

DAFTAR ISI

LEMBAR PENGESAHAN	1
LEMBAR PERNYATAAN KARYA SENDIRI.....	i
LEMBAR PERSEMBAHAN	ii
ABSTRAK	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	vi
DAFTAR TABEL	viii
BAB I	1
1.1 Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan dan Manfaat	5
1.4. Batasan Masalah	5
1.5. Kerangka Pemikiran	7
1.6. Metodologi Penelitian	8
1.7. Sistematika Penulisan	9
BAB II	11
2.1. Tinjauan Pustaka	11
2.2. Landasan Teori	17
1. Vaksinasi Covid 19.....	17
2. Analisis Sentimen.....	18
3. <i>Text Mining</i>	19
4. Klasifikasi.....	21
5. <i>Deep learning</i>	22
6. Convolutional Neural Network	22
7. <i>Web Scraping</i>	23
8. Twitter	24
9. <i>Confusion Matrix</i>	25
10. Python.....	26
11. CRISP-DM	27
BAB III	30
3.1. <i>Bussiness understanding</i>	30
3.2. <i>Data understanding</i>	33

a. <i>Data Requirement</i>	33
b. <i>Data Collection</i>	34
3.3. <i>Data Preparation</i>	37
a. <i>Labeling</i>	39
b. <i>Stopword Removing</i>	44
c. <i>Stemming</i>	44
d. <i>Tokenizing</i>	44
e. <i>Word Embedding</i>	47
3.4. <i>Modelling CNN</i>	48
3.5. <i>Evaluasi Model</i>	50
BAB IV	51
4.1. <i>Bussiness understanding</i>	51
4.2. <i>Data understanding</i>	52
4.3. <i>Data preparation</i>	53
4.4. <i>Word embedding</i>	56
4.1. <i>Modelling</i>	58
4.2. <i>Evaluasi Dataset Fase 1</i>	61
4.3. <i>Evaluasi Dataset Fase 2</i>	73
4.4. <i>Evaluasi Dataset Fase 3</i>	83
4.5. <i>Hasil Analisis Perubahan Sentimen Vaksinasi COVID-19 di Indonesia</i>	93
BAB V	97
5.1. <i>Simpulan</i>	97
5.2. <i>Saran</i>	98
REFERENSI	99

DAFTAR GAMBAR

Gambar 2. 1. Model Convolutional Neural Network [12]	23
Gambar 2. 2. Confusion Matrix [28].....	25
Gambar 2. 3 Metode CRISP-DM [30]	27
Gambar 3. 1. Flowchart Penelitian.....	32
Gambar 3. 2 Scraping URL tweet dengan snsrape.....	34
Gambar 3. 3 Hasil Scrape URL dalam bentuk .txt.....	35
Gambar 3. 4 Split URL dan siapkan elemen untuk scraping tweepy.....	36
Gambar 3. 5 Scrape dengan tweepy dan bypass API twitter limit dengan chunks	36
Gambar 3. 6 Dataset mentah hasil scraping	37
Gambar 3. 7 Jumlah data train positif dan negatif dataset fase 1	42
Gambar 3. 8. Jumlah data train positif dan negatif dataset fase 2.....	43
Gambar 3. 9. Jumlah data train positif dan negatif dataset fase 3.....	43
Gambar 3. 10. Word Emnedding FastText	48
Gambar 3. 11. Skema CNN	49
Gambar 4. 1. Visualisasi Similar Word	58
Gambar 4. 2. Build CNN Model Python.....	60
Gambar 4. 3. Split 80:20 fase 1 epochs 10.....	61
Gambar 4. 4. Split 70:30 fase 1 epochs 10.....	62
Gambar 4. 5. Split 60:40 fase 1 epochs 10.....	62
Gambar 4. 6. Classification Report 10 epoch fase 1	63
Gambar 4. 7. Model Accuracy and Loss 10 epochs fase 1	63
Gambar 4. 8. Confusion Matrix 10 epochs fase 1.....	64
Gambar 4. 9. Split 80 : 20 fase 1 epochs 50.....	65
Gambar 4. 10. Split 70 : 30 fase 1 epochs 50.....	65
Gambar 4. 11. Split 60 : 40 fase 1 epochs 50.....	65
Gambar 4. 12. Classification Report 50 epochs fase 1	65
Gambar 4. 13. Model Accuracy and Loss 50 epochs fase 1	66
Gambar 4. 14. Confusion Matrix 50 epochs fase 1	66
Gambar 4. 15. Split 80:20 fase 1 epochs 100.....	67
Gambar 4. 16. Split 70:30 fase 1 epochs 100.....	67
Gambar 4. 17. Split 60:40 fase 1 epochs 100.....	67
Gambar 4. 18. Classification Report 100 epochs fase 1	67
Gambar 4. 19. Model Accuracy and Loss 100 epochs fase 1	68
Gambar 4. 20. Confusion Matrix 100 epochs fase 1.....	69
Gambar 4. 21. Hasil Klasifikasi 489 data test fase 1	72
Gambar 4. 22. Hasil sentimen keseluruhan fase 1	72
Gambar 4. 23. Split 80:20 fase 2 epochs 10.....	73
Gambar 4. 24. Split 70:30 fase 2 epochs 10.....	74
Gambar 4. 25. Split 60:40 fase 2 epochs 10.....	74
Gambar 4. 26. Report classification 10 epochs dataset fase 2	74
Gambar 4. 27. Model Accuracy and Loss 10 epochs fase 2	75

Gambar 4. 28. Confusion matrix 10 epochs fase 2	75
Gambar 4. 29. Split 80:20 fase 2 epochs 50.....	76
Gambar 4. 30. Split 70:30 fase 2 epochs 50.....	76
Gambar 4. 31. Split 60:40 fase 2 epochs 50.....	76
Gambar 4. 32. Classification report 50 epochs dataset fase 2.....	76
Gambar 4. 33. Model Accuracy and Loss 50 epochs fase 2	77
Gambar 4. 34. Confusion Matrix 50 epochs fase 2.....	78
Gambar 4. 35. Split 80:20 fase 2 epochs 100.....	78
Gambar 4. 36. Split 70:30 fase 2 epochs 100.....	78
Gambar 4. 37. Split 60:40 fase 2 epochs 100.....	78
Gambar 4. 38. Classification report 100 epochs fase 2.....	79
Gambar 4. 39. Model Accuracy and loss 100 epochs fase 2.....	79
Gambar 4. 40. Confusion Matrix 100 epochs fase 2.....	80
Gambar 4. 41. Hasil klasifikasi 492 data test fase 2	82
Gambar 4. 42. Hasil sentimen keseluruhan fase 2	82
Gambar 4. 43. Split 80:20 fase 3 epochs 10.....	83
Gambar 4. 44. Split 70:30 fase 3 epochs 10.....	83
Gambar 4. 45. Split 60:40 fase 3 epochs 10.....	83
Gambar 4. 46. Classification Report 10 epochs fase 3	84
Gambar 4. 47. Model Accuracy and Loss 10 epochs fase 3	84
Gambar 4. 48. Confusion matrix 10 epochs fase 3	85
Gambar 4. 49. Split 80:20 fase 3 epochs 50.....	85
Gambar 4. 50. Split 70:30 fase 3 epochs 50.....	85
Gambar 4. 51. Split 60:40 fase 3 epochs 50.....	86
Gambar 4. 52. Classification Report 50 epochs fase 3.....	86
Gambar 4. 53. Model Accuracy and Loss 50 epochs fase 3	87
Gambar 4. 54. Confusion Matrix 50 epochs fase 3.....	87
Gambar 4. 55. Split 80:20 fase 3 epochs 100.....	87
Gambar 4. 56. Split 70:30 fase 3 epochs 100.....	88
Gambar 4. 57. Split 60:40 fase 3 epochs 100.....	88
Gambar 4. 58. Classification report 100 epochs fase 3.....	88
Gambar 4. 59. Model Accuracy and Loss 100 epochs fase 3	89
Gambar 4. 60. Confusion matrix 100 epochs fase 3	89
Gambar 4. 61. Hasil klasifikasi 423 data test fase 3	92
Gambar 4. 62. Hasil sentimen keseluruhan fase 3	92
Gambar 4. 63. Prediksi Sentimen oleh Model	93
Gambar 4. 64. Hasil Analisis Sentimen berdasarkan seri waktu	94
Gambar 4. 65. (kiri ke kanan) Wordcloud positif fase 1, 2, 3.....	95
Gambar 4. 66. (kiri ke kanan) Wordcloud negatif fase 1, 2, 3.....	95

DAFTAR TABEL

Tabel 2. 1. Penelitian Sebelumnya.....	13
Tabel 3. 1 Dataset fase pertama sebelum <i>labeling</i>	37
Tabel 3. 2 Dataset fase kedua sebelum <i>labeling</i>	38
Tabel 3. 3 Dataset Fase ketiga sebelum <i>labeling</i>	39
Tabel 3. 4 Dataset Fase Pertama Setelah <i>Labeling</i>	40
Tabel 3. 5 Dataset fase kedua setelah <i>labeling</i>	41
Tabel 3. 6 Dataset fase ketiga setelah <i>labeling</i>	41
Tabel 3. 7 Preprocessing stopword, <i>stemming</i> , <i>tokenizing</i>	45
Tabel 4. 1. Jumlah split data tiap fase.....	52
Tabel 4. 2. Preprocessing dataset fase 1.....	53
Tabel 4. 3. Preprocesssing dataset fase 2.....	54
Tabel 4. 4. Preprocessing dataset fase 3.....	55
Tabel 4. 5. Hasil Word Vector FastText.....	57
Tabel 4. 6. Similar Word Fast Text.....	57
Tabel 4. 7 Build CNN Model Python.....	58
Tabel 4. 8. Source Code Report Classification.....	62
Tabel 4. 9. Akurasi Model 3 Pengujian epochs fase 1.....	69
Tabel 4. 10. Loss Model 3 Pengujian epochs fase 1.....	70
Tabel 4. 11. Hasil prediksi data test fase 1.....	70
Tabel 4. 12. Save Model CNN.....	73
Tabel 4. 13. Load Model CNN.....	73
Tabel 4. 14. Akurasi Model 3 pengujian fase 2.....	80
Tabel 4. 15. Loss Model 3 pengujian fase 2.....	81
Tabel 4. 16. Hasil prediksi data test fase 2.....	81
Tabel 4. 17. Akurasi Model 3 pengujian epochs fase 3.....	90
Tabel 4. 18. Loss Model 3 pengujian fase 3.....	90
Tabel 4. 19. Hasil Prediksi Data test fase 3.....	90