|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Dataset | Experiment number | Model | Tokenizer | methods | accuracy | AUC | F1 |
| L-HSAB | Experiment 1 | arabertv02 | bert-base-arabertv02 | * Original dataset without any addition on the pipeline * deletion of the punctuation | 88% | 94.22% | 91% |
| Experiment 2 | arabertv02 | bert-base-arabertv02 | * Without removing punctuations * Without data augmentation | 88% | 94.35% | 91% |
| Experiment 3 | arabertv02 | bert-base-arabertv02 | * Without removing punctuations * Data augmentation with 2 rounds | 89% | 92.47% | 91% |
| AJCommentsClassification | Experiment 4 | arabertv02 | bert-base-arabertv02 | * Original dataset without any addition on the pipeline * deletion of the punctuation | 82.94% | 81.7% | 88% |
| Experiment 5 | arabertv02 | bert-base-arabertv02 | * without removing punctuations * without data augmentation | 82.72% | 81.42% | 89% |