Q6_change_sampling_rate

April 29, 2018

```
In [1]: # allow the notebook to access the parent directory so we can import the other modules
        # https://stackoverflow.com/a/35273613
        import os
        import sys
        nb_dir = os.path.split(os.getcwd())[0]
        if nb_dir not in sys.path:
            sys.path.append(nb_dir)
0.1 # Data Preparation
0.1.1 Constants and Folder Paths
In [2]: import os
        dataset_folder_path = os.path.join("..", "files", "dataset")
0.1.2 Load Data and Split into Test, Train/Valid
In [3]: from data.DataSet import DataSet
        dataset = DataSet()
        dataset.load(dataset_folder_path, test_set_percentage=0, validation_set_percentage=0)
In [4]: print(len(dataset.train_data))
        print(len(dataset.test_data))
3600
0
```

0.1.3 Data Preprocessing

```
7200
0
```

Test different sample rates and cross validate then compare the results to determine the optimal sampling rate

```
In [15]: NUM_SAMPLES_TO_TRY = [300, 200, 100, 75, 50, 25, 10]
        N_FOLDS = 4
        PARAM_NUM_EPOCHS = 20
        PARAM_BATCH_SIZE = 300
In [16]: import numpy as np
        import pandas as pd
        from utils.evaluation import cross_validate_model
        from models.regularized_deep_gru import NaiveRegularizedDeepGRU
        results = {}
        for num_samples in NUM_SAMPLES_TO_TRY:
            print("\n\n\n----")
            print("Evaluating Spline interpolation using %d samples" % num_samples)
            print("----")
            # setup copy of data and evaluate its spline with the currently selected number of
            data = dataset.copy()
            data.apply(partial(spline_interpolate_and_resample, num_samples=num_samples))
            x = np.array(data.train_data)
            y = np.array(data.train_labels)
            # setup the model
            mymodel = NaiveRegularizedDeepGRU(x.shape[1:])
            mymodel.batch_size = PARAM_BATCH_SIZE
            mymodel.num_epochs = PARAM_NUM_EPOCHS
            # run cross validation
            scores = cross_validate_model(x, y, mymodel, N_FOLDS)
            results[num_samples] = scores
```

```
Evaluating Spline interpolation using 300 samples
....
Cross validation fold [1]
```

```
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============ ] - 5s 3ms/step
categorical_accuracy: 95.89%
```

. . .

Cross validation fold [2]

. . .

```
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============ ] - 5s 3ms/step
categorical_accuracy: 95.00%
```

. . .

Cross validation fold [3]

. . .

```
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============ ] - 5s 3ms/step
categorical_accuracy: 91.72%
```

. . .

Cross validation fold [4]

. . .

```
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============ ] - 5s 3ms/step
categorical_accuracy: 97.44%
95.01% (+/- 2.09%)
```

Evaluating Spline interpolation using 200 samples

```
1800/1800 [============ ] - 3s 2ms/step
categorical_accuracy: 96.22%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
```

```
1800/1800 [============ ] - 3s 2ms/step
categorical_accuracy: 94.06%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
```

```
1800/1800 [============ ] - 3s 2ms/step
categorical_accuracy: 96.83%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
```

```
1800/1800 [============ ] - 3s 2ms/step
categorical_accuracy: 93.67%
95.19% (+/- 1.36%)
Evaluating Spline interpolation using 100 samples
______
Cross validation fold [1]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.8049 - categorical_accuracy:
Epoch 3/20
Epoch 4/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.3177 - categorical_accuracy:
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
```

```
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.00%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.8561 - categorical_accuracy:
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
```

```
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.89%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.7884 - categorical_accuracy:
Epoch 3/20
Epoch 4/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.1540 - categorical_accuracy:
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
```

```
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.00%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
5400/5400 [============== ] - 7s 1ms/step - loss: 1.7632 - categorical_accuracy:
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
```

```
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.44%
95.33% (+/- 0.37%)
Evaluating Spline interpolation using 75 samples
_____
Cross validation fold [1]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
```

```
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 729us/step
categorical_accuracy: 95.50%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
```

```
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 726us/step
categorical_accuracy: 96.50%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
```

```
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 722us/step
categorical_accuracy: 97.11%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
```

```
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 722us/step
categorical_accuracy: 91.00%
95.03% (+/- 2.40%)
Evaluating Spline interpolation using 50 samples
_____
Cross validation fold [1]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
```

```
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============= ] - 1s 506us/step
categorical_accuracy: 98.56%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
```

```
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============= ] - 1s 509us/step
categorical_accuracy: 96.06%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
```

```
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============= ] - 1s 507us/step
categorical_accuracy: 93.78%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
```

```
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============= ] - 1s 508us/step
categorical_accuracy: 94.33%
95.68% (+/- 1.86%)
Evaluating Spline interpolation using 25 samples
Cross validation fold [1]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
```

Epoch 11/20

```
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 291us/step
categorical_accuracy: 97.89%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
```

Epoch 7/20

```
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 291us/step
categorical_accuracy: 97.28%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
```

Epoch 7/20

```
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [=========== ] - 1s 293us/step
categorical_accuracy: 95.56%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
Epoch 4/20
Epoch 5/20
Epoch 6/20
```

Epoch 7/20

```
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
1800/1800 [============== - - 1s 290us/step
categorical_accuracy: 93.67%
96.10% (+/- 1.64%)
Evaluating Spline interpolation using 10 samples
-----
Cross validation fold [1]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
```

```
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 70.89%
Cross validation fold [2]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
```

```
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.78%
Cross validation fold [3]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
```

```
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 95.72%
Cross validation fold [4]
Train on 5400 samples, validate on 1800 samples
Epoch 1/20
Epoch 2/20
Epoch 3/20
```

```
Epoch 4/20
Epoch 5/20
Epoch 6/20
Epoch 7/20
Epoch 8/20
Epoch 9/20
Epoch 10/20
Epoch 11/20
Epoch 12/20
Epoch 13/20
Epoch 14/20
Epoch 15/20
Epoch 16/20
Epoch 17/20
Epoch 18/20
Epoch 19/20
Epoch 20/20
categorical_accuracy: 91.22%
88.40% (+/- 10.28%)
In [17]: results_df = pd.DataFrame([[key, np.mean(res), np.std(res)] for key,res in list(results
 results_df
Out[17]:
  Number of Samples Categorical Accuracy Std Deviation
  0
     300
         95.013889
             2.092157
  1
     200
         95.194444
             1.357705
  2
     100
         95.333333
             0.368514
             2.395501
  3
         95.027778
      75
```

4	50	95.680556	1.860230
5	25	96.097222	1.643532
6	10	88.402778	10.279232