

Project Report-2:

Project Milestone2:

Tirumal Reddy Ramidi

U1414933

u1414933@uemail.utah.edu

- For the second milestone, I have used the Roberta features set among the three feature sets provided.
- Got the best results when I used this feature set for training my model.
- I have used the third-party libraries as allowed one per project to train my module.
- I have used TensorFlow to train and test the dataset.
- The TensorFlow neural network machine learning module gave me the highest accuracy so far in my project.
- The Highest accuracy I got after playing around with various dense and dropout values for the algorithm is: **83.289**
- `tf.keras.layers.Dense(1,activation = "sigmoid")`
- Sigmoid is the activation function I have used for this algorithm.
- `model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])`
- I have used the above code to compile it. To get the accuracy, we modified the metrics parameter, loss value here is changed to 'binary_crossentropy' since this is a binary classifier.
-
- Also, the other submission I have made is for a different dataset on the average perceptron algorithm.
- The dataset I have used here is tfidf dataset.
- Also, the Average perceptron would give the maximum accuracy. Hence I have used this algorithm.
- I have developed a linear classifier using the perceptron variation using average weights.
- For this algorithm to find the only hyperparameter used in this algorithm, I have passed all three values to the algorithm to get the best learning rate that gives the highest accuracy.
- The hyperparameter value that is the learning rate with the highest accuracy of 75.934054 is 1
- Using this best learning rate, I've trained a linear classifier where the epoch is 20. I retained the best weights with good accuracy on the test dataset.
- With this hyperparameter, I have passed the average weights learned from the training dataset to the prediction function to predict accuracy.
- Then read the eval dataset values into a pandas data frame, passed the eval_x, and averaged weights learned earlier to the eval_predict function. This predicts the labels on the evaluation dataset.
- Evaluation accuracy using the Roberta feature set is 75.131
-
- For the next milestone, I plan on using the Naive Bayes Classifiers, Decision Trees, and SVM.
- Looking forward to starting working on building linear classifiers using some of these exciting machine learning models.
- I plan on using the other two feature sets for the next milestone, and I am eager to work on them and improve the accuracy of the models for these feature sets. This enables me to analyze what type of classifier works best on certain feature sets.