

CT24 dataset with political statements

| Partition | Model | Accuracy | Precision | f1 | Recall |
|------------------|--------------|-----------------|------------------|-----------|---------------|
| Test | random | 0.4575 | 0.2291 | 0.3071 | 0.4660 |
| | n-gram | 0.8094 | 0.7447 | 0.5185 | 0.3977 |
| | LSTM | 0.8047 | 0.6942 | 0.5443 | 0.4568 |
| Dev-Test | random | 0.5377 | 0.3818 | 0.4615 | 0.5833 |
| | n-gram | 0.8082 | 0.8406 | 0.6554 | 0.5370 |
| | LSTM | 0.8283 | 0.8602 | 0.6976 | 0.5917 |

Table 1: Performance of baselines on Test and Dev-Test

| Emdbeddings | Model | Accuracy | Precision | f1 | Recall |
|--------------------|--------------|-----------------|------------------|-----------|---------------|
| SentEmbeddings | LogReg | 0.84906 | 0.9286 | 0.7303 | 0.6019 |
| | SVM | 0.8554 | 0.9327 | 0.7326 | 0.5833 |
| | XGBoost | 0.7956 | 0.8909 | 0.6012 | 0.4537 |
| | AdaBoost | 0.8050 | 0.8833 | 0.6310 | 0.4907 |
| BoW | LogReg | 0.8019 | 0.8689 | 0.6272 | 0.4910 |
| | SVM | 0.7704 | 0.9487 | 0.5034 | 0.3426 |
| | XGBoost | 0.7610 | 0.8636 | 0.5 | 0.3519 |
| GloVe | LogReg | 0.7327 | 0.7674 | 0.4371 | 0.3056 |
| | SVM | 0.7327 | 0.8571 | 0.3529 | 0.2222 |
| | XGBoost | 0.7453 | 0.8140 | 0.4636 | 0.3241 |
| | AdaBoost | 0.7422 | 0.8250 | 0.4460 | 0.3056 |
| Word2Vec | LogReg | 0.7830 | 0.8305 | 0.5868 | 0.4537 |
| | SVM | 0.7987 | 0.8926 | 0.6098 | 0.4630 |
| | XGBoost | 0.7862 | 0.8704 | 0.5803 | 0.4352 |
| | AdaBoost | 0.7987 | 0.9074 | 0.6049 | 0.4537 |

Table 2: Performance of baselines on Dev-Test

| Emdbddings | Model | Accuracy | Precision | f1 | Recall |
|-------------------|--------------|-----------------|------------------|-----------|---------------|
| SentEmbeddings | LogReg | 0.7830 | 0.6522 | 0.4478 | 0.3409 |
| | SVM | 0.8123 | 0.7609 | 0.5224 | 0.3977 |
| | XGBoost | 0.7859 | 0.6471 | 0.4748 | 0.375 |
| | AdaBoost | 0.7801 | 0.6226 | 0.4681 | 0.375 |
| BoW | LogReg | 0.8123 | 0.7727 | 0.5151 | 0.3836 |
| | SVM | 0.7947 | 0.875 | 0.375 | 0.2386 |
| | XGBoost | 0.7771 | 0.6875 | 0.3667 | 0.25 |
| GloVe | LogReg | 0.7501 | 0.5366 | 0.3411 | 0.25 |
| | SVM | 0.7742 | 0.72 | 0.3186 | 0.2046 |
| | XGBoost | 0.7713 | 0.6389 | 0.3710 | 0.2614 |
| | AdaBoost | 0.7713 | 0.625 | 0.3906 | 0.2841 |
| Word2Vec | LogReg | 0.7771 | 0.6364 | 0.4242 | 0.3182 |
| | SVM | 0.8035 | 0.8 | 0.4553 | 0.3182 |
| | XGBoost | 0.8006 | 0.7273 | 0.4849 | 0.3636 |
| | AdaBoost | 0.8065 | 0.775 | 0.4844 | 0.3523 |

Table 3: Performance of baselines on gold-Test