PREDIKSI TINGKAT STRES DAN PERFORMA AKADEMIK MAHASISWA TEKNIK INFORMATIKA MENGGUNAKAN ALGORITMA MACHINE LEARNING BERDASARKAN ANALISIS LOG AKTIVITAS E-LEARNING

Daffa Majida Althaf
Program Studi D3 Teknik Informatika, Fakultas Teknologi Informasi
Universitas Kristen Satya Wacana

ABSTRACT

The rapid adoption of E-Learning platforms in higher education has produced large volumes of digital footprints, particularly user activity logs. At the same time, academic stress among Information Technology (IT) students remains a serious issue that can negatively affect both academic performance and psychological well-being. Conventional stress-assessment methods, such as psychometric questionnaires, are time-consuming, subjective, and only capture a single point in time. This paper proposes a conceptual approach within the field of Educational Data Mining (EDM) that uses machine learning algorithms - including Random Forest, Support Vector Machine (SVM), and Gradient Boosting - to simultaneously predict student stress levels and academic performance based on E-Learning log features such as login frequency, session duration, and click rate. The proposed model is intended to function as an early-warning framework that enables academic institutions to identify at-risk students and provide timely intervention. As a conceptual study grounded in a review of prior EDM literature, this paper aims to bridge the gap between technical data analytics and student psychological welfare in the context of Informatics education, while highlighting the need for further empirical validation using real institutional data.

Keywords: *machine learning, E-learning log, academic stress, educational data mining, IT students.*

Pendahuluan/Latar Belakang

Perkembangan teknologi digital yang sangat pesat telah membawa perubahan besar pada dunia pendidikan tinggi, salah satunya melalui penggunaan sistem E-Learning yang kini menjadi bagian tidak terpisahkan dari proses perkuliahan, baik sebagai media utama maupun pendukung pembelajaran tatap muka. Di balik fungsinya sebagai tempat mengunduh materi dan mengumpulkan tugas, sistem ini sebenarnya juga merekam jejak digital penggunanya secara otomatis dalam bentuk log data, mulai dari waktu login, durasi akses, hingga frekuensi interaksi dengan berbagai konten pembelajaran. Sayangnya, potensi besar dari data log tersebut masih belum banyak dimanfaatkan untuk memahami kondisi akademik maupun psikologis mahasiswa secara lebih mendalam.

Di sisi lain, isu stres akademik di kalangan mahasiswa bukan lagi hal baru. Beban perkuliahan yang tinggi, tugas yang menumpuk, serta tuntutan untuk terus mengikuti perkembangan keterampilan teknis menjadikan mahasiswa di bidang teknologi, termasuk Teknik Informatika, sebagai salah satu kelompok yang cukup rentan mengalami stres (Misra

& McKean, 2000; Robotham & Julian, 2006). Apabila kondisi ini tidak segera dikenali dan ditangani, dampaknya dapat meluas, mulai dari menurunnya nilai akademik, hilangnya motivasi belajar, hingga terganggunya kesehatan psikologis mahasiswa dalam jangka panjang (Pascoe et al., 2020).

Selama ini, deteksi stres pada mahasiswa umumnya dilakukan melalui kuesioner psikometri, seperti Perceived Stress Scale (PSS), Depression Anxiety Stress Scale (DASS), atau Academic Stress Scale (ASS). Meskipun instrumen-instrumen tersebut sudah teruji secara ilmiah, penggunaannya di lapangan tetap memiliki sejumlah kendala. Kuesioner hanya menggambarkan kondisi mahasiswa pada satu titik waktu tertentu sehingga tidak dapat menangkap dinamika perubahan tingkat stres dari minggu ke minggu sepanjang satu semester. Hasilnya juga sangat bergantung pada kejujuran dan kesediaan responden untuk mengisi secara objektif, sementara proses distribusi hingga analisisnya membutuhkan waktu dan tenaga yang tidak sedikit sehingga sulit diterapkan dalam skala besar. Akibatnya, mahasiswa yang sebenarnya sudah berada dalam kondisi berisiko tinggi seringkali baru teridentifikasi setelah nilainya terlanjur menurun, bahkan ada yang tidak terdeteksi sama sekali.

Kemunculan bidang Educational Data Mining (EDM) dan Learning Analytics (LA) menawarkan cara pandang baru yang lebih menjanjikan untuk menyikapi persoalan ini. Secara sederhana, EDM dapat dipahami sebagai pendekatan penambangan data yang digunakan untuk menemukan pola-pola tersembunyi dari data yang dihasilkan oleh lingkungan belajar (Baker & Yacef, 2009). Beberapa penelitian sebelumnya telah menunjukkan bahwa cara mahasiswa berinteraksi dengan sistem E-Learning dapat mencerminkan kondisi belajar mereka yang sesungguhnya. Mahasiswa yang mulai jarang login, menghabiskan waktu lebih singkat di platform, atau terlihat kurang aktif mengakses materi cenderung memiliki risiko lebih besar untuk gagal secara akademik (Agudo-Peregrina et al., 2014; Cerezo et al., 2016). Dengan demikian, pola perilaku digital ini berpotensi menjadi sinyal awal yang objektif dan terukur untuk mengindikasikan kondisi psikologis mahasiswa, termasuk stres, tanpa harus menunggu mahasiswa mengisi kuesioner terlebih dahulu.

Pendekatan machine learning, seperti Random Forest, Support Vector Machine (SVM), Gradient Boosting, dan Artificial Neural Network (ANN), telah banyak digunakan untuk membangun model prediksi yang andal dari data perilaku pengguna di berbagai bidang (Kotsiantis, 2012; Mduma et al., 2019). Dalam konteks EDM, algoritma-algoritma tersebut dapat dimanfaatkan untuk mempelajari pola dari log aktivitas E-Learning guna memprediksi berbagai hal, mulai dari nilai akhir mahasiswa, risiko dropout, hingga tanda-tanda ketidaklancaran studi. Namun, penelitian yang secara khusus memanfaatkan data log E-Learning untuk memprediksi tingkat stres mahasiswa sebagai variabel psikologis yang terukur, terutama pada mahasiswa Teknik Informatika, masih tergolong jarang ditemukan dalam literatur yang ada saat ini.

Kondisi inilah yang menjadi titik tolak penelitian ini. Penelitian-penelitian terdahulu umumnya hanya membahas prediksi performa akademik tanpa mengaitkannya dengan kondisi psikologis mahasiswa, atau sebaliknya, mengukur stres melalui kuesioner tanpa memanfaatkan data perilaku sistem secara langsung. Sejauh penelusuran yang telah dilakukan, belum banyak ditemukan penelitian yang mencoba mengubah jejak digital dari

log E-Learning menjadi dua prediksi sekaligus, yaitu tingkat stres dan performa akademik, dalam satu kerangka model machine learning yang terintegrasi. Upaya mengangkat isu psikologis ke dalam kerangka EDM yang berbasis data sistem inilah yang menjadi kebaruan (novelty) dari penelitian ini, sekaligus menjadikannya relevan dan layak dikaji dari sudut pandang keilmuan Teknik Informatika.

Berdasarkan uraian tersebut, penelitian ini diarahkan untuk merancang sebuah kerangka model prediksi tingkat stres dan performa akademik mahasiswa Teknik Informatika dengan memanfaatkan algoritma machine learning yang dilatih dari data log aktivitas E-Learning, khususnya fitur waktu login, durasi penggunaan, dan jumlah klik. Melalui pendekatan ini, diharapkan dapat dihasilkan rancangan sistem deteksi dini yang bekerja secara otomatis, objektif, dan berpotensi diterapkan dalam skala besar, sehingga pihak institusi dapat melakukan intervensi akademik maupun psikologis secara lebih cepat dan tepat sasaran dibandingkan metode konvensional yang selama ini berjalan.

Pembahasan

Sebagai sebuah kajian konseptual (conceptual study) yang disusun melalui studi pustaka, bagian ini menguraikan rancangan fitur, perbandingan algoritma, serta arsitektur sistem yang diusulkan, beserta dasar teoretisnya dari penelitian-penelitian EDM sebelumnya.

Ekstraksi dan Taksonomi Fitur Log E-Learning sebagai Refleksi Perilaku Digital

Sistem Learning Management System (LMS) berbasis web, seperti Moodle atau Canva, secara pasif merekam interaksi pengguna ke dalam basis data relasional. Dalam konteks Teknik Informatika, data mentah (raw data) tersebut tidak dapat langsung digunakan oleh algoritma machine learning, melainkan perlu melalui proses feature engineering terlebih dahulu. Aktivitas mahasiswa dapat diekstraksi ke dalam tiga dimensi fitur utama berikut.

Dimensi temporal (waktu) mencakup frekuensi login per minggu, variabilitas jam akses (misalnya kecenderungan mengakses sistem pada pukul 01.00-04.00 WIB), serta kedekatan waktu pengumpulan tugas dengan tenggat waktu (deadline). Dimensi intensitas mencakup total durasi sesi aktif (session duration) dalam satuan menit dan jumlah klik (click rate) pada halaman utama perkuliahan. Dimensi interaktivitas mencakup frekuensi mengunduh materi (file PDF/slide), intensitas menonton video pembelajaran, dan jumlah partisipasi dalam forum diskusi daring.

Secara komputasional, penurunan intensitas yang terjadi secara drastis pada dimensi temporal dan interaktivitas dalam jangka waktu dua minggu berturut-turut dapat diindikasikan sebagai anomali perilaku. Pola anomali ini secara hipotesis diduga berkorelasi dengan peningkatan beban psikologis atau stres akademik, yaitu ketika mahasiswa mulai memasuki fase burnout atau penundaan akademis akibat materi pemrograman yang semakin kompleks.

Komparasi Algoritma Machine Learning dalam Prediksi Variabel Ganda

Penelitian ini mengusulkan sebuah kerangka model terintegrasi yang dirancang untuk dapat memprediksi dua variabel luaran (Y) sekaligus, yaitu tingkat stres (kategori rendah, sedang, tinggi) dan performa akademik (kategori memuaskan, cukup, berisiko).

Untuk mencapai akurasi yang optimal, diusulkan eksperimentasi terhadap tiga algoritma supervised learning sebagai berikut.

Random Forest (RF) merupakan algoritma ensemble berbasis decision tree yang dikenal andal dalam menangani data log pendidikan yang umumnya bersifat non-linier dan memiliki outlier. RF bekerja dengan membangun sekumpulan pohon keputusan dan mengambil keputusan akhir berdasarkan hasil voting terbanyak. Support Vector Machine (SVM) merupakan algoritma yang efektif dalam mencari hyperplane optimal untuk memisahkan kelas-kelas prediktif pada ruang berdimensi tinggi; SVM dengan kernel Radial Basis Function (RBF) diusulkan untuk memisahkan pola interaksi yang kompleks. Gradient Boosting (XGBoost) merupakan pendekatan yang meminimalkan error secara bertahap dari model-model sebelumnya, sehingga umumnya menghasilkan performa komputasi yang cepat dengan akurasi tinggi pada data berskala besar.

Untuk mengukur kinerja model prediksi yang diusulkan, evaluasi nantinya dapat dilakukan menggunakan metrik akurasi (accuracy), presisi (precision), dan recall, dengan rumus sebagai berikut.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

Keterangan: TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative).

Berdasarkan kajian terhadap penelitian-penelitian EDM sejenis, kombinasi fitur log temporal dengan algoritma berbasis ensemble seperti Random Forest atau XGBoost umumnya melaporkan tingkat akurasi yang cukup tinggi dalam tugas prediksi performa mahasiswa (Kotsiantis, 2012; Mduma et al., 2019). Hal ini mengindikasikan bahwa pendekatan komputasional semacam ini berpotensi mengenali sinyal stres maupun penurunan performa akademik jauh sebelum mahasiswa menyelesaikan ujian akhir semester. Meskipun demikian, besaran akurasi yang sesungguhnya tetap perlu diuji lebih lanjut melalui eksperimen menggunakan data log E-Learning riil dari institusi terkait.

Rancangan Implementasi Early Warning System di Program Studi Teknik Informatika

Keunggulan utama dari pendekatan berbasis teknologi informasi ini adalah potensinya untuk diintegrasikan menjadi sebuah sistem deteksi dini (Early Warning System - EWS) pada bagian komputasi akademik program studi. Secara konseptual, arsitektur sistem ini dirancang untuk bekerja sebagai layanan latar belakang (background service) pada platform E-Learning kampus.

Pada setiap akhir pekan, sistem dapat menjalankan skrip otomatisasi untuk mengekstraksi data log mahasiswa. Data tersebut kemudian diumpankan ke dalam model machine learning yang telah dilatih. Apabila model mendeteksi adanya mahasiswa Teknik Informatika yang termasuk dalam klaster "stress tinggi" dan "performa berisiko", sistem dirancang untuk secara otomatis mengirimkan notifikasi melalui API kepada dosen wali maupun bagian konseling mahasiswa.

Melalui visualisasi dasbor (dashboard analytics), dosen diharapkan dapat melihat faktor apa yang menyebabkan mahasiswa terindikasi mengalami stres, misalnya akibat

terlalu sering mengumpulkan tugas pemrograman menjelang tenggat waktu (penundaan ekstrem), atau durasi membaca materi coding yang terlalu singkat namun frekuensi klik sangat tinggi (indikasi kebingungan atau frustrasi). Dengan demikian, tindakan preventif berupa bimbingan akademik atau konseling psikologis dapat diberikan secara lebih tepat sasaran.

Simpulan

Berdasarkan kajian konseptual yang telah dipaparkan, dapat disimpulkan bahwa pemanfaatan data log aktivitas E-Learning melalui pendekatan Educational Data Mining (EDM) memiliki potensi besar untuk memprediksi tingkat stres dan performa akademik mahasiswa Teknik Informatika. Jejak digital seperti dimensi temporal, intensitas, dan interaktivitas mahasiswa di dalam sistem bukan sekadar data sampah, melainkan dapat menjadi indikator yang relatif objektif untuk merefleksikan kondisi psikologis dan akademik mahasiswa secara berkelanjutan.

Algoritma machine learning seperti Random Forest dan XGBoost diproyeksikan dapat memberikan performa prediksi yang baik sekaligus mengurangi kelemahan metode kuesioner konvensional yang bersifat subjektif dan tidak dinamis. Apabila rancangan model ini diimplementasikan dalam bentuk Early Warning System di program studi Teknik Informatika, pendekatan ini dapat menjadi jembatan inovatif yang menyatukan analisis data teknis dengan upaya peningkatan kesejahteraan psikologis mahasiswa di era digital. Penelitian lanjutan dengan data log riil dan pengujian model secara factual tetap diperlukan untuk memvalidasi efektivitas pendekatan yang diusulkan.

Daftar Pustaka

Agudo-Peregrina, Á. H., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for business school students and its relation with performance. Computers in Human Behavior, 31, 542-550.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3-17.

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS log data and their relationship with achievement. Spanish Journal of Psychology, 19, E78.

Kotsiantis, S. B. (2012). Use of machine learning techniques for educational data mining. Applied Artificial Intelligence, 26(3), 262-286.

Mduma, N., Kalegele, K., & Machve, J. (2019). Machine learning approach for predicting student dropout in e-learning platforms. International Journal of Computer Applications, 177(28), 13-19.

Misra, R., & McKean, M. (2000). College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction. American Journal of Health Studies, 16(1), 41-51.

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. International Journal of Adolescent Medicine and Health, 32(2).

Robotham, D., & Julian, C. (2006). Stress and the higher education student: A critical review of the literature. Journal of Further and Higher Education, 30(2), 107-117.