

## ABSTRAK

Nama : Ria Arafiyah  
Program Studi : Ilmu Komputer  
Judul : Personalisasi Pembelajaran Berbasis Kemajuan Belajar Siswa  
Menggunakan *Machine Learning*

Kegagalan belajar siswa dapat dihindari, jika pembelajaran diperbaiki sedini mungkin dengan adanya prediksi kemajuan belajar. Data mengenai pembelajaran siswa di Indonesia cukup banyak dan tersimpan diberbagai tempat. Data Ini dapat digunakan untuk menjawab permasalahan ini. Diperlukan upaya untuk mengumpulkan, menyusun, dan membangun model dari dataset siswa yang dapat digunakan digunakan untuk mencegah kegagalan siswa dalam belajar.

Penelitian ini membahas penyusunan dataset kemajuan belajar siswa, yang akan digunakan untuk membuat model kemajuan belajar (*learning progress*) dengan menggunakan *machine learning*. Selanjutnya model ini akan digunakan untuk mendesain *Early Warning System (EWS)* yang menunjang personalisasi pembelajaran.

Dataset disusun dari berbagai sumber, yaitu data nilai rapor dan latar belakang siswa yang tersimpan di database DAPODIK, data nilai standar akreditasi sekolah yang tersimpan di BAN S/M, dan data nilai proses belajar siswa yang tersimpan pada database e\_Rapor. Ketiga sumber data dihubungkan dengan menggunakan NISN (Nomor Induk Siswa Nasional) dan NPSN (Nomor Pokok Sekolah Nasional). Hasil penggabungan diperoleh dataset Kemajuan belajar siswa selama di SMA yang berukuran 11.319 baris, 61 kolom. Data ini menjadi bahan untuk membuat dataset kemajuan belajar siswa. Dilakukan preprocessing untuk mendapatkan dataset yang siap dimodelkan dengan machine learning. Dalam membuat model dilakukan *feature extraction* dan *feature selection* dengan menggunakan PCA, K-Mean, dan Random Forest sehingga diperoleh model kemajuan belajar dan faktor-faktor yang memengaruhinya.

Pembangunan model menggunakan pendekatan *hybrid*, yaitu *unsupervised* dan *supervised*. Pemodelan *unsupervised* menggunakan metode pengklasteran algoritma K-Mean dengan validasi metode elbow dan silhouette score. Hasil kluster digunakan sebagai target pada model klasifikasi yang dibangun menggunakan metode *supervised* algoritma Random Forest dengan validasi menggunakan *K-Fold Cross Validation*. Model yang dihasilkan digunakan untuk mendesain EWS yang dapat digunakan untuk memonitor kemajuan belajar dan mempersonalisasi pembelajaran. Desain EWS divalidasi menggunakan metode Delphi.

Dengan menggunakan model tersebut dapat diprediksi kluster kemajuan belajar, faktor yang memengaruhinya dan saran untuk mempersonalisasi pembelajaran. Ketiga model yang berhasil dibangun memiliki rata-rata *Area Under*

*Curve* (AUC) lebih dari 99%. Desain EWS berbasis model kemajuan belajar yang akan menjadi dasar untuk mengembangkan personalisasi pembelajaran sehingga semua pembelajar tidak tertinggal dan lulus matapelajaran yang diikutinya. Desain ini menggunakan model kemajuan belajar untuk memonitor pembelajar dan memberikan saran tindakan (*suggested respond*). Validasi desain menggunakan metode Delphi 1 putaran.

Kata Kunci: Dataset, Model Kemajuan belajar, *Machine Learning*, *Early Warning Systems*, Monitoring, Personalisasi Pembelajaran

## ABSTRACT

Name : Ria Arafiyah  
Study Program : Computer Science  
Title : *Personalized Learning Based on Student Learning Progress Using Machine Learning*  
Keywords : Dataset, Modeling Learning Progress, Machine Learning, Early Warning Systems, Monitoring, Personalized Learning

Student learning failures can be avoided if learning is corrected as early as possible with predictions of learning progress. Data regarding student learning in Indonesia is quite a lot and stored in various places. This data can be used to answer this problem. Efforts are needed to collect, compile, and build models from student datasets that can be used to prevent student failure in learning.

This study discusses the preparation of student learning progress datasets, which will be used to create learning progress models using machine learning. Furthermore, this model will be used to design the EWS (Early Warning System) to support learning monitoring automation.

From the value data and student background data stored in the DAPODIK database, the school accreditation data stored in the BAN S/M obtained data measuring 11,319 rows, and 61 columns. This data is used as material to create a dataset of student learning progress. Preprocessing is done to get a dataset that is ready to be modeled with machine learning. In making the model, feature extraction and feature selection were carried out using PCA, K-Mean, and Random Forest in order to obtain a model of learning progress and the factors that influence it.

The model development uses a hybrid approach, namely unsupervised and supervised. Unsupervised modeling uses the K-Mean algorithm clustering method with elbow and silhouette score method validation. The results of the cluster are used as targets in the classification model built using the supervised method of the Random Forest algorithm with validation using K-Fold Cross Validation. The resulting model is used to design an Early Warning System (EWS) that can be used to monitor learning progress. The EWS design is validated using the Delphi methods.

By using this model, clusters of learning progress can be predicted, the factors that influence them, and suggestions for personalizing learning. The three successfully built models have an average Area Under Curve (AUC) of more than 99%. The EWS design is based on the learning progress model which will be the basis for developing personalized learning so that all students are not left behind and pass the subjects they take. This design uses a learning progress model to monitor students and provides suggested responses—design validation using the Delphi method 1 round.

Keywords:

Dataset, Learning Progress Model, Machine Learning, Monitoring