.18 seconds.
-- Epoch 94
Norm: 12.26, NNZs: 2298, Bias: -13.093099, T: 2506698, Avg. loss: 0.681167
Total training time: 20.40 seconds.
-- Epoch 95
Norm: 12.19, NNZs: 2298, Bias: -13.053846, T: 2533365, Avg. loss: 0.680411
Total training time: 20.61 seconds.
-- Epoch 96
Norm: 12.15, NNZs: 2298, Bias: -12.986840, T: 2560032, Avg. loss: 0.682798
Total training time: 20.83 seconds.
-- Epoch 97
Norm: 12.08, NNZs: 2298, Bias: -12.946820, T: 2586699, Avg. loss: 0.680822
Total training time: 21.06 seconds.
Convergence after 97 epochs took 21.06 seconds
-- Epoch 1
Norm: 15.06, NNZs: 2298, Bias: -7.654296, T: 26666, Avg. loss: 24.178301
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 9.36, NNZs: 2298, Bias: -6.259588, T: 53332, Avg. loss: 3.249852
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 6.77, NNZs: 2298, Bias: -5.535284, T: 79998, Avg. loss: 1.864617
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 5.67, NNZs: 2298, Bias: -4.971578, T: 106664, Avg. loss: 1.317552
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 4.94, NNZs: 2298, Bias: -4.615239, T: 133330, Avg. loss: 1.040464
Total training time: 1.08 seconds.
-- Epoch 6
Norm: 4.52, NNZs: 2298, Bias: -4.329992, T: 159996, Avg. loss: 0.930182
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 4.12, NNZs: 2298, Bias: -4.134154, T: 186662, Avg. loss: 0.845872
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 3.86, NNZs: 2298, Bias: -3.941613, T: 213328, Avg. loss: 0.780769
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 3.81, NNZs: 2298, Bias: -3.737112, T: 239994, Avg. loss: 0.730541
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 3.56, NNZs: 2298, Bias: -3.646715, T: 266660, Avg. loss: 0.718415
Total training time: 2.17 seconds.
-- Epoch 11
Norm: 3.42, NNZs: 2298, Bias: -3.547643, T: 293326, Avg. loss: 0.673403
Total training time: 2.38 seconds.
-- Epoch 12
Norm: 3.30, NNZs: 2298, Bias: -3.453841, T: 319992, Avg. loss: 0.677807
```

```
Total training time: 2.61 seconds.
-- Epoch 13
Norm: 3.22, NNZs: 2298, Bias: -3.353975, T: 346658, Avg. loss: 0.644708
Total training time: 2.83 seconds.
-- Epoch 14
Norm: 3.15, NNZs: 2298, Bias: -3.273491, T: 373324, Avg. loss: 0.634988
Total training time: 3.04 seconds.
-- Epoch 15
Norm: 3.06, NNZs: 2298, Bias: -3.204685, T: 399990, Avg. loss: 0.634137
Total training time: 3.26 seconds.
-- Epoch 16
Norm: 2.99, NNZs: 2298, Bias: -3.147291, T: 426656, Avg. loss: 0.618684
Total training time: 3.48 seconds.
-- Epoch 17
Norm: 2.92, NNZs: 2298, Bias: -3.096467, T: 453322, Avg. loss: 0.605630
Total training time: 3.69 seconds.
-- Epoch 18
Norm: 2.88, NNZs: 2298, Bias: -3.041659, T: 479988, Avg. loss: 0.608772
Total training time: 3.91 seconds.
-- Epoch 19
Norm: 2.84, NNZs: 2298, Bias: -2.988749, T: 506654, Avg. loss: 0.601927
Total training time: 4.13 seconds.
-- Epoch 20
Norm: 2.80, NNZs: 2298, Bias: -2.939851, T: 533320, Avg. loss: 0.599235
Total training time: 4.35 seconds.
-- Epoch 21
Norm: 2.75, NNZs: 2298, Bias: -2.905140, T: 559986, Avg. loss: 0.596676
Total training time: 4.57 seconds.
-- Epoch 22
Norm: 2.72, NNZs: 2298, Bias: -2.863166, T: 586652, Avg. loss: 0.588108
Total training time: 4.79 seconds.
-- Epoch 23
Norm: 2.71, NNZs: 2298, Bias: -2.819054, T: 613318, Avg. loss: 0.587871
Total training time: 5.01 seconds.
-- Epoch 24
Norm: 2.68, NNZs: 2298, Bias: -2.785351, T: 639984, Avg. loss: 0.583209
Total training time: 5.22 seconds.
-- Epoch 25
Norm: 2.63, NNZs: 2298, Bias: -2.768468, T: 666650, Avg. loss: 0.586044
Total training time: 5.44 seconds.
-- Epoch 26
Norm: 2.61, NNZs: 2298, Bias: -2.728404, T: 693316, Avg. loss: 0.574879
Total training time: 5.66 seconds.
-- Epoch 27
Norm: 2.59, NNZs: 2298, Bias: -2.696118, T: 719982, Avg. loss: 0.575595
Total training time: 5.88 seconds.
-- Epoch 28
Norm: 2.56, NNZs: 2298, Bias: -2.676897, T: 746648, Avg. loss: 0.578284
Total training time: 6.10 seconds.
-- Epoch 29
Norm: 2.55, NNZs: 2298, Bias: -2.644704, T: 773314, Avg. loss: 0.572052
Total training time: 6.31 seconds.
-- Epoch 30
Norm: 2.52, NNZs: 2298, Bias: -2.621617, T: 799980, Avg. loss: 0.572329
Total training time: 6.53 seconds.
-- Epoch 31
Norm: 2.50, NNZs: 2298, Bias: -2.599354, T: 826646, Avg. loss: 0.568914
Total training time: 6.75 seconds.
-- Epoch 32
Norm: 2.48, NNZs: 2298, Bias: -2.574024, T: 853312, Avg. loss: 0.569868
Total training time: 6.97 seconds.
-- Epoch 33
Norm: 2.47, NNZs: 2298, Bias: -2.549672, T: 879978, Avg. loss: 0.565282
Total training time: 7.18 seconds.
-- Epoch 34
Norm: 2.47, NNZs: 2298, Bias: -2.524930, T: 906644, Avg. loss: 0.564996
Total training time: 7.40 seconds.
-- Epoch 35
Norm: 2.45, NNZs: 2298, Bias: -2.507393, T: 933310, Avg. loss: 0.568229
Total training time: 7.61 seconds.
-- Epoch 36
Norm: 2.43, NNZs: 2298, Bias: -2.496290, T: 959976, Avg. loss: 0.568083
Total training time: 7.83 seconds.
-- Epoch 37
Norm: 2.41, NNZs: 2298, Bias: -2.475777, T: 986642, Avg. loss: 0.560729
Total training time: 8.05 seconds.
-- Epoch 38
```

```
Norm: 2.40, NNZs: 2298, Bias: -2.454790, T: 1013308, Avg. loss: 0.560218
Total training time: 8.26 seconds.
-- Epoch 39
Norm: 2.39, NNZs: 2298, Bias: -2.439805, T: 1039974, Avg. loss: 0.563153
Total training time: 8.49 seconds.
-- Epoch 40
Norm: 2.38, NNZs: 2298, Bias: -2.423905, T: 1066640, Avg. loss: 0.560482
Total training time: 8.71 seconds.
-- Epoch 41
Norm: 2.37, NNZs: 2298, Bias: -2.404236, T: 1093306, Avg. loss: 0.557754
Total training time: 8.93 seconds.
-- Epoch 42
Norm: 2.36, NNZs: 2298, Bias: -2.390920, T: 1119972, Avg. loss: 0.561540
Total training time: 9.14 seconds.
-- Epoch 43
Norm: 2.35, NNZs: 2298, Bias: -2.379267, T: 1146638, Avg. loss: 0.560263
Total training time: 9.36 seconds.
-- Epoch 44
Norm: 2.33, NNZs: 2298, Bias: -2.364712, T: 1173304, Avg. loss: 0.557797
Total training time: 9.58 seconds.
-- Epoch 45
Norm: 2.32, NNZs: 2298, Bias: -2.352075, T: 1199970, Avg. loss: 0.560038
Total training time: 9.80 seconds.
-- Epoch 46
Norm: 2.32, NNZs: 2298, Bias: -2.336059, T: 1226636, Avg. loss: 0.556078
Total training time: 10.02 seconds.
-- Epoch 47
Norm: 2.31, NNZs: 2298, Bias: -2.320931, T: 1253302, Avg. loss: 0.554300
Total training time: 10.23 seconds.
-- Epoch 48
Norm: 2.30, NNZs: 2298, Bias: -2.310400, T: 1279968, Avg. loss: 0.555478
Total training time: 10.46 seconds.
-- Epoch 49
Norm: 2.29, NNZs: 2298, Bias: -2.298745, T: 1306634, Avg. loss: 0.555475
Total training time: 10.68 seconds.
-- Epoch 50
Norm: 2.29, NNZs: 2298, Bias: -2.284201, T: 1333300, Avg. loss: 0.553669
Total training time: 10.90 seconds.
-- Epoch 51
Norm: 2.28, NNZs: 2298, Bias: -2.270457, T: 1359966, Avg. loss: 0.553530
Total training time: 11.11 seconds.
-- Epoch 52
Norm: 2.27, NNZs: 2298, Bias: -2.262558, T: 1386632, Avg. loss: 0.557291
Total training time: 11.33 seconds.
Convergence after 52 epochs took 11.33 seconds
-- Epoch 1
Norm: 16.18, NNZs: 2297, Bias: -7.984441, T: 26667, Avg. loss: 24.243338
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 11.14, NNZs: 2297, Bias: -6.308754, T: 53334, Avg. loss: 3.289223
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 7.00, NNZs: 2297, Bias: -5.774339, T: 80001, Avg. loss: 2.334682
Total training time: 0.64 seconds.
-- Epoch 4
Norm: 5.89, NNZs: 2297, Bias: -5.165896, T: 106668, Avg. loss: 1.343813
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 5.43, NNZs: 2297, Bias: -4.752427, T: 133335, Avg. loss: 1.349150
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 4.74, NNZs: 2297, Bias: -4.489194, T: 160002, Avg. loss: 0.971870
Total training time: 1.31 seconds.
-- Epoch 7
Norm: 4.43, NNZs: 2297, Bias: -4.226057, T: 186669, Avg. loss: 0.844186
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 4.09, NNZs: 2297, Bias: -4.044408, T: 213336, Avg. loss: 0.791025
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 3.85, NNZs: 2297, Bias: -3.890624, T: 240003, Avg. loss: 0.745592
Total training time: 1.96 seconds.
-- Epoch 10
Norm: 3.66, NNZs: 2297, Bias: -3.751343, T: 266670, Avg. loss: 0.710028
Total training time: 2.18 seconds.
-- Epoch 11
Norm: 3.49, NNZs: 2297, Bias: -3.648353, T: 293337, Avg. loss: 0.677722
```

```
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 3.36, NNZs: 2297, Bias: -3.536128, T: 320004, Avg. loss: 0.669416
Total training time: 2.61 seconds.
-- Epoch 13
Norm: 3.47, NNZs: 2297, Bias: -3.402021, T: 346671, Avg. loss: 0.665500
Total training time: 2.82 seconds.
-- Epoch 14
Norm: 3.26, NNZs: 2297, Bias: -3.356262, T: 373338, Avg. loss: 0.690379
Total training time: 3.04 seconds.
-- Epoch 15
Norm: 3.18, NNZs: 2297, Bias: -3.285250, T: 400005, Avg. loss: 0.635154
Total training time: 3.27 seconds.
-- Epoch 16
Norm: 3.09, NNZs: 2297, Bias: -3.225926, T: 426672, Avg. loss: 0.624746
Total training time: 3.48 seconds.
-- Epoch 17
Norm: 3.00, NNZs: 2297, Bias: -3.173013, T: 453339, Avg. loss: 0.618474
Total training time: 3.70 seconds.
-- Epoch 18
Norm: 2.95, NNZs: 2297, Bias: -3.114337, T: 480006, Avg. loss: 0.606131
Total training time: 3.92 seconds.
-- Epoch 19
Norm: 2.91, NNZs: 2297, Bias: -3.053323, T: 506673, Avg. loss: 0.600655
Total training time: 4.14 seconds.
-- Epoch 20
Norm: 2.83, NNZs: 2297, Bias: -3.030828, T: 533340, Avg. loss: 0.603914
Total training time: 4.35 seconds.
-- Epoch 21
Norm: 2.80, NNZs: 2297, Bias: -2.978613, T: 560007, Avg. loss: 0.584992
Total training time: 4.57 seconds.
-- Epoch 22
Norm: 2.78, NNZs: 2297, Bias: -2.924556, T: 586674, Avg. loss: 0.587154
Total training time: 4.78 seconds.
-- Epoch 23
Norm: 2.72, NNZs: 2297, Bias: -2.908174, T: 613341, Avg. loss: 0.596822
Total training time: 5.00 seconds.
-- Epoch 24
Norm: 2.70, NNZs: 2297, Bias: -2.863506, T: 640008, Avg. loss: 0.581274
Total training time: 5.22 seconds.
-- Epoch 25
Norm: 2.66, NNZs: 2297, Bias: -2.831204, T: 666675, Avg. loss: 0.582252
Total training time: 5.43 seconds.
-- Epoch 26
Norm: 2.64, NNZs: 2297, Bias: -2.791329, T: 693342, Avg. loss: 0.584638
Total training time: 5.64 seconds.
-- Epoch 27
Norm: 2.62, NNZs: 2297, Bias: -2.763554, T: 720009, Avg. loss: 0.578912
Total training time: 5.86 seconds.
-- Epoch 28
Norm: 2.58, NNZs: 2297, Bias: -2.742289, T: 746676, Avg. loss: 0.580378
Total training time: 6.09 seconds.
-- Epoch 29
Norm: 2.57, NNZs: 2297, Bias: -2.708752, T: 773343, Avg. loss: 0.571428
Total training time: 6.30 seconds.
-- Epoch 30
Norm: 2.55, NNZs: 2297, Bias: -2.683140, T: 800010, Avg. loss: 0.576511
Total training time: 6.52 seconds.
-- Epoch 31
Norm: 2.53, NNZs: 2297, Bias: -2.657274, T: 826677, Avg. loss: 0.571398
Total training time: 6.74 seconds.
-- Epoch 32
Norm: 2.51, NNZs: 2297, Bias: -2.629464, T: 853344, Avg. loss: 0.568261
Total training time: 6.96 seconds.
-- Epoch 33
Norm: 2.50, NNZs: 2297, Bias: -2.608984, T: 880011, Avg. loss: 0.568389
Total training time: 7.17 seconds.
-- Epoch 34
Norm: 2.48, NNZs: 2297, Bias: -2.585960, T: 906678, Avg. loss: 0.570758
Total training time: 7.39 seconds.
-- Epoch 35
Norm: 2.46, NNZs: 2297, Bias: -2.572224, T: 933345, Avg. loss: 0.571895
Total training time: 7.61 seconds.
-- Epoch 36
Norm: 2.44, NNZs: 2297, Bias: -2.551796, T: 960012, Avg. loss: 0.564350
Total training time: 7.83 seconds.
-- Epoch 37
```

```
Norm: 2.43, NNZs: 2297, Bias: -2.531356, T: 986679, Avg. loss: 0.563354
Total training time: 8.05 seconds.
-- Epoch 38
Norm: 2.42, NNZs: 2297, Bias: -2.511332, T: 1013346, Avg. loss: 0.565000
Total training time: 8.27 seconds.
-- Epoch 39
Norm: 2.41, NNZs: 2297, Bias: -2.492448, T: 1040013, Avg. loss: 0.564043
Total training time: 8.48 seconds.
-- Epoch 40
Norm: 2.41, NNZs: 2297, Bias: -2.471797, T: 1066680, Avg. loss: 0.571761
Total training time: 8.70 seconds.
-- Epoch 41
Norm: 2.39, NNZs: 2297, Bias: -2.461841, T: 1093347, Avg. loss: 0.567849
Total training time: 8.92 seconds.
Convergence after 41 epochs took 8.92 seconds
-- Epoch 1
Norm: 20.68, NNZs: 2298, Bias: -12.535884, T: 26667, Avg. loss: 22.138863
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 9.93, NNZs: 2298, Bias: -9.928950, T: 53334, Avg. loss: 3.971006
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 7.67, NNZs: 2298, Bias: -8.610846, T: 80001, Avg. loss: 1.779018
Total training time: 0.65 seconds.
-- Epoch 4
Norm: 7.31, NNZs: 2298, Bias: -7.729630, T: 106668, Avg. loss: 1.525916
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 6.20, NNZs: 2298, Bias: -7.142724, T: 133335, Avg. loss: 1.151114
Total training time: 1.08 seconds.
-- Epoch 6
Norm: 5.60, NNZs: 2298, Bias: -6.714303, T: 160002, Avg. loss: 0.922894
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 5.28, NNZs: 2298, Bias: -6.290202, T: 186669, Avg. loss: 0.819209
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 5.23, NNZs: 2298, Bias: -5.915983, T: 213336, Avg. loss: 0.882826
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 4.92, NNZs: 2298, Bias: -5.688440, T: 240003, Avg. loss: 0.754385
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 4.66, NNZs: 2298, Bias: -5.492517, T: 266670, Avg. loss: 0.721833
Total training time: 2.17 seconds.
-- Epoch 11
Norm: 4.46, NNZs: 2298, Bias: -5.317284, T: 293337, Avg. loss: 0.670187
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 4.30, NNZs: 2298, Bias: -5.155496, T: 320004, Avg. loss: 0.653903
Total training time: 2.60 seconds.
-- Epoch 13
Norm: 4.17, NNZs: 2298, Bias: -5.013470, T: 346671, Avg. loss: 0.642878
Total training time: 2.83 seconds.
-- Epoch 14
Norm: 4.07, NNZs: 2298, Bias: -4.864921, T: 373338, Avg. loss: 0.621401
Total training time: 3.04 seconds.
-- Epoch 15
Norm: 3.99, NNZs: 2298, Bias: -4.744446, T: 400005, Avg. loss: 0.624864
Total training time: 3.26 seconds.
-- Epoch 16
Norm: 3.87, NNZs: 2298, Bias: -4.657671, T: 426672, Avg. loss: 0.611765
Total training time: 3.48 seconds.
-- Epoch 17
Norm: 3.82, NNZs: 2298, Bias: -4.565028, T: 453339, Avg. loss: 0.650597
Total training time: 3.71 seconds.
-- Epoch 18
Norm: 3.73, NNZs: 2298, Bias: -4.480850, T: 480006, Avg. loss: 0.606747
Total training time: 3.92 seconds.
-- Epoch 19
Norm: 3.67, NNZs: 2298, Bias: -4.389650, T: 506673, Avg. loss: 0.592519
Total training time: 4.14 seconds.
-- Epoch 20
Norm: 3.60, NNZs: 2298, Bias: -4.317769, T: 533340, Avg. loss: 0.589418
Total training time: 4.35 seconds.
-- Epoch 21
Norm: 3.55, NNZs: 2298, Bias: -4.238421, T: 560007, Avg. loss: 0.581226
```

```
Total training time: 4.57 seconds.
-- Epoch 22
Norm: 3.49, NNZs: 2298, Bias: -4.170832, T: 586674, Avg. loss: 0.581895
Total training time: 4.80 seconds.
-- Epoch 23
Norm: 3.44, NNZs: 2298, Bias: -4.113618, T: 613341, Avg. loss: 0.582799
Total training time: 5.01 seconds.
-- Epoch 24
Norm: 3.40, NNZs: 2298, Bias: -4.048529, T: 640008, Avg. loss: 0.574104
Total training time: 5.23 seconds.
-- Epoch 25
Norm: 3.34, NNZs: 2298, Bias: -4.008444, T: 666675, Avg. loss: 0.580090
Total training time: 5.45 seconds.
-- Epoch 26
Norm: 3.32, NNZs: 2298, Bias: -3.943485, T: 693342, Avg. loss: 0.565187
Total training time: 5.67 seconds.
-- Epoch 27
Norm: 3.27, NNZs: 2298, Bias: -3.905666, T: 720009, Avg. loss: 0.576911
Total training time: 5.89 seconds.
-- Epoch 28
Norm: 3.23, NNZs: 2298, Bias: -3.859327, T: 746676, Avg. loss: 0.568065
Total training time: 6.10 seconds.
-- Epoch 29
Norm: 3.21, NNZs: 2298, Bias: -3.805000, T: 773343, Avg. loss: 0.561024
Total training time: 6.32 seconds.
-- Epoch 30
Norm: 3.17, NNZs: 2298, Bias: -3.769616, T: 800010, Avg. loss: 0.568737
Total training time: 6.54 seconds.
-- Epoch 31
Norm: 3.14, NNZs: 2298, Bias: -3.733346, T: 826677, Avg. loss: 0.565886
Total training time: 6.76 seconds.
-- Epoch 32
Norm: 3.12, NNZs: 2298, Bias: -3.686200, T: 853344, Avg. loss: 0.558470
Total training time: 6.98 seconds.
-- Epoch 33
Norm: 3.09, NNZs: 2298, Bias: -3.649255, T: 880011, Avg. loss: 0.560985
Total training time: 7.19 seconds.
-- Epoch 34
Norm: 3.07, NNZs: 2298, Bias: -3.616326, T: 906678, Avg. loss: 0.561003
Total training time: 7.41 seconds.
-- Epoch 35
Norm: 3.04, NNZs: 2298, Bias: -3.586077, T: 933345, Avg. loss: 0.561432
Total training time: 7.63 seconds.
-- Epoch 36
Norm: 3.02, NNZs: 2298, Bias: -3.550785, T: 960012, Avg. loss: 0.556711
Total training time: 7.85 seconds.
-- Epoch 37
Norm: 3.00, NNZs: 2298, Bias: -3.519852, T: 986679, Avg. loss: 0.556587
Total training time: 8.06 seconds.
-- Epoch 38
Norm: 2.98, NNZs: 2298, Bias: -3.489972, T: 1013346, Avg. loss: 0.559875
Total training time: 8.28 seconds.
-- Epoch 39
Norm: 2.96, NNZs: 2298, Bias: -3.459624, T: 1040013, Avg. loss: 0.556483
Total training time: 8.49 seconds.
-- Epoch 40
Norm: 2.95, NNZs: 2298, Bias: -3.434177, T: 1066680, Avg. loss: 0.574325
Total training time: 8.72 seconds.
-- Epoch 41
Norm: 2.92, NNZs: 2298, Bias: -3.414750, T: 1093347, Avg. loss: 0.562018
Total training time: 8.93 seconds.
Convergence after 41 epochs took 8.93 seconds
-- Epoch 1
Norm: 1.55, NNZs: 2298, Bias: -0.930551, T: 26666, Avg. loss: 2.832193
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 1.03, NNZs: 2298, Bias: -0.847560, T: 53332, Avg. loss: 0.690954
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 0.87, NNZs: 2298, Bias: -0.798830, T: 79998, Avg. loss: 0.617022
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 0.80, NNZs: 2298, Bias: -0.767643, T: 106664, Avg. loss: 0.597767
Total training time: 0.89 seconds.
-- Epoch 5
Norm: 0.76, NNZs: 2298, Bias: -0.749926, T: 133330, Avg. loss: 0.589112
Total training time: 1.11 seconds.
```

```
-- Epoch 6
Norm: 0.75, NNZs: 2298, Bias: -0.730850, T: 159996, Avg. loss: 0.580679
Total training time: 1.32 seconds.
-- Epoch 7
Norm: 0.73, NNZs: 2298, Bias: -0.716724, T: 186662, Avg. loss: 0.578205
Total training time: 1.54 seconds.
-- Epoch 8
Norm: 0.72, NNZs: 2298, Bias: -0.706785, T: 213328, Avg. loss: 0.578335
Total training time: 1.76 seconds.
-- Epoch 9
Norm: 0.71, NNZs: 2298, Bias: -0.695265, T: 239994, Avg. loss: 0.571735
Total training time: 1.98 seconds.
-- Epoch 10
Norm: 0.71, NNZs: 2298, Bias: -0.687290, T: 266660, Avg. loss: 0.572797
Total training time: 2.20 seconds.
-- Epoch 11
Norm: 0.70, NNZs: 2298, Bias: -0.681481, T: 293326, Avg. loss: 0.574750
Total training time: 2.42 seconds.
-- Epoch 12
Norm: 0.70, NNZs: 2298, Bias: -0.674691, T: 319992, Avg. loss: 0.573067
Total training time: 2.64 seconds.
-- Epoch 13
Norm: 0.69, NNZs: 2298, Bias: -0.668503, T: 346658, Avg. loss: 0.570369
Total training time: 2.85 seconds.
-- Epoch 14
Norm: 0.69, NNZs: 2298, Bias: -0.663383, T: 373324, Avg. loss: 0.570517
Total training time: 3.07 seconds.
-- Epoch 15
Norm: 0.69, NNZs: 2298, Bias: -0.658978, T: 399990, Avg. loss: 0.570800
Total training time: 3.29 seconds.
-- Epoch 16
Norm: 0.69, NNZs: 2298, Bias: -0.655184, T: 426656, Avg. loss: 0.571103
Total training time: 3.52 seconds.
-- Epoch 17
Norm: 0.68, NNZs: 2298, Bias: -0.651639, T: 453322, Avg. loss: 0.570610
Total training time: 3.73 seconds.
-- Epoch 18
Norm: 0.68, NNZs: 2298, Bias: -0.648021, T: 479988, Avg. loss: 0.570091
Total training time: 3.95 seconds.
Convergence after 18 epochs took 3.95 seconds
-- Epoch 1
Norm: 1.56, NNZs: 2297, Bias: -0.234808, T: 26667, Avg. loss: 2.598657
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 1.14, NNZs: 2297, Bias: -0.277199, T: 53334, Avg. loss: 0.721493
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 0.82, NNZs: 2297, Bias: -0.305057, T: 80001, Avg. loss: 0.693620
Total training time: 0.64 seconds.
-- Epoch 4
Norm: 0.75, NNZs: 2297, Bias: -0.316386, T: 106668, Avg. loss: 0.625245
Total training time: 0.86 seconds.
-- Epoch 5
Norm: 0.72, NNZs: 2297, Bias: -0.326217, T: 133335, Avg. loss: 0.612908
Total training time: 1.08 seconds.
-- Epoch 6
Norm: 0.72, NNZs: 2297, Bias: -0.334702, T: 160002, Avg. loss: 0.616705
Total training time: 1.29 seconds.
-- Epoch 7
Norm: 0.70, NNZs: 2297, Bias: -0.339536, T: 186669, Avg. loss: 0.603071
Total training time: 1.51 seconds.
-- Epoch 8
Norm: 0.68, NNZs: 2297, Bias: -0.343767, T: 213336, Avg. loss: 0.597362
Total training time: 1.73 seconds.
-- Epoch 9
Norm: 0.68, NNZs: 2297, Bias: -0.348063, T: 240003, Avg. loss: 0.596814
Total training time: 1.94 seconds.
-- Epoch 10
Norm: 0.67, NNZs: 2297, Bias: -0.351620, T: 266670, Avg. loss: 0.594357
Total training time: 2.16 seconds.
-- Epoch 11
Norm: 0.66, NNZs: 2297, Bias: -0.355648, T: 293337, Avg. loss: 0.593438
Total training time: 2.38 seconds.
-- Epoch 12
Norm: 0.66, NNZs: 2297, Bias: -0.356883, T: 320004, Avg. loss: 0.588944
Total training time: 2.59 seconds.
-- Epoch 13
```

```
Norm: 0.65, NNZs: 2297, Bias: -0.359711, T: 346671, Avg. loss: 0.590804
Total training time: 2.81 seconds.
-- Epoch 14
Norm: 0.65, NNZs: 2297, Bias: -0.361762, T: 373338, Avg. loss: 0.588353
Total training time: 3.03 seconds.
-- Epoch 15
Norm: 0.65, NNZs: 2297, Bias: -0.363535, T: 400005, Avg. loss: 0.588396
Total training time: 3.24 seconds.
-- Epoch 16
Norm: 0.65, NNZs: 2297, Bias: -0.365699, T: 426672, Avg. loss: 0.588208
Total training time: 3.46 seconds.
-- Epoch 17
Norm: 0.65, NNZs: 2297, Bias: -0.367342, T: 453339, Avg. loss: 0.587179
Total training time: 3.68 seconds.
-- Epoch 18
Norm: 0.65, NNZs: 2297, Bias: -0.368927, T: 480006, Avg. loss: 0.586820
Total training time: 3.89 seconds.
-- Epoch 19
Norm: 0.65, NNZs: 2297, Bias: -0.369834, T: 506673, Avg. loss: 0.584511
Total training time: 4.11 seconds.
-- Epoch 20
Norm: 0.65, NNZs: 2297, Bias: -0.371376, T: 533340, Avg. loss: 0.586540
Total training time: 4.34 seconds.
-- Epoch 21
Norm: 0.64, NNZs: 2297, Bias: -0.373053, T: 560007, Avg. loss: 0.586269
Total training time: 4.56 seconds.
-- Epoch 22
Norm: 0.64, NNZs: 2297, Bias: -0.373941, T: 586674, Avg. loss: 0.584112
Total training time: 4.78 seconds.
-- Epoch 23
Norm: 0.64, NNZs: 2297, Bias: -0.375480, T: 613341, Avg. loss: 0.586227
Total training time: 4.99 seconds.
-- Epoch 24
Norm: 0.64, NNZs: 2297, Bias: -0.376434, T: 640008, Avg. loss: 0.583677
Total training time: 5.21 seconds.
Convergence after 24 epochs took 5.21 seconds
-- Epoch 1
Norm: 1.99, NNZs: 2298, Bias: -0.177352, T: 26667, Avg. loss: 2.570074
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 0.93, NNZs: 2298, Bias: -0.221952, T: 53334, Avg. loss: 0.818987
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 0.79, NNZs: 2298, Bias: -0.245395, T: 80001, Avg. loss: 0.644085
Total training time: 0.65 seconds.
-- Epoch 4
Norm: 0.73, NNZs: 2298, Bias: -0.265192, T: 106668, Avg. loss: 0.629630
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 0.70, NNZs: 2298, Bias: -0.277604, T: 133335, Avg. loss: 0.615868
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.68, NNZs: 2298, Bias: -0.288969, T: 160002, Avg. loss: 0.612172
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 0.67, NNZs: 2298, Bias: -0.294168, T: 186669, Avg. loss: 0.601801
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 0.67, NNZs: 2298, Bias: -0.295775, T: 213336, Avg. loss: 0.596223
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 0.66, NNZs: 2298, Bias: -0.300802, T: 240003, Avg. loss: 0.599023
Total training time: 1.96 seconds.
-- Epoch 10
Norm: 0.66, NNZs: 2298, Bias: -0.305631, T: 266670, Avg. loss: 0.598437
Total training time: 2.17 seconds.
-- Epoch 11
Norm: 0.65, NNZs: 2298, Bias: -0.309964, T: 293337, Avg. loss: 0.597625
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 0.65, NNZs: 2298, Bias: -0.313267, T: 320004, Avg. loss: 0.594363
Total training time: 2.61 seconds.
-- Epoch 13
Norm: 0.65, NNZs: 2298, Bias: -0.317189, T: 346671, Avg. loss: 0.595411
Total training time: 2.83 seconds.
-- Epoch 14
Norm: 0.65, NNZs: 2298, Bias: -0.320423, T: 373338, Avg. loss: 0.593384
```

```
Total training time: 3.04 seconds.
-- Epoch 15
Norm: 0.64, NNZs: 2298, Bias: -0.322054, T: 400005, Avg. loss: 0.590481
Total training time: 3.26 seconds.
-- Epoch 16
Norm: 0.64, NNZs: 2298, Bias: -0.324631, T: 426672, Avg. loss: 0.591956
Total training time: 3.48 seconds.
-- Epoch 17
Norm: 0.64, NNZs: 2298, Bias: -0.327385, T: 453339, Avg. loss: 0.592220
Total training time: 3.70 seconds.
-- Epoch 18
Norm: 0.64, NNZs: 2298, Bias: -0.329562, T: 480006, Avg. loss: 0.591404
Total training time: 3.91 seconds.
-- Epoch 19
Norm: 0.64, NNZs: 2298, Bias: -0.331121, T: 506673, Avg. loss: 0.588904
Total training time: 4.13 seconds.
-- Epoch 20
Norm: 0.64, NNZs: 2298, Bias: -0.332558, T: 533340, Avg. loss: 0.588365
Total training time: 4.34 seconds.
-- Epoch 21
Norm: 0.64, NNZs: 2298, Bias: -0.334622, T: 560007, Avg. loss: 0.589884
Total training time: 4.56 seconds.
-- Epoch 22
Norm: 0.64, NNZs: 2298, Bias: -0.335912, T: 586674, Avg. loss: 0.587296
Total training time: 4.78 seconds.
-- Epoch 23
Norm: 0.64, NNZs: 2298, Bias: -0.336856, T: 613341, Avg. loss: 0.586619
Total training time: 5.00 seconds.
-- Epoch 24
Norm: 0.64, NNZs: 2298, Bias: -0.337994, T: 640008, Avg. loss: 0.587268
Total training time: 5.22 seconds.
-- Epoch 25
Norm: 0.64, NNZs: 2298, Bias: -0.339494, T: 666675, Avg. loss: 0.588539
Total training time: 5.43 seconds.
-- Epoch 26
Norm: 0.64, NNZs: 2298, Bias: -0.340538, T: 693342, Avg. loss: 0.586283
Total training time: 5.66 seconds.
-- Epoch 27
Norm: 0.64, NNZs: 2298, Bias: -0.341893, T: 720009, Avg. loss: 0.588048
Total training time: 5.88 seconds.
Convergence after 27 epochs took 5.88 seconds
-- Epoch 1
Norm: 0.24, NNZs: 2298, Bias: -0.430293, T: 26666, Avg. loss: 0.709373
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.21, NNZs: 2298, Bias: -0.411325, T: 53332, Avg. loss: 0.581230
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 0.21, NNZs: 2298, Bias: -0.400531, T: 79998, Avg. loss: 0.577783
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 0.20, NNZs: 2298, Bias: -0.393152, T: 106664, Avg. loss: 0.577996
Total training time: 0.88 seconds.
-- Epoch 5
Norm: 0.20, NNZs: 2298, Bias: -0.387935, T: 133330, Avg. loss: 0.579988
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.20, NNZs: 2298, Bias: -0.383638, T: 159996, Avg. loss: 0.579165
Total training time: 1.32 seconds.
-- Epoch 7
Norm: 0.20, NNZs: 2298, Bias: -0.380148, T: 186662, Avg. loss: 0.579670
Total training time: 1.54 seconds.
-- Epoch 8
Norm: 0.20, NNZs: 2298, Bias: -0.377185, T: 213328, Avg. loss: 0.580479
Total training time: 1.76 seconds.
Convergence after 8 epochs took 1.76 seconds
-- Epoch 1
Norm: 0.37, NNZs: 2297, Bias: 1.325448, T: 26667, Avg. loss: 1.346246
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.34, NNZs: 2297, Bias: 1.186164, T: 53334, Avg. loss: 1.088899
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 0.32, NNZs: 2297, Bias: 1.110506, T: 80001, Avg. loss: 1.043129
Total training time: 0.65 seconds.
-- Epoch 4
Norm: 0.32, NNZs: 2297, Bias: 1.059176, T: 106668, Avg. loss: 1.017096
```

```
Total training time: 0.86 seconds.
-- Epoch 5
Norm: 0.31, NNZs: 2297, Bias: 1.020672, T: 133335, Avg. loss: 0.999599
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.30, NNZs: 2297, Bias: 0.990094, T: 160002, Avg. loss: 0.986164
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 0.30, NNZs: 2297, Bias: 0.964827, T: 186669, Avg. loss: 0.974886
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 0.29, NNZs: 2297, Bias: 0.943354, T: 213336, Avg. loss: 0.965397
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 0.29, NNZs: 2297, Bias: 0.924725, T: 240003, Avg. loss: 0.958739
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 0.29, NNZs: 2297, Bias: 0.908334, T: 266670, Avg. loss: 0.951320
Total training time: 2.16 seconds.
-- Epoch 11
Norm: 0.28, NNZs: 2297, Bias: 0.893659, T: 293337, Avg. loss: 0.947083
Total training time: 2.38 seconds.
-- Epoch 12
Norm: 0.28, NNZs: 2297, Bias: 0.880459, T: 320004, Avg. loss: 0.940773
Total training time: 2.60 seconds.
-- Epoch 13
Norm: 0.28, NNZs: 2297, Bias: 0.868461, T: 346671, Avg. loss: 0.936030
Total training time: 2.81 seconds.
-- Epoch 14
Norm: 0.28, NNZs: 2297, Bias: 0.857458, T: 373338, Avg. loss: 0.931910
Total training time: 3.04 seconds.
-- Epoch 15
Norm: 0.28, NNZs: 2297, Bias: 0.847307, T: 400005, Avg. loss: 0.928572
Total training time: 3.25 seconds.
-- Epoch 16
Norm: 0.27, NNZs: 2297, Bias: 0.837903, T: 426672, Avg. loss: 0.924883
Total training time: 3.47 seconds.
-- Epoch 17
Norm: 0.27, NNZs: 2297, Bias: 0.829149, T: 453339, Avg. loss: 0.921293
Total training time: 3.69 seconds.
-- Epoch 18
Norm: 0.27, NNZs: 2297, Bias: 0.820949, T: 480006, Avg. loss: 0.918865
Total training time: 3.90 seconds.
-- Epoch 19
Norm: 0.27, NNZs: 2297, Bias: 0.813265, T: 506673, Avg. loss: 0.915184
Total training time: 4.12 seconds.
-- Epoch 20
Norm: 0.27, NNZs: 2297, Bias: 0.806018, T: 533340, Avg. loss: 0.913399
Total training time: 4.34 seconds.
-- Epoch 21
Norm: 0.27, NNZs: 2297, Bias: 0.799169, T: 560007, Avg. loss: 0.910747
Total training time: 4.56 seconds.
-- Epoch 22
Norm: 0.27, NNZs: 2297, Bias: 0.792685, T: 586674, Avg. loss: 0.908141
Total training time: 4.77 seconds.
-- Epoch 23
Norm: 0.27, NNZs: 2297, Bias: 0.786536, T: 613341, Avg. loss: 0.905851
Total training time: 5.00 seconds.
-- Epoch 24
Norm: 0.26, NNZs: 2297, Bias: 0.780675, T: 640008, Avg. loss: 0.904025
Total training time: 5.21 seconds.
-- Epoch 25
Norm: 0.26, NNZs: 2297, Bias: 0.775089, T: 666675, Avg. loss: 0.901950
Total training time: 5.43 seconds.
-- Epoch 26
Norm: 0.26, NNZs: 2297, Bias: 0.769748, T: 693342, Avg. loss: 0.900174
Total training time: 5.65 seconds.
-- Epoch 27
Norm: 0.26, NNZs: 2297, Bias: 0.764634, T: 720009, Avg. loss: 0.898324
Total training time: 5.86 seconds.
-- Epoch 28
Norm: 0.26, NNZs: 2297, Bias: 0.759740, T: 746676, Avg. loss: 0.896130
Total training time: 6.09 seconds.
-- Epoch 29
Norm: 0.26, NNZs: 2297, Bias: 0.755033, T: 773343, Avg. loss: 0.894955
Total training time: 6.30 seconds.
-- Epoch 30
```

```
Norm: 0.26, NNZs: 2297, Bias: 0.750510, T: 800010, Avg. loss: 0.893202
Total training time: 6.52 seconds.
-- Epoch 31
Norm: 0.26, NNZs: 2297, Bias: 0.746151, T: 826677, Avg. loss: 0.891933
Total training time: 6.74 seconds.
-- Epoch 32
Norm: 0.26, NNZs: 2297, Bias: 0.741953, T: 853344, Avg. loss: 0.890073
Total training time: 6.95 seconds.
-- Epoch 33
Norm: 0.26, NNZs: 2297, Bias: 0.737900, T: 880011, Avg. loss: 0.888816
Total training time: 7.17 seconds.
-- Epoch 34
Norm: 0.26, NNZs: 2297, Bias: 0.733985, T: 906678, Avg. loss: 0.887455
Total training time: 7.38 seconds.
-- Epoch 35
Norm: 0.26, NNZs: 2297, Bias: 0.730196, T: 933345, Avg. loss: 0.886355
Total training time: 7.60 seconds.
-- Epoch 36
Norm: 0.26, NNZs: 2297, Bias: 0.726529, T: 960012, Avg. loss: 0.884993
Total training time: 7.82 seconds.
-- Epoch 37
Norm: 0.25, NNZs: 2297, Bias: 0.722979, T: 986679, Avg. loss: 0.883457
Total training time: 8.05 seconds.
-- Epoch 38
Norm: 0.25, NNZs: 2297, Bias: 0.719534, T: 1013346, Avg. loss: 0.882521
Total training time: 8.27 seconds.
-- Epoch 39
Norm: 0.25, NNZs: 2297, Bias: 0.716193, T: 1040013, Avg. loss: 0.881247
Total training time: 8.49 seconds.
-- Epoch 40
Norm: 0.25, NNZs: 2297, Bias: 0.712945, T: 1066680, Avg. loss: 0.880392
Total training time: 8.70 seconds.
-- Epoch 41
Norm: 0.25, NNZs: 2297, Bias: 0.709789, T: 1093347, Avg. loss: 0.879121
Total training time: 8.92 seconds.
-- Epoch 42
Norm: 0.25, NNZs: 2297, Bias: 0.706722, T: 1120014, Avg. loss: 0.877840
Total training time: 9.14 seconds.
-- Epoch 43
Norm: 0.25, NNZs: 2297, Bias: 0.703734, T: 1146681, Avg. loss: 0.877086
Total training time: 9.36 seconds.
-- Epoch 44
Norm: 0.25, NNZs: 2297, Bias: 0.700827, T: 1173348, Avg. loss: 0.875895
Total training time: 9.58 seconds.
-- Epoch 45
Norm: 0.25, NNZs: 2297, Bias: 0.697991, T: 1200015, Avg. loss: 0.875197
Total training time: 9.79 seconds.
-- Epoch 46
Norm: 0.25, NNZs: 2297, Bias: 0.695226, T: 1226682, Avg. loss: 0.874275
Total training time: 10.01 seconds.
-- Epoch 47
Norm: 0.25, NNZs: 2297, Bias: 0.692529, T: 1253349, Avg. loss: 0.873283
Total training time: 10.23 seconds.
-- Epoch 48
Norm: 0.25, NNZs: 2297, Bias: 0.689897, T: 1280016, Avg. loss: 0.872228
Total training time: 10.45 seconds.
-- Epoch 49
Norm: 0.25, NNZs: 2297, Bias: 0.687326, T: 1306683, Avg. loss: 0.871543
Total training time: 10.66 seconds.
-- Epoch 50
Norm: 0.25, NNZs: 2297, Bias: 0.684816, T: 1333350, Avg. loss: 0.870423
Total training time: 10.88 seconds.
-- Epoch 51
Norm: 0.25, NNZs: 2297, Bias: 0.682362, T: 1360017, Avg. loss: 0.869642
Total training time: 11.10 seconds.
-- Epoch 52
Norm: 0.25, NNZs: 2297, Bias: 0.679960, T: 1386684, Avg. loss: 0.868975
Total training time: 11.32 seconds.
-- Epoch 53
Norm: 0.25, NNZs: 2297, Bias: 0.677610, T: 1413351, Avg. loss: 0.868321
Total training time: 11.54 seconds.
-- Epoch 54
Norm: 0.25, NNZs: 2297, Bias: 0.675310, T: 1440018, Avg. loss: 0.867582
Total training time: 11.75 seconds.
-- Epoch 55
Norm: 0.25, NNZs: 2297, Bias: 0.673059, T: 1466685, Avg. loss: 0.866600
Total training time: 11.97 seconds.
```

```
Total training time: 11.97 seconds.
Convergence after 55 epochs took 11.97 seconds
-- Epoch 1
Norm: 0.38, NNZs: 2298, Bias: 1.262792, T: 26667, Avg. loss: 1.296632
Total training time: 0.21 seconds.
-- Epoch 2
Norm: 0.33, NNZs: 2298, Bias: 1.130087, T: 53334, Avg. loss: 1.072740
Total training time: 0.43 seconds.
-- Epoch 3
Norm: 0.31, NNZs: 2298, Bias: 1.057082, T: 80001, Avg. loss: 1.021192
Total training time: 0.65 seconds.
-- Epoch 4
Norm: 0.31, NNZs: 2298, Bias: 1.007730, T: 106668, Avg. loss: 0.997879
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 0.30, NNZs: 2298, Bias: 0.970763, T: 133335, Avg. loss: 0.980730
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.29, NNZs: 2298, Bias: 0.941323, T: 160002, Avg. loss: 0.968747
Total training time: 1.31 seconds.
-- Epoch 7
Norm: 0.29, NNZs: 2298, Bias: 0.917009, T: 186669, Avg. loss: 0.957442
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 0.28, NNZs: 2298, Bias: 0.896380, T: 213336, Avg. loss: 0.948026
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 0.28, NNZs: 2298, Bias: 0.878474, T: 240003, Avg. loss: 0.941639
Total training time: 1.96 seconds.
-- Epoch 10
Norm: 0.28, NNZs: 2298, Bias: 0.862723, T: 266670, Avg. loss: 0.935152
Total training time: 2.18 seconds.
-- Epoch 11
Norm: 0.28, NNZs: 2298, Bias: 0.848636, T: 293337, Avg. loss: 0.930950
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 0.27, NNZs: 2298, Bias: 0.835951, T: 320004, Avg. loss: 0.925139
Total training time: 2.61 seconds.
-- Epoch 13
Norm: 0.27, NNZs: 2298, Bias: 0.824426, T: 346671, Avg. loss: 0.920870
Total training time: 2.83 seconds.
-- Epoch 14
Norm: 0.27, NNZs: 2298, Bias: 0.813854, T: 373338, Avg. loss: 0.917233
Total training time: 3.05 seconds.
-- Epoch 15
Norm: 0.27, NNZs: 2298, Bias: 0.804098, T: 400005, Avg. loss: 0.913838
Total training time: 3.26 seconds.
-- Epoch 16
Norm: 0.27, NNZs: 2298, Bias: 0.795066, T: 426672, Avg. loss: 0.910327
Total training time: 3.48 seconds.
-- Epoch 17
Norm: 0.27, NNZs: 2298, Bias: 0.786661, T: 453339, Avg. loss: 0.906971
Total training time: 3.70 seconds.
-- Epoch 18
Norm: 0.26, NNZs: 2298, Bias: 0.778789, T: 480006, Avg. loss: 0.904771
Total training time: 3.92 seconds.
-- Epoch 19
Norm: 0.26, NNZs: 2298, Bias: 0.771407, T: 506673, Avg. loss: 0.901558
Total training time: 4.14 seconds.
-- Epoch 20
Norm: 0.26, NNZs: 2298, Bias: 0.764457, T: 533340, Avg. loss: 0.898920
Total training time: 4.35 seconds.
-- Epoch 21
Norm: 0.26, NNZs: 2298, Bias: 0.757875, T: 560007, Avg. loss: 0.897189
Total training time: 4.57 seconds.
-- Epoch 22
Norm: 0.26, NNZs: 2298, Bias: 0.751654, T: 586674, Avg. loss: 0.894223
Total training time: 4.79 seconds.
-- Epoch 23
Norm: 0.26, NNZs: 2298, Bias: 0.745751, T: 613341, Avg. loss: 0.891901
Total training time: 5.00 seconds.
-- Epoch 24
Norm: 0.26, NNZs: 2298, Bias: 0.740121, T: 640008, Avg. loss: 0.890639
Total training time: 5.22 seconds.
-- Epoch 25
Norm: 0.26, NNZs: 2298, Bias: 0.734758, T: 666675, Avg. loss: 0.888563
Total training time: 5.44 seconds.
-- Epoch 26
```

```
-- Epoch 26
Norm: 0.26, NNZs: 2298, Bias: 0.729633, T: 693342, Avg. loss: 0.886626
Total training time: 5.65 seconds.
-- Epoch 27
Norm: 0.25, NNZs: 2298, Bias: 0.724726, T: 720009, Avg. loss: 0.885123
Total training time: 5.87 seconds.
-- Epoch 28
Norm: 0.25, NNZs: 2298, Bias: 0.720023, T: 746676, Avg. loss: 0.883406
Total training time: 6.09 seconds.
-- Epoch 29
Norm: 0.25, NNZs: 2298, Bias: 0.715499, T: 773343, Avg. loss: 0.882056
Total training time: 6.30 seconds.
-- Epoch 30
Norm: 0.25, NNZs: 2298, Bias: 0.711165, T: 800010, Avg. loss: 0.879732
Total training time: 6.52 seconds.
-- Epoch 31
Norm: 0.25, NNZs: 2298, Bias: 0.706980, T: 826677, Avg. loss: 0.879131
Total training time: 6.74 seconds.
-- Epoch 32
Norm: 0.25, NNZs: 2298, Bias: 0.702949, T: 853344, Avg. loss: 0.877411
Total training time: 6.96 seconds.
-- Epoch 33
Norm: 0.25, NNZs: 2298, Bias: 0.699061, T: 880011, Avg. loss: 0.875880
Total training time: 7.17 seconds.
-- Epoch 34
Norm: 0.25, NNZs: 2298, Bias: 0.695302, T: 906678, Avg. loss: 0.874783
Total training time: 7.39 seconds.
-- Epoch 35
Norm: 0.25, NNZs: 2298, Bias: 0.691661, T: 933345, Avg. loss: 0.874102
Total training time: 7.62 seconds.
-- Epoch 36
Norm: 0.25, NNZs: 2298, Bias: 0.688144, T: 960012, Avg. loss: 0.872247
Total training time: 7.84 seconds.
-- Epoch 37
Norm: 0.25, NNZs: 2298, Bias: 0.684737, T: 986679, Avg. loss: 0.870973
Total training time: 8.06 seconds.
-- Epoch 38
Norm: 0.25, NNZs: 2298, Bias: 0.681427, T: 1013346, Avg. loss: 0.870495
Total training time: 8.28 seconds.
-- Epoch 39
Norm: 0.25, NNZs: 2298, Bias: 0.678221, T: 1040013, Avg. loss: 0.868808
Total training time: 8.49 seconds.
-- Epoch 40
Norm: 0.25, NNZs: 2298, Bias: 0.675103, T: 1066680, Avg. loss: 0.868159
Total training time: 8.71 seconds.
-- Epoch 41
Norm: 0.25, NNZs: 2298, Bias: 0.672073, T: 1093347, Avg. loss: 0.867087
Total training time: 8.93 seconds.
-- Epoch 42
Norm: 0.24, NNZs: 2298, Bias: 0.669129, T: 1120014, Avg. loss: 0.865855
Total training time: 9.15 seconds.
-- Epoch 43
Norm: 0.24, NNZs: 2298, Bias: 0.666260, T: 1146681, Avg. loss: 0.865060
Total training time: 9.37 seconds.
-- Epoch 44
Norm: 0.24, NNZs: 2298, Bias: 0.663472, T: 1173348, Avg. loss: 0.863808
Total training time: 9.58 seconds.
-- Epoch 45
Norm: 0.24, NNZs: 2298, Bias: 0.660749, T: 1200015, Avg. loss: 0.863246
Total training time: 9.81 seconds.
-- Epoch 46
Norm: 0.24, NNZs: 2298, Bias: 0.658094, T: 1226682, Avg. loss: 0.862411
Total training time: 10.03 seconds.
-- Epoch 47
Norm: 0.24, NNZs: 2298, Bias: 0.655507, T: 1253349, Avg. loss: 0.861169
Total training time: 10.24 seconds.
-- Epoch 48
Norm: 0.24, NNZs: 2298, Bias: 0.652983, T: 1280016, Avg. loss: 0.860213
Total training time: 10.46 seconds.
-- Epoch 49
Norm: 0.24, NNZs: 2298, Bias: 0.650515, T: 1306683, Avg. loss: 0.859888
Total training time: 10.68 seconds.
-- Epoch 50
Norm: 0.24, NNZs: 2298, Bias: 0.648106, T: 1333350, Avg. loss: 0.858612
Total training time: 10.90 seconds.
-- Epoch 51
Norm: 0.24, NNZs: 2298, Bias: 0.645748, T: 1360017, Avg. loss: 0.858254
Total training time: 11.12 seconds.
```

```
Total training time: 11.12 seconds.
-- Epoch 52
Norm: 0.24, NNZs: 2298, Bias: 0.643445, T: 1386684, Avg. loss: 0.857090
Total training time: 11.33 seconds.
-- Epoch 53
Norm: 0.24, NNZs: 2298, Bias: 0.641191, T: 1413351, Avg. loss: 0.856571
Total training time: 11.55 seconds.
-- Epoch 54
Norm: 0.24, NNZs: 2298, Bias: 0.638982, T: 1440018, Avg. loss: 0.856140
Total training time: 11.78 seconds.
-- Epoch 55
Norm: 0.24, NNZs: 2298, Bias: 0.636822, T: 1466685, Avg. loss: 0.855021
Total training time: 12.00 seconds.
-- Epoch 56
Norm: 0.24, NNZs: 2298, Bias: 0.634707, T: 1493352, Avg. loss: 0.854292
Total training time: 12.21 seconds.
-- Epoch 57
Norm: 0.24, NNZs: 2298, Bias: 0.632631, T: 1520019, Avg. loss: 0.854035
Total training time: 12.43 seconds.
-- Epoch 58
Norm: 0.24, NNZs: 2298, Bias: 0.630597, T: 1546686, Avg. loss: 0.853305
Total training time: 12.65 seconds.
-- Epoch 59
Norm: 0.24, NNZs: 2298, Bias: 0.628605, T: 1573353, Avg. loss: 0.852148
Total training time: 12.87 seconds.
-- Epoch 60
Norm: 0.24, NNZs: 2298, Bias: 0.626652, T: 1600020, Avg. loss: 0.851555
Total training time: 13.09 seconds.
-- Epoch 61
Norm: 0.24, NNZs: 2298, Bias: 0.624732, T: 1626687, Avg. loss: 0.851295
Total training time: 13.31 seconds.
-- Epoch 62
Norm: 0.24, NNZs: 2298, Bias: 0.622849, T: 1653354, Avg. loss: 0.850474
Total training time: 13.53 seconds.
-- Epoch 63
Norm: 0.24, NNZs: 2298, Bias: 0.621000, T: 1680021, Avg. loss: 0.849965
Total training time: 13.74 seconds.
-- Epoch 64
Norm: 0.24, NNZs: 2298, Bias: 0.619184, T: 1706688, Avg. loss: 0.849332
Total training time: 13.96 seconds.
Convergence after 64 epochs took 13.96 seconds
-- Epoch 1
Norm: 0.05, NNZs: 2298, Bias: -0.203372, T: 26666, Avg. loss: 0.619385
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.05, NNZs: 2298, Bias: -0.201633, T: 53332, Avg. loss: 0.612918
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 0.05, NNZs: 2298, Bias: -0.200612, T: 79998, Avg. loss: 0.612617
Total training time: 0.65 seconds.
-- Epoch 4
Norm: 0.05, NNZs: 2298, Bias: -0.199899, T: 106664, Avg. loss: 0.613121
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 0.05, NNZs: 2298, Bias: -0.199356, T: 133330, Avg. loss: 0.613332
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.05, NNZs: 2298, Bias: -0.198916, T: 159996, Avg. loss: 0.613452
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 0.05, NNZs: 2298, Bias: -0.198549, T: 186662, Avg. loss: 0.613558
Total training time: 1.52 seconds.
Convergence after 7 epochs took 1.52 seconds
-- Epoch 1
Norm: 0.06, NNZs: 2297, Bias: 1.606424, T: 26667, Avg. loss: 1.638309
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.06, NNZs: 2297, Bias: 1.583885, T: 53334, Avg. loss: 1.598169
Total training time: 0.44 seconds.
-- Epoch 3
Norm: 0.06, NNZs: 2297, Bias: 1.570781, T: 80001, Avg. loss: 1.585440
Total training time: 0.66 seconds.
-- Epoch 4
Norm: 0.06, NNZs: 2297, Bias: 1.561530, T: 106668, Avg. loss: 1.577421
Total training time: 0.87 seconds.
-- Epoch 5
Norm: 0.06, NNZs: 2297, Bias: 1.554380, T: 133335, Avg. loss: 1.571605
Total training time: 1.09 seconds.
```

```
Total training time: 1.09 seconds.
-- Epoch 6
Norm: 0.06, NNZs: 2297, Bias: 1.548556, T: 160002, Avg. loss: 1.566996
Total training time: 1.30 seconds.
-- Epoch 7
Norm: 0.06, NNZs: 2297, Bias: 1.543641, T: 186669, Avg. loss: 1.563060
Total training time: 1.52 seconds.
-- Epoch 8
Norm: 0.06, NNZs: 2297, Bias: 1.539391, T: 213336, Avg. loss: 1.559707
Total training time: 1.74 seconds.
-- Epoch 9
Norm: 0.06, NNZs: 2297, Bias: 1.535651, T: 240003, Avg. loss: 1.556906
Total training time: 1.95 seconds.
-- Epoch 10
Norm: 0.06, NNZs: 2297, Bias: 1.532310, T: 266670, Avg. loss: 1.554323
Total training time: 2.17 seconds.
-- Epoch 11
Norm: 0.06, NNZs: 2297, Bias: 1.529293, T: 293337, Avg. loss: 1.552128
Total training time: 2.39 seconds.
-- Epoch 12
Norm: 0.06, NNZs: 2297, Bias: 1.526541, T: 320004, Avg. loss: 1.550033
Total training time: 2.61 seconds.
-- Epoch 13
Norm: 0.06, NNZs: 2297, Bias: 1.524013, T: 346671, Avg. loss: 1.548123
Total training time: 2.82 seconds.
-- Epoch 14
Norm: 0.06, NNZs: 2297, Bias: 1.521674, T: 373338, Avg. loss: 1.546411
Total training time: 3.03 seconds.
-- Epoch 15
Norm: 0.06, NNZs: 2297, Bias: 1.519500, T: 400005, Avg. loss: 1.544812
Total training time: 3.25 seconds.
-- Epoch 16
Norm: 0.06, NNZs: 2297, Bias: 1.517468, T: 426672, Avg. loss: 1.543299
Total training time: 3.47 seconds.
-- Epoch 17
Norm: 0.06, NNZs: 2297, Bias: 1.515560, T: 453339, Avg. loss: 1.541891
Total training time: 3.69 seconds.
-- Epoch 18
Norm: 0.06, NNZs: 2297, Bias: 1.513764, T: 480006, Avg. loss: 1.540593
Total training time: 3.91 seconds.
-- Epoch 19
Norm: 0.06, NNZs: 2297, Bias: 1.512065, T: 506673, Avg. loss: 1.539316
Total training time: 4.12 seconds.
-- Epoch 20
Norm: 0.06, NNZs: 2297, Bias: 1.510455, T: 533340, Avg. loss: 1.538172
Total training time: 4.33 seconds.
-- Epoch 21
Norm: 0.06, NNZs: 2297, Bias: 1.508925, T: 560007, Avg. loss: 1.537044
Total training time: 4.56 seconds.
-- Epoch 22
Norm: 0.06, NNZs: 2297, Bias: 1.507467, T: 586674, Avg. loss: 1.535978
Total training time: 4.78 seconds.
-- Epoch 23
Norm: 0.06, NNZs: 2297, Bias: 1.506075, T: 613341, Avg. loss: 1.534982
Total training time: 4.99 seconds.
-- Epoch 24
Norm: 0.06, NNZs: 2297, Bias: 1.504742, T: 640008, Avg. loss: 1.534019
Total training time: 5.21 seconds.
-- Epoch 25
Norm: 0.06, NNZs: 2297, Bias: 1.503466, T: 666675, Avg. loss: 1.533085
Total training time: 5.43 seconds.
-- Epoch 26
Norm: 0.06, NNZs: 2297, Bias: 1.502239, T: 693342, Avg. loss: 1.532194
Total training time: 5.66 seconds.
-- Epoch 27
Norm: 0.06, NNZs: 2297, Bias: 1.501060, T: 720009, Avg. loss: 1.531345
Total training time: 5.87 seconds.
Convergence after 27 epochs took 5.87 seconds
-- Epoch 1
Norm: 0.06, NNZs: 2298, Bias: 1.601844, T: 26667, Avg. loss: 1.633998
Total training time: 0.22 seconds.
-- Epoch 2
Norm: 0.06, NNZs: 2298, Bias: 1.579318, T: 53334, Avg. loss: 1.595175
Total training time: 0.45 seconds.
-- Epoch 3
Norm: 0.06, NNZs: 2298, Bias: 1.566213, T: 80001, Avg. loss: 1.582367
Total training time: 0.67 seconds.
-- Epoch 4
```

```
-- Epoch 4
Norm: 0.06, NNZs: 2298, Bias: 1.556972, T: 106668, Avg. loss: 1.574493
Total training time: 0.88 seconds.
-- Epoch 5
Norm: 0.06, NNZs: 2298, Bias: 1.549832, T: 133335, Avg. loss: 1.568587
Total training time: 1.10 seconds.
-- Epoch 6
Norm: 0.06, NNZs: 2298, Bias: 1.544014, T: 160002, Avg. loss: 1.564058
Total training time: 1.32 seconds.
-- Epoch 7
Norm: 0.06, NNZs: 2298, Bias: 1.539104, T: 186669, Avg. loss: 1.560115
Total training time: 1.54 seconds.
-- Epoch 8
Norm: 0.06, NNZs: 2298, Bias: 1.534856, T: 213336, Avg. loss: 1.556766
Total training time: 1.75 seconds.
-- Epoch 9
Norm: 0.06, NNZs: 2298, Bias: 1.531118, T: 240003, Avg. loss: 1.553989
Total training time: 1.97 seconds.
-- Epoch 10
Norm: 0.06, NNZs: 2298, Bias: 1.527780, T: 266670, Avg. loss: 1.551434
Total training time: 2.19 seconds.
-- Epoch 11
Norm: 0.06, NNZs: 2298, Bias: 1.524765, T: 293337, Avg. loss: 1.549232
Total training time: 2.41 seconds.
-- Epoch 12
Norm: 0.06, NNZs: 2298, Bias: 1.522015, T: 320004, Avg. loss: 1.547141
Total training time: 2.63 seconds.
-- Epoch 13
Norm: 0.06, NNZs: 2298, Bias: 1.519490, T: 346671, Avg. loss: 1.545276
Total training time: 2.85 seconds.
-- Epoch 14
Norm: 0.06, NNZs: 2298, Bias: 1.517154, T: 373338, Avg. loss: 1.543558
Total training time: 3.06 seconds.
-- Epoch 15
Norm: 0.06, NNZs: 2298, Bias: 1.514981, T: 400005, Avg. loss: 1.541938
Total training time: 3.28 seconds.
-- Epoch 16
Norm: 0.06, NNZs: 2298, Bias: 1.512951, T: 426672, Avg. loss: 1.540450
Total training time: 3.50 seconds.
-- Epoch 17
Norm: 0.06, NNZs: 2298, Bias: 1.511046, T: 453339, Avg. loss: 1.539040
Total training time: 3.71 seconds.
-- Epoch 18
Norm: 0.06, NNZs: 2298, Bias: 1.509252, T: 480006, Avg. loss: 1.537749
Total training time: 3.93 seconds.
-- Epoch 19
Norm: 0.06, NNZs: 2298, Bias: 1.507555, T: 506673, Avg. loss: 1.536487
Total training time: 4.15 seconds.
-- Epoch 20
Norm: 0.06, NNZs: 2298, Bias: 1.505947, T: 533340, Avg. loss: 1.535309
Total training time: 4.37 seconds.
-- Epoch 21
Norm: 0.06, NNZs: 2298, Bias: 1.504418, T: 560007, Avg. loss: 1.534220
Total training time: 4.59 seconds.
-- Epoch 22
Norm: 0.06, NNZs: 2298, Bias: 1.502962, T: 586674, Avg. loss: 1.533123
Total training time: 4.80 seconds.
-- Epoch 23
Norm: 0.06, NNZs: 2298, Bias: 1.501571, T: 613341, Avg. loss: 1.532107
Total training time: 5.02 seconds.
-- Epoch 24
Norm: 0.06, NNZs: 2298, Bias: 1.500240, T: 640008, Avg. loss: 1.531168
Total training time: 5.24 seconds.
-- Epoch 25
Norm: 0.06, NNZs: 2298, Bias: 1.498964, T: 666675, Avg. loss: 1.530249
Total training time: 5.46 seconds.
-- Epoch 26
Norm: 0.06, NNZs: 2298, Bias: 1.497739, T: 693342, Avg. loss: 1.529352
Total training time: 5.67 seconds.
-- Epoch 27
Norm: 0.06, NNZs: 2298, Bias: 1.496561, T: 720009, Avg. loss: 1.528515
Total training time: 5.89 seconds.
-- Epoch 28
Norm: 0.06, NNZs: 2298, Bias: 1.495427, T: 746676, Avg. loss: 1.527695
Total training time: 6.10 seconds.
Convergence after 28 epochs took 6.10 seconds
```

```
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed:  2.8min finished
```

```
-- Epoch 1
Norm: 1.67, NNZs: 2298, Bias: -2.352059, T: 40000, Avg. loss: 2.077412
Total training time: 0.34 seconds.
-- Epoch 2
Norm: 1.26, NNZs: 2298, Bias: -1.974342, T: 80000, Avg. loss: 0.584295
Total training time: 0.67 seconds.
-- Epoch 3
Norm: 1.14, NNZs: 2298, Bias: -1.784647, T: 120000, Avg. loss: 0.565022
Total training time: 1.01 seconds.
-- Epoch 4
Norm: 1.04, NNZs: 2298, Bias: -1.675912, T: 160000, Avg. loss: 0.551150
Total training time: 1.34 seconds.
-- Epoch 5
Norm: 1.00, NNZs: 2298, Bias: -1.587582, T: 200000, Avg. loss: 0.540340
Total training time: 1.68 seconds.
-- Epoch 6
Norm: 0.95, NNZs: 2298, Bias: -1.525381, T: 240000, Avg. loss: 0.546370
Total training time: 2.02 seconds.
-- Epoch 7
Norm: 0.93, NNZs: 2298, Bias: -1.471632, T: 280000, Avg. loss: 0.540995
Total training time: 2.35 seconds.
-- Epoch 8
Norm: 0.91, NNZs: 2298, Bias: -1.427152, T: 320000, Avg. loss: 0.543602
Total training time: 2.68 seconds.
-- Epoch 9
Norm: 0.89, NNZs: 2298, Bias: -1.391177, T: 360000, Avg. loss: 0.544649
Total training time: 3.02 seconds.
-- Epoch 10
Norm: 0.87, NNZs: 2298, Bias: -1.360481, T: 400000, Avg. loss: 0.546160
Total training time: 3.35 seconds.
Convergence after 10 epochs took 3.36 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=-1,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
0	24.511982	2.279984	0.057655	0.004150	0.001	{'alpha': 0.001}	0.669730	0.648811
1	9.999835	1.135979	0.053826	0.001591	0.01	{'alpha': 0.01}	0.694401	0.683721

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
2	5.291111	0.804470	0.053583	0.000492	0.1	{'alpha': 0.1}	0.702428	0.701264
3	9.526056	5.362641	0.053209	0.002267	1	{'alpha': 1}	0.692823	0.689266
4	4.773052	2.106325	0.055878	0.003150	10	{'alpha': 10}	0.673368	0.662757

In [ ]:

```
clf = classifier.best_estimator_
```

In [ ]:

```
# clf = SGDClassifier(loss = 'log',alpha = 0.001, class_weight= 'balanced', n_jobs = -1)
```

In [ ]:

```
clf.fit(X_train_churn,y_train_churn)
```

```
-- Epoch 1
Norm: 1.67, NNZs: 2298, Bias: -2.352059, T: 40000, Avg. loss: 2.077412
Total training time: 0.34 seconds.
-- Epoch 2
Norm: 1.26, NNZs: 2298, Bias: -1.974342, T: 80000, Avg. loss: 0.584295
Total training time: 0.67 seconds.
-- Epoch 3
Norm: 1.14, NNZs: 2298, Bias: -1.784647, T: 120000, Avg. loss: 0.565022
Total training time: 1.01 seconds.
-- Epoch 4
Norm: 1.04, NNZs: 2298, Bias: -1.675912, T: 160000, Avg. loss: 0.551150
Total training time: 1.35 seconds.
-- Epoch 5
Norm: 1.00, NNZs: 2298, Bias: -1.587582, T: 200000, Avg. loss: 0.540340
Total training time: 1.70 seconds.
-- Epoch 6
Norm: 0.95, NNZs: 2298, Bias: -1.525381, T: 240000, Avg. loss: 0.546370
Total training time: 2.04 seconds.
-- Epoch 7
Norm: 0.93, NNZs: 2298, Bias: -1.471632, T: 280000, Avg. loss: 0.540995
Total training time: 2.37 seconds.
-- Epoch 8
Norm: 0.91, NNZs: 2298, Bias: -1.427152, T: 320000, Avg. loss: 0.543602
Total training time: 2.71 seconds.
-- Epoch 9
Norm: 0.89, NNZs: 2298, Bias: -1.391177, T: 360000, Avg. loss: 0.544649
Total training time: 3.04 seconds.
-- Epoch 10
Norm: 0.87, NNZs: 2298, Bias: -1.360481, T: 400000, Avg. loss: 0.546160
Total training time: 3.38 seconds.
Convergence after 10 epochs took 3.39 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.1, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=-1, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn)[:,1]
```

In [ ]:

```
ln_train_auc_score_churn = roc_auc_score(y_train_churn, y_train_churn_pred)
```

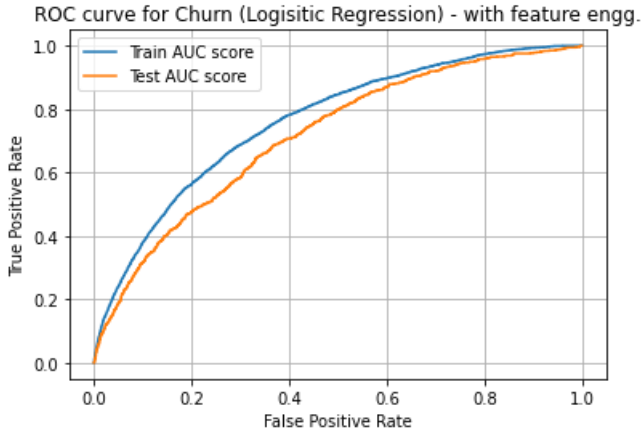
```
lr_train_auc_score_churn = roc_auc_score(y_train_churn, y_train_churn_pred)
lr_test_auc_score_churn = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Logisitic Regression) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
lr_churn_score = ['Logistic Regresion (Churn)', lr_train_auc_score_churn, lr_test_auc_score_churn]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_score)
```

In [ ]:

```
print(score_table)
```

```
+----------------------------+--------------------+--------------------+
|           Model            |     Train AUC      |      Test AUC      |
+----------------------------+--------------------+--------------------+
| Logistic Regresion (Churn) | 0.7617842109476075 | 0.7156584883588942 |
+----------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced',n_jobs = -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= True)
```

In [ ]:

```
classifier.fit(X_train_churn,y_train_churn)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits

[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  90 out of  90 | elapsed: 68.1min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	2.816484	1.171615	0.190589	0.009535	3	10	{'max_dep 3, 'n_estima 10}
1	3.335069	0.024760	0.181797	0.000750	3	20	{'max_dep 3, 'n_estima 20}
2	7.469988	0.041231	0.185245	0.001837	3	50	{'max_dep 3, 'n_estima 50}
							{'max_dep 3

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
3	14.189169	0.083805	0.283356	0.001565	3	100	'n_estima 100}
4	28.687862	0.859474	0.389389	0.001628	3	200	{'max_dep 3, 'n_estima 200}
5	72.287622	0.150396	0.853877	0.089694	3	500	{'max_dep 3, 'n_estima 500}
6	2.886007	0.031357	0.182726	0.000639	5	10	{'max_dep 5, 'n_estima 10}
7	5.079081	0.013666	0.183004	0.001254	5	20	{'max_dep 5, 'n_estima 20}
8	11.736913	0.018062	0.215554	0.047162	5	50	{'max_dep 5, 'n_estima 50}
9	22.902870	0.021217	0.284173	0.000865	5	100	{'max_dep 5, 'n_estima 100}
10	45.379531	0.306231	0.488938	0.002861	5	200	{'max_dep 5, 'n_estima 200}
11	112.297792	0.338774	1.095100	0.000825	5	500	{'max_dep 5, 'n_estima 500}
12	3.617024	0.066220	0.184153	0.000527	7	10	{'max_dep 7, 'n_estima 10}
13	6.512194	0.025545	0.182949	0.001213	7	20	{'max_dep 7, 'n_estima 20}
14	15.284030	0.077638	0.283650	0.000309	7	50	{'max_dep 7, 'n_estima 50}
15	30.866552	1.198951	0.441183	0.078496	7	100	{'max_dep 7, 'n_estima 100}
16	59.199829	0.248937	0.590966	0.000706	7	200	{'max_dep 7, 'n_estima 200}
17	146.984344	0.561459	1.298243	0.001783	7	500	{'max_dep 7, 'n_estima 500}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	{'max_dep pa
18	4.397127	0.089600	0.183865	0.001106	10	10	10, 'n_estima 10}
19	8.116001	0.042157	0.182463	0.001228	10	20	{'max_dep 10, 'n_estima 20}
20	19.240777	0.164159	0.283287	0.000568	10	50	{'max_dep 10, 'n_estima 50}
21	38.011190	0.241738	0.485912	0.001115	10	100	{'max_dep 10, 'n_estima 100}
22	75.372448	0.522107	0.794573	0.003083	10	200	{'max_dep 10, 'n_estima 200}
23	187.433179	1.743006	1.665156	0.044979	10	500	{'max_dep 10, 'n_estima 500}
24	5.112669	0.117266	0.182720	0.000172	15	10	{'max_dep 15, 'n_estima 10}
25	9.646317	0.106730	0.181697	0.000308	15	20	{'max_dep 15, 'n_estima 20}
26	22.971080	0.259187	0.318797	0.047239	15	50	{'max_dep 15, 'n_estima 50}
27	45.310350	0.385102	0.524775	0.044164	15	100	{'max_dep 15, 'n_estima 100}
28	89.868462	1.248058	0.893397	0.000506	15	200	{'max_dep 15, 'n_estima 200}
29	223.636395	3.509909	2.106362	0.004747	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
clf = RandomForestClassifier(n_estimators= 200,max_depth=7, n_jobs= -1, verbose=1, class_weight= 'balanced')
# clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn,y_train_churn)
```

```
[Parallel(n_jobs=-1)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:   14.9s
[Parallel(n_jobs=-1)]: Done 196 tasks      | elapsed:  1.1min
[Parallel(n_jobs=-1)]: Done 200 out of 200 | elapsed:  1.1min finished
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=7, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=200,
                       n_jobs=-1, oob_score=False, random_state=None, verbose=1,
                       warm_start=False)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn)[:,1]
```

```
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.4s
[Parallel(n_jobs=2)]: Done 196 tasks      | elapsed:    1.5s
[Parallel(n_jobs=2)]: Done 200 out of 200 | elapsed:    1.5s finished
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.1s
[Parallel(n_jobs=2)]: Done 196 tasks      | elapsed:    0.4s
[Parallel(n_jobs=2)]: Done 200 out of 200 | elapsed:    0.4s finished
```

In [ ]:

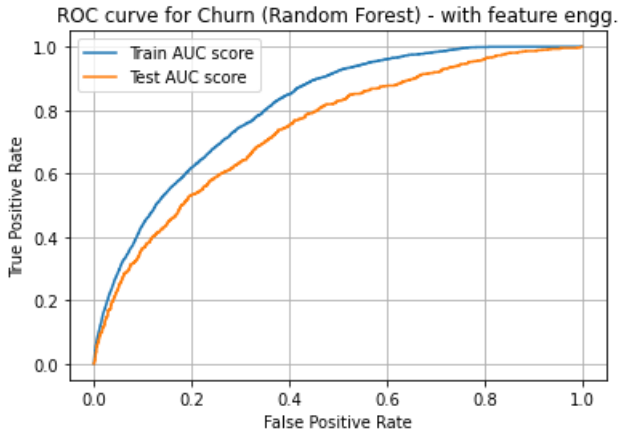
```
rf_train_auc_score_churn = roc_auc_score(y_train_churn, y_train_churn_pred)
rf_test_auc_score_churn = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Random Forest) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_churn_score = ['Random Forest (Churn)', rf_train_auc_score_churn, rf_test_auc_score_churn]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_churn_score)
```

In [ ]:

```
print(score_table)
```

```
+----------------------+--------------------+--------------------+
|         Model        |     Train AUC      |      Test AUC      |
+----------------------+--------------------+--------------------+
| Random Forest (Churn) | 0.8093789557791758 | 0.7410214954793564 |
+----------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_churn, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs=-1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,250,500], 'max_depth' : [1,2,3,4]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= True)
```

In [ ]:

```
classifier.fit(X_train_churn,y_train_churn)
```

```
Fitting 3 folds for each of 24 candidates, totalling 72 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  72 out of  72 | elapsed: 200.7min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=12.614703880190605,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=None,
```

```
            param_grid={'max_depth': [1, 2, 3, 4],
                        'n_estimators': [10, 20, 50, 100, 250, 500]},
            pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
            scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	10.672447	0.883436	0.400495	0.237986	1	10	{'max_dep 1, 'n_estima 10}
1	15.030084	0.027284	0.243297	0.012071	1	20	{'max_dep 1, 'n_estima 20}
2	30.019935	0.093895	0.249385	0.011356	1	50	{'max_dep 1, 'n_estima 50}
3	55.219683	0.030190	0.265350	0.014716	1	100	{'max_dep 1, 'n_estima 100}
4	130.436486	0.699106	0.280439	0.019627	1	250	{'max_dep 1, 'n_estima 250}
5	254.676241	0.820080	0.297604	0.014110	1	500	{'max_dep 1, 'n_estima 500}
6	13.976642	0.134332	0.254993	0.031300	2	10	{'max_dep 2, 'n_estima 10}
7	22.655317	0.062736	0.248956	0.010368	2	20	{'max_dep 2, 'n_estima 20}
8	48.766472	0.049187	0.260112	0.014468	2	50	{'max_dep 2, 'n_estima 50}
9	92.060174	0.193564	0.272114	0.014573	2	100	{'max_dep 2, 'n_estima 100}
10	221.124851	0.133734	0.290794	0.017784	2	250	{'max_dep 2, 'n_estima

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	250} pa
11	433.931621	0.953384	0.308113	0.016782	2	500	{'max_dep 2, 'n_estima 500}
12	17.945625	0.141139	0.249697	0.014662	3	10	{'max_dep 3, 'n_estima 10}
13	30.307209	0.131131	0.252653	0.015589	3	20	{'max_dep 3, 'n_estima 20}
14	67.979127	0.542777	0.263083	0.014127	3	50	{'max_dep 3, 'n_estima 50}
15	128.924748	1.056484	0.273283	0.016392	3	100	{'max_dep 3, 'n_estima 100}
16	309.020513	0.372798	0.292669	0.018180	3	250	{'max_dep 3, 'n_estima 250}
17	611.052146	1.399712	0.328404	0.018502	3	500	{'max_dep 3, 'n_estima 500}
18	22.003828	0.180153	0.245551	0.011659	4	10	{'max_dep 4, 'n_estima 10}
19	38.475384	0.237389	0.252831	0.016357	4	20	{'max_dep 4, 'n_estima 20}
20	86.498904	0.015125	0.264284	0.014941	4	50	{'max_dep 4, 'n_estima 50}
21	164.795103	0.867015	0.275436	0.014807	4	100	{'max_dep 4, 'n_estima 100}
22	399.184558	2.150260	0.315915	0.040691	4	250	{'max_dep 4, 'n_estima 250}
23	788.472999	2.299754	0.389472	0.057074	4	500	{'max_dep 4, 'n_estima 500}

In [ ]:

```
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn,y_train_churn)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=1,
              min_child_weight=1, missing=None, n_estimators=250, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=12.614703880190605,
              seed=None, silent=None, subsample=1, verbosity=1)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn)[:,1]
```

In [ ]:

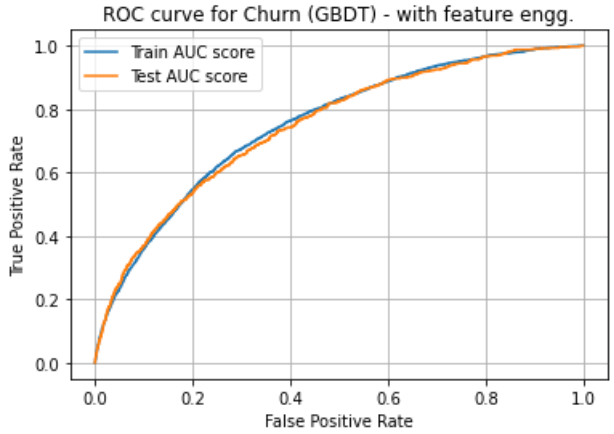
```
gbdt_train_auc_score_churn = roc_auc_score(y_train_churn, y_train_churn_pred)
gbdt_test_auc_score_churn = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (GBDT) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
gbdt_churn_score = ['GBDT (Churn)', gbdt_train_auc_score_churn, gbdt_test_auc_score_churn]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_churn_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------+--------------------+--------------------+
|    Model     |     Train AUC      |      Test AUC      |
+--------------+--------------------+--------------------+
| GBDT (Churn) | 0.7508474199893695 | 0.7469584240765366 |
+--------------+--------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_score)
score_table.add_row(rf_churn_score)
score_table.add_row(gbdt_churn_score)

print(score_table)
```

```
+----------------------------+--------------------+--------------------+
|           Model            |     Train AUC      |      Test AUC      |
+----------------------------+--------------------+--------------------+
| Logistic Regresion (Churn) | 0.7617842109476075 | 0.7156584883588942 |
|   Random Forest (Churn)    | 0.8093789557791758 | 0.7410214954793564 |
|        GBDT (Churn)        | 0.7508474199893695 | 0.7469584240765366 |
+----------------------------+--------------------+--------------------+
```

Observation:

- **Both Random Forest and GBDT are performing equally on Churn test dataset. However, Random Forest is a little overfitting.**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.1, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 200, max_depth= 7, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_churn, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 250, max_depth= 1, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.0001,0.001,0.01,0.1,1]}
```

In [ ]:

```
```

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_churn,y_train_churn)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits

[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 154.7min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
                                                                       eta0=0.0,
                                                                       fit_intercept=True,
                                                                       l1_ratio=0.15,
                                                                       learning_rate='optimal',
                                                                       loss='log',
                                                                       max_iter=1000,
                                                                       n_iter_no_change=5,
                                                                       n_jobs=-1,
                                                                       penalty='12',
                                                                       power_t=0.5,
                                                                       random_state=None,
                                                                       shuffle=True,
                                                                       t...
                                                                          validation_fraction=0.1,
                                                                          verbose=0,
                                                                          warm_start=False),
                                                    shuffle=True,
                                                    store_train_meta_features=False,
                                                    stratify=True, use_clones=True,
                                                    use_features_in_secondary=False,
                                                    use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.0001, 0.001, 0.01, 0.1,
                                                    1]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	604.985582	4.714214	1.078161	0.077233	0.0001	{'meta_classifier__al 0.0001}
1	637.654474	49.915948	1.092374	0.068677	0.001	{'meta_classifier__al 0.001}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	par
2	645.484345	34.640610	1.301048	0.187253	0.01	{'meta_classifier__al 0.01}
3	594.641006	1.866836	1.245446	0.050865	0.1	{'meta_classifier__al 0.1}
4	593.655848	4.691946	1.286672	0.419109	1	{'meta_classifier__al 1}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_churn,y_train_churn)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.1, average=False,
                                                class_weight='balanced',
                                                early_stopping=False,
                                                epsilon=0.1, eta0=0.0,
                                                fit_intercept=True,
                                                l1_ratio=0.15,
                                                learning_rate='optimal',
                                                loss='log', max_iter=1000,
                                                n_iter_no_change=5, n_jobs=-1,
                                                penalty='l2', power_t=0.5,
                                                random_state=None, shuffle=True,
                                                tol=0.001,
                                                validation_fraction=0.1,
                                                verbose=0, w...
                                                   fit_intercept=True,
                                                   l1_ratio=0.15,
                                                   learning_rate='optimal',
                                                   loss='log', max_iter=1000,
                                                   n_iter_no_change=5,
                                                   n_jobs=-1, penalty='l2',
                                                   power_t=0.5,
                                                   random_state=None,
                                                   shuffle=True, tol=0.001,
                                                   validation_fraction=0.1,
                                                   verbose=0,
                                                   warm_start=False),
                          shuffle=True, store_train_meta_features=False,
                          stratify=True, use_clones=True,
                          use_features_in_secondary=False, use_probas=True,
                          verbose=0)
```

In [ ]:

```
y_train_churn_pred = clf.predict_proba(X_train_churn)[:,1]
y_test_churn_pred = clf.predict_proba(X_test_churn)[:,1]
```

In [ ]:

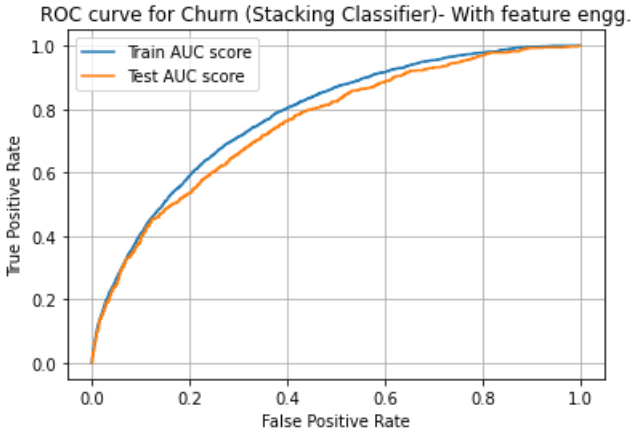
```
stack_train_auc_score_churn = roc_auc_score(y_train_churn, y_train_churn_pred)
stack_test_auc_score_churn = roc_auc_score(y_test_churn, y_test_churn_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_churn,y_train_churn_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_churn,y_test_churn_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Churn (Stacking Classifier)- With feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_churn_score = ['Stacking Classifier (churn)', stack_train_auc_score_churn, stack_test_auc_score_c
hurn]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(stack_churn_score)
```

In [ ]:

```
print(score_table)
```

```
+------------------------------+--------------------+--------------------+
|            Model             |     Train AUC      |      Test AUC      |
+------------------------------+--------------------+--------------------+
| Stacking Classifier (churn) | 0.7798586744365475 | 0.7527890485916988 |
+------------------------------+--------------------+--------------------+
```

Upselling

Logistic Regression

In [ ]:

```
sgd = SGDClassifier(loss = 'log',class_weight='balanced', verbose=1,random_state = 13)
```

In [ ]:

```
param_grid = {'alpha': [0.001,0.01,0.1,1,10,]}
```

In [ ]:

```
classifier = GridSearchCV(sgd, param_grid, scoring='roc_auc', n_jobs= -1, cv = 3, verbose=1, return_tra
in_score= True)
```

In [ ]:

```
classifier.fit(X_train_upselling, y_train_upselling)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  15 out of  15 | elapsed:  1.8min finished
```

```
-- Epoch 1
Norm: 3.99, NNZs: 2298, Bias: 6.682274, T: 40000, Avg. loss: 5.586482
Total training time: 0.33 seconds.
-- Epoch 2
Norm: 3.31, NNZs: 2298, Bias: 5.282280, T: 80000, Avg. loss: 1.210823
Total training time: 0.66 seconds.
-- Epoch 3
Norm: 2.96, NNZs: 2298, Bias: 4.574902, T: 120000, Avg. loss: 0.995503
Total training time: 0.98 seconds.
-- Epoch 4
Norm: 2.72, NNZs: 2298, Bias: 4.133747, T: 160000, Avg. loss: 0.901428
Total training time: 1.31 seconds.
-- Epoch 5
Norm: 2.56, NNZs: 2298, Bias: 3.811687, T: 200000, Avg. loss: 0.862150
Total training time: 1.64 seconds.
-- Epoch 6
Norm: 2.43, NNZs: 2298, Bias: 3.564883, T: 240000, Avg. loss: 0.824526
Total training time: 1.97 seconds.
-- Epoch 7
Norm: 2.33, NNZs: 2298, Bias: 3.369293, T: 280000, Avg. loss: 0.796398
Total training time: 2.30 seconds.
-- Epoch 8
Norm: 2.24, NNZs: 2298, Bias: 3.206629, T: 320000, Avg. loss: 0.778419
Total training time: 2.63 seconds.
-- Epoch 9
Norm: 2.17, NNZs: 2298, Bias: 3.068041, T: 360000, Avg. loss: 0.762984
Total training time: 2.96 seconds.
-- Epoch 10
Norm: 2.11, NNZs: 2298, Bias: 2.949136, T: 400000, Avg. loss: 0.748797
Total training time: 3.28 seconds.
-- Epoch 11
Norm: 2.05, NNZs: 2298, Bias: 2.844489, T: 440000, Avg. loss: 0.738467
Total training time: 3.62 seconds.
-- Epoch 12
Norm: 2.00, NNZs: 2298, Bias: 2.751517, T: 480000, Avg. loss: 0.729867
Total training time: 3.94 seconds.
-- Epoch 13
Norm: 1.96, NNZs: 2298, Bias: 2.668232, T: 520000, Avg. loss: 0.721407
Total training time: 4.27 seconds.
-- Epoch 14
Norm: 1.92, NNZs: 2298, Bias: 2.592354, T: 560000, Avg. loss: 0.715431
Total training time: 4.61 seconds.
-- Epoch 15
Norm: 1.89, NNZs: 2298, Bias: 2.524141, T: 600000, Avg. loss: 0.705614
Total training time: 4.93 seconds.
-- Epoch 16
Norm: 1.85, NNZs: 2298, Bias: 2.460776, T: 640000, Avg. loss: 0.703210
Total training time: 5.26 seconds.
-- Epoch 17
Norm: 1.82, NNZs: 2298, Bias: 2.402965, T: 680000, Avg. loss: 0.695040
Total training time: 5.59 seconds.
-- Epoch 18
Norm: 1.80, NNZs: 2298, Bias: 2.349455, T: 720000, Avg. loss: 0.690571
Total training time: 5.92 seconds.
-- Epoch 19
Norm: 1.77, NNZs: 2298, Bias: 2.299259, T: 760000, Avg. loss: 0.687680
Total training time: 6.24 seconds.
-- Epoch 20
Norm: 1.75, NNZs: 2298, Bias: 2.252697, T: 800000, Avg. loss: 0.682487
Total training time: 6.58 seconds.
-- Epoch 21
Norm: 1.72, NNZs: 2298, Bias: 2.209252, T: 840000, Avg. loss: 0.677731
Total training time: 6.91 seconds.
```

```
-- Epoch 22
Norm: 1.70, NNZs: 2298, Bias: 2.167949, T: 880000, Avg. loss: 0.676482
Total training time: 7.24 seconds.
-- Epoch 23
Norm: 1.68, NNZs: 2298, Bias: 2.129316, T: 920000, Avg. loss: 0.671920
Total training time: 7.58 seconds.
-- Epoch 24
Norm: 1.66, NNZs: 2298, Bias: 2.092608, T: 960000, Avg. loss: 0.669881
Total training time: 7.90 seconds.
-- Epoch 25
Norm: 1.65, NNZs: 2298, Bias: 2.057842, T: 1000000, Avg. loss: 0.666625
Total training time: 8.23 seconds.
-- Epoch 26
Norm: 1.63, NNZs: 2298, Bias: 2.025230, T: 1040000, Avg. loss: 0.662053
Total training time: 8.56 seconds.
-- Epoch 27
Norm: 1.61, NNZs: 2298, Bias: 1.993958, T: 1080000, Avg. loss: 0.660988
Total training time: 8.89 seconds.
-- Epoch 28
Norm: 1.60, NNZs: 2298, Bias: 1.964133, T: 1120000, Avg. loss: 0.658448
Total training time: 9.22 seconds.
-- Epoch 29
Norm: 1.58, NNZs: 2298, Bias: 1.935560, T: 1160000, Avg. loss: 0.656795
Total training time: 9.55 seconds.
-- Epoch 30
Norm: 1.57, NNZs: 2298, Bias: 1.908413, T: 1200000, Avg. loss: 0.653627
Total training time: 9.88 seconds.
-- Epoch 31
Norm: 1.56, NNZs: 2298, Bias: 1.882350, T: 1240000, Avg. loss: 0.651622
Total training time: 10.20 seconds.
-- Epoch 32
Norm: 1.54, NNZs: 2298, Bias: 1.857565, T: 1280000, Avg. loss: 0.648540
Total training time: 10.53 seconds.
-- Epoch 33
Norm: 1.53, NNZs: 2298, Bias: 1.833403, T: 1320000, Avg. loss: 0.649038
Total training time: 10.86 seconds.
-- Epoch 34
Norm: 1.52, NNZs: 2298, Bias: 1.810368, T: 1360000, Avg. loss: 0.646145
Total training time: 11.19 seconds.
-- Epoch 35
Norm: 1.51, NNZs: 2298, Bias: 1.788028, T: 1400000, Avg. loss: 0.645557
Total training time: 11.51 seconds.
-- Epoch 36
Norm: 1.50, NNZs: 2298, Bias: 1.766690, T: 1440000, Avg. loss: 0.642214
Total training time: 11.85 seconds.
-- Epoch 37
Norm: 1.49, NNZs: 2298, Bias: 1.745973, T: 1480000, Avg. loss: 0.641567
Total training time: 12.17 seconds.
-- Epoch 38
Norm: 1.48, NNZs: 2298, Bias: 1.726076, T: 1520000, Avg. loss: 0.639706
Total training time: 12.50 seconds.
-- Epoch 39
Norm: 1.47, NNZs: 2298, Bias: 1.706901, T: 1560000, Avg. loss: 0.637300
Total training time: 12.83 seconds.
-- Epoch 40
Norm: 1.46, NNZs: 2298, Bias: 1.688467, T: 1600000, Avg. loss: 0.635259
Total training time: 13.16 seconds.
-- Epoch 41
Norm: 1.45, NNZs: 2298, Bias: 1.670316, T: 1640000, Avg. loss: 0.636403
Total training time: 13.49 seconds.
-- Epoch 42
Norm: 1.44, NNZs: 2298, Bias: 1.652888, T: 1680000, Avg. loss: 0.633769
Total training time: 13.82 seconds.
-- Epoch 43
Norm: 1.43, NNZs: 2298, Bias: 1.635993, T: 1720000, Avg. loss: 0.632140
Total training time: 14.15 seconds.
-- Epoch 44
Norm: 1.43, NNZs: 2298, Bias: 1.619637, T: 1760000, Avg. loss: 0.630862
Total training time: 14.48 seconds.
-- Epoch 45
Norm: 1.42, NNZs: 2298, Bias: 1.603711, T: 1800000, Avg. loss: 0.630215
Total training time: 14.81 seconds.
-- Epoch 46
Norm: 1.41, NNZs: 2298, Bias: 1.588227, T: 1840000, Avg. loss: 0.628807
Total training time: 15.14 seconds.
-- Epoch 47
Norm: 1.40, NNZs: 2298, Bias: 1.573154, T: 1880000, Avg. loss: 0.627949
```

```
Total training time: 15.46 seconds.
-- Epoch 48
Norm: 1.40, NNZs: 2298, Bias: 1.558508, T: 1920000, Avg. loss: 0.626720
Total training time: 15.80 seconds.
-- Epoch 49
Norm: 1.39, NNZs: 2298, Bias: 1.544251, T: 1960000, Avg. loss: 0.625817
Total training time: 16.13 seconds.
-- Epoch 50
Norm: 1.38, NNZs: 2298, Bias: 1.530441, T: 2000000, Avg. loss: 0.623990
Total training time: 16.46 seconds.
-- Epoch 51
Norm: 1.38, NNZs: 2298, Bias: 1.516853, T: 2040000, Avg. loss: 0.624248
Total training time: 16.79 seconds.
-- Epoch 52
Norm: 1.37, NNZs: 2298, Bias: 1.503699, T: 2080000, Avg. loss: 0.622406
Total training time: 17.11 seconds.
-- Epoch 53
Norm: 1.36, NNZs: 2298, Bias: 1.490881, T: 2120000, Avg. loss: 0.621370
Total training time: 17.44 seconds.
-- Epoch 54
Norm: 1.36, NNZs: 2298, Bias: 1.478340, T: 2160000, Avg. loss: 0.620661
Total training time: 17.78 seconds.
-- Epoch 55
Norm: 1.35, NNZs: 2298, Bias: 1.466186, T: 2200000, Avg. loss: 0.618979
Total training time: 18.11 seconds.
-- Epoch 56
Norm: 1.35, NNZs: 2298, Bias: 1.454191, T: 2240000, Avg. loss: 0.619445
Total training time: 18.44 seconds.
-- Epoch 57
Norm: 1.34, NNZs: 2298, Bias: 1.442488, T: 2280000, Avg. loss: 0.618248
Total training time: 18.77 seconds.
-- Epoch 58
Norm: 1.34, NNZs: 2298, Bias: 1.431034, T: 2320000, Avg. loss: 0.617514
Total training time: 19.09 seconds.
-- Epoch 59
Norm: 1.33, NNZs: 2298, Bias: 1.419828, T: 2360000, Avg. loss: 0.616853
Total training time: 19.42 seconds.
-- Epoch 60
Norm: 1.32, NNZs: 2298, Bias: 1.408927, T: 2400000, Avg. loss: 0.615495
Total training time: 19.75 seconds.
-- Epoch 61
Norm: 1.32, NNZs: 2298, Bias: 1.398224, T: 2440000, Avg. loss: 0.614970
Total training time: 20.08 seconds.
-- Epoch 62
Norm: 1.31, NNZs: 2298, Bias: 1.387782, T: 2480000, Avg. loss: 0.613870
Total training time: 20.41 seconds.
-- Epoch 63
Norm: 1.31, NNZs: 2298, Bias: 1.377485, T: 2520000, Avg. loss: 0.613982
Total training time: 20.74 seconds.
-- Epoch 64
Norm: 1.31, NNZs: 2298, Bias: 1.367505, T: 2560000, Avg. loss: 0.612023
Total training time: 21.07 seconds.
-- Epoch 65
Norm: 1.30, NNZs: 2298, Bias: 1.357671, T: 2600000, Avg. loss: 0.611966
Total training time: 21.40 seconds.
-- Epoch 66
Norm: 1.30, NNZs: 2298, Bias: 1.348056, T: 2640000, Avg. loss: 0.610953
Total training time: 21.73 seconds.
-- Epoch 67
Norm: 1.29, NNZs: 2298, Bias: 1.338499, T: 2680000, Avg. loss: 0.611946
Total training time: 22.06 seconds.
-- Epoch 68
Norm: 1.29, NNZs: 2298, Bias: 1.329271, T: 2720000, Avg. loss: 0.609439
Total training time: 22.39 seconds.
-- Epoch 69
Norm: 1.28, NNZs: 2298, Bias: 1.320141, T: 2760000, Avg. loss: 0.609769
Total training time: 22.72 seconds.
-- Epoch 70
Norm: 1.28, NNZs: 2298, Bias: 1.311193, T: 2800000, Avg. loss: 0.608972
Total training time: 23.04 seconds.
-- Epoch 71
Norm: 1.27, NNZs: 2298, Bias: 1.302504, T: 2840000, Avg. loss: 0.607101
Total training time: 23.37 seconds.
-- Epoch 72
Norm: 1.27, NNZs: 2298, Bias: 1.293827, T: 2880000, Avg. loss: 0.608267
Total training time: 23.70 seconds.
-- Epoch 73
```

```
Norm: 1.27, NNZs: 2298, Bias: 1.285335, T: 2920000, Avg. loss: 0.607406
Total training time: 24.03 seconds.
-- Epoch 74
Norm: 1.26, NNZs: 2298, Bias: 1.277065, T: 2960000, Avg. loss: 0.605743
Total training time: 24.36 seconds.
-- Epoch 75
Norm: 1.26, NNZs: 2298, Bias: 1.268902, T: 3000000, Avg. loss: 0.605744
Total training time: 24.69 seconds.
-- Epoch 76
Norm: 1.26, NNZs: 2298, Bias: 1.260893, T: 3040000, Avg. loss: 0.605005
Total training time: 25.02 seconds.
-- Epoch 77
Norm: 1.25, NNZs: 2298, Bias: 1.253038, T: 3080000, Avg. loss: 0.604185
Total training time: 25.35 seconds.
-- Epoch 78
Norm: 1.25, NNZs: 2298, Bias: 1.245292, T: 3120000, Avg. loss: 0.604077
Total training time: 25.68 seconds.
-- Epoch 79
Norm: 1.24, NNZs: 2298, Bias: 1.237649, T: 3160000, Avg. loss: 0.603733
Total training time: 26.01 seconds.
Convergence after 79 epochs took 26.01 seconds
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=SGDClassifier(alpha=0.0001, average=False,
                                     class_weight='balanced',
                                     early_stopping=False, epsilon=0.1,
                                     eta0=0.0, fit_intercept=True,
                                     l1_ratio=0.15, learning_rate='optimal',
                                     loss='log', max_iter=1000,
                                     n_iter_no_change=5, n_jobs=None,
                                     penalty='l2', power_t=0.5, random_state=13,
                                     shuffle=True, tol=0.001,
                                     validation_fraction=0.1, verbose=1,
                                     warm_start=False),
             iid='deprecated', n_jobs=-1,
             param_grid={'alpha': [0.001, 0.01, 0.1, 1, 10]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test
0	32.321732	5.915073	0.079165	0.001856	0.001	{'alpha': 0.001}	0.758196	0.761336
1	16.046879	1.694171	0.071004	0.008779	0.01	{'alpha': 0.01}	0.789979	0.798309
2	8.873898	2.105626	0.072296	0.009965	0.1	{'alpha': 0.1}	0.802341	0.805826
3	7.061776	0.631439	0.071217	0.009563	1	{'alpha': 1}	0.798020	0.796515
4	2.920274	0.108986	0.085156	0.032400	10	{'alpha': 10}	0.764738	0.760448

In [ ]:

```
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling,y_train_upselling)
```

```
-- Epoch 1
Norm: 3.99, NNZs: 2298, Bias: 6.682274, T: 40000, Avg. loss: 5.586482
Total training time: 0.33 seconds.
-- Epoch 2
Norm: 3.31, NNZs: 2298, Bias: 5.282280, T: 80000, Avg. loss: 1.210823
Total training time: 0.66 seconds.
-- Epoch 3
Norm: 2.96, NNZs: 2298, Bias: 4.574902, T: 120000, Avg. loss: 0.995503
Total training time: 0.99 seconds.
-- Epoch 4
Norm: 2.72, NNZs: 2298, Bias: 4.133747, T: 160000, Avg. loss: 0.901428
Total training time: 1.33 seconds.
-- Epoch 5
Norm: 2.56, NNZs: 2298, Bias: 3.811687, T: 200000, Avg. loss: 0.862150
Total training time: 1.66 seconds.
-- Epoch 6
Norm: 2.43, NNZs: 2298, Bias: 3.564883, T: 240000, Avg. loss: 0.824526
Total training time: 1.99 seconds.
-- Epoch 7
Norm: 2.33, NNZs: 2298, Bias: 3.369293, T: 280000, Avg. loss: 0.796398
Total training time: 2.33 seconds.
-- Epoch 8
Norm: 2.24, NNZs: 2298, Bias: 3.206629, T: 320000, Avg. loss: 0.778419
Total training time: 2.66 seconds.
-- Epoch 9
Norm: 2.17, NNZs: 2298, Bias: 3.068041, T: 360000, Avg. loss: 0.762984
Total training time: 2.99 seconds.
-- Epoch 10
Norm: 2.11, NNZs: 2298, Bias: 2.949136, T: 400000, Avg. loss: 0.748797
Total training time: 3.32 seconds.
-- Epoch 11
Norm: 2.05, NNZs: 2298, Bias: 2.844489, T: 440000, Avg. loss: 0.738467
Total training time: 3.65 seconds.
-- Epoch 12
Norm: 2.00, NNZs: 2298, Bias: 2.751517, T: 480000, Avg. loss: 0.729867
Total training time: 3.98 seconds.
-- Epoch 13
Norm: 1.96, NNZs: 2298, Bias: 2.668232, T: 520000, Avg. loss: 0.721407
Total training time: 4.32 seconds.
-- Epoch 14
Norm: 1.92, NNZs: 2298, Bias: 2.592354, T: 560000, Avg. loss: 0.715431
Total training time: 4.64 seconds.
-- Epoch 15
Norm: 1.89, NNZs: 2298, Bias: 2.524141, T: 600000, Avg. loss: 0.705614
Total training time: 4.97 seconds.
-- Epoch 16
Norm: 1.85, NNZs: 2298, Bias: 2.460776, T: 640000, Avg. loss: 0.703210
Total training time: 5.31 seconds.
-- Epoch 17
Norm: 1.82, NNZs: 2298, Bias: 2.402965, T: 680000, Avg. loss: 0.695040
Total training time: 5.63 seconds.
-- Epoch 18
Norm: 1.80, NNZs: 2298, Bias: 2.349455, T: 720000, Avg. loss: 0.690571
Total training time: 5.96 seconds.
-- Epoch 19
Norm: 1.77, NNZs: 2298, Bias: 2.299259, T: 760000, Avg. loss: 0.687680
Total training time: 6.29 seconds.
-- Epoch 20
Norm: 1.75, NNZs: 2298, Bias: 2.252697, T: 800000, Avg. loss: 0.682487
Total training time: 6.62 seconds.
-- Epoch 21
Norm: 1.72, NNZs: 2298, Bias: 2.209252, T: 840000, Avg. loss: 0.677731
Total training time: 6.94 seconds.
-- Epoch 22
Norm: 1.70, NNZs: 2298, Bias: 2.167949, T: 880000, Avg. loss: 0.676482
Total training time: 7.28 seconds.
-- Epoch 23
Norm: 1.68, NNZs: 2298, Bias: 2.129316, T: 920000, Avg. loss: 0.671920
Total training time: 7.60 seconds.
-- Epoch 24
```

```
Norm: 1.66, NNZs: 2298, Bias: 2.092608, T: 960000, Avg. loss: 0.669881
Total training time: 7.93 seconds.
-- Epoch 25
Norm: 1.65, NNZs: 2298, Bias: 2.057842, T: 1000000, Avg. loss: 0.666625
Total training time: 8.27 seconds.
-- Epoch 26
Norm: 1.63, NNZs: 2298, Bias: 2.025230, T: 1040000, Avg. loss: 0.662053
Total training time: 8.59 seconds.
-- Epoch 27
Norm: 1.61, NNZs: 2298, Bias: 1.993958, T: 1080000, Avg. loss: 0.660988
Total training time: 8.92 seconds.
-- Epoch 28
Norm: 1.60, NNZs: 2298, Bias: 1.964133, T: 1120000, Avg. loss: 0.658448
Total training time: 9.25 seconds.
-- Epoch 29
Norm: 1.58, NNZs: 2298, Bias: 1.935560, T: 1160000, Avg. loss: 0.656795
Total training time: 9.58 seconds.
-- Epoch 30
Norm: 1.57, NNZs: 2298, Bias: 1.908413, T: 1200000, Avg. loss: 0.653627
Total training time: 9.90 seconds.
-- Epoch 31
Norm: 1.56, NNZs: 2298, Bias: 1.882350, T: 1240000, Avg. loss: 0.651622
Total training time: 10.23 seconds.
-- Epoch 32
Norm: 1.54, NNZs: 2298, Bias: 1.857565, T: 1280000, Avg. loss: 0.648540
Total training time: 10.56 seconds.
-- Epoch 33
Norm: 1.53, NNZs: 2298, Bias: 1.833403, T: 1320000, Avg. loss: 0.649038
Total training time: 10.89 seconds.
-- Epoch 34
Norm: 1.52, NNZs: 2298, Bias: 1.810368, T: 1360000, Avg. loss: 0.646145
Total training time: 11.22 seconds.
-- Epoch 35
Norm: 1.51, NNZs: 2298, Bias: 1.788028, T: 1400000, Avg. loss: 0.645557
Total training time: 11.56 seconds.
-- Epoch 36
Norm: 1.50, NNZs: 2298, Bias: 1.766690, T: 1440000, Avg. loss: 0.642214
Total training time: 11.89 seconds.
-- Epoch 37
Norm: 1.49, NNZs: 2298, Bias: 1.745973, T: 1480000, Avg. loss: 0.641567
Total training time: 12.22 seconds.
-- Epoch 38
Norm: 1.48, NNZs: 2298, Bias: 1.726076, T: 1520000, Avg. loss: 0.639706
Total training time: 12.55 seconds.
-- Epoch 39
Norm: 1.47, NNZs: 2298, Bias: 1.706901, T: 1560000, Avg. loss: 0.637300
Total training time: 12.88 seconds.
-- Epoch 40
Norm: 1.46, NNZs: 2298, Bias: 1.688467, T: 1600000, Avg. loss: 0.635259
Total training time: 13.20 seconds.
-- Epoch 41
Norm: 1.45, NNZs: 2298, Bias: 1.670316, T: 1640000, Avg. loss: 0.636403
Total training time: 13.53 seconds.
-- Epoch 42
Norm: 1.44, NNZs: 2298, Bias: 1.652888, T: 1680000, Avg. loss: 0.633769
Total training time: 13.86 seconds.
-- Epoch 43
Norm: 1.43, NNZs: 2298, Bias: 1.635993, T: 1720000, Avg. loss: 0.632140
Total training time: 14.19 seconds.
-- Epoch 44
Norm: 1.43, NNZs: 2298, Bias: 1.619637, T: 1760000, Avg. loss: 0.630862
Total training time: 14.52 seconds.
-- Epoch 45
Norm: 1.42, NNZs: 2298, Bias: 1.603711, T: 1800000, Avg. loss: 0.630215
Total training time: 14.84 seconds.
-- Epoch 46
Norm: 1.41, NNZs: 2298, Bias: 1.588227, T: 1840000, Avg. loss: 0.628807
Total training time: 15.17 seconds.
-- Epoch 47
Norm: 1.40, NNZs: 2298, Bias: 1.573154, T: 1880000, Avg. loss: 0.627949
Total training time: 15.50 seconds.
-- Epoch 48
Norm: 1.40, NNZs: 2298, Bias: 1.558508, T: 1920000, Avg. loss: 0.626720
Total training time: 15.83 seconds.
-- Epoch 49
Norm: 1.39, NNZs: 2298, Bias: 1.544251, T: 1960000, Avg. loss: 0.625817
Total training time: 16.16 seconds.
```

```
-- Epoch 50
Norm: 1.38, NNZs: 2298, Bias: 1.530441, T: 2000000, Avg. loss: 0.623990
Total training time: 16.49 seconds.
-- Epoch 51
Norm: 1.38, NNZs: 2298, Bias: 1.516853, T: 2040000, Avg. loss: 0.624248
Total training time: 16.82 seconds.
-- Epoch 52
Norm: 1.37, NNZs: 2298, Bias: 1.503699, T: 2080000, Avg. loss: 0.622406
Total training time: 17.14 seconds.
-- Epoch 53
Norm: 1.36, NNZs: 2298, Bias: 1.490881, T: 2120000, Avg. loss: 0.621370
Total training time: 17.47 seconds.
-- Epoch 54
Norm: 1.36, NNZs: 2298, Bias: 1.478340, T: 2160000, Avg. loss: 0.620661
Total training time: 17.80 seconds.
-- Epoch 55
Norm: 1.35, NNZs: 2298, Bias: 1.466186, T: 2200000, Avg. loss: 0.618979
Total training time: 18.12 seconds.
-- Epoch 56
Norm: 1.35, NNZs: 2298, Bias: 1.454191, T: 2240000, Avg. loss: 0.619445
Total training time: 18.45 seconds.
-- Epoch 57
Norm: 1.34, NNZs: 2298, Bias: 1.442488, T: 2280000, Avg. loss: 0.618248
Total training time: 18.78 seconds.
-- Epoch 58
Norm: 1.34, NNZs: 2298, Bias: 1.431034, T: 2320000, Avg. loss: 0.617514
Total training time: 19.11 seconds.
-- Epoch 59
Norm: 1.33, NNZs: 2298, Bias: 1.419828, T: 2360000, Avg. loss: 0.616853
Total training time: 19.44 seconds.
-- Epoch 60
Norm: 1.32, NNZs: 2298, Bias: 1.408927, T: 2400000, Avg. loss: 0.615495
Total training time: 19.76 seconds.
-- Epoch 61
Norm: 1.32, NNZs: 2298, Bias: 1.398224, T: 2440000, Avg. loss: 0.614970
Total training time: 20.09 seconds.
-- Epoch 62
Norm: 1.31, NNZs: 2298, Bias: 1.387782, T: 2480000, Avg. loss: 0.613870
Total training time: 20.42 seconds.
-- Epoch 63
Norm: 1.31, NNZs: 2298, Bias: 1.377485, T: 2520000, Avg. loss: 0.613982
Total training time: 20.75 seconds.
-- Epoch 64
Norm: 1.31, NNZs: 2298, Bias: 1.367505, T: 2560000, Avg. loss: 0.612023
Total training time: 21.07 seconds.
-- Epoch 65
Norm: 1.30, NNZs: 2298, Bias: 1.357671, T: 2600000, Avg. loss: 0.611966
Total training time: 21.40 seconds.
-- Epoch 66
Norm: 1.30, NNZs: 2298, Bias: 1.348056, T: 2640000, Avg. loss: 0.610953
Total training time: 21.73 seconds.
-- Epoch 67
Norm: 1.29, NNZs: 2298, Bias: 1.338499, T: 2680000, Avg. loss: 0.611946
Total training time: 22.05 seconds.
-- Epoch 68
Norm: 1.29, NNZs: 2298, Bias: 1.329271, T: 2720000, Avg. loss: 0.609439
Total training time: 22.38 seconds.
-- Epoch 69
Norm: 1.28, NNZs: 2298, Bias: 1.320141, T: 2760000, Avg. loss: 0.609769
Total training time: 22.71 seconds.
-- Epoch 70
Norm: 1.28, NNZs: 2298, Bias: 1.311193, T: 2800000, Avg. loss: 0.608972
Total training time: 23.04 seconds.
-- Epoch 71
Norm: 1.27, NNZs: 2298, Bias: 1.302504, T: 2840000, Avg. loss: 0.607101
Total training time: 23.36 seconds.
-- Epoch 72
Norm: 1.27, NNZs: 2298, Bias: 1.293827, T: 2880000, Avg. loss: 0.608267
Total training time: 23.69 seconds.
-- Epoch 73
Norm: 1.27, NNZs: 2298, Bias: 1.285335, T: 2920000, Avg. loss: 0.607406
Total training time: 24.02 seconds.
-- Epoch 74
Norm: 1.26, NNZs: 2298, Bias: 1.277065, T: 2960000, Avg. loss: 0.605743
Total training time: 24.34 seconds.
-- Epoch 75
Norm: 1.26, NNZs: 2298, Bias: 1.268902, T: 3000000, Avg. loss: 0.605744
```

```
Total training time: 24.67 seconds.
-- Epoch 76
Norm: 1.26, NNZs: 2298, Bias: 1.260893, T: 3040000, Avg. loss: 0.605005
Total training time: 25.00 seconds.
-- Epoch 77
Norm: 1.25, NNZs: 2298, Bias: 1.253038, T: 3080000, Avg. loss: 0.604185
Total training time: 25.32 seconds.
-- Epoch 78
Norm: 1.25, NNZs: 2298, Bias: 1.245292, T: 3120000, Avg. loss: 0.604077
Total training time: 25.65 seconds.
-- Epoch 79
Norm: 1.24, NNZs: 2298, Bias: 1.237649, T: 3160000, Avg. loss: 0.603733
Total training time: 25.98 seconds.
Convergence after 79 epochs took 25.98 seconds
```

Out[ ]:

```
SGDClassifier(alpha=0.1, average=False, class_weight='balanced',
              early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True,
              l1_ratio=0.15, learning_rate='optimal', loss='log', max_iter=1000,
              n_iter_no_change=5, n_jobs=None, penalty='l2', power_t=0.5,
              random_state=13, shuffle=True, tol=0.001, validation_fraction=0.1,
              verbose=1, warm_start=False)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling)[:,1]
```

In [ ]:

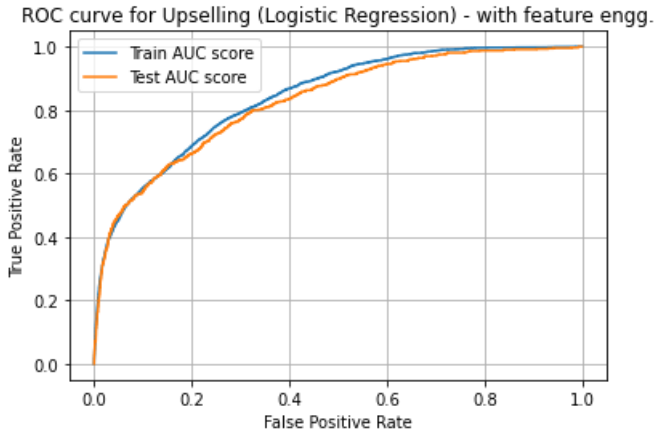
```
lr_train_auc_score_upselling = roc_auc_score(y_train_upselling, y_train_upselling_pred)
lr_test_auc_score_upselling = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling (Logistic Regression) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
lr_upselling_score = ['Logistic Regression (Upselling)', lr_train_auc_score_upselling, lr_test_auc_scor
e_upselling]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_score)
```

In [ ]:

```
print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logistic Regression (Upselling) | 0.8421514026005824 | 0.8277341403375384 |
+---------------------------------+--------------------+--------------------+
```

Random Forest

In [ ]:

```
clf = RandomForestClassifier(class_weight='balanced', n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,200,500], 'max_depth' : [3,5,7,10,15] }
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= Tr
ue)
```

In [ ]:

```
classifier.fit(X_train_upselling,y_train_upselling)
```

```
Fitting 3 folds for each of 30 candidates, totalling 90 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  90 out of  90 | elapsed: 66.4min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
                                              class_weight='balanced',
                                              criterion='gini', max_depth=None,
                                              max_features='auto',
                                              max_leaf_nodes=None,
                                              max_samples=None,
                                              min_impurity_decrease=0.0,
                                              min_impurity_split=None,
                                              min_samples_leaf=1,
                                              min_samples_split=2,
                                              min_weight_fraction_leaf=0.0,
                                              n_estimators=100, n_jobs=-1,
                                              oob_score=False,
                                              random_state=None, verbose=0,
                                              warm_start=False),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [3, 5, 7, 10, 15],
                         'n_estimators': [10, 20, 50, 100, 200, 500]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
```

```
                scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	2.444959	0.641424	0.197162	0.010282	3	10	{'max_dep 3, 'n_estima 10}
1	3.306671	0.064756	0.182341	0.002095	3	20	{'max_dep 3, 'n_estima 20}
2	7.418089	0.016896	0.183176	0.001400	3	50	{'max_dep 3, 'n_estima 50}
3	14.081718	0.090619	0.284850	0.002109	3	100	{'max_dep 3, 'n_estima 100}
4	29.128346	1.038089	0.461555	0.108696	3	200	{'max_dep 3, 'n_estima 200}
5	69.488508	0.140207	0.727324	0.050542	3	500	{'max_dep 3, 'n_estima 500}
6	2.760450	0.042927	0.183940	0.002671	5	10	{'max_dep 5, 'n_estima 10}
7	4.874164	0.040067	0.181667	0.000348	5	20	{'max_dep 5, 'n_estima 20}
8	11.365791	0.048163	0.183616	0.002582	5	50	{'max_dep 5, 'n_estima 50}
9	22.099782	0.073988	0.287698	0.004498	5	100	{'max_dep 5, 'n_estima 100}
10	44.436979	1.203356	0.487545	0.003310	5	200	{'max_dep 5, 'n_estima 200}
							{'max_dep

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
11	109.876870	1.363548	0.976868	0.060930	5	500	5, 'n_estima 500}
12	3.506583	0.043966	0.195413	0.019488	7	10	{'max_dep 7, 'n_estima 10}
13	6.251997	0.068199	0.181583	0.000676	7	20	{'max_dep 7, 'n_estima 20}
14	15.177426	0.512120	0.294970	0.013931	7	50	{'max_dep 7, 'n_estima 50}
15	29.084190	0.143730	0.386971	0.003515	7	100	{'max_dep 7, 'n_estima 100}
16	57.484583	0.545145	0.525488	0.045841	7	200	{'max_dep 7, 'n_estima 200}
17	142.336437	0.677838	1.128318	0.048087	7	500	{'max_dep 7, 'n_estima 500}
18	4.180973	0.076147	0.184982	0.004394	10	10	{'max_dep 10, 'n_estima 10}
19	7.739408	0.096961	0.184123	0.004619	10	20	{'max_dep 10, 'n_estima 20}
20	18.608244	0.226279	0.281379	0.001941	10	50	{'max_dep 10, 'n_estima 50}
21	36.860946	0.443662	0.386345	0.004465	10	100	{'max_dep 10, 'n_estima 100}
22	73.180238	0.711513	0.689272	0.000530	10	200	{'max_dep 10, 'n_estima 200}
23	181.977756	2.184308	1.428785	0.046902	10	500	{'max_dep 10, 'n_estima 500}
24	5.042654	0.154657	0.182693	0.001772	15	10	{'max_dep 15, 'n_estima 10}
25	9.490875	0.139593	0.181748	0.001451	15	20	{'max_dep 15, 'n_estima 20}

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	{'max_dep 15, 'n_estima 50}
26	22.689226	0.368734	0.287093	0.002540	15	50	
27	44.847532	0.754033	0.488069	0.000443	15	100	{'max_dep 15, 'n_estima 100}
28	88.652604	0.972435	0.792469	0.002037	15	200	{'max_dep 15, 'n_estima 200}
29	221.105839	2.111766	1.783049	0.024875	15	500	{'max_dep 15, 'n_estima 500}

In [ ]:

```
clf = RandomForestClassifier(n_estimators= 100,max_depth= 7, n_jobs= -1, verbose=1, class_weight= 'bala
nced')
# clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling,y_train_upselling)
```

```
[Parallel(n_jobs=-1)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  46 tasks      | elapsed:   18.5s
[Parallel(n_jobs=-1)]: Done 100 out of 100 | elapsed:   39.7s finished
```

Out[ ]:

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight='balanced',
                       criterion='gini', max_depth=7, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=100,
                       n_jobs=-1, oob_score=False, random_state=None, verbose=1,
                       warm_start=False)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling)[:,1]
```

```
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.3s
[Parallel(n_jobs=2)]: Done 100 out of 100 | elapsed:    0.6s finished
[Parallel(n_jobs=2)]: Using backend ThreadingBackend with 2 concurrent workers.
[Parallel(n_jobs=2)]: Done  46 tasks      | elapsed:    0.1s
[Parallel(n_jobs=2)]: Done 100 out of 100 | elapsed:    0.2s finished
```

In [ ]:

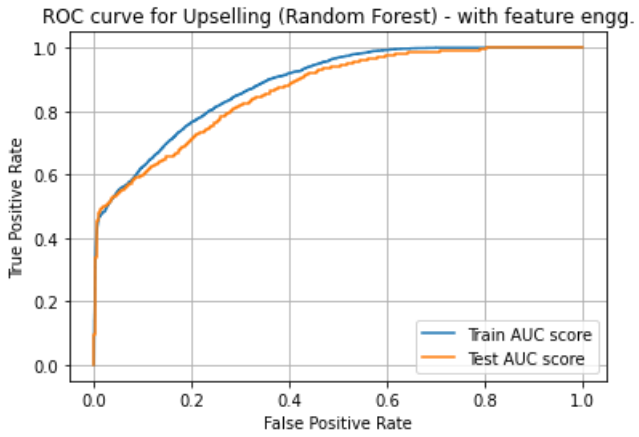
```
rf_train_auc_score_upselling = roc_auc_score(y_train_upselling, y_train_upselling_pred)
rf_test_auc_score_upselling = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling (Random Forest) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
rf_upselling_score = ['Random Forest (Upselling)', rf_train_auc_score_upselling, rf_test_auc_score_upse
lling]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(rf_upselling_score)
```

In [ ]:

```
print(score_table)
```

```
+---------------------------+--------------------+--------------------+
|           Model           |     Train AUC      |      Test AUC      |
+---------------------------+--------------------+--------------------+
| Random Forest (Upselling) | 0.8852810952298714 | 0.8648483552508073 |
+---------------------------+--------------------+--------------------+
```

GBDT

In [ ]:

```
neg, pos = np.unique(y_train_upselling, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf = XGBClassifier(scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
param_grid = {'n_estimators': [10,20,50,100,300], 'max_depth' : [1,2,3]}
```

In [ ]:

```
classifier = GridSearchCV(clf, param_grid, scoring='roc_auc', cv = 3, verbose=1, return_train_score= True)
```

In [ ]:

```
classifier.fit(X_train_upselling,y_train_upselling)
```

```
Fitting 3 folds for each of 15 candidates, totalling 45 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  45 out of  45 | elapsed: 51.4min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=XGBClassifier(base_score=0.5, booster='gbtree',
                                     colsample_bylevel=1, colsample_bynode=1,
                                     colsample_bytree=1, gamma=0,
                                     learning_rate=0.1, max_delta_step=0,
                                     max_depth=3, min_child_weight=1,
                                     missing=None, n_estimators=100, n_jobs=-1,
                                     nthread=None, objective='binary:logistic',
                                     random_state=0, reg_alpha=0, reg_lambda=1,
                                     scale_pos_weight=12.577732518669382,
                                     seed=None, silent=None, subsample=1,
                                     verbosity=1),
             iid='deprecated', n_jobs=None,
             param_grid={'max_depth': [1, 2, 3],
                         'n_estimators': [10, 20, 50, 100, 300]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(classifier.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	pa
0	7.585863	0.343870	0.265843	0.025938	1	10	{'max_dep 1, 'n_estima 10}
1	11.210710	0.028899	0.234933	0.001797	1	20	{'max_dep 1, 'n_estima 20}
2	22.806027	0.060801	0.241112	0.008489	1	50	{'max_dep 1, 'n_estima 50}
3	42.253057	0.158778	0.242062	0.002682	1	100	{'max_dep 1, 'n_estima 100}
							{'max_dep 1

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_max_depth	param_n_estimators	
4	120.644930	0.777198	0.318584	0.039943	1	300	'n_estima 300}
5	10.103747	0.045827	0.235885	0.004445	2	10	{'max_dep 2, 'n_estima 10}
6	16.645641	0.055360	0.236648	0.004839	2	20	{'max_dep 2, 'n_estima 20}
7	36.717714	0.145200	0.241906	0.001692	2	50	{'max_dep 2, 'n_estima 50}
8	70.759754	0.837103	0.274778	0.033302	2	100	{'max_dep 2, 'n_estima 100}
9	202.782579	1.526075	0.364835	0.120024	2	300	{'max_dep 2, 'n_estima 300}
10	13.572557	0.408581	0.247615	0.002847	3	10	{'max_dep 3, 'n_estima 10}
11	22.739791	0.106026	0.250511	0.004578	3	20	{'max_dep 3, 'n_estima 20}
12	51.061770	0.297573	0.261198	0.006378	3	50	{'max_dep 3, 'n_estima 50}
13	98.755730	0.637752	0.386418	0.163581	3	100	{'max_dep 3, 'n_estima 100}
14	286.812290	0.775812	0.403382	0.116631	3	300	{'max_dep 3, 'n_estima 300}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = classifier.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling,y_train_upselling)
```

Out[ ]:

```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0,
              learning_rate=0.1, max_delta_step=0, max_depth=2,
              min_child_weight=1, missing=None, n_estimators=100, n_jobs=-1,
              nthread=None, objective='binary:logistic', random_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=12.577732518669382,
              seed=None, silent=None, subsample=1, verbosity=1)
```

seed=None, silent=None, subsample=1, verbosity=1)

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling)[:,1]
```

In [ ]:

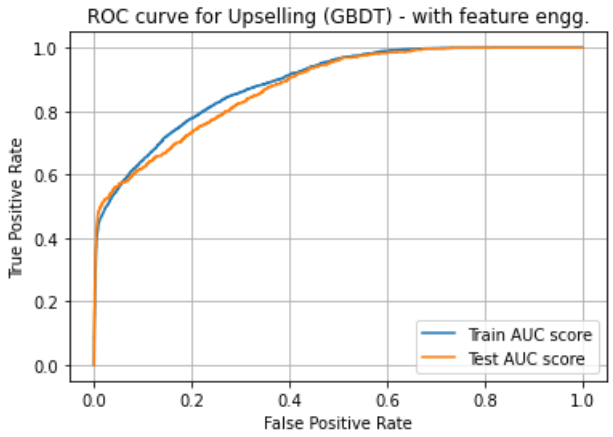
```
gbdt_train_auc_score_upselling = roc_auc_score(y_train_upselling, y_train_upselling_pred)
gbdt_test_auc_score_upselling = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
plt.title('ROC curve for Upselling (GBDT) - with feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
gbdt_upselling_score = ['GBDT (Upselling)', gbdt_train_auc_score_upselling, gbdt_test_auc_score_upselli
ng]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(gbdt_upselling_score)
```

In [ ]:

```
print(score_table)
```

```
+------------------+-------------------+--------------------+
|      Model       |     Train AUC     |      Test AUC      |
+------------------+-------------------+--------------------+
| GBDT (Upselling) | 0.888045120548638 | 0.8769405118926935 |
+------------------+-------------------+--------------------+
```

In [ ]:

```
lr_upselling_score = ['Logisitic Regression (Upselling)', '0.8421514026005824', '0.8277341403375384']
rf_upselling_score = ['Random Forest (Churn)', '0.8852810952298714' ,'0.8648483552508073']
```

In [ ]:

```
lr_upselling_score
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_score)
score_table.add_row(rf_upselling_score)
score_table.add_row(gbdt_upselling_score)

print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logisitic Regression (Upselling) | 0.8421514026005824 | 0.8277341403375384 |
|      Random Forest (Churn)      | 0.8852810952298714 | 0.8648483552508073 |
|         GBDT (Upselling)        | 0.888045120548638  | 0.8769405118926935 |
+---------------------------------+--------------------+--------------------+
```

Observation:

- **Out of all the models, GBDT is performing well on Upselling dataset**

Stacking Classifier

In [ ]:

```
clf1 = SGDClassifier(loss = 'log', alpha = 0.1, n_jobs= -1, class_weight= 'balanced')
```

In [ ]:

```
clf2 = RandomForestClassifier(n_estimators= 100, max_depth= 7, n_jobs= -1, class_weight= 'balanced',)
```

In [ ]:

```
neg, pos = np.unique(y_train_upselling, return_counts=True)[1]
weights = neg/pos
```

In [ ]:

```
clf3 = XGBClassifier(n_estimators= 100, max_depth= 2, scale_pos_weight= weights, n_jobs= -1)
```

In [ ]:

```
classifiers = [clf1, clf2, clf3]
```

In [ ]:

```
params = {"meta_classifier__alpha": [0.0001,0.001,0.01,0.1,1]}
```

In [ ]:

```
stack_classifier = StackingCVClassifier(classifiers, meta_classifier= SGDClassifier(loss = 'log', class
_weight= 'balanced', n_jobs=-1), use_probas= True, cv=3, stratify= True )
```

In [ ]:

```
gridcv = GridSearchCV(stack_classifier, params, scoring= 'roc_auc', cv =3,verbose =1,return_train_score
=True)
```

In [ ]:

```
gridcv.fit(X_train_upselling,y_train_upselling)
```

```
Fitting 3 folds for each of 5 candidates, totalling 15 fits
```

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done  15 out of  15 | elapsed: 92.2min finished
```

Out[ ]:

```
GridSearchCV(cv=3, error_score=nan,
             estimator=StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.01,
                                                                       average=False,
                                                                       class_weight='balanced',
                                                                       early_stopping=False,
                                                                       epsilon=0.1,
                                                                       eta0=0.0,
                                                                       fit_intercept=True,
                                                                       l1_ratio=0.15,
                                                                       learning_rate='optimal',
                                                                       loss='log',
                                                                       max_iter=1000,
                                                                       n_iter_no_change=5,
                                                                       n_jobs=-1,
                                                                       penalty='12',
                                                                       power_t=0.5,
                                                                       random_state=None,
                                                                       shuffle=True,...
                                                                          validation_fraction=0.1,
                                                                          verbose=0,
                                                                          warm_start=False),
                                                    shuffle=True,
                                                    store_train_meta_features=False,
                                                    stratify=True, use_clones=True,
                                                    use_features_in_secondary=False,
                                                    use_probas=True, verbose=0),
             iid='deprecated', n_jobs=None,
             param_grid={'meta_classifier__alpha': [0.0001, 0.001, 0.01, 0.1,
                                                    1]},
             pre_dispatch='2*n_jobs', refit=True, return_train_score=True,
             scoring='roc_auc', verbose=1)
```

In [ ]:

```
results = pd.DataFrame.from_dict(gridcv.cv_results_)
```

In [ ]:

```
results
```

Out[ ]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	pa
0	398.460164	2.025761	0.825692	0.080265	0.0001	{'meta_classifier__al 0.0001}
1	365.705780	9.344722	0.680506	0.088611	0.001	{'meta_classifier__al 0.001}
2	355.591579	3.434394	0.624763	0.022040	0.01	{'meta_classifier__al 0.01}
3	356.322114	1.594288	0.632244	0.021618	0.1	{'meta_classifier__al

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_meta_classifier__alpha	0.1} pa
4	357.643215	5.325836	0.665074	0.048263	1	{'meta_classifier__al 1}

In [ ]:

```
# clf = RandomForestClassifier(n_estimators= 500, n_jobs= -1, verbose=1, class_weight= 'balanced')
clf = gridcv.best_estimator_
```

In [ ]:

```
clf.fit(X_train_upselling,y_train_upselling)
```

Out[ ]:

```
StackingCVClassifier(classifiers=[SGDClassifier(alpha=0.01, average=False,
                                                class_weight='balanced',
                                                early_stopping=False,
                                                epsilon=0.1, eta0=0.0,
                                                fit_intercept=True,
                                                l1_ratio=0.15,
                                                learning_rate='optimal',
                                                loss='log', max_iter=1000,
                                                n_iter_no_change=5, n_jobs=-1,
                                                penalty='l2', power_t=0.5,
                                                random_state=None, shuffle=True,
                                                tol=0.001,
                                                validation_fraction=0.1,
                                                verbose=0,...
                                                   fit_intercept=True,
                                                   l1_ratio=0.15,
                                                   learning_rate='optimal',
                                                   loss='log', max_iter=1000,
                                                   n_iter_no_change=5,
                                                   n_jobs=-1, penalty='l2',
                                                   power_t=0.5,
                                                   random_state=None,
                                                   shuffle=True, tol=0.001,
                                                   validation_fraction=0.1,
                                                   verbose=0,
                                                   warm_start=False),
                          shuffle=True, store_train_meta_features=False,
                          stratify=True, use_clones=True,
                          use_features_in_secondary=False, use_probas=True,
                          verbose=0)
```

In [ ]:

```
y_train_upselling_pred = clf.predict_proba(X_train_upselling)[:,1]
y_test_upselling_pred = clf.predict_proba(X_test_upselling)[:,1]
```

In [ ]:

```
stack_train_auc_score_upselling = roc_auc_score(y_train_upselling, y_train_upselling_pred)
stack_test_auc_score_upselling = roc_auc_score(y_test_upselling, y_test_upselling_pred)
```

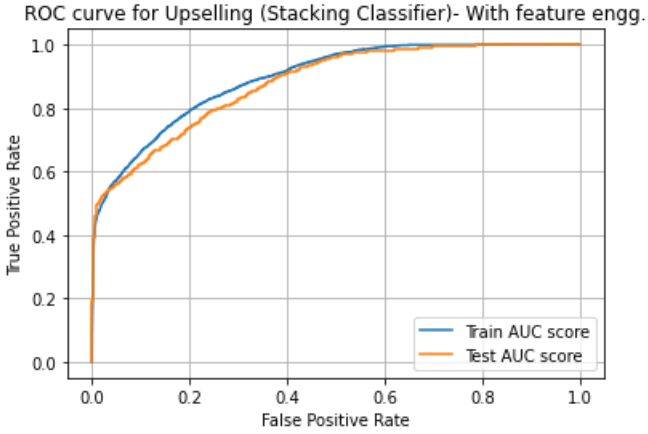
In [ ]:

```
tr_fpr, tr_tpr, _ = roc_curve(y_train_upselling,y_train_upselling_pred)
te_fpr, te_tpr, _ = roc_curve(y_test_upselling,y_test_upselling_pred)
```

In [ ]:

```
plt.plot(tr_fpr, tr_tpr, label='Train AUC score')
plt.plot(te_fpr, te_tpr, label='Test AUC score')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.legend()
```

```
plt.title('ROC curve for Upselling (Stacking Classifier)- With feature engg.')
plt.grid()
plt.show()
```



In [ ]:

```
columns = ['Model', 'Train AUC', 'Test AUC']
stack_upselling_score = ['Stacking Classifier (upselling)', stack_train_auc_score_upselling, stack_test
_auc_score_upselling]
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(stack_upselling_score)
```

In [ ]:

```
print(score_table)
```

```
+--------------------------------+--------------------+--------------------+
|              Model             |     Train AUC      |      Test AUC      |
+--------------------------------+--------------------+--------------------+
| Stacking Classifier (upselling) | 0.8938367816134917 | 0.8780133593339341 |
+--------------------------------+--------------------+--------------------+
```

Finding Best Model

Appetency

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_vanilla_score)
score_table.add_row(rf_appetency_vanilla_score)
score_table.add_row(gbdt_appetency_vanilla_score)

print(score_table)
```

```
+---------------------------------------+--------------------+--------------------+
|                 Model                 |     Train AUC      |      Test AUC      |
+---------------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency Vanilla) | 0.8004914800871239 | 0.8022376961601908 |
|    Random Forest (Appetency Vanilla)    | 0.8760928373360423 | 0.8213869803856969 |
|         GBDT (Appetency Vanilla)        | 0.8614878188496815 | 0.8419659260682851 |
+---------------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
| Stacking Classifier (Appetency Vanilla) | 0.8689946497086338 | 0.8438709020566075 |
+-----------------------------------------+--------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_appetency_score)
score_table.add_row(rf_appetency_score)
score_table.add_row(gbdt_appetency_score)

print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logistic Regression (Appetency) | 0.8866793102619893 | 0.8190018280448158 |
|    Random Forest (Appetency)    | 0.8883534212350627 | 0.8459632011604309 |
|         GBDT (Appetency)        | 0.863205203607357  | 0.8496999398278114 |
+---------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Stacking Classifier (apptency)  | 0.8777009204857703 | 0.8510881328089429 |
+---------------------------------+--------------------+--------------------+
```

Observation:

- **Stacking Classifier (with FE) is giving the best results on Appetency.**

Churn

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_vanilla_score)
score_table.add_row(rf_churn_vanilla_score)
score_table.add_row(gbdt_churn_vanilla_score)

print(score_table)
```

```
+-----------------------------------+--------------------+--------------------+
|               Model               |     Train AUC      |      Test AUC      |
+-----------------------------------+--------------------+--------------------+
| Logistic Regression (Churn Vanilla) | 0.6842109623015381 | 0.6967547407503685 |
|    Random Forest (Churn Vanilla)    | 0.7946363330829114 | 0.7351618321589403 |
|         GBDT (Churn Vanilla)        | 0.7486085630837572 |  0.75076905048547  |
+-----------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------+--------------------+--------------------+
|               Model               |     Train AUC      |      Test AUC      |
+-----------------------------------+--------------------+--------------------+
| Stacking Classifier (churn Vanilla) | 0.7569440059210848 | 0.7516441109891073 |
+-----------------------------------+--------------------+--------------------+
```

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_churn_score)
score_table.add_row(rf_churn_score)
score_table.add_row(gbdt_churn_score)

print(score_table)
```

```
+---------------------------+--------------------+--------------------+
|           Model           |     Train AUC      |      Test AUC      |
+---------------------------+--------------------+--------------------+
| Logistic Regresion (Churn) | 0.7617842109476075 | 0.7156584883588942 |
|   Random Forest (Churn)    | 0.8093789557791758 | 0.7410214954793564 |
|        GBDT (Churn)        | 0.7508474199893695 | 0.7469584240765366 |
+---------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+----------------------------+--------------------+--------------------+
|           Model            |     Train AUC      |      Test AUC      |
+----------------------------+--------------------+--------------------+
| Stacking Classifier (churn) | 0.7798586744365475 | 0.7527890485916988 |
+----------------------------+--------------------+--------------------+
```

Observation:

- **Stacking Classifier (with FE) is performing best (0.752 Test AUC) on Churn.**

Upselling

In [ ]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_vanilla_score)
score_table.add_row(rf_upselling_vanilla_score)
score_table.add_row(gbdt_upselling_vanilla_score)

print(score_table)
```

```
+----------------------------------------+--------------------+--------------------+
|                 Model                  |     Train AUC      |      Test AUC      |
+----------------------------------------+--------------------+--------------------+
| Logistic Regression (Upselling Vanilla) | 0.8049619771089853 | 0.8160040972065781 |
|    Random Forest (Upselling Vanilla)    | 0.8049619771089853 | 0.8160040972065781 |
|         GBDT (Upselling Vanilla)        | 0.8779311865389684 | 0.8797164954804011 |
+----------------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+----------------------------------------+--------------------+--------------------+
|                 Model                  |     Train AUC      |      Test AUC      |
+----------------------------------------+--------------------+--------------------+
| Stacking Classifier (upselling Vanilla) | 0.8842637363329957 | 0.8795341187485919 |
```

```
+----------------------------------------+--------------------+--------------------+
```

In [ ]:

```
lr_upselling_score
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(lr_upselling_score)
score_table.add_row(rf_upselling_score)
score_table.add_row(gbdt_upselling_score)

print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Logisitic Regression (Upselling) | 0.8421514026005824 | 0.8277341403375384 |
|      Random Forest (Churn)      | 0.8852810952298714 | 0.8648483552508073 |
|         GBDT (Upselling)        | 0.888045120548638  | 0.8769405118926935 |
+---------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+---------------------------------+--------------------+--------------------+
|              Model              |     Train AUC      |      Test AUC      |
+---------------------------------+--------------------+--------------------+
| Stacking Classifier (upselling) | 0.8938367816134917 | 0.8780133593339341 |
+---------------------------------+--------------------+--------------------+
```

Observation:

- **Stacking Classifier (Vanilla) is giving best Test AUC score (0.879) on Upselling**

Mean AUC Score

In [ ]:

```
print(score_table)
```

```
+--------------------------------+--------------------+--------------------+
|             Model              |     Train AUC      |      Test AUC      |
+--------------------------------+--------------------+--------------------+
| Stacking Classifier (apptency) | 0.8777009204857703 | 0.8510881328089429 |
+--------------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------+--------------------+--------------------+
|            Model            |     Train AUC      |      Test AUC      |
+-----------------------------+--------------------+--------------------+
| Stacking Classifier (churn) | 0.7798586744365475 | 0.7527890485916988 |
+-----------------------------+--------------------+--------------------+
```

In [ ]:

```
print(score_table)
```

```
+-----------------------------------------+--------------------+--------------------+
|                  Model                  |     Train AUC      |      Test AUC      |
+-----------------------------------------+--------------------+--------------------+
```

```
| Stacking Classifier (upselling Vanilla) | 0.8842637363329957 | 0.8795341187485919 |
+-----------------------------------------+--------------------+--------------------+
```

In [1]:

```
stack_train_auc_score_appetency, stack_test_auc_score_appetency = 0.8777009204857703, 0.851088132808942
9
stack_train_auc_score_churn, stack_test_auc_score_churn = 0.7798586744365475 , 0.7527890485916988
stack_train_auc_score_upselling_vanilla, stack_test_auc_score_upselling_vanilla =  0.8842637363329957,
0.8795341187485919
```

In [2]:

```
mean_train_auc_score = (stack_train_auc_score_appetency + stack_train_auc_score_churn + stack_train_auc
_score_upselling_vanilla)/3
mean_test_auc_score = (stack_test_auc_score_appetency + stack_test_auc_score_churn + stack_test_auc_sco
re_upselling_vanilla)/3
```

In [4]:

```
columns = ['Model', 'Avg Train AUC', 'Avg Test AUC']
mean_auc = ['Stacking Classifier', mean_train_auc_score, mean_test_auc_score]
```

In [7]:

```
score_table = PrettyTable()
score_table.field_names = columns
score_table.add_row(mean_auc)
print(score_table)
```

```
+---------------------+--------------------+--------------------+
|        Model        |   Avg Train AUC    |    Avg Test AUC    |
+---------------------+--------------------+--------------------+
| Stacking Classifier | 0.8472744437517713 | 0.8278037667164112 |
+---------------------+--------------------+--------------------+
```

Observation:

- **Avg Train AUC - 0.847**
- **Avg Test AUC - 0.827**

List of all the experiments that did and did not go well

- **Class based mean imputation and response encoding**
  - **It didn't perform well on test dataset.**
- **Mean imputation and frequency encoding**
  - **It gave better results on both train as well as test dataset.**
- **For feature engg.:**
  - **Count of NaNs for a given datapoint**
  - **Binary indicator for each feature indicating NaN**
  - **Cluster label**
  - **Average value of each feature group (Feature group 50, Feature group 200).**
  - **Average value per category for each categorical feature.**
  - **Polynomial features for numerical features.**

The feature engg dataset is performing well in most cases as compared to vanilla (without feature engg.) dataset

- **SMOTE + undersampling**
  - **We didn't continue with this approach as:**
    - **resampling needs to be done before feature engg.**
    - **SMOTE doesn't accept NaNs in data.**

> **So, in order to resample using SMOTE, we'll have to impute before feature engg**
> **meaning we'll lose 2 features (Count of NaN, Binary indicator)**

meaning we'll lose 2 features (Count of NaN, Binary indicator)

- **MLPClassifier**
  - **It wasn't giving results as good as traditional ML models.**
- **Stacking Classifier**
  - **Stacking classifier is performing as good as or better than other traditional ML models.**