

LIST OF TABLES

| | |
|--|----|
| Table 2.1 Comparing the method, parameters, performance, and weakness of 7 kinds of research in the text classification problem..... | 25 |
| Table 3.1 Example of a comment using a target account name that classified as hate speech..... | 29 |
| Table 3.2 Example of a comment using a direct name that classified as hate speech | 29 |
| Table 3.3 Example of a comment using the target account name and direct name that classified as hate speech..... | 31 |
| Table 3.4 Number of comment extracted from 21 political hashtags..... | 32 |
| Table 3.5 Number of comment extracted from 12 random hashtags | 33 |
| Table 3.6 Comment before pre-processing..... | 36 |
| Table 3.7 Comment after pre-processing..... | 36 |
| Table 3.8 Example of Indonesian language comments labeled as hate speech by the rule | 38 |
| Table 3.9 Comment before and after the lowercase process | 39 |
| Table 3.10 Comment before and after the stopword removal process | 40 |
| Table 3.11 Summary of parameter generated by the modified TextCNN..... | 45 |
| Table 3.12 Testing on dummy dataset..... | 48 |
| Table 4.1 Composition of hate and not hate speech text in imbalance dataset | 50 |
| Table 4.2 Result of modified TextCNN without balancing method on imbalance dataset experiment..... | 52 |
| Table 4.3 Result of modified TextCNN using 10-fold cross-validation on imbalance dataset experiment..... | 54 |
| Table 4.4 Composition of hate and not hate label before and after oversampling | 55 |
| Table 4.5 Result of modified TextCNN using random oversampling on imbalance dataset experiment..... | 56 |
| Table 4.6 Composition of hate and not hate label before and after undersampling | 58 |
| Table 4.7 Result of modified TextCNN using random undersampling on imbalance dataset experiment..... | 58 |
| Table 4.8 Value of weight added to the class | 61 |
| Table 4.9 Result of modified TextCNN using class weight value on imbalance dataset experiment | 61 |
| Table 4.10 Data composition in the proposed model and baseline model | 63 |
| Table 4.11 Performance comparison of the proposed model and baseline model | 64 |
| Table 4.12 Performances of modified TextCNN using Baseline Dataset | 65 |
| Table 4.13 Fasttext with 33 hashtags dataset result..... | 67 |
| Table 4.14 Performance comparison between the proposed model, baseline, and TextCNN using baseline dataset result | 68 |