

Assignment 4 Report

Anoop (2015CS10265)

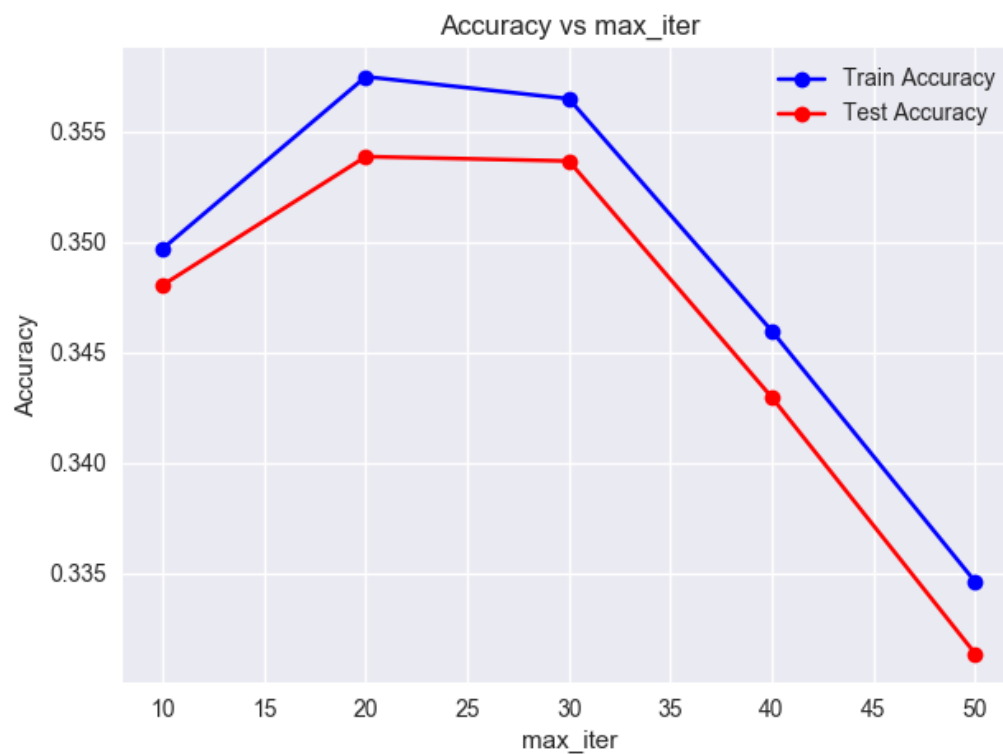
May 2, 2018

K-Means with Scikit-Learn

Parts (i) and (ii)

n_init	= 10	Train Accuracy	0.34279
n_clusters	= 20	Test Accuracy	= 0.34437
max_iter	= 300		

Part (iii)



The accuracy peaks when max_iter is set to 20. This is one of the local minima that the K-Means converges to. We can infer this from the fact that later on when max_iter is set to 300 (in the previous part), the accuracy increases as compared to other settings.

SVM with Dimensionality Reduction using PCA

Parameters:

kernel = 'rbf'
gamma = [1e-2, 1e-3]
C = [0.01, 0.1, 1, 10, 100]

Best Setting - C = 10 and gamma = 1e-2

Train Accuracy = 0.69376
Test Accuracy = 0.64714

Parameters:

kernel = 'linear'
C = [0.01, 0.1, 1, 10, 100]

Best Setting - C = 1

Train Accuracy = 0.68145
Test Accuracy = 0.63538

Hyperparameter Tuning - 3-Fold Grid Search Cross-Validation

SVM takes too long to train even after dimensionality reduction. Normalization is necessary for fast convergence. SVM gives decent performance on both train and test data.

Neural Networks with Keras

Parameters:

Number of Hidden Units = [2, 5, 10, 20, 50, 100, 200, 300, 500, 1000]
Number of Epochs = 50
Batch Size = 64

Hyperparameter Tuning - 3-Fold Grid Search Cross-Validation

Best Setting - Number of Hidden Units = 5
Best Model is trained for 100 more epochs after Cross-Validation

Train Accuracy = 0.81200
Test Accuracy = 0.66855

Convolutional Neural Networks with Keras

Parameters:

Number of Filters (3 X 3) = 32
Number of Hidden Units = 50 (Fully Connected Layer)
Number of Epochs = 100
Batch Size = 32

Train Accuracy = 0.95317
Test Accuracy = 0.81761

Comparison of Models

- K-Means
 - Training Time - 3s per model (Approx. 20s for CV)
 - Poor training and test accuracies.
- SVM with PCA
 - Training Time - 15 mins per model (Approx. 4 hrs for CV)
 - Decent training and test accuracies.
- Neural Networks
 - Training Time - 2s per epoch (Approx. 1 hr for CV) + 5 mins for extra 100 epochs [Used GPU]
 - Very good training accuracy but comparatively poor test accuracy. Significant Overfitting.
- Convolutional Neural Networks
 - Training Time - 5s per epoch (Approx. 10 mins for 100 epochs) [Used GPU]
 - Very good training accuracy but comparatively poor test accuracy. Significant Overfitting.

CNN performs well on the test data though it overfits. It takes comparatively lesser time to train than an SVM. Hyperparameter tuning in CNNs is difficult as compared to other models.

Best Model (Submitted in the Competition)

Key ideas used:

- Ensemble of CNNs is used to predict the test labels. Prediction of the ensemble is simply the most frequent prediction.
- Individual CNN is trained for 25 epochs. After 5 epochs, a snapshot of the model along with the test predictions are stored. This gives 5 different models per architecture.
Note: Only those architectures are considered which give very good train accuracy (≈ 0.90) after first 5 epochs.
- Data Augmentation using in-built Image Preprocessing techniques in Keras in one of the architectures (Not utilized completely).

Note: Individual architectures are not described here. Check the code for complete details of the architectures. Individual Test Accuracies are not calculated.

Train Accuracy = 0.93114

Test Accuracy = 0.92502

[Trained Models Link \(Only Best Models\)](#)