

BAB V

ANALISIS DAN PEMBAHASAN

5.1 Analisis Algoritma Klasifikasi Kelulusan Mahasiswa TI UNTIRTA

Algoritma yang digunakan dalam pembuatan aplikasi evaluasi mahasiswa TI UNTIRTA terdapat 8 jenis algoritma yaitu *Decision Tree*, *Random Forest*, *Support Vector Machine*, *Naive Bayes*, KNN (*K-Nearest Neighbor*), *XGBoost*, *MLP Classifier*, dan *Logistic Regression*. Kedelapan algoritma tersebut merupakan algoritma klasifikasi yang terdapat pada *Machine Learning* (Buslim, 2019). Algoritma terbaik dipilih berdasarkan nilai ROC-AUC tertinggi untuk setiap semester yang terdiri dari semester 1, 2, 3, 4, 5, 6, dan 7. Algoritma terbaik pada proses *modelling* digunakan untuk membuat model prediksi pada proses *deployment* sehingga prediksi yang dihasilkan mampu menginterpretasikan data dan hasil prediksi mendekati situasi secara nyata atau *real*. ROC-AUC digunakan untuk mengukur sejauh mana model benar dalam melakukan pengklasifikasian sehingga dapat memberikan gambaran umum tentang kinerja model (Saifuddin, 2015).

Dalam semester 1, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Random Forest* dengan nilai hasil evaluasi sebesar 0,718. Nilai hasil evaluasi dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Dalam semester 2, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Support Vector Machine* dengan nilai hasil evaluasi sebesar 0,795. Nilai hasil evaluasi dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Dalam semester 3, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *XGBoost* dengan nilai hasil evaluasi sebesar 0,756. Nilai hasil evaluasi

dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Dalam semester 4, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Logistic Regression* dengan nilai hasil evaluasi sebesar 0,718. Nilai hasil evaluasi dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Dalam semester 5, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Random Forest* dengan nilai hasil evaluasi sebesar 0,836. Nilai hasil evaluasi dengan rentang 0,8 hingga 0,9 memiliki kualitas *good classification* yaitu model memiliki performa yang baik dan cukup handal untuk digunakan meskipun beberapa keadaan tidak diklasifikasikan dengan benar.

Dalam semester 6, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Decision Tree* dengan nilai hasil evaluasi sebesar 0,756. Nilai hasil evaluasi dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Dalam semester 7, algoritma yang memiliki nilai ROC-AUC tertinggi yaitu algoritma *Random Forest* dengan nilai hasil evaluasi sebesar 0,798. Nilai hasil evaluasi dengan rentang 0,7 hingga 0,8 memiliki kualitas *fair classification* yaitu model memiliki performa yang cukup dan memungkinkan membuat beberapa kesalahan dalam klasifikasi namun model masih dapat digunakan.

Berdasarkan beberapa model terbaik yang digunakan pada setiap semester, algoritma *Random Forest* menjadi algoritma yang paling banyak digunakan dalam ketujuh semester. Hal tersebut dikarenakan *Random Forest* mampu menghasilkan klasifikasi yang akurat, mampu mengatasi data yang hilang atau tidak lengkap, mampu mengidentifikasi banyak variabel dengan baik, dan cocok digunakan pada variabel yang saling berhubungan (Kember, 2011). Selain itu, *Random Forest* mampu mengelompokkan variabel-variabel ke dalam data nilai diskrit (Rafrastara, 2023). Data nilai diskrit berupa data nilai mata kuliah wajib Jurusan Teknik Industri Universitas Sultan Ageng Tirtayasa. *Random Forest* mampu mengatasi masalah

ketidakseimbangan kelas sehingga dapat menangani kelas minoritas dan hasil klasifikasi cenderung lebih stabil (Erlin, 2022). Beberapa penelitian yang telah dilakukan menyimpulkan bahwa algoritma *Random Forest* merupakan algoritma terbaik dibandingkan dengan beberapa algoritma lainnya seperti dalam penelitian Adnyana (2021) didapatkan akurasi terbesar yaitu 0,77 dengan membandingkan algoritma *Naïve Bayes*, *J48*, dan *Random Forest* menggunakan data nilai 41 mata kuliah dari 1580 mahasiswa dan penelitian Mu'tashim *et al* (2023) didapatkan akurasi terbesar 0,82 dengan membandingkan algoritma *Gradien Boosting* dan *Random Forest* menggunakan data 10 variabel dari 590 mahasiswa.

5.2 Analisis Aplikasi Evaluasi Mahasiswa TI UNTIRTA

Pada penelitian ini, aplikasi evaluasi mahasiswa TI UNTIRTA dirancang untuk mahasiswa Jurusan Teknik Industri UNTIRTA dan aplikasi diberi nama “GoLulus”. Aplikasi ini dibuat menggunakan Streamlit. Streamlit adalah *framework open-source* bagi para pengembang untuk membuat aplikasi *website* dengan cepat dan mudah diakses dalam bahasa pemrograman Python (Ferdyadi, 2022). Tahapan perancangan aplikasi menggunakan dengan editor Google Collaboratory dimulai dari *business understanding*, *analytic approach*, *data requirements*, *data collection*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment*. Aplikasi dibuat menjadi tujuh bagian yaitu bagian semester 1, semester 2, semester 3, semester 4, semester 5, semester 6, dan semester 7 dimana setiap semester terdiri dari *input* nilai-nilai mata kuliah berdasarkan Mata Kuliah Kurikulum 2022.

Aplikasi yang dibuat pada penelitian berupa aplikasi *website* dengan Streamlit dan menggunakan bahasa pemrograman Python untuk pembuatan model hingga aplikasi. Penelitian ini berbeda dengan penelitian Alifya (2023) dimana pengembangan aplikasi prediksi kelulusan mahasiswa Fakultas Teknik Universitas Negeri Makassar dengan mengembangkan bahasa pemrograman Python untuk pembuatan model algoritma SVM dan bahasa pemrograman php dan html untuk mengembangkan aplikasi *website*. Penggunaan algoritma SVM pada penelitian tersebut dikarenakan variabel non linear yang digunakan berupa jenis kelamin, jurusan, dan Klasifikasi Kelulusan (Alifya, 2023). Penelitian ini juga berbeda dengan penelitian Purwati (2021) dimana prediksi tingkat kelulusan di Perguruan

Tinggi XYZ menggunakan metode *Naïve Bayes* dengan mengembangkan bahasa pemrograman php dan MySQL untuk membuat *database*. Penggunaan algoritma *Naïve Bayes* pada penelitian tersebut dikarenakan variabel yang digunakan variabel independen atau sesama variabel tidak bergantung pada variabel lainnya seperti jenis kelamin, umur, asal, kelas, jurusan, IPK, dan Klasifikasi Kelulusan (Purwati, 2021).

Php dan html dapat digunakan bersamaan dalam pengembangan *web* dengan penggunaan php untuk mengolah data sedangkan html untuk mengatur tampilan dan struktur halaman (Alifya, 2023). Namun, dalam penggunaan php dan HTML memiliki kelemahan yaitu masalah keamanan sehingga mudah dicuri atau dipalsukan sehingga tidak cocok jika peneliti ingin menjaga kerahasiaan *source code*, memiliki kompleksitas tinggi sehingga cocok untuk aplikasi data besar, dan peneliti harus memiliki pengetahuan mendalam terhadap pemahaman bahasa pemrograman tersebut (Sutiono, 2017). Penggunaan bahasa pemrograman Python dan Streamlit dalam Aplikasi GoLulus mampu dengan mudah membangun aplikasi *web* yang dinamis dan menyederhanakan proses pembuatan *coding* dengan cepat. Selama eksplorasi yang peneliti lakukan, belum ditemukan aplikasi klasifikasi kelulusan mahasiswa menggunakan Streamlit sehingga aplikasi ini dapat terus dikembangkan.