

ABSTRAK

Transportasi ojek online telah berkembang pesat di Indonesia sebagai bagian dari transformasi digital ekonomi. Status penyedia layanan ojek online dengan driver ojek online merupakan mitra sebagaimana diatur dalam Pasal 26 (huruf F) Undang-Undang Nomor 20 Tahun 2008 tentang Usaha Mikro, Kecil dan Menengah. Data tahun 2017 menunjukkan tingkat adopsi layanan ojek online melampaui 70%, dengan perkiraan 17.000 hingga 22.000 mitra aktif di ibu kota. Akibat perkembangannya yang begitu cepat, jumlah mitra yang membludak menimbulkan masalah persaingan ketat antar mitra untuk mendapatkan pesanan. Penelitian ini bertujuan untuk membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbors (KNN) dalam mengklasifikasikan lokasi penarikan pesanan yang optimal berdasarkan pola waktu. Penelitian ini menggunakan dataset “Online Taxis Pontianak Indonesia” dari Kaggle dengan fokus pada pengolahan fitur temporal untuk menghindari kebocoran data dari fitur koordinat. Berdasarkan hasil pengujian, KNN terbukti lebih unggul dengan akurasi tertinggi sebesar 27,45% pada skenario pembagian data 90:10 dengan nilai $k=3$. Hal ini menunjukkan bahwa KNN lebih efektif dalam menangkap pola lokal hubungan antara waktu dan lokasi jika didukung data latih yang besar. Sebaliknya, Naïve Bayes menunjukkan performa yang stagnan dengan akurasi 25% pada seluruh skenario akibat bias terhadap kelas mayoritas. Penelitian ini menyimpulkan bahwa model KNN lebih direkomendasikan sebagai landasan sistem rekomendasi bagi mitra ojek online dibandingkan probabilitas acak yang hanya sebesar 16,6% karena mampu memberikan prediksi yang lebih presisi.

Kata kunci—Naïve Bayes, K-Nearest Neighbor, Transportasi, Ojek Online, Dataset, Klasifikasi, Lokasi.

ABSTRACT

Online motorcycle taxi services (ojek online) have grown rapidly in Indonesia as part of the digital economic transformation. The partnership status between service providers and drivers is regulated under Article 26 (letter F) of Law Number 20 of 2008 concerning Micro, Small, and Medium Enterprises (MSMEs). Data collected in 2017 indicated that the adoption rate of online motorcycle taxi services exceeded 70%, estimating between 17,000 to 22,000 active partners in the capital city. Due to this rapid development and the open registration system, the surging number of partners has created intense internal competition among drivers to secure ride orders. This study aims to compare the performance of Naïve Bayes and K-Nearest Neighbors (KNN) algorithms in classifying optimal locations and times for order pickups by drivers. The study utilizes the "Online Taxis Pontianak Indonesia" dataset from Kaggle, which contains geospatial information (latitude and longitude). Test results demonstrate that the Naïve Bayes algorithm is superior, achieving a highest accuracy of 63% at a 70:30 training data ratio, due to its ability to stably capture spatial-temporal probability patterns. Conversely, the performance of KNN tends to be lower and highly dependent on training data density, where the highest result at $k=3$ with a 90:10 ratio only yielded an average F1-score of 37%. This study concludes that the Naïve Bayes model is more recommended as the foundation for a location and time recommendation system to assist partners in increasing their chances of securing orders.

Keywords —*Naïve Bayes, K-Nearest Neighbor, Transportation, Online Ojek, Dataset, Latitude, Longitude.*