

Spam Detection HW

Read complete instructions before starting the HW

Q1: Load the dataset (1 Point)

- For this Hw you will use spam dataset from kaggle which can be found from [this link](#). You can download this data and either upload it in google drive or in colab workspace. Load the data in pandas dataframe.
- There are only two useful columns. These columns are related to (1) label (ham and spam) and the (2) text of email.
- Rename columns as label and message
- Find the % ham and spam in the data.

```
In [ ]: import sys
if 'google.colab' in str(get_ipython()):
    from google.colab import files

    uploaded = files.upload()

    for fn in uploaded.keys():
        print('User uploaded file "{name}" with length {length} bytes'.format(
            name=fn, length=len(uploaded[fn])))

# Then move kaggle.json into the folder where the API expects to find it.

!mkdir -p ~/.kaggle/ && mv kaggle.json ~/.kaggle/ && chmod 600 ~/.kaggle
!kaggle datasets download "uciml/sms-spam-collection-dataset"
!unzip sms-spam-collection-dataset.zip
```

```
In [ ]: import pandas as pd
if 'google.colab' in str(get_ipython()):
    spam_df = pd.read_csv('spam.csv', encoding='latin-1')
else:
    spam_df = pd.read_csv('/Users/harikrishnadev/Library/CloudStorage/Google
spam_df = spam_df[['v1', 'v2']]
spam_df.columns = ['label', 'message']
spam_df.head()
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/1337418617.py:1: DeprecationWarning:
Pyarrow will become a required dependency of pandas in the next major release of pandas (pandas 3.0),
(to allow more performant data types, such as the Arrow string type, and better interoperability with other libraries)
but was not found to be installed on your system.
If this would cause problems for you,
please provide us feedback at https://github.com/pandas-dev/pandas/issues/54466

import pandas as pd
```

```
Out[ ]:      label      message
0    ham  Go until jurong point, crazy.. Available only ...
1    ham                Ok lar... Joking wif u oni...
2  spam  Free entry in 2 a wkly comp to win FA Cup fina...
3    ham  U dun say so early hor... U c already then say...
4    ham  Nah I don't think he goes to usf, he lives aro...
```

```
In [ ]: spam_df['label'].value_counts()
```

```
Out[ ]: label
ham      4825
spam      747
Name: count, dtype: int64
```

```
In [ ]: spam_df['label'].value_counts(normalize=True)*100
```

```
Out[ ]: label
ham      86.593683
spam     13.406317
Name: proportion, dtype: float64
```

```
In [ ]: spam_df['label'].replace({'ham':0, 'spam':1}, inplace=True)
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/1359867004.py:1: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.
```

For example, when doing `df[col].method(value, inplace=True)`, try using `df.method({col: value}, inplace=True)` or `df[col] = df[col].method(value)` instead, to perform the operation inplace on the original object.

```
spam_df['label'].replace({'ham':0, 'spam':1}, inplace=True)
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/1359867004.py:1: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
spam_df['label'].replace({'ham':0, 'spam':1}, inplace=True)
```

```
In [ ]: spam_df['label'].value_counts(normalize=True)*100
```

```
Out[ ]: label
0      86.593683
1      13.406317
Name: proportion, dtype: float64
```

Q2 : Provide the metric for evaluating model (1 Point)

As you will notice, the data is highly imbalanced (most messages are labelled as ham and only few are labelled as spam). Always predicting ham will give us very good accuracy (close to 90%). So you need to choose a different metric.

Task: Provide the metric you will choose to evaluate your model. Explain why this is an appropriate metric for this case.

Ans :Using F1 Score for the minority class would help us. As it provides balance between Precision and recall. In this experiment, my trial with using recall led to decline in accuracy.

Q3 : Classification Pipelines (18 Points)

In the previous lectures you learned Data processing, Featurization such as CountVectorizer, TfidfVectorizer, and also Feature Engineering.

- You will now use following methods to create features which you can use in your model.

1. Sparse Embeddings (TF-IDF) (6 Points)
2. Feature Engineering (see examples below) (6 Points)
3. Sparse Embeddings (TF-IDF) + Feature Engineering (6 Points)

Approach:

***Use a smaller subset of dataset (e.g. 5-10 %) to evaluate the three pipelines . Based on your analysis (e.g. model score, learning curves) , choose one pipeline from the three. Provide your rationale for choosing the pipeline. Train only the final pipeline on randomly selected larger subset (e.g. 40%) of the data.**

Requirements:

1. You can use any ML model (Logistic Regression, XgBoost) for the classification. You will need to tune the **model for imbalanced dataset** (The link on XGBoost tutorial for imbalanced data: <https://machinelearningmastery.com/xgboost-for-imbalanced-classification/>).
2. For feature engineering, you can choose from the examples below. You do not have to use all of them. You can add other features as well. Think about what features can distinguish a spam from a regular email. Some examples :

Count of following (Words, characters, digits, exclamation marks, numbers, Nouns, ProperNouns, AUX, VERBS, Adjectives, named entities, spelling mistakes (see the link on how to get spelling mistakes <https://pypi.org/project/pyspellchecker/>).
3. For Sparse embeddings you will use **tfidf vectorization**. You need to choose appropriate parameters e.g. min_df, max_df, max_features, n-grams etc.).
4. Think carefully about the pre-processing you will do.

Tip: **Using GridSearch for hyperparameter tuning might take a lot of time. Try using RandomizedSearch.** You can also explore faster implementation of Gridsearch and RandomizedSearch in sklearn:

1. [Halving Grid Search](#)
2. [HalvingRandomSearchCV](#)

Required Submissions:

1. Submit two colab/jupyter notebooks

- (analysis with smaller subset and all three pipelines)
 - (analysis with bigger subset and only final pipeline)
2. Pdf version of the notebooks (HWs will not be graded if pdf version is not provided.)
 3. **The notebooks and pdf files should have the output.**
 4. **Name files as follows : FirstName_file1_hw2, FirstName_file2_h2**

```
In [ ]: spam_df['label'].value_counts(normalize=True)*100
```

```
Out[ ]: label
0      86.593683
1      13.406317
Name: proportion, dtype: float64
```

```
In [ ]: import sklearn
from sklearn.model_selection import train_test_split
X = spam_df['message'].values
y = spam_df['label'].values

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, stratify=y)
print(X_train[:5])
print(y_train[:5])
print(X_test[:5])
print(y_test[:5])
```

```
['Today my system sh get ready.all is well and i am also in the deep well'
'Its a valentine game. . . Send dis msg to all ur friends. .. If 5 answers
r d same then someone really loves u. Ques- which colour suits me the best?r
ply me'
'What is important is that you prevent dehydration by giving her enough flu
ids'
'Nowadays people are notixiquating the laxinorficated opportunity for bambl
ing of entropication.... Have you ever oblisngately opted ur books for the
masteriastering amplikater of fidalfication? It is very champlaxigating, i t
hink it is atrocious.. Wotz Ur Opinion???? Junna'
'Pls she needs to dat slowly or she will vomit more.']*
[0 0 0 0 0]
['What Today-sunday..sunday is holiday..so no work..'
'I am in hospital da. .I will return home in evening'
'Do you want bold 2 or bb torch'
'I am great princess! What are you thinking about me? :)'
'i can call in &#gt; min if thats ok']
[0 0 0 0 0]
```

```
In [ ]: import sys
if 'google.colab' in str(get_ipython()):
    from google.colab import drive
    drive.mount('/content/drive')

    !pip install -U nltk -qq
    !pip install -U spacy -qq
    !pip install -U scikit-optimize -qq
    !python -m spacy download en_core_web_sm -qq
    !pip install pyspellchecker -qq
    !pip install optuna -qq
```

```

    basepath = '/content/drive/MyDrive/Colab_Notebooks/BUAN_6342_Applied_Nat
    sys.path.append('/content/drive/MyDrive/Colab_Notebooks/BUAN_6342_Applie
else:
    basepath = '/Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harik
    sys.path.append(
        '/Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harikrish060

```

```

In [ ]: from pathlib import Path
        base_folder = Path(basepath)
        data_folder = base_folder/'0_Data_Folder'
        model_folder = base_folder/'models/spam'
        custom_functions = base_folder/'custom-functions'

```

```

In [ ]: import CustomPreprocessorSpacy as cp
        from FeaturizerSpacy_w_SpellChecker import ManualFeatures
        from plot_learning_curve import plot_learning_curve

```

```

In [ ]: import spacy
        nlp = spacy.load('en_core_web_sm')

```

```

In [ ]: cpp = cp.SpacyPreprocessor(model = 'en_core_web_sm', batch_size=500)

```

```

In [ ]: X_train_cleaned_bow = cpp.transform(X_train)
        X_test_cleaned_bow = cpp.transform(X_test)

```

```

/Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harikrish0607@gmail.c
om/My Drive/Colab_Notebooks/BUAN_6342_Applied_Natural_Language_Processing/0_
Custom_files/CustomPreprocessorSpacy.py:83: MarkupResemblesLocatorWarning: T
he input looks more like a filename than markup. You may want to open this f
ile and pass the filehandle into BeautifulSoup.
    soup = BeautifulSoup(text, "html.parser")

```

```

In [ ]: # Import the joblib library for saving and loading models
        import joblib
        import os

        file_X_train_cleaned_bow = data_folder / 'x_train_cleaned_bow_small.pkl'
        os.makedirs(os.path.dirname(file_X_train_cleaned_bow), exist_ok=True)
        joblib.dump(X_train_cleaned_bow, file_X_train_cleaned_bow)

        file_X_test_cleaned_bow = data_folder / 'x_test_cleaned_bow_small.pkl'
        os.makedirs(os.path.dirname(file_X_test_cleaned_bow), exist_ok=True)
        joblib.dump(X_test_cleaned_bow, file_X_test_cleaned_bow)

```

```

Out[ ]: ['/Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harikrish0607@gmai
l.com/My Drive/Colab_Notebooks/BUAN_6342_Applied_Natural_Language_Processin
g/0_Data_Folder/x_test_cleaned_bow_small.pkl']

```

```

In [ ]: import numpy as np
        unique, counts = np.unique(y_train, return_counts=True)
        print(dict(zip(unique, counts/sum(counts))))

```

```

{0: 0.8658290329818263, 1: 0.13417096701817366}

```

```
In [ ]: weights = dict(zip(unique[::-1], counts/sum(counts)))
weights
```

```
Out[ ]: {1: 0.8658290329818263, 0: 0.13417096701817366}
```

```
In [ ]: X_train_cleaned_basic = cp.SpacyPreprocessor(model='en_core_web_sm',
                                                    lemmatize=False, lower=False,
                                                    remove_stop=False, remove_punctuation=False,
                                                    remove_email=False, remove_url=False,
                                                    add_user_mention_prefix=True,
                                                    basic_clean_only=True).
```

```
In [ ]: # save this to a file
file_X_train_cleaned_basic = data_folder / 'x_train_cleaned_basic_small.pkl'
joblib.dump(X_train_cleaned_basic, file_X_train_cleaned_basic)
```

```
Out[ ]: ['Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harikrish0607@gmail.com/My Drive/Colab Notebooks/BUAN_6342_Applied Natural Language Processing/0_Data_Folder/x_train_cleaned_basic_small.pkl']
```

```
In [ ]: featurizer = ManualFeatures(spacy_model='en_core_web_sm', batch_size = 1000)
```

```
In [ ]: X_test_cleaned_basic = cp.SpacyPreprocessor(model='en_core_web_sm',
                                                    lemmatize=False, lower=False,
                                                    remove_stop=False, remove_punctuation=False,
                                                    remove_email=False, remove_url=False,
                                                    add_user_mention_prefix=True,
                                                    basic_clean_only=True).
file_X_test_cleaned_basic = data_folder / 'x_test_cleaned_basic_small.pkl'
joblib.dump(X_test_cleaned_basic, file_X_test_cleaned_basic)
```

```
Out[ ]: ['Users/harikrishnadev/Library/CloudStorage/GoogleDrive-harikrish0607@gmail.com/My Drive/Colab Notebooks/BUAN_6342_Applied Natural Language Processing/0_Data_Folder/x_test_cleaned_basic_small.pkl']
```

```
In [ ]: featurizer = ManualFeatures(spacy_model='en_core_web_sm', batch_size = 1000)
```

```
In [ ]: X_train_features, feature_names = featurizer.fit_transform(X_train_cleaned_b
```

```
In [ ]: X_train_features.shape
```

```
Out[ ]: (4457, 12)
```

```
In [ ]: X_train_final = pd.concat((pd.DataFrame(X_train_cleaned_bow, columns=['cleaned_bow']),
pd.DataFrame(X_train_features, columns=feature_names))
```

```
In [ ]: X_train_final.head()
```

Out []:	cleaned_text	count_words	count_characters	count_characters_no_space	avg_woi
0	today system sh deep	16.0	71.0		56.0
1	valentine game send dis msg ur friend 5 answer...	30.0	156.0		124.0
2	important prevent dehydration give fluid	13.0	77.0		65.0
3	nowadays people notixiquate laxinorfficate oppo...	37.0	274.0		238.0
4	pls need dat slowly vomit	11.0	51.0		41.0

Sparse Embeddings (TF-IDF) + Feature Engineering

```
In [ ]: from sklearn.base import TransformerMixin, BaseEstimator
from scipy.sparse import csr_matrix
class SparseTransformer(TransformerMixin, BaseEstimator):
    def __init__(self):
        pass

    def fit(self, X, y=None):
        return self

    def transform(self, X, y=None):
        return csr_matrix(X)
```

```
In [ ]: from sklearn.pipeline import Pipeline
from sklearn.feature_extraction.text import TfidfVectorizer
sparse_features = Pipeline([('sparse', SparseTransformer()), ])
vectorizer = Pipeline([('tfidf', TfidfVectorizer(max_features=5)), ])
```

```
In [ ]: from sklearn.compose import ColumnTransformer
combined_features = ColumnTransformer(
    transformers=[
        ('tfidf', vectorizer, 'cleaned_text'),
    ], remainder=sparse_features
)
```



```
In [ ]: test = combined_features.fit_transform(X_train_final[0:10])
```

```
In [ ]: test[:5]
```

```
Out[ ]: array([[ 0.          ,  0.          ,  0.          ,  0.          ,
                 0.          , 16.          , 71.          , 56.          ,
                 3.29411765,  0.          ,  1.          ,  3.          ,
                 2.          ,  1.          ,  2.          ,  1.          ,
                 0.          ],
               [ 0.          ,  0.          ,  0.          ,  0.          ,
                 1.          , 30.          , 156.         , 124.         ,
                 4.          ,  1.          ,  3.          , 11.          ,
                 0.          ,  4.          ,  2.          ,  2.          ,
                 0.          ],
               [ 0.          ,  0.          ,  0.          ,  0.          ,
                 0.          , 13.          , 77.          , 65.          ,
                 4.64285714,  0.          ,  1.          ,  2.          ,
                 2.          ,  2.          ,  2.          ,  0.          ,
                 0.          ],
               [ 0.          ,  0.          ,  0.          ,  0.          ,
                 1.          , 37.          , 274.         , 238.         ,
                 6.26315789,  0.          ,  3.          , 10.          ,
                 4.          ,  7.          ,  1.          ,  2.          ,
                 0.          ],
               [ 0.          ,  0.          ,  0.          ,  1.          ,
                 0.          , 11.          , 51.          , 41.          ,
                 3.41666667,  0.          ,  1.          ,  0.          ,
                 1.          ,  3.          ,  1.          ,  0.          ,
                 0.          ]])
```

```
In [ ]: import optuna
        from sklearn.pipeline import Pipeline
        from sklearn.linear_model import LogisticRegression
        classifier_3 = Pipeline([('combined_features', combined_features),
                                ('classifier', LogisticRegression(max_iter=10000)),
                                ])
        classifier_3.get_params().keys()
```

```
/opt/homebrew/anaconda3/envs/nlp/lib/python3.11/site-packages/tqdm/auto.py:2
1: TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. S
ee https://ipywidgets.readthedocs.io/en/stable/user_install.html
    from .autonotebook import tqdm as notebook_tqdm
```

```
Out[ ]: dict_keys(['memory', 'steps', 'verbose', 'combined_features', 'classifier',
'combined_features__n_jobs', 'combined_features__remainder__memory', 'combined_features__remainder__steps', 'combined_features__remainder__verbose', 'combined_features__remainder__sparse', 'combined_features__remainder', 'combined_features__sparse_threshold', 'combined_features__transformer_weights', 'combined_features__transformers', 'combined_features__verbose', 'combined_features__verbose_feature_names_out', 'combined_features__tfidf', 'combined_features__tfidf__memory', 'combined_features__tfidf__steps', 'combined_features__tfidf__verbose', 'combined_features__tfidf__tfidf', 'combined_features__tfidf__tfidf__analyzer', 'combined_features__tfidf__tfidf__binary', 'combined_features__tfidf__tfidf__decode_error', 'combined_features__tfidf__tfidf__dtype', 'combined_features__tfidf__tfidf__encoding', 'combined_features__tfidf__tfidf__input', 'combined_features__tfidf__tfidf__lowercase', 'combined_features__tfidf__tfidf__max_df', 'combined_features__tfidf__tfidf__max_features', 'combined_features__tfidf__tfidf__min_df', 'combined_features__tfidf__tfidf__ngram_range', 'combined_features__tfidf__tfidf__norm', 'combined_features__tfidf__tfidf__preprocessor', 'combined_features__tfidf__tfidf__smooth_idf', 'combined_features__tfidf__tfidf__stop_words', 'combined_features__tfidf__tfidf__strip_accents', 'combined_features__tfidf__tfidf__sublinear_tf', 'combined_features__tfidf__tfidf__token_pattern', 'combined_features__tfidf__tfidf__tokenizer', 'combined_features__tfidf__tfidf__use_idf', 'combined_features__tfidf__tfidf__vocabulary', 'classifier__C', 'classifier__class_weight', 'classifier__dual', 'classifier__fit_intercept', 'classifier__intercept_scaling', 'classifier__l1_ratio', 'classifier__max_iter', 'classifier__multi_class', 'classifier__n_jobs', 'classifier__penalty', 'classifier__random_state', 'classifier__solver', 'classifier__tol', 'classifier__verbose', 'classifier__warm_start'])
```

```
In [ ]: import optuna
from sklearn.compose import ColumnTransformer
from sklearn.model_selection import cross_val_score
from sklearn.metrics import make_scorer, f1_score
from sklearn.feature_extraction.text import TfidfVectorizer

def objective(trial):
    # Define the hyperparameters to search over
    C = trial.suggest_loguniform('C', 1e-5, 100)
    combined_features__tfidf__tfidf__max_features = trial.suggest_int('combined_features__tfidf__tfidf__max_features', 10, 1000)
    weight = trial.suggest_float('weight', 0.1, 0.9)

    class_weight = {0: 1 - weight, 1: weight}

    # Use cross-validation to evaluate the classifier with the current hyperparameters
    # Use f1_score as the scoring metric
    scores = cross_val_score(classifier_3, X_train_final, y_train, cv=5, scoring='f1')
    return scores.mean()
```

```
In [ ]: study = optuna.create_study(direction='maximize')
study.optimize(objective, n_trials=100)
```

```

[I 2024-02-13 14:12:30,549] A new study created in memory with name: no-name
-e5246495-13ff-474a-8f0e-f218c15eca36
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:33,405] Trial 0 finished with value: 0.8082041480574598
and parameters: {'C': 0.12535992345832675, 'combined_features__tfidf__tfidf__
_max_features': 9108, 'weight': 0.7748605176191444}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:35,550] Trial 1 finished with value: 0.8082041480574598
and parameters: {'C': 0.02816735834224162, 'combined_features__tfidf__tfidf__
_max_features': 6576, 'weight': 0.8162611537967875}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:37,559] Trial 2 finished with value: 0.8082041480574598
and parameters: {'C': 0.0003835050598253934, 'combined_features__tfidf__tfidf__
_max_features': 3125, 'weight': 0.5099614372916251}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:40,240] Trial 3 finished with value: 0.8082041480574598
and parameters: {'C': 1.4215112400525138, 'combined_features__tfidf__tfidf__
_max_features': 6839, 'weight': 0.5639219148097253}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:42,442] Trial 4 finished with value: 0.8082041480574598
and parameters: {'C': 0.0007855621695125993, 'combined_features__tfidf__tfidf__
_max_features': 282, 'weight': 0.41505466501866106}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:44,288] Trial 5 finished with value: 0.8082041480574598
and parameters: {'C': 2.8521856752136223, 'combined_features__tfidf__tfidf__
_max_features': 6844, 'weight': 0.5450669164667569}. Best is trial 0 with val
ue: 0.8082041480574598.

```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:46,259] Trial 6 finished with value: 0.8082041480574598
and parameters: {'C': 9.515030461361165e-05, 'combined_features__tfidf__tfidf__max_features': 4001, 'weight': 0.8941612749436085}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:49,387] Trial 7 finished with value: 0.8082041480574598
and parameters: {'C': 3.6753446991782576e-05, 'combined_features__tfidf__tfidf__max_features': 7306, 'weight': 0.6303058043916587}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:51,416] Trial 8 finished with value: 0.8082041480574598
and parameters: {'C': 0.010686384251606888, 'combined_features__tfidf__tfidf__max_features': 1713, 'weight': 0.8553542450790907}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:53,412] Trial 9 finished with value: 0.8082041480574598
and parameters: {'C': 4.909455150329548e-05, 'combined_features__tfidf__tfidf__max_features': 8414, 'weight': 0.621456557185048}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:55,589] Trial 10 finished with value: 0.8082041480574598
and parameters: {'C': 93.55167614660019, 'combined_features__tfidf__tfidf__max_features': 9726, 'weight': 0.20245029403825088}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:12:57,892] Trial 11 finished with value: 0.8082041480574598
and parameters: {'C': 0.08700482283140826, 'combined_features__tfidf__tfidf__max_features': 5335, 'weight': 0.7569247264195956}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This

feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use `suggest_float(..., log=True)` instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:00,476] Trial 12 finished with value: 0.8082041480574598
and parameters: {'C': 0.03382291031010231, 'combined_features__tfidf__tfidf__
_max_features': 9963, 'weight': 0.7483895717712827}. Best is trial 0 with va
lue: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:02,431] Trial 13 finished with value: 0.8082041480574598
and parameters: {'C': 0.4821421039841941, 'combined_features__tfidf__tfidf__
_max_features': 5546, 'weight': 0.7489216068989537}. Best is trial 0 with val
ue: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:04,997] Trial 14 finished with value: 0.8082041480574598
and parameters: {'C': 0.003864763998694529, 'combined_features__tfidf__tfidf__
_max_features': 8480, 'weight': 0.36715120528952655}. Best is trial 0 with
value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:06,897] Trial 15 finished with value: 0.8082041480574598
and parameters: {'C': 19.8186342708331, 'combined_features__tfidf__tfidf__ma
x_features': 8122, 'weight': 0.8320961951764776}. Best is trial 0 with valu
e: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:10,736] Trial 16 finished with value: 0.8082041480574598
and parameters: {'C': 0.20022041281607458, 'combined_features__tfidf__tfidf__
_max_features': 6052, 'weight': 0.6771404316921531}. Best is trial 0 with va
lue: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:13,390] Trial 17 finished with value: 0.8082041480574598
and parameters: {'C': 0.00474658383983427, 'combined_features__tfidf__tfidf__
_max_features': 3827, 'weight': 0.20184857942747125}. Best is trial 0 with v
alue: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```



```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:16,017] Trial 18 finished with value: 0.8082041480574598
and parameters: {'C': 0.03530016153252745, 'combined_features__tfidf__tfidf__
_max_features': 9286, 'weight': 0.40682507687707076}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:18,465] Trial 19 finished with value: 0.8082041480574598
and parameters: {'C': 3.2519985916942127, 'combined_features__tfidf__tfidf__
_max_features': 7391, 'weight': 0.7918833838559434}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:21,256] Trial 20 finished with value: 0.8082041480574598
and parameters: {'C': 0.0010110338801684678, 'combined_features__tfidf__tfidf
f__max_features': 4720, 'weight': 0.6781872339802852}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:23,327] Trial 21 finished with value: 0.8082041480574598
and parameters: {'C': 0.0003894377910425727, 'combined_features__tfidf__tfidf
f__max_features': 2646, 'weight': 0.30946698616219204}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:25,416] Trial 22 finished with value: 0.8082041480574598
and parameters: {'C': 0.3686843304746577, 'combined_features__tfidf__tfidf__
_max_features': 2602, 'weight': 0.464913936333059}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:27,525] Trial 23 finished with value: 0.8082041480574598
and parameters: {'C': 0.00017573012704188954, 'combined_features__tfidf__tfi
df__max_features': 4037, 'weight': 0.11502716464300755}. Best is trial 0 wit
h value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:30,107] Trial 24 finished with value: 0.8082041480574598
```

```
and parameters: {'C': 0.00824205583600414, 'combined_features__tfidf__tfidf__max_features': 2794, 'weight': 0.684181739335301}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:33,082] Trial 25 finished with value: 0.8082041480574598 and parameters: {'C': 0.0018319355423340143, 'combined_features__tfidf__tfidf__max_features': 6223, 'weight': 0.5161403087906254}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:35,297] Trial 26 finished with value: 0.8082041480574598 and parameters: {'C': 0.11348090186461945, 'combined_features__tfidf__tfidf__max_features': 1666, 'weight': 0.8198854648264735}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:37,262] Trial 27 finished with value: 0.8082041480574598 and parameters: {'C': 0.01856906957486446, 'combined_features__tfidf__tfidf__max_features': 9159, 'weight': 0.8954559417062513}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:39,793] Trial 28 finished with value: 0.8082041480574598 and parameters: {'C': 1.3571411105986436e-05, 'combined_features__tfidf__tfidf__max_features': 4534, 'weight': 0.7210262382507454}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:41,922] Trial 29 finished with value: 0.8082041480574598 and parameters: {'C': 0.9714374260686041, 'combined_features__tfidf__tfidf__max_features': 7616, 'weight': 0.6325980510111315}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:44,656] Trial 30 finished with value: 0.8082041480574598 and parameters: {'C': 0.0623109153377088, 'combined_features__tfidf__tfidf__max_features': 6404, 'weight': 0.47289251811001354}. Best is trial 0 with va
```

```

lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:47,505] Trial 31 finished with value: 0.8082041480574598
and parameters: {'C': 1.836569161031432, 'combined_features__tfidf__tfidf__m
ax_features': 175, 'weight': 0.5670363153663872}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:49,744] Trial 32 finished with value: 0.8082041480574598
and parameters: {'C': 8.680509952936461, 'combined_features__tfidf__tfidf__m
ax_features': 6778, 'weight': 0.5765016002627704}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:51,967] Trial 33 finished with value: 0.8082041480574598
and parameters: {'C': 6.126410108574272, 'combined_features__tfidf__tfidf__m
ax_features': 1158, 'weight': 0.33144288150869405}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:54,505] Trial 34 finished with value: 0.8082041480574598
and parameters: {'C': 0.8759960511577186, 'combined_features__tfidf__tfidf__
max_features': 7898, 'weight': 0.5137437560649084}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:56,733] Trial 35 finished with value: 0.8082041480574598
and parameters: {'C': 0.0003056061280742204, 'combined_features__tfidf__tfidf__
max_features': 7035, 'weight': 0.433634370605286}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:13:58,832] Trial 36 finished with value: 0.8082041480574598
and parameters: {'C': 0.21094118342655827, 'combined_features__tfidf__tfidf__
max_features': 3213, 'weight': 0.6022480995532941}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.

```



```
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:01,604] Trial 37 finished with value: 0.8082041480574598
and parameters: {'C': 0.002328295947075327, 'combined_features__tfidf__tfidf__
__max_features': 8906, 'weight': 0.8725036794645741}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:04,078] Trial 38 finished with value: 0.8082041480574598
and parameters: {'C': 26.666541834394494, 'combined_features__tfidf__tfidf__
__max_features': 5812, 'weight': 0.807688984871497}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:06,379] Trial 39 finished with value: 0.8082041480574598
and parameters: {'C': 7.852410488936692e-05, 'combined_features__tfidf__tfidf__
__max_features': 5120, 'weight': 0.5423198653269747}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:08,918] Trial 40 finished with value: 0.8082041480574598
and parameters: {'C': 0.013689983593705445, 'combined_features__tfidf__tfidf__
__max_features': 6817, 'weight': 0.7122166997702477}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:11,295] Trial 41 finished with value: 0.8082041480574598
and parameters: {'C': 0.0007680257545426, 'combined_features__tfidf__tfidf__
__max_features': 1529, 'weight': 0.3869727086132387}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:13,535] Trial 42 finished with value: 0.8082041480574598
and parameters: {'C': 1.3131781470636957e-05, 'combined_features__tfidf__tfi
df__max_features': 197, 'weight': 0.4691130094701239}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
```

```

ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:16,264] Trial 43 finished with value: 0.8082041480574598
and parameters: {'C': 2.785106166532424e-05, 'combined_features__tfidf__tfidf__max_features': 1078, 'weight': 0.3199348604623423}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:18,563] Trial 44 finished with value: 0.8082041480574598
and parameters: {'C': 0.0001374461264247581, 'combined_features__tfidf__tfidf__max_features': 8586, 'weight': 0.2694803780892126}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:20,923] Trial 45 finished with value: 0.8082041480574598
and parameters: {'C': 0.05207594940955501, 'combined_features__tfidf__tfidf__max_features': 734, 'weight': 0.7717434925047152}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:24,216] Trial 46 finished with value: 0.8082041480574598
and parameters: {'C': 0.006144895001048448, 'combined_features__tfidf__tfidf__max_features': 5525, 'weight': 0.8452395769558682}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:26,845] Trial 47 finished with value: 0.8082041480574598
and parameters: {'C': 0.1799783937501248, 'combined_features__tfidf__tfidf__max_features': 9788, 'weight': 0.6347806588431858}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:29,491] Trial 48 finished with value: 0.8082041480574598
and parameters: {'C': 0.0007358507765726429, 'combined_features__tfidf__tfidf__max_features': 2305, 'weight': 0.4133371095826031}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)

```

```
[I 2024-02-13 14:14:32,420] Trial 49 finished with value: 0.8082041480574598
and parameters: {'C': 0.020496720539932397, 'combined_features__tfidf__tfidf__
__max_features': 3371, 'weight': 0.3645201553203793}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:35,006] Trial 50 finished with value: 0.8082041480574598
and parameters: {'C': 0.0023356147848115115, 'combined_features__tfidf__tfidf__
__max_features': 4740, 'weight': 0.25691368322616875}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:38,494] Trial 51 finished with value: 0.8082041480574598
and parameters: {'C': 1.8796889175820957, 'combined_features__tfidf__tfidf__
__max_features': 7743, 'weight': 0.5363299116509108}. Best is trial 0 with val
ue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:41,072] Trial 52 finished with value: 0.8082041480574598
and parameters: {'C': 0.4951156605423794, 'combined_features__tfidf__tfidf__
__max_features': 6512, 'weight': 0.49010347124534603}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:43,338] Trial 53 finished with value: 0.8082041480574598
and parameters: {'C': 58.13612742854314, 'combined_features__tfidf__tfidf__m
ax_features': 8084, 'weight': 0.4304940974174647}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:46,038] Trial 54 finished with value: 0.8082041480574598
and parameters: {'C': 5.154817571619107, 'combined_features__tfidf__tfidf__m
ax_features': 5887, 'weight': 0.6024780349101867}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:48,114] Trial 55 finished with value: 0.8082041480574598
and parameters: {'C': 10.89168408535852, 'combined_features__tfidf__tfidf__m
```

```
ax_features': 4125, 'weight': 0.7794513850508227}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:50,569] Trial 56 finished with value: 0.8082041480574598 and parameters: {'C': 0.0003389465015629418, 'combined_features__tfidf__tfidf__max_features': 7246, 'weight': 0.7379272016262715}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:53,197] Trial 57 finished with value: 0.8082041480574598 and parameters: {'C': 2.7170405787080196, 'combined_features__tfidf__tfidf__max_features': 9469, 'weight': 0.6744255704544767}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:55,244] Trial 58 finished with value: 0.8082041480574598 and parameters: {'C': 1.0364101726461141, 'combined_features__tfidf__tfidf__max_features': 2249, 'weight': 0.447371687299072}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:14:57,604] Trial 59 finished with value: 0.8082041480574598 and parameters: {'C': 0.3817800205558645, 'combined_features__tfidf__tfidf__max_features': 8322, 'weight': 0.6533802498124281}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:00,662] Trial 60 finished with value: 0.8082041480574598 and parameters: {'C': 0.08806096623922874, 'combined_features__tfidf__tfidf__max_features': 7334, 'weight': 0.554653763173761}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:02,728] Trial 61 finished with value: 0.8082041480574598 and parameters: {'C': 3.749641688211663e-05, 'combined_features__tfidf__tfidf__max_features': 3677, 'weight': 0.8856846547576213}. Best is trial 0 with value: 0.8082041480574598.
```


/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:05,052] Trial 62 finished with value: 0.8082041480574598
and parameters: {'C': 8.94395964616018e-05, 'combined_features__tfidf__tfidf__max_features': 4389, 'weight': 0.8277620693413501}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:07,675] Trial 63 finished with value: 0.8082041480574598
and parameters: {'C': 0.00020976061185031428, 'combined_features__tfidf__tfidf__max_features': 5405, 'weight': 0.8441780790990838}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:10,320] Trial 64 finished with value: 0.8082041480574598
and parameters: {'C': 0.0012504478572374518, 'combined_features__tfidf__tfidf__max_features': 3311, 'weight': 0.86555276650749}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:12,517] Trial 65 finished with value: 0.8082041480574598
and parameters: {'C': 5.345391202109184e-05, 'combined_features__tfidf__tfidf__max_features': 6171, 'weight': 0.8061361409445048}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:15,281] Trial 66 finished with value: 0.8082041480574598
and parameters: {'C': 0.0005067029019382086, 'combined_features__tfidf__tfidf__max_features': 5057, 'weight': 0.49154841202067856}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use suggest_float(..., log=True) instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:18,385] Trial 67 finished with value: 0.8082041480574598
and parameters: {'C': 0.028707323863205715, 'combined_features__tfidf__tfidf__max_features': 6521, 'weight': 0.8982226805383152}. Best is trial 0 with value: 0.8082041480574598.
```

/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This

feature will be removed in v6.0.0. See <https://github.com/optuna/optuna/releases/tag/v3.0.0>. Use `suggest_float(..., log=True)` instead.

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:20,579] Trial 68 finished with value: 0.8082041480574598
and parameters: {'C': 0.0039421080051805875, 'combined_features__tfidf__tfidf__max_features': 2942, 'weight': 0.5145086109218826}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:22,986] Trial 69 finished with value: 0.8082041480574598
and parameters: {'C': 2.0377962765885815e-05, 'combined_features__tfidf__tfidf__max_features': 2034, 'weight': 0.7373097959051723}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:25,130] Trial 70 finished with value: 0.8082041480574598
and parameters: {'C': 16.597908823279973, 'combined_features__tfidf__tfidf__max_features': 3768, 'weight': 0.6062330782451363}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:27,483] Trial 71 finished with value: 0.8082041480574598
and parameters: {'C': 0.00010291658869177527, 'combined_features__tfidf__tfidf__max_features': 6876, 'weight': 0.574524175635228}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:30,250] Trial 72 finished with value: 0.8082041480574598
and parameters: {'C': 5.529300824289393e-05, 'combined_features__tfidf__tfidf__max_features': 8849, 'weight': 0.6921460004077524}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:32,405] Trial 73 finished with value: 0.8082041480574598
and parameters: {'C': 0.00018588405618423555, 'combined_features__tfidf__tfidf__max_features': 5778, 'weight': 0.5296477924731243}. Best is trial 0 with value: 0.8082041480574598.
```

```
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
```

```
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:34,418] Trial 74 finished with value: 0.8082041480574598
and parameters: {'C': 1.009165055318639e-05, 'combined_features__tfidf__tfidf__max_features': 7580, 'weight': 0.5859441212769989}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:37,084] Trial 75 finished with value: 0.8082041480574598
and parameters: {'C': 0.7782176386739278, 'combined_features__tfidf__tfidf__max_features': 7031, 'weight': 0.6496920275889052}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:39,284] Trial 76 finished with value: 0.8082041480574598
and parameters: {'C': 0.008595253703718848, 'combined_features__tfidf__tfidf__max_features': 4853, 'weight': 0.7645330698899404}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:41,697] Trial 77 finished with value: 0.8082041480574598
and parameters: {'C': 0.22400029167624347, 'combined_features__tfidf__tfidf__max_features': 8047, 'weight': 0.7983177575902075}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:43,706] Trial 78 finished with value: 0.8082041480574598
and parameters: {'C': 0.0005245999515568128, 'combined_features__tfidf__tfidf__max_features': 8480, 'weight': 0.36800208012839863}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:46,392] Trial 79 finished with value: 0.8082041480574598
and parameters: {'C': 1.9743533603982137e-05, 'combined_features__tfidf__tfidf__max_features': 6327, 'weight': 0.5517882386010373}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:48,871] Trial 80 finished with value: 0.8082041480574598
```

```
and parameters: {'C': 3.6827876272032087, 'combined_features__tfidf__tfidf__max_features': 549, 'weight': 0.4783987043776179}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:51,328] Trial 81 finished with value: 0.8082041480574598 and parameters: {'C': 0.010168539996728128, 'combined_features__tfidf__tfidf__max_features': 1815, 'weight': 0.8580292699498105}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:53,890] Trial 82 finished with value: 0.8082041480574598 and parameters: {'C': 0.03848320603118535, 'combined_features__tfidf__tfidf__max_features': 666, 'weight': 0.8778453884174487}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:56,348] Trial 83 finished with value: 0.8082041480574598 and parameters: {'C': 0.0019501168721069958, 'combined_features__tfidf__tfidf__max_features': 1447, 'weight': 0.8423680496677542}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:15:58,876] Trial 84 finished with value: 0.8082041480574598 and parameters: {'C': 0.12846136258102672, 'combined_features__tfidf__tfidf__max_features': 1298, 'weight': 0.785112831321534}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:01,577] Trial 85 finished with value: 0.8082041480574598 and parameters: {'C': 0.059469360485617455, 'combined_features__tfidf__tfidf__max_features': 2862, 'weight': 0.8586927864575638}. Best is trial 0 with value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This feature will be removed in v6.0.0. See https://github.com/optuna/optuna/releases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
  C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:04,083] Trial 86 finished with value: 0.8082041480574598 and parameters: {'C': 0.000128287352125335, 'combined_features__tfidf__tfidf__max_features': 6587, 'weight': 0.5070599691946317}. Best is trial 0 with v
```



```

alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:06,444] Trial 87 finished with value: 0.8082041480574598
and parameters: {'C': 0.00026087257859987466, 'combined_features__tfidf__tfidf__
__max_features': 4250, 'weight': 0.44635917087461974}. Best is trial 0 wit
h value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:08,947] Trial 88 finished with value: 0.8082041480574598
and parameters: {'C': 0.005679038610205349, 'combined_features__tfidf__tfidf__
__max_features': 986, 'weight': 0.8195674072594301}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:11,335] Trial 89 finished with value: 0.8082041480574598
and parameters: {'C': 1.7058222657530886, 'combined_features__tfidf__tfidf__
__max_features': 2512, 'weight': 0.127399387547672}. Best is trial 0 with valu
e: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:13,391] Trial 90 finished with value: 0.8082041480574598
and parameters: {'C': 0.016174589024055535, 'combined_features__tfidf__tfidf__
__max_features': 7167, 'weight': 0.7510794121304802}. Best is trial 0 with v
alue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:16,028] Trial 91 finished with value: 0.8082041480574598
and parameters: {'C': 6.631753420678186e-05, 'combined_features__tfidf__tfidf__
__max_features': 9065, 'weight': 0.6184367212765192}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:18,508] Trial 92 finished with value: 0.8082041480574598
and parameters: {'C': 3.471936537085266e-05, 'combined_features__tfidf__tfidf__
__max_features': 9415, 'weight': 0.7060478786480028}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.

```

```
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:21,120] Trial 93 finished with value: 0.8082041480574598
and parameters: {'C': 0.0010370918828543185, 'combined_features__tfidf__tfid
f__max_features': 8740, 'weight': 0.5619815511146921}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:23,786] Trial 94 finished with value: 0.8082041480574598
and parameters: {'C': 0.0029917388350430936, 'combined_features__tfidf__tfid
f__max_features': 7495, 'weight': 0.5867833635026698}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:26,122] Trial 95 finished with value: 0.8082041480574598
and parameters: {'C': 3.689453781340131e-05, 'combined_features__tfidf__tfid
f__max_features': 8287, 'weight': 0.5346017965906953}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:28,746] Trial 96 finished with value: 0.8082041480574598
and parameters: {'C': 2.1726684207321142e-05, 'combined_features__tfidf__tfi
df__max_features': 6000, 'weight': 0.643238832284229}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:31,173] Trial 97 finished with value: 0.8082041480574598
and parameters: {'C': 0.0001421505631425961, 'combined_features__tfidf__tfid
f__max_features': 7883, 'weight': 0.8258582681444592}. Best is trial 0 with
value: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.
    C = trial.suggest_loguniform('C', 1e-5, 100)
[I 2024-02-13 14:16:33,673] Trial 98 finished with value: 0.8082041480574598
and parameters: {'C': 0.026501099494174996, 'combined_features__tfidf__tfidf
__max_features': 392, 'weight': 0.6247974495428718}. Best is trial 0 with va
lue: 0.8082041480574598.
/var/folders/3d/fk1zpy415g31bg07yrkc5qbm0000gn/T/ipykernel_70486/3904844807.
py:9: FutureWarning: suggest_loguniform has been deprecated in v3.0.0. This
feature will be removed in v6.0.0. See https://github.com/optuna/optuna/rele
```

```
ases/tag/v3.0.0. Use suggest_float(..., log=True) instead.  
C = trial.suggest_loguniform('C', 1e-5, 100)  
[I 2024-02-13 14:16:35,848] Trial 99 finished with value: 0.8082041480574598  
and parameters: {'C': 0.5822984584448512, 'combined_features__tfidf__tfidf__  
max_features': 889, 'weight': 0.6681704091284701}. Best is trial 0 with valu  
e: 0.8082041480574598.
```

```
In [ ]: study.best_params
```

```
Out[ ]: {'C': 0.12535992345832675,  
        'combined_features__tfidf__tfidf__max_features': 9108,  
        'weight': 0.7748605176191444}
```

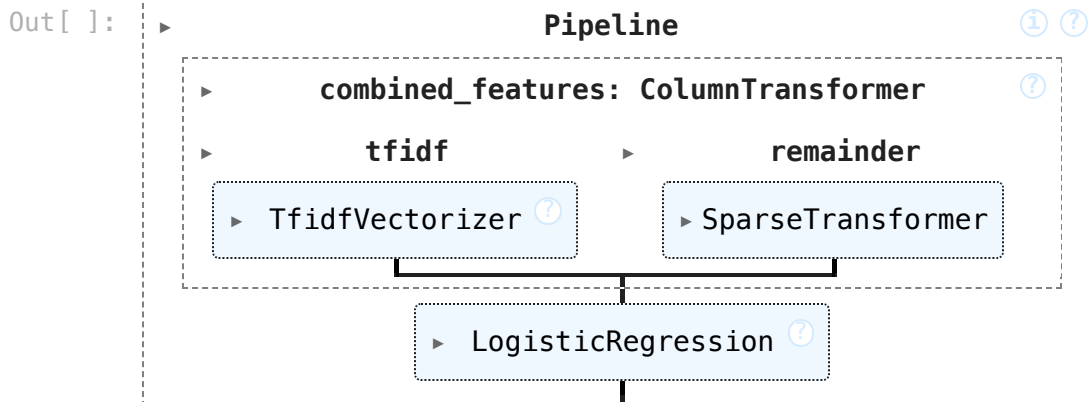
```
In [ ]: study.best_value
```

```
Out[ ]: 0.8082041480574598
```

```
In [ ]: # Get the best hyperparameters found by Optuna  
best_max_features = study.best_params['combined_features__tfidf__tfidf__max_  
best_C = study.best_params['C']  
best_weight = study.best_params['weight']
```

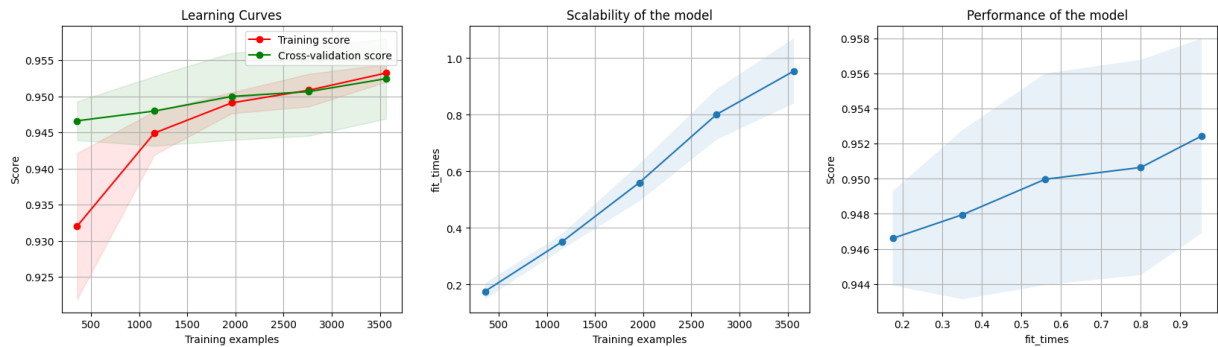
```
In [ ]: best_classifier_3 = Pipeline([('combined_features',  
                                     ColumnTransformer(transformers=[('tfidf',  
                                                                       TfidfVector  
                                                                       remainder=sparse_features))  
                                     ('classifier', LogisticRegression(max_iter=100  
                                                                       class_weight
```

```
In [ ]: best_classifier_3.fit(X_train_final, y_train)
```



```
In [ ]: # plot learning curves  
plot_learning_curve(best_classifier_3, 'Learning Curves',  
                    X_train_final, y_train, n_jobs=-1)
```

```
Out[ ]: <module 'matplotlib.pyplot' from '/opt/homebrew/anaconda3/envs/nlp/lib/pyth  
on3.11/site-packages/matplotlib/pyplot.py'>
```



```
In [ ]: # Final Pipeline
def final_pipeline(text):
    X_features, feature_names = featurizer.fit_transform(text)
    cleaned_text = X_test_cleaned_bow
    X_final = pd.concat((pd.DataFrame(cleaned_text, columns=['cleaned_text']),
                             pd.DataFrame(X_features, columns=feature_names)), axis=1)
    predictions = best_classifier_3.predict(X_final)
    return predictions
y_test_pred = final_pipeline(X_test)
```

```
In [ ]: np.unique(y_test_pred, return_counts=True)
```

```
Out[ ]: (array([0, 1]), array([934, 181]))
```

```
In [ ]: from sklearn import metrics
print('\nTest set classification report:\n\n',
      metrics.classification_report(y_test, y_test_pred))
```

Test set classification report:

	precision	recall	f1-score	support
0	0.98	0.95	0.97	966
1	0.74	0.90	0.81	149
accuracy			0.94	1115
macro avg	0.86	0.93	0.89	1115
weighted avg	0.95	0.94	0.95	1115