

## ABSTRACT

MARTIN BINSARTUA (01679220004)

### **IDENTIFICATION OF SIGNIFICANT FACTORS USING NOMOGRAM VISUALISATION AND MACHINE LEARNING MODELLING FROM DIAGNOSTIC TEST RESULTS OF PRIMARY SCHOOL STUDENTS**

(xiii + 46 pages; 9 figures; 6 table; 1 attachment)

The objective of this research is to utilize nomograms for visualizing data trends and to apply machine learning models—Logistic Regression, Naïve Bayes, Neural Network, and Random Forest—to evaluate the differences in accuracy when predicting primary school students' diagnostic test outcomes. This identification is crucial for optimizing the learning process of primary school pupils through diagnostic assessments. The goal of this identification is to support the learning process, thereby enabling students to achieve their maximum potential.

This research commences with the collection of datasets containing information on diagnostic test results and student backgrounds. In the subsequent stage, the data is analyzed using ROC analysis and performance curves to assess the efficacy of various machine learning models. In the final stage, nomograms are employed to model and visualize the predictor variables and the probabilities associated with achieving the passing grade.

The purpose of this research is to compare the performance of Logistic Regression, Naïve Bayes, Neural Networks, and Random Forest machine learning models. The results obtained using the Orange application indicate that the Random Forest model performs the best, with an AUC score of 0.80, a classification accuracy of 0.72, and an F1 score of 0.72. After this evaluation, the final stage of the research involves visualizing the results with a nomogram model. These findings offer insights into the parameters that influence the diagnostic test outcomes for elementary school students.

Keywords: *Logistic Regression, Naïve Bayes, Neural Network, Random Forest*  
Nomogram, , Identification, Diagnostic Tests

References 69 (2003 – 2023)

## ABSTRAK

MARTIN BINSARTUA (01679220004)

### IDENTIFIKASI FAKTOR SIGNIFIKAN DENGAN MENGGUNAKAN VISUALISASI NOMOGRAM DAN PEMODELAN MACHINE LEARNING DARI HASIL TES DIAGNOSTIK SISWA SEKOLAH DASAR

(xiii + 46 halaman; 9 gambar; 6 tabel; 1 lampiran)

Tujuan dari penelitian ini adalah menggunakan nomogram untuk visualisasi tren data dan menerapkan model *machine learning*; *Logistic Regression*, *Naïve Bayes*, *Neural Network*, dan *Random Forest* untuk mendapatkan perbedaan akurasi dalam melakukan prediksi hasil tes diagnostik siswa sekolah dasar. Identifikasi ini sangat penting untuk memaksimalkan pembelajaran siswa melalui tes diagnostik pada siswa sekolah dasar. Tujuan dilakukan identifikasi adalah untuk membantu proses pembelajaran siswa agar dapat mencapai hasil siswa yang maksimal.

Penelitian ini diawali dengan mengumpulkan *data set* yang berisikan informasi terkait dengan hasil tes diagnostik beserta data informasi terkait latar belakang siswa. Pada tahap selanjutnya dilakukan pemeriksaan data dengan melakukan evaluasi menggunakan *ROC analysis* dan *performance curve* untuk melakukan evaluasi terhadap penggunaan *machine learning* yang digunakan. Pada tahap akhir, dilakukan pemodelan dengan menggunakan nomogram untuk mendapatkan visualisasi terkait dengan *variabel predictor* dengan peluang yang dihasilkan berdasarkan kelulusan.

Hasil dari penelitian ini adalah membandingkan hasil kinerja model *machine learning* *Logistic Regression*, *Naïve Bayes*, *Neural Network*, dan *Random Forest*. Dari hasil yang dilakukan dengan menggunakan aplikasi *orange* didapatkan bahwa performa dari model *Random Forest* lebih baik. Hasil ini ditunjukkan melalui skor AUC 0.80, *Classification Accuracy* 0.72, dan *F1 Score* 0.72. Setelah dilakukan evaluasi, tahap akhir penelitian adalah melakukan visualisasi dengan menerapkan model nomogram. Dari hasil tersebut memberikan informasi terkait dengan parameter – parameter yang mempengaruhi hasil tes diagnostik pada siswa sekolah dasar.

Kata Kunci: *Logistic Regression*, *Naïve Bayes*, *Neural Network*, *Random Forest*  
Nomogram, Identifikasi, Tes Diagnostik

Referensi 69 (2003 – 2023)