Perbandingan Performa Algoritma Dalam Klasifikasi Sentimen Ulasan Hotel

ATHALLAH ANARGYA Y.

athallahanargya@student.telkomuniversity.ac.id Data Science School of Computing Telkom University

AZKA NADHIRA

azkanadhiraa@student.telkomuniversity.ac.id Data Science School of Computing Telkom University

CETRIN AZAHRA

cetrinazahra@student.telkomuniversity.ac.id Data Science School of Computing Telkom University

Dr. WARIH MAHARANI, S.T., M.T.

wmaharani@telkomuniversity.ac.id Data Science School of Computing Telkom University

1

Latar Belakang

Ulasan hotel online telah menjadi sumber informasi yang berharga bagi wisatawan dalam membuat keputusan pemesanan hotel. Analisis sentimen merupakan salah satu cara memberikan pemahaman mengenai kinerja hotel.



5

Dataset



Sumber Kaggle 20.492 Data (Bahasa Inggris) skala 1 - 5.



Luaran









Referensi

[1] A. H. Hasugian, M. Fakhriza, and D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naïve Bayes," Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD, vol. 6, no. 1, pp. 98–107, 2023, [Online]. Available: https://ojs.trigunadharma.ac.id/index.php/jsk/index

[2] V. W. D. Thomas and F. Rumaisa, "Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 6, no. 3, pp. 1767–1774, Jul. 2022, doi: 10.30865/mib.v6i3.4218.

2 Tujuan



Membandingkan performa 2 algoritma Machine Learning dalam mengklasifikasikan sentimen ulasan hotel untuk mengidentifikasi faktor-faktor utama yang mempengaruhi sentimen pelanggan terhadap layanan dan fasilitas hotel.



Metode

Model dilatih menggunakan algoritma XGBoost dan Logistic Regression. Kinerja model dievaluasi dengan menggunakan Confusion Matrix untuk mengukur dan memvalidasi hasil prediksi.





Hasil dan Diskusi

Berikut hasil evaluasi dari algoritma yang digunakan:



Berdasarkan visualisasi tersebut, algoritma XGBoost memiliki akurasi tertinggi, yaitu 91,29%. Hal ini berarti bahwa algoritma XGBoost lebih akurat dalam memprediksi hasil dibandingkan algoritma lainnya.