

Q1. The accuracies for the different models and different thresholds is as follows:

Skip-Gram

Threshold is: 0.4	Accuracy Score: 0.9076923076923077
Threshold is: 0.5	Accuracy Score: 0.8307692307692308
Threshold is: 0.6	Accuracy Score: 0.6461538461538462
Threshold is: 0.7	Accuracy Score: 0.5538461538461539
Threshold is: 0.8	Accuracy Score: 0.7692307692307693

CBOW

Threshold is: 0.4	Accuracy Score: 0.8076923076923077
Threshold is: 0.5	Accuracy Score: 0.7
Threshold is: 0.6	Accuracy Score: 0.5769230769230769
Threshold is: 0.7	Accuracy Score: 0.5538461538461539
Threshold is: 0.8	Accuracy Score: 0.7769230769230769

Fasttext

Threshold is: 0.4	Accuracy Score: 0.8153846153846154
Threshold is: 0.5	Accuracy Score: 0.7435897435897436
Threshold is: 0.6	Accuracy Score: 0.5743589743589743
Threshold is: 0.7	Accuracy Score: 0.5538461538461539
Threshold is: 0.8	Accuracy Score: 0.7794871794871795

Glove

Threshold is: 0.4	Accuracy Score: 0.8461538461538461
Threshold is: 0.5	Accuracy Score: 0.7230769230769231
Threshold is: 0.6	Accuracy Score: 0.6307692307692307
Threshold is: 0.7	Accuracy Score: 0.5846153846153846
Threshold is: 0.8	Accuracy Score: 0.8153846153846154

The corresponding CSV files are present in the code folder

Q2. The train F1 on fine tuning is: 0.69

The test F1 score is: 0.48

Q3. From the plots we can see that it follows Zipfian Distribution