

# End-to-end Machine Learning dan Visualisasi Berbasis Web Untuk Mahasiswa Drop Out Menggunakan Data dari Pangkalan Data Pendidikan Tinggi (PDDikti) = Web-based End-to-end Machine Learning and Visualization for Undergraduate Student Drop Out by Utilizing Higher Education Data from Pangkalan Data Pendidikan Tinggi (PDDikti)

Reynard Adha Ryanda, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920521908&lokasi=lokal>

---

## Abstrak

Mahasiswa drop out memiliki dampak negatif untuk mahasiswa serta perguruan tinggi. Mahasiswa dikatakan drop out apabila mahasiswa tersebut belum dapat menyelesaikan masa studinya dalam rentang waktu yang telah ditentukan. Data mengenai status penyelesaian serta data akademis mahasiswa terekap pada Pangkalan Data Pendidikan Tinggi (PDDikti). Berdasarkan data tersebut, mahasiswa drop out pada tahun 2019 mencapai 602.208 mahasiswa atau 7% dari total mahasiswa. Penelitian menggunakan data PDDikti untuk memprediksi drop out telah dilakukan untuk mahasiswa yang telah mencapai tujuh tahun masa studi. Namun, belum terdapat sistem berbasis web yang dapat memprediksi mahasiswa drop out menggunakan data semester yang lebih sedikit dan visualisasi yang menggambarkan mahasiswa drop out berdasarkan data yang diperoleh dari PDDikti melalui API tertentu. Penelitian ini membandingkan empat model pembelajaran mesin untuk memprediksi drop out dimana model CatBoost dengan teknik undersampling edited nearest neighbors merupakan classifier terbaik untuk memprediksi drop out dengan f1-score sebesar 64.23%. Selain itu, penelitian ini berhasil mengimplementasi sistem berbasis web yang dapat digunakan untuk melakukan visualisasi data berdasarkan API yang digunakan untuk memperoleh data dari PDDikti dan juga prediksi mahasiswa yang berpotensi drop out berdasarkan data dari PDDikti. Visualisasi mahasiswa drop out berhasil divisualisasi dengan menggunakan diagram sankey, diagram geo, dan diagram bar. Perolehan data dapat dilakukan menggunakan query data dengan API yang dibuat menggunakan Express.js dan Flask.

.....Dropped out student giving negative impact to the student itself and also university. A student is said to have dropped out if they can't complete their studies within the specified timeframe. Data regarding completion status as well as student academic data are recorded in Higher Education Database (PDDikti). Based on these data, the drop out students in 2019 reached 602,208 students or 7% of the total student. Research using PDDikti data to predict drop out has been conducted for students who have reached seven years of study. However, there is no web-based system that could predict drop out student using data with fewer semesters and visualizations portraying dropout students based on PDDikti data through particular API. This study compares four machine learning models to predict drop outs where CatBoost model with undersampling edited nearest neighbors technique is the best classifier to predict drop outs with an f1-score of 64.23%. Other than that, this study succeeded to implement web-based system that could visualize PDDikti data through API and to predict potential students dropping out based on PDDikti data. The visualization of drop out students was successfully visualized using Sankey diagrams, geo diagrams, and bar

charts. Data retrieval can be done using data queries with APIs created using Express.js and Flask.