

BAB 1. PENDAHULUAN

1.1 1.1 Latar Belakang

Transformasi digital dalam dunia pendidikan tinggi mendorong perguruan tinggi untuk mengintegrasikan teknologi informasi ke dalam proses pembelajaran secara sistematis. Salah satu manifestasi nyata dari integrasi ini adalah penggunaan *Learning Management System* (LMS) sebagai infrastruktur pembelajaran *daring* yang memungkinkan interaksi antara dosen dan mahasiswa tanpa batasan ruang dan waktu (Pratama & Cahyono, 2022). Menurut data Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (2024), sebanyak 92% perguruan tinggi negeri di Indonesia telah mengimplementasikan LMS sebagai bagian integral dari proses pembelajaran. Keberhasilan implementasi LMS tidak hanya ditentukan oleh kelengkapan fitur teknis, tetapi juga oleh tingkat kepuasan dan persepsi penggunakhususnya mahasiswa sebagai pihak yang paling intensif berinteraksi dengan sistem tersebut (Rahman & Santoso, 2022).

Politeknik Negeri Jember (POLIJE) merupakan institusi pendidikan vokasi yang mengembangkan LMS milik sendiri bernama S-Learn, yang dirancang dan dikelola oleh *Teaching Factory* (TEFA) Jurusan Teknologi Informasi. Hingga tahun ajaran 2025/2026, S-Learn telah diimplementasikan pada delapan mata kuliah dan digunakan oleh empat program studi di Jurusan Teknologi Informasi, yaitu Teknik Informatika, Manajemen Informatika, Teknik Komputer, dan Bisnis Digital. Sebagai *platform* yang dikembangkan secara internal, S-Learn memiliki keunggulan berupa fleksibilitas pengembangan yang tinggi, namun juga memerlukan mekanisme evaluasi berkelanjutan yang berbasis data untuk memastikan kualitasnya sesuai kebutuhan pengguna.

Mahasiswa Jurusan Teknologi Informasi, sebagai pengguna utama dengan literasi digital yang tinggi, secara aktif memberikan umpan balik terkait pengalaman mereka dalam menggunakan S-Learn. Umpan balik ini umumnya dituangkan dalam bentuk ulasan teks bebas yang mengandung opini, keluhan, dan apresiasi jenis data yang kaya akan informasi semantik namun sulit diolah secara manual dalam jumlah besar (Sari & Watrianthos, 2023). Pengolahan manual atas

ratusan atau ribuan ulasan teks tidak hanya tidak efisien, tetapi juga rentan terhadap subjektivitas dan inkonsistensi penilaian oleh manusia.

Analisis sentimen merupakan pendekatan komputasional yang terbukti efektif untuk mengotomatiskan proses ekstraksi opini dari data teks dalam skala besar. Dari berbagai algoritma yang tersedia, *Naive Bayes* telah menunjukkan keunggulan konsisten dalam tugas klasifikasi teks berbahasa Indonesia berkat kecepatan komputasi, kesederhanaan implementasi, dan kinerjanya yang kompetitif bahkan pada dataset berukuran terbatas (Hidayat et al., 2022). Penelitian oleh Dewi et al. (2023) yang menerapkan *Naive Bayes* pada data ulasan sistem informasi akademik kampus mencapai akurasi di atas 85%, mengindikasikan potensi besar metode ini untuk konteks serupa.

Namun demikian, terdapat beberapa celah metodologis yang belum terjawab dalam penelitian-penelitian terdahulu. Pertama, sebagian besar studi menggunakan data yang diperoleh melalui *website scraping* dari media sosial atau toko aplikasi, bukan data primer yang dikumpulkan langsung dari pengguna *platform target*. Kedua, proses pelabelan data sentimen umumnya dilakukan secara manual tanpa mekanisme validasi linguistik yang terstandar, sehingga konsistensi label tidak dapat dijamin. Ketiga, hasil klasifikasi jarang diintegrasikan ke dalam sistem manajemen informasi yang dapat diakses secara langsung oleh pengelola *platform* untuk pengambilan keputusan berbasis data.

Penelitian ini dirancang untuk menjawab celah tersebut melalui tiga kontribusi utama. Pertama, pengumpulan data *primer* secara langsung dari mahasiswa pengguna S-Learn melalui kuesioner yang disematkan dalam *platform* selama periode Ujian Akhir Semester (UAS) waktu yang dipilih karena mahasiswa telah berinteraksi intensif dengan *platform* selama hampir satu semester. Kedua, pelabelan sentimen otomatis berbasis leksikon *InSet* yang telah divalidasi oleh ahli bahasa, dengan pengujian empat metode determinasi label untuk memilih yang paling akurat. Ketiga, integrasi hasil klasifikasi ke dalam *dashboard* berbasis *web* yang memungkinkan pengelola S-Learn memantau sentimen pengguna secara terstruktur dan mengidentifikasi area prioritas perbaikan.

Berdasarkan pemaparan di atas, maka dibutuhkan sebuah analisis untuk mengukur dan menganalisis tingkat kenyamanan serta kelayakan S-LEARN sebagai *platform* dalam pelaksanaan ujian dan tugas di Jurusan Teknologi Informasi Politeknik Negeri Jember dengan menerapkan metode *Naive Bayes*. Melalui penelitian ini, diharapkan S-LEARN dapat terus berkembang menjadi *platform* LMS yang lebih baik dan lebih sesuai dengan kebutuhan akademik mahasiswa dan dosen. Temuan yang dihasilkan dari analisis data akan memberikan wawasan berharga bagi institusi dalam mengevaluasi dan meningkatkan kualitas layanan *e-learning*, sehingga mampu mendukung proses pembelajaran secara lebih optimal di masa depan. Berdasarkan pertimbangan tersebut, penelitian ini mengangkat judul: “Penerapan Algoritma *Naive Bayes* Untuk Analisis Sentimen Aplikasi S-LEARN Pada Jurusan Teknologi Informasi Politeknik Negeri Jember”.

1.2 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan *pipeline* praproses teks mencakup pembersihan teks, normalisasi kata tidak baku berbasis kamus yang divalidasi ahli bahasa, dan penanganan negasi dapat meningkatkan kualitas representasi fitur untuk klasifikasi sentimen ulasan pengguna S-Learn?
2. Bagaimana tingkat kinerja *algoritma Multinomial Naive Bayes* dibandingkan *Bernoulli Naive Bayes* dalam mengklasifikasikan sentimen ulasan pengguna S-Learn, diukur menggunakan *metrik akurasi, presisi, recall, dan F1-Score*?
3. Bagaimana hasil klasifikasi sentimen dapat divisualisasikan dalam *dashboard* berbasis *website* yang informatif untuk mendukung identifikasi kategori keluhan teknis dan pengambilan keputusan perbaikan oleh pengelola S-Learn?

1.3 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini ditetapkan sebagai berikut:

1. Penelitian mencakup penggunaan S-Learn di lingkungan Jurusan Teknologi Informasi Politeknik Negeri Jember pada periode semester ganjil tahun ajaran 2025/2026.
2. Kategori analisis sentimen untuk tahap pemodelan dibatasi pada klasifikasi *biner*, yaitu sentimen positif dan negatif. Data berlabel netral dieliminasi dari dataset pemodelan karena mengandung ambiguitas semantik yang tinggi.
3. Data yang digunakan berasal dari kolom ulasan teks bebas *kuesioner* kepuasan pengguna yang disebarakan melalui *Google Form* dan disematkan langsung dalam *platform* S-Learn pada saat Ujian Akhir Semester (UAS) Desember 2025.
4. Metode pelabelan sentimen menggunakan pendekatan *lexicon-based* dengan leksikon *InSet* (Indonesia *Sentiment Lexicon*) yang dikembangkan oleh Koto & Rahmanningrum (2017), diperkaya dengan kamus normalisasi slang dari *Kaggle* yang telah divalidasi oleh ahli bahasa.
5. Algoritma klasifikasi yang dibandingkan adalah *Multinomial Naive Bayes* dan *Bernoulli Naive Bayes* dengan empat konfigurasi vektorisasi (*CountVectorizer Unigram/Bigram* dan *TF-IDF Unigram/Bigram*).
6. Keluaran penelitian berupa *dashboard* interaktif berbasis *website* (*framework Laravel*) yang menampilkan hasil analisis sentimen, *confusion matrix*, dan kategorisasi keluhan teknis.
7. Penelitian tidak melakukan modifikasi atau perbaikan fitur pada aplikasi S-Learn; rekomendasi yang dihasilkan bersifat analitis berdasarkan persepsi pengguna.

1.4 1.4 Tujuan Penelitian

Terdapat beberapa tujuan dari penelitian ini antara lain sebagai berikut:

1. Membangun *pipeline* praproses teks bahasa Indonesia yang *komprehensif* mencakup pembersihan teks, normalisasi slang berbasis kamus tervalidasi, dan penanganan negasi untuk meningkatkan kualitas representasi fitur dataset ulasan S-Learn.

2. Membandingkan kinerja *Multinomial Naive Bayes* dan *Bernoulli Naive Bayes* pada empat konfigurasi *vektorisasi* untuk menentukan kombinasi model-*vektorisasi* terbaik dalam klasifikasi sentimen ulasan pengguna S-Learn.
3. Mengukur performa model klasifikasi terbaik menggunakan *metrik akurasi, presisi, recall, dan F1-Score* pada data uji yang independen.
4. Mengembangkan *dashboard* interaktif berbasis *website* sebagai media visualisasi hasil analisis sentimen yang dilengkapi fitur kategorisasi keluhan teknis dan rekomendasi perbaikan.
5. Memberikan rekomendasi berbasis data kepada pengelola S-Learn di Jurusan Teknologi Informasi POLIJE untuk mendukung pengembangan *platform* yang berorientasi pada kebutuhan pengguna.

1.5 1.5 Manfaat Penelitian

Berdasarkan tujuan penelitian di atas, manfaat yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Bagi Jurusan Teknologi Informasi POLIJE: menyediakan sistem analisis sentimen berbasis *dashboard* yang dapat digunakan sebagai instrumen monitoring kepuasan pengguna S-Learn secara berkelanjutan, termasuk identifikasi otomatis kategori masalah teknis seperti performa sistem, *bug/error*, dan antarmuka.
2. Bagi komunitas penelitian: memberikan kontribusi metodologis dalam hal pipeline praproses teks bahasa Indonesia informal yang mencakup validasi ahli bahasa, serta perbandingan *empiris* antara *Multinomial dan Bernoulli Naive Bayes* pada konteks *dataset* LMS kampus.
3. Bagi pengembangan keilmuan: menjadi referensi bagi penelitian selanjutnya terkait analisis sentimen pada platform *e-learning* berbahasa Indonesia, khususnya yang menggunakan data primer dari pengguna aktif.

