

ABSTRAK

Jeane Maya Parinding, 2024. Implementasi Metode Naive Bayes Untuk Memprediksi Hasil Penerimaan Peserta Didik Baru di SMA Negeri 3 Yogyakarta Tahun 2023. Program Studi Pendidikan Matematika. Jurusan Pendidikan Matematika dan Ilmu Pengetahuan Alam. Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sanata Dharma

Penelitian ini bertujuan untuk menerapkan metode *Naive Bayes* dalam membentuk sebuah model prediksi untuk data PPDB di SMA Negeri 3 Yogyakarta berdasarkan nilai ASPD. Selain itu, penelitian bertujuan untuk mengetahui faktor-faktor yang mempengaruhi hasil PPDB di SMA Negeri 3 Yogyakarta.

Penelitian ini menerapkan jenis penelitian studi kasus dengan menggunakan metode *Naive Bayes*. Data yang digunakan berasal dari Dinas Pendidikan, Pemuda, dan Olahraga DIY. Teknik analisis yang digunakan adalah pemilihan data, pembersihan data, transformasi data, *data mining*, dan evaluasi. Penelitian ini menggunakan tiga seleksi fitur yang sama pada setiap jalur pendaftaran yaitu tanpa menggunakan seleksi fitur, menggunakan *Select K-Best*, dan menggunakan PCA.

Model terbaik berdasarkan jalur zonasi diperoleh akurasi sebesar 0.94 tanpa seleksi fitur dan menggunakan *Select K-Best* dengan faktor yang mempengaruhinya yaitu zona, pilihan, nilai prestasi, literasi numerik, literasi baca Bahasa Indonesia, literasi baca Bahasa Inggris, literasi sains, akreditasi sekolah asal, dan rata-rata raport 4 mata pelajaran. Jalur afirmasi diperoleh akurasi sebesar 0.90 menggunakan seleksi fitur PCA dengan faktor yang mempengaruhinya yaitu zona, pilihan, nilai prestasi, literasi numerik, literasi baca Bahasa Indonesia, literasi baca Bahasa Inggris, literasi sains, akreditasi sekolah asal, dan rata-rata raport 4 mata pelajaran. Jalur prestasi diperoleh akurasi sebesar 0.83 menggunakan fitur seleksi PCA dengan faktor yang mempengaruhinya yaitu pilihan, nilai prestasi, literasi numerik, literasi baca Bahasa Indonesia, literasi baca Bahasa Inggris, literasi sains, akreditasi sekolah asal, dan rata-rata raport 4 mata pelajaran. Jalur perpindahan tugas orang tua/wali diperoleh akurasi sebesar 1 tanpa menggunakan seleksi fitur dengan faktor yang mempengaruhinya yaitu pilihan, nilai prestasi, literasi numerik, literasi baca Bahasa Indonesia, literasi baca Bahasa Inggris, literasi sains, akreditasi sekolah asal, dan rata-rata raport 4 mata pelajaran.

Kata kunci: *data mining, naive bayes, ppdb*

ABSTRACT

Jeane Maya Parinding, 2024. ImplementationOf Naive Bayes Method To Predict The Result Of Admission New Students In SMA Negeri 3 Yogyakarta Year 2023. Mathematics Education Study Program. Department of Mathematics and Natural Sciences Education. Faculty of Teacher Training and Education, Sanata Dharma University.

This research aims to apply the Naive Bayes method in forming a prediction model for PPDB data at SMA Negeri 3 Yogyakarta based on ASPD scores. In addition, the research aims to determine the factors that influence the results of PPDB at SMA Negeri 3 Yogyakarta.

This research applies a type of case study research using the Naive Bayes method. The data used comes from the Yogyakarta Education, Youth and Sports Office. The analysis techniques used are data selection, data cleaning, data transformation, data mining, and evaluation. This research uses the same three feature selections in each registration pathway, namely without using feature selection, using Select K-Best, and using PCA.

The best model based on the zoning pathway obtained an accuracy of 94% without feature selection and using Select K-Best with the influencing factors of zone, choice, achievement score, numerical literacy, Indonesian reading literacy, English reading literacy, science literacy, accreditation of the school of origin, and average report card of 4 subjects. The affirmation pathway obtained an accuracy of 90% using PCA feature selection with influencing factors namely zone, choice, achievement score, numerical literacy, Indonesian reading literacy, English reading literacy, science literacy, accreditation of the school of origin, and average report card of 4 subjects. The achievement pathway obtained an accuracy of 83% using the PCA selection feature with the influencing factors being choice, achievement score, numerical literacy, Indonesian reading literacy, English reading literacy, science literacy, accreditation of the school of origin, and average report card of 4 subjects. The parent/guardian assignment transfer pathway obtained an accuracy of 100% without using feature selection with the influencing factors being choice, achievement score, numerical literacy, Indonesian reading literacy, English reading literacy, science literacy, accreditation of the school of origin, and average report card of 4 subjects.

Keywords: data mining, naive bayes, ppdb