

DAFTAR GAMBAR

Gambar I-1. Kota dengan peringkat kualitas udara terburuk di Indonesia.	1
Gambar II-1. Logo PT Nafas Indonesia.	7
Gambar II-2. Negara dengan pengguna Twitter terbanyak.	8
Gambar II-3. Tahapan KDD.	9
Gambar II-4. Jenis-jenis <i>machine learning</i> (Yao & Zheng, 2023).	12
Gambar II-5. Ilustrasi algoritma SVM.	13
Gambar II-6. Ilustrasi <i>oversampling</i> , <i>undersampling</i> , dan <i>hybrid sampling</i>	22
Gambar II-7. <i>Confusion matrix</i>	26
Gambar II-8. Contoh implementasi kurva ROC pada algoritma SVM (Meng dkk., 2014).	28
Gambar III-1. Sistematika penyelesaian masalah.	39
Gambar IV-1. <i>Script crawling</i> data komentar Twitter.	47
Gambar IV-2. Hasil pelabelan data komentar.	50
Gambar IV-3 <i>Script Case Folding</i>	51
Gambar IV-4. <i>Script cleansing</i>	52
Gambar IV-5. <i>Script tokenizing</i>	54
Gambar IV-6. <i>Script normalization</i>	55
Gambar IV-7. <i>Script Normalization 2</i>	56
Gambar IV-8. <i>Script normalization 3</i>	56
Gambar IV-9. <i>Script stopword removal</i>	58
Gambar IV-10. <i>Script stemming</i>	59
Gambar IV-11. Frekuensi AQI per kategori.	61
Gambar IV-12. Script TF-IDF.	63
Gambar IV-13. <i>Script encoding</i>	65
Gambar IV-14. <i>Script data splitting</i>	66
Gambar IV-15. <i>Script imbalance data handling</i>	66
Gambar IV-16. Visualisasi perbandingan distribusi sentimen antara (a) sebelum dan (b) setelah penerapan metode SMOTE-Tomek.	67
Gambar IV-17. <i>Script algoritma SVM</i>	68
Gambar IV-18. <i>Script confusion matrix</i>	70
Gambar IV-19. <i>Script evaluation metric</i>	71

Gambar IV-20. <i>Script</i> kurva ROC.	72
Gambar V-1. <i>Confusion matrix</i>	73
Gambar V-2. Kurva ROC.	76
Gambar V-3. Distribusi NSS.....	77
Gambar V-4. Distribusi sentimen pada uji Kruskal-Wallis.	83
Gambar V-5. <i>Unigram</i> sentimen positif.....	86
Gambar V-6. <i>Unigram</i> sentimen negatif.....	86
Gambar V-7. <i>Bigram</i> sentimen positif.	87
Gambar V-8. <i>Bigram</i> sentimen negatif.	88