

ABSTRAK

Di kalangan mahasiswa, depresi termasuk gangguan mental yang paling sering ditemukan. Dampaknya bisa dirasakan langsung pada kualitas hidup maupun prestasi akademik. Langkah deteksi dini sebenarnya sudah diketahui penting untuk mencegah efek jangka panjang. Hanya saja, di lapangan masih ada kendala berupa terbatasnya tenaga kesehatan dan minimnya kesadaran dari mahasiswa itu sendiri. Ada empat tujuan utama dalam penelitian ini. Pertama, menerapkan algoritma Random Forest untuk klasifikasi tingkat depresi mahasiswa. Kedua, mengukur akurasi, presisi, dan recall dari model yang dibangun. Ketiga, mencari tahu faktor-faktor mana yang paling dominan sebagai penyebab depresi. Keempat, membandingkan kinerja Random Forest dengan beberapa metode ensemble lainnya. Penelitian ini menggunakan pendekatan kuantitatif eksperimental. Data diperoleh dari 153 mahasiswa Universitas Prima Indonesia Medan melalui kuesioner online. Algoritma Random Forest dijalankan dengan parameter yang akurat hasil Grid Search: $n_estimators=50$, $max_depth=10$. Data dibagi dengan metode stratified split 80:20 untuk keperluan latih dan uji. Evaluasi model menggunakan confusion matrix, kurva ROC, dan cross validation 5-lipat. Sebagai pembanding, digunakan Logistic Regression, Naive Bayes, Decision Tree, SVM, dan KNN. Hasilnya, model Random Forest mencapai akurasi 93,55%, presisi 85,71%, recall 85,71%, F1-Score 85,71%, dan nilai AUC 87,80%. Angka-angka ini termasuk dalam kategori sangat baik. Dari cross validation 5-lipat, rata-rata akurasi yang diperoleh adalah 92,11% dengan standar deviasi 5,00%. Artinya, model ini tergolong stabil dan konsisten. Berdasarkan feature importance, tiga faktor utama penyebab depresi adalah kondisi kesehatan (52,33%), IPK (13,91%), dan biaya per bulan (11,31%). Menariknya, meskipun Logistic Regression mencatat akurasi tertinggi (96,69%), Random Forest tetap unggul dari sisi interpretabilitas dan stabilitas. Sebagai penutup, algoritma Random Forest terbukti efektif untuk memprediksi tingkat depresi mahasiswa. Akurasinya tergolong tinggi dan mampu menunjukkan faktor-faktor dominan penyebab depresi. Model ini layak dipertimbangkan sebagai alat bantu skrining awal di lingkungan perguruan tinggi. Untuk penelitian berikutnya, sebaiknya cakupan data diperluas dan variabel seperti dukungan sosial serta tantangan sosial ikut dianalisis.

Kata Kunci: Depresi Mahasiswa, Random Forest, Machine Learning, Klasifikasi, Feature Importance, Deteksi Dini