

ABSTRAK

Satvika, Gd. Aditya Jana (2024), Meningkatkan Kinerja *K-Nearest Neighbor* Menggunakan *Permutation Feature Importance* Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I: Dr. I Nyoman Sukajaya, M.T. dan Pembimbing II: Dr. I Gede Aris Gunadi, S.Si., M.Kom.

Kata-kata kunci: Klasifikasi, *K-Nearest Neighbor*, *Permutation Feature Importance*

Penelitian ini bertujuan untuk meningkatkan kinerja algoritma *K-Nearest Neighbor* (KNN) dalam memprediksi kelulusan tepat waktu mahasiswa. Penggunaan dataset dari data kelulusan mahasiswa program studi Pendidikan Teknik Informatika Universitas Pendidikan Ganesha (PTI Undiksha) angkatan 2014-2017 sebanyak 460 data membutuhkan proses data *pre-processing* untuk memperoleh data yang lebih berkualitas. Penelitian ini membandingkan kinerja KNN dengan penerapan metode *permutation feature importance* (PFI) untuk meningkatkan kinerja model. Hasil penelitian menunjukkan implementasi PFI pada model KNN mencapai kinerja terbaik saat menerapkan *oversampling* pada *training* data menggunakan rasio jumlah sampel kelas minoritas sebesar 25% dari jumlah sampel kelas mayoritas. *Accuracy score* meningkat dari 90,22% menjadi 92,39%, sementara *recall score* tetap 55,56%, *precision score* meningkat dari 50,00% menjadi 62,50%, dan *f1-score* meningkat dari 52,63% menjadi 58,82%. Dari analisis PFI, teridentifikasi bahwa IPS 1 memiliki kontribusi terendah terhadap model yang digunakan. Berdasarkan penelitian ini, terdapat rekomendasi praktis. Pertama, lakukan pengujian terhadap PFI secara berulang dengan *n-shuffle* yang berbeda untuk mendapatkan pola kontribusi fitur yang lebih konsisten. Kedua, evaluasi lebih lanjut diperlukan untuk menentukan fitur dengan *importance score* terendah yang perlu dihapus guna mengoptimalkan kinerja model.

ABSTRACT

Satvika, Gd. Aditya Jana (2024), *Optimizing K-Nearest Neighbor Performance Using Permutation Feature Importance For On-time Graduation Prediction of Universitas Pendidikan Ganesha Students.*

This thesis has been approved and reviewed by Advisor I: Dr. I Nyoman Sukajaya, M.T. and Advisor II: Dr. I Gede Aris Gunadi, S.Si., M.Kom.

Keywords: Classification, K-Nearest Neighbor, Permutation Feature Importance

This study aimed to improve the performance of the K-Nearest Neighbor (KNN) algorithm in predicting students' on-time graduation. The usage of datasets from the 2014–2017 Informatics Engineering Education program, Universitas Pendidikan Ganesha (Pendidikan Teknik Informatika or PTI Undiksha) student graduation data, comprising 460 data points, required data pre-processing to obtain higher-quality data. This study compared the performance of KNN with the application of the permutation feature importance (PFI) method to improve model performance. The results of this study revealed that the implementation of PFI in the KNN model achieved the best performance when oversampling the training data by using a minority class sample ratio of 25% of the number of majority class samples. The accuracy score increased from 90,22% to 92,39%, while the recall score remained at 55,56%, the precision score increased from 50,00% to 62,50%, and the f1-score increased from 52,63% to 58,82%. From the PFI analysis, it was identified that IPS 1 had the lowest contribution to the model used. Based on this research, practical recommendations were suggested. Firstly, it was recommended to repeatedly use PFI with different n-shuffles to obtain more consistent feature contribution patterns. Secondly, further evaluation was suggested to determine which features with the lowest importance score should be removed to optimize model performance.