

PREDIKSI TINGKAT PROKRASTINASI MAHASISWA BERDASARKAN POLA PENGGUNAAN MEDIA SOSIAL MENGUNAKAN ALGORITMA *CATEGORICAL NAÏVE BAYES*

Hanif Ihsan Fatoro

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah
Sukabumi, Sukabumi, Indonesia

e-mail: hanifihsan44831@gmail.com

ABSTRAK

Prokrastinasi akademik merupakan perilaku menunda penyelesaian tugas akademik yang masih banyak ditemukan pada mahasiswa. Penelitian ini bertujuan membangun model prediksi tingkat prokrastinasi akademik mahasiswa berdasarkan pola penggunaan media sosial menggunakan algoritma *Categorical Naïve Bayes*, mengevaluasi kinerja model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta mengidentifikasi variabel yang paling berkontribusi terhadap hasil prediksi. Penelitian menerapkan metode *Cross Industry Standard Process for Data Mining (CRISP-DM)*. Data diperoleh melalui penyebaran kuesioner kepada 300 mahasiswa aktif Program Studi Teknik Informatika Universitas Muhammadiyah Sukabumi. Variabel prediktor meliputi angkatan, jenis kelamin, durasi penggunaan media sosial, frekuensi akses, waktu dominan penggunaan, platform media sosial yang paling sering digunakan, serta kebiasaan mengecek notifikasi saat belajar. Hasil penelitian menunjukkan bahwa model *Categorical Naïve Bayes* memperoleh *accuracy* sebesar 35%. Evaluasi menggunakan *precision*, *recall*, dan *F1-score* menunjukkan bahwa model masih mengalami kesulitan dalam membedakan kategori tingkat prokrastinasi akademik. Analisis *Permutation Importance* menunjukkan bahwa variabel notifikasi, waktu penggunaan, dan platform media sosial memberikan kontribusi terbesar terhadap proses prediksi. Hasil penelitian menunjukkan bahwa pola penggunaan media sosial dapat dimanfaatkan sebagai variabel dalam membangun model prediksi tingkat prokrastinasi akademik mahasiswa.

Kata kunci: Prokrastinasi Akademik; Media Sosial; *Categorical Naïve Bayes*; *Machine Learning*; Klasifikasi.

ABSTRACT

Academic procrastination is the tendency to delay completing academic tasks and remains a common issue among university students. This study aims to develop a prediction model for students' academic procrastination levels based on social media usage patterns using the Categorical Naïve Bayes algorithm, evaluate the model performance using accuracy, precision, recall, and F1-score, and identify the variables that contribute the most to the prediction results. This research employed the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. Data were collected through questionnaires distributed to 300 active students of the Informatics Engineering Study Program at Universitas Muhammadiyah Sukabumi. The predictor variables included academic year, gender, duration of social media usage, frequency of social media access, dominant time of use, the most frequently used social media platform, and the habit of checking social media notifications while studying. The results showed that the Categorical Naïve Bayes model achieved an accuracy of 35%. Evaluation using precision, recall, and F1-score indicated that the model still had difficulty distinguishing among the academic procrastination categories. Furthermore, Permutation Importance analysis revealed that notification-checking habits, time of social media usage, and social media platform were the variables contributing the most to the prediction process. These findings indicate that social

media usage patterns can be utilized as predictor variables for developing an academic procrastination prediction model.

Keywords: *Academic Procrastination; Social Media; Categorical Naïve Bayes; Machine Learning; Classification.*

1. PENDAHULUAN

Prokrastinasi akademik merupakan perilaku menunda penyelesaian tugas akademik yang masih sering ditemukan di kalangan mahasiswa. Perilaku tersebut dapat berdampak pada keterlambatan penyelesaian tugas, menurunnya kualitas hasil belajar, serta meningkatnya stres akademik. Salah satu faktor yang diduga memengaruhi munculnya prokrastinasi akademik adalah pola penggunaan media sosial. Saat ini, hampir seluruh mahasiswa menggunakan media sosial sebagai sarana komunikasi maupun hiburan sehingga menghabiskan waktu berjam-jam untuk mengakses berbagai platform digital. Kondisi tersebut dapat mengurangi waktu belajar dan menyebabkan tugas akademik dikerjakan mendekati batas waktu pengumpulan atau bahkan terlambat diselesaikan [1],[2][3].

Berbagai penelitian di Indonesia menunjukkan adanya keterkaitan antara penggunaan media sosial dengan tingkat prokrastinasi akademik mahasiswa. Mahasiswa yang memiliki kontrol diri yang baik dalam menggunakan media sosial cenderung memiliki tingkat prokrastinasi yang lebih rendah. Sebaliknya, mahasiswa yang memiliki intensitas penggunaan media sosial tinggi lebih berisiko menunda penyelesaian tugas akademik. Penelitian lain juga menunjukkan bahwa tingginya penggunaan media sosial dapat mengurangi konsentrasi belajar, mengganggu pengelolaan waktu, serta meningkatkan kecenderungan menunda penyelesaian tugas akademik [4],[5],[6],[7], [8], [9].

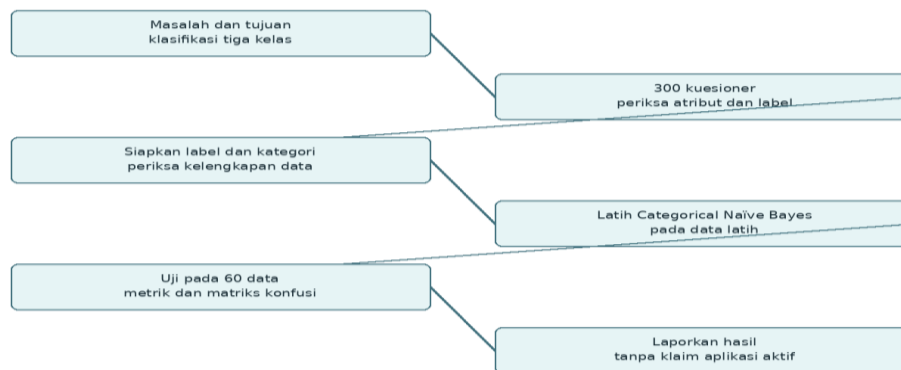
Perkembangan *machine learning* telah dimanfaatkan dalam berbagai penelitian untuk menyelesaikan permasalahan klasifikasi pada bidang pendidikan. Salah satu algoritma yang banyak digunakan adalah *Naïve Bayes* karena memiliki proses komputasi yang sederhana, efisien, serta mampu mengolah data kategorikal dengan baik. Beberapa penelitian sebelumnya berhasil menerapkan algoritma *Naïve Bayes* untuk mengklasifikasikan perilaku maupun memprediksi berbagai kondisi mahasiswa dengan tingkat akurasi yang baik [10],[11],[12]. Namun, penelitian mengenai prokrastinasi akademik masih didominasi oleh pendekatan deskriptif dan korelasional sehingga belum banyak yang membangun model prediksi berdasarkan pola penggunaan media sosial mahasiswa [13],[14], [15].

Berdasarkan kondisi tersebut, penelitian ini membangun model prediksi tingkat prokrastinasi akademik mahasiswa menggunakan algoritma *Categorical Naïve Bayes* berdasarkan pola penggunaan media sosial. Variabel prediktor yang digunakan meliputi angkatan, jenis kelamin, durasi penggunaan media sosial, frekuensi akses media sosial, waktu dominan penggunaan media sosial, platform media sosial yang paling sering digunakan, serta kebiasaan mengecek notifikasi media sosial saat belajar. Proses penelitian menerapkan kerangka kerja *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. Penelitian ini bertujuan membangun model prediksi tingkat prokrastinasi akademik mahasiswa, mengevaluasi kinerja model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta mengidentifikasi variabel yang memberikan kontribusi terbesar terhadap hasil prediksi melalui analisis *Permutation Importance*. [8], [16], [17].

2. METODE PENELITIAN

Penelitian ini menggunakan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* sebagai kerangka kerja dalam membangun model prediksi tingkat prokrastinasi akademik mahasiswa. Metode ini dipilih karena menyediakan tahapan yang sistematis, mulai dari identifikasi permasalahan hingga evaluasi model klasifikasi [18]. Data penelitian diperoleh melalui penyebaran kuesioner kepada 300 mahasiswa aktif Program Studi Teknik Informatika Universitas Muhammadiyah Sukabumi. Variabel prediktor yang digunakan meliputi angkatan, jenis kelamin, durasi penggunaan media sosial, frekuensi akses media sosial, waktu dominan penggunaan media sosial, platform media sosial yang paling sering digunakan, serta kebiasaan mengecek notifikasi media sosial saat belajar. Variabel target berupa tingkat prokrastinasi akademik yang diklasifikasikan ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi berdasarkan skor total kuesioner.[15], [19].

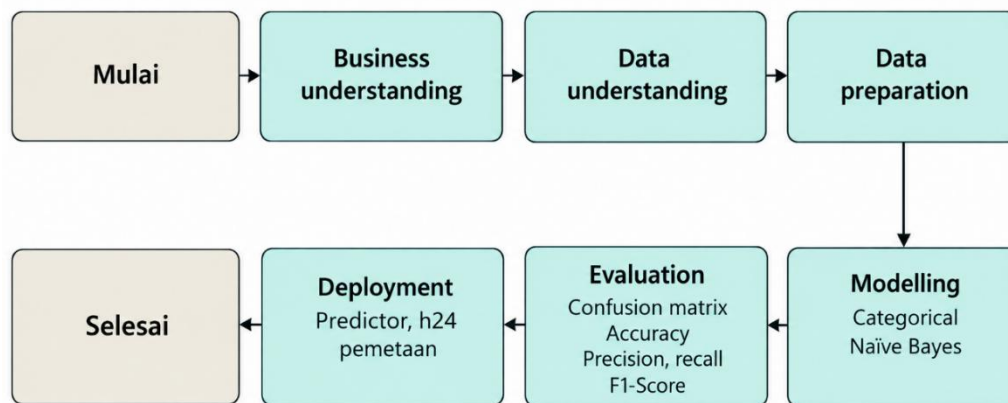
Tahapan penelitian mengikuti enam fase *CRISP-DM*, yaitu *Business Understanding* untuk mengidentifikasi permasalahan dan tujuan penelitian, *Data Understanding* untuk memahami karakteristik data dan variabel yang digunakan, *Data Preparation* untuk melakukan pemeriksaan kelengkapan data, data *cleaning*, pembentukan variabel target, dan transformasi data kategorikal, *Modeling* menggunakan algoritma *Categorical Naïve Bayes*, *Evaluation* menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta *Deployment* sebagai tahap penyajian hasil penelitian [18]. Rincian batas skor untuk pembentukan kategori rendah, sedang, dan tinggi serta teknik *encoding* yang digunakan perlu disesuaikan dengan prosedur pengolahan data asli penulis. [19], [20]



Gambar 1. Tahapan *CRISP-DM*

Penerapan *CRISP-DM* dalam penelitian ini tidak hanya digunakan sebagai alur konseptual, tetapi sebagai kerangka kerja yang menghubungkan setiap tahap penelitian dengan proses pemodelan. *Business Understanding* berfokus pada perumusan kebutuhan prediksi tingkat prokrastinasi akademik berdasarkan pola penggunaan media sosial. *Data Understanding* digunakan untuk memahami 300 data responden dan atribut prediktor yang tersedia. *Data Preparation* mencakup pemeriksaan data, pembersihan, pembentukan label target, dan transformasi atribut kategorikal. *Modeling* dilakukan dengan melatih *Categorical Naïve Bayes* menggunakan data *training*, *Evaluation*

dilakukan pada data *testing* menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*, sedangkan *Deployment* diarahkan pada penyajian model dan hasil prediksi.



Gambar 2. Tahapan *CRISP-DM*

Model klasifikasi dibangun menggunakan bahasa pemrograman Python dengan pustaka Scikit-learn. *Dataset* sebanyak 300 responden dibagi menjadi data *training* sebanyak 240 data (80%) dan data *testing* sebanyak 60 data (20%). Pembagian dilakukan sebelum *oversampling* sehingga data *testing* tetap merepresentasikan distribusi data asli. Selanjutnya, *RandomOverSampler* diterapkan hanya pada data *training* untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas. Data *training* hasil *oversampling* digunakan untuk melatih *Categorical Naïve Bayes*, sedangkan data *testing* yang tidak di-*oversampling* digunakan untuk mengukur kemampuan generalisasi model. Hasil prediksi kemudian dievaluasi berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, analisis *Permutation Importance* dilakukan untuk mengidentifikasi variabel yang memberikan kontribusi terbesar terhadap hasil prediksi model. [15], [17], [20], [21], [22].

3. HASIL DAN PEMBAHASAN

Analisis dan Perancangan

Analisis dilakukan untuk memahami permasalahan penelitian, yaitu tingkat prokrastinasi akademik mahasiswa yang diduga berkaitan dengan pola penggunaan media sosial. Berdasarkan hasil analisis, dibangun model klasifikasi menggunakan algoritma *Categorical Naïve Bayes* untuk memprediksi tingkat prokrastinasi akademik mahasiswa berdasarkan data hasil kuesioner. Variabel prediktor yang digunakan terdiri atas angkatan, jenis kelamin, durasi penggunaan media sosial, frekuensi akses media sosial, waktu dominan penggunaan media sosial, platform media sosial yang paling sering digunakan, serta kebiasaan mengecek notifikasi saat belajar. Variabel target berupa tingkat prokrastinasi akademik yang dikelompokkan ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi.

Pada tahap perancangan, dilakukan proses *data preparation* yang meliputi pemeriksaan kelengkapan data, data *cleaning*, pembentukan kategori tingkat

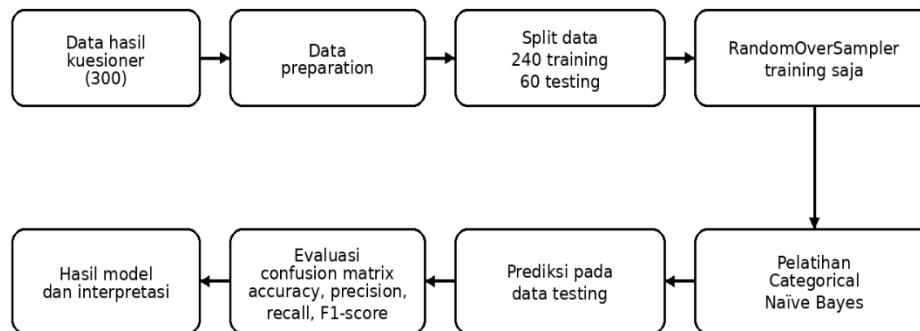
prokrastinasi akademik berdasarkan skor kuesioner, serta transformasi atribut kategorikal agar dapat diproses oleh *Categorical Naïve Bayes*. Jenis *encoding* dan batas skor kategorisasi harus mengikuti prosedur pengolahan data asli penulis dan belum dapat dispesifikasikan sebelum data/kode sumber dikonfirmasi. Setelah *data preparation* selesai, *dataset* dibagi menjadi data *training* dan data *testing* untuk proses pembangunan model klasifikasi. *Categorical Naïve Bayes* dirancang untuk fitur yang bersifat kategorikal; dalam implementasi scikit-learn, kategori perlu direpresentasikan sebagai kode numerik yang sesuai dengan kategori fitur sebelum pemodelan [15].

Model klasifikasi dirancang menggunakan algoritma *Categorical Naïve Bayes* dengan bahasa pemrograman Python dan pustaka Scikit-learn. *Dataset* terlebih dahulu dibagi menjadi data *training* dan data *testing* dengan komposisi 80% (240 data) dan 20% (60 data). Setelah pembagian data, *RandomOverSampler* diterapkan hanya pada data *training* untuk mengatasi ketidakseimbangan kelas sebelum proses pelatihan model. Data *testing* tidak diubah melalui *oversampling*. Selanjutnya, model dilatih menggunakan data *training* hasil *oversampling* dan digunakan untuk menghasilkan prediksi pada data *testing*. Alur perancangan model ditunjukkan pada Gambar 2.



Gambar 3. Alur Pelatihan dan Pengujian Model *Categorical Naïve Bayes*

Gambar 2 memperlihatkan alur teknis pemodelan setelah tahap *CRISP-DM*, sehingga fungsinya berbeda dengan Gambar 1. Gambar 1 menunjukkan kerangka kerja penelitian secara umum, sedangkan Gambar 2 memperlihatkan tahapan reproduktif pada pemodelan, yaitu data hasil *preparation* dibagi menjadi data *training* dan *testing*, *RandomOverSampler* diterapkan pada data *training*, model *Categorical Naïve Bayes* dilatih menggunakan data *training* hasil *oversampling*, kemudian model menghasilkan prediksi terhadap data *testing*. Kelas aktual pada data *testing* berasal dari variabel target tingkat prokrastinasi akademik yang telah dibentuk dari skor total kuesioner. Hasil prediksi dibandingkan dengan kelas aktual untuk menghasilkan *confusion matrix* dan *classification report*.

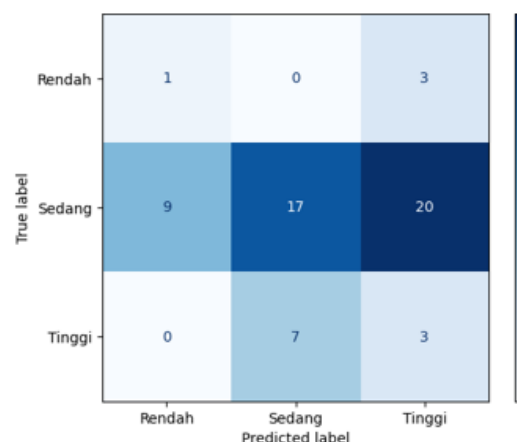


Gambar 4. Alur Perancangan Model *Categorical Naïve Bayes*

Implementasi dan Pembahasan

Implementasi penelitian dilakukan dengan membangun model klasifikasi menggunakan algoritma *Categorical Naïve Bayes* berdasarkan data hasil kuesioner yang telah melalui tahap *data preparation*. *Dataset* sebanyak 300 responden dibagi menjadi 240 data *training* dan 60 data *testing*. *Random Over Sampler* diterapkan setelah pembagian data dan hanya pada data *training* untuk mengurangi dampak ketidakseimbangan kelas dalam proses pembelajaran. Data *testing* tetap menggunakan distribusi asli dan digunakan tanpa *oversampling* untuk mengukur kemampuan model pada data yang belum dimodifikasi. Model kemudian dilatih menggunakan data *training* hasil *oversampling* dan digunakan untuk memprediksi kategori tingkat prokrastinasi akademik pada data *testing*.

Model yang telah dilatih selanjutnya digunakan untuk memprediksi tiga kategori tingkat prokrastinasi akademik, yaitu rendah, sedang, dan tinggi, pada data *testing*. Kelas aktual (*true label*) berasal dari kategori tingkat prokrastinasi akademik masing-masing responden dalam data *testing*, sedangkan *predicted label* berasal dari hasil prediksi model. Hasil pengujian divisualisasikan menggunakan *confusion matrix* untuk membandingkan kelas aktual dengan kelas hasil prediksi. Visualisasi ini memberikan gambaran mengenai distribusi prediksi yang benar maupun yang mengalami kesalahan klasifikasi pada setiap kategori prokrastinasi akademik.



Gambar 5. *Confusion Matrix* Hasil Pengujian Model

Model menghasilkan beberapa kesalahan klasifikasi pada masing-masing kategori sehingga memengaruhi kemampuan model dalam membedakan tingkat prokrastinasi akademik mahasiswa. Oleh karena itu, dilakukan evaluasi lebih lanjut menggunakan *classification report* untuk mengetahui nilai *accuracy*, *precision*, *recall*, dan *F1-score* pada setiap kategori. Nilai *support* pada *classification report* menunjukkan jumlah data aktual pada masing-masing kelas di data *testing*, sehingga Rendah = 4, Sedang = 46, dan Tinggi = 10 dengan total 60 data *testing*.

Tabel 1. *Classification Report Model Categorical Naïve Bayes*

Kategori	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Rendah	0,10	0,25	0,14	4
Sedang	0,71	0,37	0,49	46
Tinggi	0,12	0,30	0,17	10
<i>Accuracy</i>	0,35			60
<i>Macro Average</i>	0,31	0,31	0,27	60
<i>Weighted Average</i>	0,57	0,35	0,41	60

Model *Categorical Naïve Bayes* memperoleh nilai *accuracy* sebesar 35%. Kategori Sedang menunjukkan nilai *precision*, *recall*, dan *F1-score* yang lebih baik dibandingkan kategori Rendah dan Tinggi. Distribusi kelas pada data *testing* menunjukkan ketidakseimbangan yang cukup ekstrem, yaitu hanya 4 data untuk kelas Rendah, 46 data untuk kelas Sedang, dan 10 data untuk kelas Tinggi. Kondisi tersebut membuat ukuran *precision* dan *recall* pada kelas minoritas lebih sensitif terhadap kesalahan klasifikasi. Meskipun *RandomOverSampler* telah diterapkan pada data *training*, data *testing* tetap mempertahankan distribusi aslinya agar evaluasi menggambarkan performa model pada data yang tidak dimodifikasi. Perbedaan *weighted average precision* sebesar 0,57 dengan *macro average precision* sebesar 0,31 juga menunjukkan bahwa performa model lebih banyak dipengaruhi oleh kelas dengan jumlah data lebih besar, terutama kelas Sedang. Hal ini mengindikasikan kecenderungan model untuk lebih baik mengenali kelas mayoritas dibandingkan membedakan ketiga kategori secara seimbang. Oleh karena itu, performa model secara keseluruhan belum optimal dan nilai *accuracy* 35% perlu dibaca bersama metrik per kelas. [22], [23]

Untuk mengetahui kontribusi masing-masing variabel terhadap hasil prediksi, dilakukan analisis menggunakan metode *Permutation Importance*. Analisis ini bertujuan mengidentifikasi perubahan kinerja model ketika informasi suatu variabel diacak, sehingga dapat memberikan gambaran mengenai kontribusi variabel terhadap prediksi. Interpretasi *permutation importance* perlu dilakukan secara hati-hati karena nilai tersebut menggambarkan kontribusi variabel terhadap kinerja model yang digunakan, bukan hubungan kausal antara variabel dan prokrastinasi akademik. [16], [17]

Tabel 2. Nilai *Permutation Importance*

No	Variabel	Nilai <i>Permutation Importance</i>
1	Notifikasi	0,0194
2	Waktu	0,0130
3	Platform	0,0058
4	Frekuensi	-0,0150

5	Durasi	-0,0474
6	Jenis Kelamin	-0,0765
7	Angkatan	-0,0808

Hasil analisis menunjukkan bahwa variabel Notifikasi (0,0194), Waktu (0,0130), dan Platform (0,0058) memiliki nilai *permutation importance* positif dan merupakan tiga nilai tertinggi. Sebaliknya, Frekuensi (-0,0150), Durasi (-0,0474), Jenis Kelamin (-0,0765), dan Angkatan (-0,0808) memiliki nilai negatif. Dalam konteks *permutation importance*, nilai negatif menunjukkan bahwa pengacakan variabel tersebut pada evaluasi yang digunakan tidak menurunkan kinerja model dan bahkan dapat menghasilkan kinerja yang lebih tinggi, sehingga variabel tersebut perlu dipertimbangkan kembali karena kemungkinan kontribusinya terhadap model relatif rendah atau menambah variasi yang tidak membantu prediksi. Selain itu, nilai *importance* tertinggi hanya 0,0194. Nilai yang relatif kecil ini menunjukkan bahwa tidak terdapat satu variabel yang secara dominan menentukan kinerja prediksi model. Temuan tersebut sejalan dengan *accuracy* keseluruhan yang masih rendah. Dengan demikian, penelitian lanjutan dapat mempertimbangkan *feature selection*, penambahan variabel prediktor yang lebih relevan, serta pengujian algoritma alternatif. [17]

Secara keseluruhan, hasil implementasi menunjukkan bahwa algoritma *Categorical Naïve Bayes* dapat digunakan untuk membangun model prediksi tingkat prokrastinasi akademik mahasiswa berdasarkan pola penggunaan media sosial. Hasil *accuracy* 35%, *classification report*, *confusion matrix*, dan *permutation importance* pada penelitian ini berasal dari pengujian model terhadap 60 data *testing* yang dipisahkan dari 300 responden dan tidak mengalami *oversampling*. Nilai *accuracy* tersebut menunjukkan bahwa performa model masih perlu ditingkatkan melalui penambahan jumlah data, pengembangan variabel prediktor, penanganan ketidakseimbangan kelas yang lebih tepat, maupun penggunaan algoritma klasifikasi lain pada penelitian selanjutnya.

4. KESIMPULAN

Berdasarkan hasil penelitian, algoritma *Categorical Naïve Bayes* berhasil diterapkan untuk membangun model prediksi tingkat prokrastinasi akademik mahasiswa berdasarkan pola penggunaan media sosial. Model yang dibangun menggunakan variabel angkatan, jenis kelamin, durasi penggunaan media sosial, frekuensi akses media sosial, waktu dominan penggunaan media sosial, platform media sosial yang paling sering digunakan, serta kebiasaan mengecek notifikasi saat belajar.

Hasil evaluasi menunjukkan bahwa model memperoleh nilai *accuracy* sebesar 35%. Berdasarkan nilai *precision*, *recall*, dan *F1-score*, model masih mengalami kesulitan dalam membedakan kategori tingkat prokrastinasi akademik. Ketidakseimbangan distribusi kelas pada data *testing* juga perlu dipertimbangkan dalam membaca hasil, karena kelas Rendah hanya memiliki 4 data aktual, sedangkan kelas Sedang memiliki 46 data. Oleh karena itu, hasil penelitian ini menunjukkan kemampuan prediksi model pada *dataset* yang digunakan dan tidak dimaksudkan sebagai bukti hubungan sebab-akibat antara penggunaan media sosial dan prokrastinasi akademik. Analisis menggunakan *Permutation Importance* menunjukkan bahwa variabel kebiasaan mengecek notifikasi, waktu dominan penggunaan media sosial, dan platform media sosial merupakan variabel dengan nilai *importance* tertinggi dalam model. Namun, nilai *importance* tertinggi yang relatif kecil dan adanya nilai negatif pada empat variabel lainnya menunjukkan bahwa kontribusi prediktor belum kuat dan merata. Hasil penelitian

ini menunjukkan bahwa pola penggunaan media sosial dapat dimanfaatkan sebagai variabel dalam membangun model prediksi tingkat prokrastinasi akademik mahasiswa, meskipun performa model masih memerlukan pengembangan lebih lanjut.

DAFTAR PUSTAKA

- [1] D. Azmizahra, D. Rakhmawati, dan Y. D. Setyoningsih, “Kecanduan Media Sosial terhadap Prokrastinasi Akademik Mahasiswa,” *Observasi: Jurnal Publikasi Ilmu Psikologi*, vol. 4, no. 3, hlm. 125–142, 2026, doi: 10.61132/observasi.v4i3.2393.
- [2] B. Purba. U dan H. R. Damra, “Pengaruh Kontrol Diri Terhadap Prokrastinasi Akademik Pada Mahasiswa Tingkat Akhir Pengguna Media Sosial Tiktok di Universitas Negeri Padang,” *J. Ilmu Multidisiplin*, 2025.
- [3] N. Azimah dan H. Nirwana, “The Relationship Between Academic Procrastination and Academic Stress Among Students,” *TOFEDU: The Future of Education Journal*, vol. 5, no. 1, hlm. 724–732, 2026, doi: 10.61445/tofedu.v5i1.1633.
- [4] Ricka Yoseva dan Ratna Yunita Setiyani Subardjo, “Hubungan Intensitas Penggunaan Media Sosial TikTok dengan Prokrastinasi Akademik Pada Mahasiswa,” *EduInovasi: Journal of Basic Educational Studies*, vol. 5, no. 2, 2025, doi: 10.47467/edu.v5i2.7553.
- [5] D. Christian dan Y. Budiarto, “Hubungan Durasi Tiktok Dan Rentang Perhatian Pada Pengguna Aktif di Usia Dewasa Muda,” *J. Ilmu Pendidik. Dan Psikol.*, 2024.
- [6] X. Zhu, T. Zheng, L. Ding, X. Zhang, Z. Li, dan H. Jiang, “Exploring associations between social media addiction, social media fatigue, fear of missing out and sleep quality among university students: A cross-section study,” *PLOS ONE*, vol. 18, no. 10, hlm. e0292429, 2023, doi: 10.1371/journal.pone.0292429.
- [7] M. H. Sidqi, I. D. Nurdiansah, dan D. Darmawan, “Faktor-faktor yang Mempengaruhi Prokrastinasi Akademik Mahasiswa: Kejenuhan Akademik, Kecanduan Media Sosial, dan Stres Akademik,” *FONDATIA*, vol. 9, no. 2, hlm. 437–460, 2025, doi: 10.36088/fondatia.v9i2.5749.
- [8] T. Touloupis dan M. Campbell, “The role of academic context-related factors and problematic social media use in academic procrastination: A cross-sectional study of students in elementary, secondary, and tertiary education,” *Soc. Psychol. Educ.*, vol. 27, hlm. 175–214, 2024, doi: 10.1007/s11218-023-09817-8.
- [9] S. Üztemur dan E. Dinç, “Academic procrastination using social media: A cross-sectional study with middle school students on the buffering and moderator roles of self-control and gender,” *Psychol. Sch.*, vol. 60, hlm. 1060–1082, 2023, doi: 10.1002/pits.22818.
- [10] Khoirunnisa, A. Susanti, F. Azima, dan F. S. Nugraha, “Pembelajaran Daring Menggunakan Algoritma Naive Bayes,” *Semin. Nas. Amikom Surak. Semnasa*, 2024.
- [11] Irwansyahputra dan Nurdin, “Perbandingan Algoritma Naive Bayes dan Modified Naive Bayes dalam Klasifikasi Tingkat Stres Mahasiswa,” *JUTECH: Journal Education and Technology*, vol. 7, no. 1, hlm. 105–119, 2026, doi: 10.31932/jutech.v7i1.6627.
- [12] T. Gori, A. Sunyoto, dan H. Al Fatta, “Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 1, hlm. 215–224, 2024, doi: 10.25126/jtiik.20241118074.

- [13] A. R. Husain, T. Wantu, dan M. R. Pautina, “Perilaku Prokrastinasi Akademik dan Faktor Penyebabnya pada Mahasiswa,” *Stud. J. Guid. Couns.*, 2023, doi: <https://doi.org/10.37411/sjgc.v2i2.1907>.
- [14] N. Putri, S. Hanida Tri, dan A. Rohmani, “Hubungan Antara Burnout Akademik Dengan Tingkat Kecemasan Pada Mahasiswa Fakultas Kedokteran,” 2025.
- [15] C. Herdian, “Perbandingan Algoritma Naive Bayes di dalam Scikit-Learn Python Library dengan Murni Algoritma Naive Bayes: Studi Kasus Klasifikasi Email Berbahaya,” *Techno Xplore: Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 9, no. 1, hlm. 1–10, 2024, doi: 10.36805/technoxplore.v9i1.5824.
- [16] X. Zhou, F. Yang, Y. Chen, dan Y. Gao, “The correlation between mobile phone addiction and procrastination in students: A meta-analysis,” *J. Affect. Disord.*, vol. 348, hlm. 317–328, 2024, doi: 10.1016/j.jad.2023.11.020.
- [17] A. Khan, A. Ali, J. Khan, F. Ullah, dan M. Faheem, “Using Permutation-Based Feature Importance for Improved Machine Learning Model Performance at Reduced Costs,” *IEEE Access*, vol. 13, hlm. 36421–36435, 2025, doi: 10.1109/access.2025.3544625.
- [18] A. Rianti, N. Wachid, A. Majid, dan A. Fauzi, “Sistem, P., Informasi, T., Telekomunikasi, S., Upi Di Purwakarta, K., & Edu, A. CRISP-DM: Metodologi Proyek Data Science,” *Pros. Semin. Nas. Teknol. Inf. Dan Bisnis SENATIB*, 2023, doi: <https://ojs.udb.ac.id/index.php/Senatib/article/view/3015>.
- [19] S. Peker dan Ö. Kart, “Transactional data-based customer segmentation applying CRISP-DM methodology: A systematic review,” *Journal of Data, Information and Management*, vol. 5, no. 1–2, hlm. 1–21, 2023, doi: 10.1007/s42488-023-00085-x.
- [20] S. Rahimi, N. C. Hall, dan F. Sticca, “Understanding academic procrastination: A Longitudinal analysis of procrastination and emotions in undergraduate and graduate students,” *Motivation and Emotion*, vol. 47, no. 4, hlm. 554–574, 2023, doi: 10.1007/s11031-023-10010-9.
- [21] D. Yustianisa, F. Wajidi, W. Firgiawan, dan A. G. Sholeha, “Performance Comparison of K-Nearest Neighbors and Decision Tree Algorithms with Random Oversampling for Imbalanced Heart Disease Classification,” *Jurnal Teknik Informatika (Jutif)*, vol. 7, no. 3, hlm. 2633–2645, 2026, doi: 10.52436/1.jutif.2026.7.3.5626.
- [22] P. Billion-Polak dan T. M. Khoshgoftaar, “Evaluating techniques from low-shot learning on traditional imbalanced classification tasks,” *Journal of Big Data*, vol. 12, no. 1, 2025, doi: 10.1186/s40537-025-01171-0.
- [23] A. de la Cruz Huayanay, J. L. Bazán, dan C. M. Russo, “Performance of evaluation metrics for classification in imbalanced data,” *Computational Statistics*, 2024, doi: 10.1007/s00180-024-01539-5.