

---

## LIST OF FIGURES

1.1	Difference between Lemmatization and Stemming . . . . .	2
1.2	General Research Design . . . . .	4
2.1	Stemming Classification . . . . .	19
3.1	HA Algorithm . . . . .	23
3.2	Step of Developing Modify Algorithm . . . . .	25
3.3	Research Design . . . . .	26
3.4	Research Design for Modify Algorithm . . . . .	27
3.5	Checking Dictionary HA Algorithm . . . . .	30
3.6	Checking Dictionary Proposed Method . . . . .	31
3.7	Preprocessing Dataset . . . . .	32
3.8	Output in CSV File . . . . .	34
4.1	Experiment Scenario . . . . .	37
4.2	Dataset in PDF File . . . . .	38
4.3	Conversion Dataset from PDF to Text . . . . .	39
4.4	Ready-processed Dataset . . . . .	39
4.5	Accuracy . . . . .	40
4.6	Distribution Lemmatization Result of Markesot Bertutur . . . . .	42
4.7	Distribution Lemmatization Result of Laskar Pelangi . . . . .	42
4.8	Distribution Lemmatization Result of Muhammad . . . . .	43
4.9	Distribution Lemmatization Result of Supernova . . . . .	43
4.10	Distribution Lemmatization Result of Kau, Aku, dan Sepucuk Angpau Merah . . . . .	44
4.11	Distribution Lemmatization Result of La Tahzan . . . . .	44
4.12	Distribution Lemmatization Result of Riyadush Shalihin . . . . .	45
4.13	Distribution Lemmatization Result of News . . . . .	45
4.14	Comparison Accuracy of HA Algorithm and Sastrawi . . . . .	47
4.15	Comparison Accuracy of Proposed Method and HA Algorithm . . . . .	48
4.16	Comparison Accuracy of Proposed Method and Sastrawi . . . . .	51
4.17	Foreign Language . . . . .	58
4.18	Name . . . . .	58
4.19	Acronym . . . . .	58
4.20	Two Words in One Phrase . . . . .	59
4.21	Not in Dictionary . . . . .	59
4.22	Not Word . . . . .	59
4.23	Not Raw . . . . .	60