

Analisis Sentimen Ulasan Pengguna Aplikasi Pembelajaran Menggunakan Metode TF-IDF dan Support Vector Machine: Sebuah Kajian Literatur

Rian Aditia^{1,a,*}, Onis Alamsyah^{2,b}, Abiel Mulya Cipta^{3,c}
^aProgram Studi Ekonomi Pembangunan, Universitas Islam Al Azhar Mataram, Indonesia
^bProgram Studi Informatika, Universitas Bumi Gora Mataram, Indonesia
^cProgram Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram
¹rianmardinata03@gmail.com*, ²onisalamsyah@gmail.com, ³gameabel05@gmail.com
*Corresponding author

ARTICLE INFO	ABSTRACT
Article history Received July 14, 2026 Revised July 15, 2026 Accepted July 30, 2026 Published July 31, 2026 Keywords Sentiment Analysis TF-IDF Support Vector Machine Machine Learning Educational Applications  License by CC-BY-SA Copyright © 2026, The Author(s).	The rapid development of digital learning applications has significantly expanded public access to technology-based educational services. As the number of users continues to grow, user reviews published on application distribution platforms have become an important source of information for evaluating service quality and user satisfaction. <i>Sentiment analysis</i> has emerged as one of the most widely adopted approaches for automatically identifying user opinions through <i>text mining</i> and <i>machine learning</i> techniques. This study aims to review the application of the <i>Term Frequency–Inverse Document Frequency</i> (TF-IDF) method and the <i>Support Vector Machine</i> (SVM) algorithm for sentiment analysis of user reviews on educational applications based on previous studies. The research employed a literature review approach by examining scientific articles published between 2020 and 2025 and indexed in Google Scholar, IEEE Xplore, ScienceDirect, SpringerLink, and Scopus. The analysis focused on comparing the methods, text preprocessing techniques, feature extraction approaches, and classification performance reported in previous studies. The review indicates that the combination of TF-IDF and SVM is among the most widely used approaches for sentiment analysis due to its effective feature representation and stable classification performance in text-based datasets. Furthermore, the effectiveness of the classification process is influenced by text preprocessing, dataset characteristics, and model parameter selection. This study is expected to serve as a valuable reference for researchers and application developers in selecting appropriate sentiment analysis methods for digital learning applications.
How to cite: Aditia, R., Alamsyah, O., & Cipta, A. M. (2026). Analisis Sentimen Ulasan Pengguna Aplikasi Pembelajaran Menggunakan Metode TF-IDF dan Support Vector Machine: Sebuah Kajian Literatur. <i>Journal of Science and Technology: Alpha</i>, 2(3), 101-109. doi: https://doi.org/10.70716/alpha.v2i3.753	

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi digital pada berbagai sektor, termasuk bidang pendidikan. Pemanfaatan aplikasi pembelajaran digital memungkinkan proses belajar berlangsung secara fleksibel tanpa dibatasi ruang dan waktu. Berbagai platform pembelajaran menyediakan materi, latihan, evaluasi, serta komunikasi antara pengajar dan peserta didik secara daring. Meningkatnya penggunaan aplikasi tersebut menyebabkan jumlah ulasan pengguna pada platform distribusi aplikasi seperti Google Play Store terus bertambah dan menjadi sumber informasi penting untuk mengevaluasi kualitas layanan serta pengalaman pengguna.

Ulasan pengguna merupakan bentuk umpan balik (*feedback*) yang mencerminkan tingkat kepuasan, kritik, saran, maupun harapan terhadap suatu aplikasi. Informasi tersebut dapat dimanfaatkan oleh pengembang untuk mengetahui kelebihan dan kelemahan sistem sehingga pengembangan aplikasi dapat dilakukan secara lebih tepat sasaran. Namun, jumlah ulasan yang sangat besar menyebabkan proses analisis

secara manual menjadi tidak efisien dan berpotensi menghasilkan interpretasi yang subjektif. Oleh karena itu, diperlukan pendekatan otomatis yang mampu mengolah data teks secara cepat, akurat, dan konsisten.

Salah satu pendekatan yang banyak digunakan untuk menganalisis opini pengguna adalah analisis sentimen (*sentiment analysis*). Analisis sentimen merupakan bagian dari *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi polaritas suatu teks ke dalam kategori positif, negatif, maupun netral. Dalam bidang pendidikan, analisis sentimen dimanfaatkan untuk memahami persepsi peserta didik terhadap proses pembelajaran, sistem informasi akademik, maupun aplikasi pembelajaran digital sehingga dapat mendukung pengambilan keputusan berbasis data.

Berbagai penelitian menunjukkan bahwa analisis sentimen telah diterapkan pada beragam domain, seperti layanan transportasi, aplikasi keuangan, e-commerce, pelayanan publik, hingga aplikasi berbasis kecerdasan buatan. Meskipun objek penelitian berbeda, sebagian besar penelitian memiliki tahapan yang serupa, yaitu pengumpulan data ulasan, *text preprocessing*, ekstraksi fitur, proses klasifikasi, dan evaluasi model menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

Dalam proses analisis sentimen, representasi fitur menjadi salah satu faktor yang menentukan keberhasilan model klasifikasi. Metode *Term Frequency–Inverse Document Frequency* (TF-IDF) masih menjadi teknik pembobotan fitur yang paling banyak digunakan karena mampu mengubah data teks menjadi representasi numerik yang informatif dengan kompleksitas komputasi yang relatif rendah. TF-IDF juga terbukti kompatibel dengan berbagai algoritma *machine learning*, termasuk *Support Vector Machine* (SVM).

Di sisi lain, *Support Vector Machine* merupakan algoritma klasifikasi yang banyak digunakan pada penelitian analisis sentimen karena memiliki kemampuan yang baik dalam menangani data berdimensi tinggi. Berbagai penelitian melaporkan bahwa SVM mampu memberikan performa yang stabil pada data teks serta sering menghasilkan akurasi yang kompetitif dibandingkan algoritma klasifikasi tradisional lainnya. Oleh karena itu, kombinasi TF-IDF dan SVM masih menjadi salah satu pendekatan yang paling banyak digunakan dalam penelitian analisis sentimen.

Walaupun demikian, perkembangan *deep learning* dan model berbasis *transformer* seperti IndoBERT mulai menawarkan alternatif baru dalam analisis sentimen. Beberapa penelitian menunjukkan bahwa model berbasis *transformer* mampu memahami konteks kalimat dengan lebih baik dibandingkan metode tradisional. Namun, kebutuhan komputasi yang lebih tinggi menjadikan pendekatan TF-IDF dan SVM tetap relevan, terutama untuk penelitian yang membutuhkan efisiensi komputasi dan implementasi yang lebih sederhana.

Hasil telaah terhadap berbagai penelitian menunjukkan bahwa penerapan TF-IDF dan SVM telah banyak dilakukan pada ulasan aplikasi transportasi, aplikasi layanan publik, aplikasi e-commerce, maupun aplikasi berbasis kecerdasan buatan. Akan tetapi, kajian yang secara khusus membahas penerapan kedua metode tersebut pada aplikasi pembelajaran digital masih relatif lebih sedikit dibandingkan domain lainnya. Kondisi ini menunjukkan bahwa masih terdapat ruang untuk melakukan sintesis pengetahuan mengenai efektivitas pendekatan tersebut pada konteks pendidikan digital.

Selain itu, penelitian-penelitian terdahulu masih menunjukkan variasi pada tahapan *text preprocessing*, teknik pelabelan data, pemilihan parameter model, maupun metode evaluasi yang digunakan. Perbedaan tersebut menyebabkan hasil penelitian sulit dibandingkan secara langsung sehingga diperlukan suatu kajian yang merangkum karakteristik metode, kelebihan, keterbatasan, dan kecenderungan hasil penelitian yang telah dipublikasikan.

Berdasarkan kondisi tersebut, terdapat *research gap* berupa belum banyaknya kajian literatur yang secara khusus membahas penerapan TF-IDF dan SVM pada analisis sentimen ulasan aplikasi pembelajaran digital. Sebagian besar publikasi masih berfokus pada implementasi eksperimen terhadap satu objek penelitian, sedangkan pembahasan komprehensif mengenai tren penelitian, efektivitas metode, serta tantangan implementasinya masih terbatas.

Novelty dari artikel ini terletak pada penyusunan sintesis literatur yang mengintegrasikan berbagai penelitian mengenai penerapan TF-IDF dan SVM dalam analisis sentimen ulasan aplikasi pembelajaran. Kajian ini tidak hanya membahas tahapan metode yang digunakan, tetapi juga mengidentifikasi perkembangan penelitian, kelebihan dan keterbatasan pendekatan, serta peluang pengembangan pada penelitian selanjutnya berdasarkan publikasi ilmiah terbaru.

Berdasarkan uraian tersebut, penelitian ini bertujuan mengkaji penerapan metode TF-IDF dan algoritma *Support Vector Machine* pada analisis sentimen ulasan pengguna aplikasi pembelajaran melalui pendekatan kajian literatur. Hasil kajian diharapkan dapat memberikan gambaran mengenai perkembangan penelitian, efektivitas metode yang digunakan, serta menjadi referensi bagi peneliti dan pengembang aplikasi dalam memilih pendekatan yang sesuai untuk pengembangan sistem analisis sentimen pada bidang pendidikan digital.

METODE PENELITIAN

Penelitian ini menggunakan metode kajian literatur (*Literature Review*) dengan pendekatan deskriptif kualitatif. Metode ini dipilih karena mampu memberikan gambaran yang komprehensif mengenai perkembangan penelitian terkait penerapan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) dan algoritma *Support Vector Machine* (SVM) dalam analisis sentimen ulasan pengguna aplikasi pembelajaran. Melalui kajian literatur, berbagai hasil penelitian terdahulu dapat dibandingkan, dianalisis, serta disintesis sehingga menghasilkan pemahaman yang lebih mendalam mengenai efektivitas metode, kelebihan, keterbatasan, serta peluang pengembangannya pada masa mendatang. Pendekatan ini juga memungkinkan peneliti untuk mengidentifikasi kecenderungan penelitian yang telah dilakukan oleh berbagai peneliti baik di tingkat nasional maupun internasional.

Sumber data dalam penelitian ini berupa artikel ilmiah yang diperoleh dari berbagai basis data akademik yang memiliki kredibilitas tinggi, yaitu Google Scholar, IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, dan Scopus. Pemilihan basis data tersebut dilakukan karena menyediakan artikel ilmiah yang telah melalui proses *peer review* sehingga kualitas informasi yang diperoleh lebih dapat dipertanggungjawabkan secara akademik. Artikel yang digunakan dalam penelitian ini merupakan publikasi yang diterbitkan pada rentang waktu 2020 hingga 2025 agar hasil kajian mampu menggambarkan perkembangan terkini mengenai analisis sentimen, *Natural Language Processing* (NLP), *text mining*, serta penerapan algoritma *machine learning* dalam pengolahan data teks.

Proses pencarian literatur dilakukan menggunakan kombinasi beberapa kata kunci yang berkaitan dengan topik penelitian. Kata kunci tersebut meliputi *sentiment analysis*, *text mining*, *Natural Language Processing*, *Term Frequency–Inverse Document Frequency*, *TF-IDF*, *Support Vector Machine*, *machine learning*, *educational application*, *learning application reviews*, serta *Google Play Store reviews*. Penggunaan kombinasi kata kunci bertujuan untuk memperoleh artikel yang memiliki relevansi tinggi terhadap fokus penelitian sehingga proses kajian dapat dilakukan secara lebih efektif dan sistematis.

Tahapan pencarian artikel dilakukan secara bertahap dengan memanfaatkan fasilitas pencarian lanjutan (*advanced search*) pada masing-masing basis data. Pada tahap awal diperoleh sejumlah artikel yang memiliki keterkaitan dengan topik penelitian berdasarkan judul, kata kunci, maupun abstrak. Selanjutnya dilakukan proses penyaringan awal (*screening*) untuk menghilangkan artikel yang tidak relevan, artikel duplikat, serta publikasi yang tidak memiliki akses terhadap naskah lengkap (*full text*). Tahapan ini bertujuan untuk memastikan bahwa seluruh artikel yang dianalisis benar-benar sesuai dengan ruang lingkup penelitian.

Setelah proses penyaringan awal selesai dilakukan, artikel yang tersisa dievaluasi kembali berdasarkan kriteria inklusi dan kriteria eksklusi. Kriteria inklusi meliputi artikel yang membahas analisis sentimen menggunakan metode TF-IDF dan algoritma SVM, artikel yang dipublikasikan pada jurnal atau prosiding ilmiah bereputasi, tersedia dalam bentuk teks lengkap, serta ditulis menggunakan bahasa Indonesia atau bahasa Inggris. Sebaliknya, artikel yang tidak membahas analisis sentimen, menggunakan metode yang tidak

relevan dengan fokus penelitian, tidak tersedia secara lengkap, atau merupakan artikel duplikat dikeluarkan dari proses analisis.

Artikel yang memenuhi seluruh kriteria kemudian dianalisis secara mendalam melalui proses ekstraksi data. Pada tahap ini, setiap artikel diidentifikasi berdasarkan beberapa aspek penting, antara lain nama penulis, tahun publikasi, tujuan penelitian, sumber data yang digunakan, jumlah dataset, metode *text preprocessing*, teknik ekstraksi fitur, algoritma klasifikasi, metode evaluasi, hasil penelitian, serta kelebihan dan keterbatasan yang dilaporkan. Informasi tersebut selanjutnya disusun dalam bentuk tabel sintesis sehingga memudahkan proses perbandingan antarpengelitian.

Proses analisis data dilakukan menggunakan teknik analisis isi (*content analysis*). Teknik ini digunakan untuk mengidentifikasi pola, persamaan, serta perbedaan antarpengelitian yang telah dipublikasikan. Seluruh artikel dikelompokkan berdasarkan karakteristik penelitian, objek yang dikaji, metode ekstraksi fitur, algoritma klasifikasi yang digunakan, serta hasil evaluasi yang diperoleh. Selanjutnya dilakukan proses interpretasi terhadap hasil-hasil penelitian tersebut sehingga dapat diperoleh gambaran mengenai perkembangan metode TF-IDF dan SVM dalam analisis sentimen ulasan pengguna aplikasi pembelajaran.

Selain melakukan identifikasi terhadap metode yang digunakan, penelitian ini juga menganalisis berbagai tantangan yang sering ditemukan dalam implementasi analisis sentimen. Beberapa tantangan yang banyak dilaporkan dalam penelitian terdahulu meliputi penggunaan bahasa tidak baku, singkatan, istilah gaul, campuran bahasa, ketidakseimbangan jumlah data (*class imbalance*), serta variasi kualitas data ulasan yang berasal dari pengguna aplikasi. Faktor-faktor tersebut menjadi perhatian penting karena dapat memengaruhi kualitas representasi fitur maupun performa model klasifikasi yang dihasilkan.

Selanjutnya dilakukan proses sintesis hasil penelitian dengan membandingkan efektivitas penerapan TF-IDF dan SVM pada berbagai objek penelitian. Analisis tidak hanya difokuskan pada hasil akhir yang dilaporkan oleh masing-masing penelitian, tetapi juga mempertimbangkan tahapan *text preprocessing*, karakteristik dataset, teknik pembobotan fitur, metode evaluasi, serta konfigurasi algoritma yang digunakan. Dengan demikian, kajian ini diharapkan mampu memberikan gambaran yang lebih menyeluruh mengenai kondisi penelitian analisis sentimen pada aplikasi pembelajaran digital.

Untuk menjaga objektivitas hasil penelitian, seluruh artikel dianalisis secara sistematis berdasarkan informasi yang disampaikan oleh masing-masing penulis tanpa melakukan perubahan terhadap hasil penelitian yang telah dipublikasikan. Setiap temuan dibandingkan dengan penelitian lain yang memiliki topik serupa sehingga dapat diidentifikasi kecenderungan (*research trend*), kesenjangan penelitian (*research gap*), serta peluang pengembangan penelitian pada masa mendatang.

Berdasarkan keseluruhan tahapan tersebut, penelitian ini menghasilkan suatu sintesis literatur mengenai penerapan metode TF-IDF dan algoritma SVM dalam analisis sentimen ulasan pengguna aplikasi pembelajaran. Hasil sintesis diharapkan dapat memberikan kontribusi akademik berupa pemetaan perkembangan penelitian, identifikasi keunggulan dan keterbatasan metode yang digunakan, serta rekomendasi bagi peneliti maupun pengembang aplikasi dalam memilih pendekatan analisis sentimen yang sesuai dengan karakteristik data dan tujuan penelitian.

HASIL DAN PEMBAHASAN

Perkembangan teknologi informasi telah mendorong meningkatnya penggunaan aplikasi pembelajaran digital sebagai media pendukung proses belajar mengajar. Berbagai aplikasi pembelajaran menyediakan layanan berupa materi pembelajaran, latihan soal, evaluasi, forum diskusi, hingga komunikasi antara pendidik dan peserta didik. Seiring dengan meningkatnya jumlah pengguna, jumlah ulasan yang diberikan pengguna pada platform distribusi aplikasi juga terus bertambah. Ulasan tersebut menjadi salah satu sumber informasi penting yang dapat dimanfaatkan untuk mengetahui persepsi pengguna terhadap kualitas layanan yang diberikan.

Hasil kajian menunjukkan bahwa analisis sentimen menjadi salah satu pendekatan yang paling banyak digunakan dalam mengolah data ulasan pengguna. Melalui analisis sentimen, opini pengguna dapat dikelompokkan ke dalam kategori sentimen positif, negatif, maupun netral sehingga memudahkan pengembang dalam mengevaluasi kualitas aplikasi. Pendekatan ini dinilai lebih efektif dibandingkan analisis manual karena mampu mengolah data dalam jumlah besar secara otomatis serta menghasilkan informasi yang lebih objektif.

Selain digunakan untuk mengevaluasi aplikasi pembelajaran, analisis sentimen juga telah diterapkan pada berbagai bidang lain seperti e-commerce, media sosial, pelayanan publik, kesehatan, dan transportasi. Hal tersebut menunjukkan bahwa analisis sentimen telah berkembang menjadi salah satu teknik penting dalam proses pengambilan keputusan berbasis data (*data-driven decision making*).

Penerapan Metode TF-IDF

Berdasarkan hasil kajian terhadap berbagai penelitian, metode *Term Frequency–Inverse Document Frequency* (TF-IDF) masih menjadi salah satu metode ekstraksi fitur yang paling banyak digunakan dalam penelitian analisis sentimen. Metode ini mengubah data teks menjadi representasi numerik melalui proses pembobotan kata berdasarkan frekuensi kemunculan pada dokumen dan tingkat kepentingannya terhadap seluruh kumpulan dokumen.

Keunggulan TF-IDF terletak pada kemudahan implementasi, efisiensi komputasi, serta kemampuannya menghasilkan representasi fitur yang cukup baik untuk berbagai algoritma klasifikasi. Oleh karena itu, metode ini masih menjadi pilihan utama dalam penelitian yang menggunakan pendekatan *machine learning* klasik.

Meskipun demikian, TF-IDF memiliki beberapa keterbatasan. Metode ini belum mampu memahami hubungan semantik antar kata maupun konteks kalimat secara menyeluruh. Kata-kata yang memiliki makna sama tetapi ditulis dalam bentuk berbeda tetap dianggap sebagai fitur yang berbeda. Kondisi tersebut menyebabkan kualitas proses *text preprocessing* menjadi faktor yang sangat menentukan keberhasilan proses klasifikasi.

Penerapan Support Vector Machine

Algoritma *Support Vector Machine* (SVM) merupakan salah satu algoritma klasifikasi yang paling sering digunakan dalam penelitian analisis sentimen. SVM memiliki kemampuan yang baik dalam menangani data berdimensi tinggi sehingga sangat sesuai diterapkan pada data teks yang dihasilkan dari proses pembobotan TF-IDF. Hasil kajian menunjukkan bahwa sebagian besar penelitian menggunakan *kernel linear* karena lebih sesuai untuk data teks yang memiliki jumlah fitur besar. Selain menghasilkan proses komputasi yang lebih cepat, *kernel linear* juga mampu menghasilkan performa klasifikasi yang stabil pada berbagai objek penelitian. Keunggulan lain dari SVM adalah kemampuannya membentuk batas pemisah (*hyperplane*) yang optimal antar kelas sehingga proses klasifikasi dapat dilakukan secara efektif. Hal tersebut menjadikan SVM sebagai salah satu algoritma yang masih banyak digunakan hingga saat ini meskipun telah berkembang berbagai metode berbasis *deep learning*.

Pengaruh Tahapan Text Preprocessing

Hasil kajian menunjukkan bahwa tahapan *text preprocessing* merupakan salah satu faktor yang paling berpengaruh terhadap kualitas analisis sentimen. Hampir seluruh penelitian menerapkan tahapan *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming* sebelum proses ekstraksi fitur dilakukan. Tahapan tersebut bertujuan menghilangkan informasi yang tidak relevan sehingga representasi data menjadi lebih bersih dan konsisten. Dengan demikian, algoritma klasifikasi mampu mempelajari pola data secara lebih baik dibandingkan apabila data diproses tanpa tahapan *preprocessing*. Selain meningkatkan kualitas representasi data, *preprocessing* juga berperan dalam mengurangi jumlah fitur yang tidak penting sehingga proses komputasi menjadi lebih efisien.

Kelebihan dan Keterbatasan Pendekatan TF-IDF dan SVM

Berdasarkan hasil sintesis literatur, kombinasi TF-IDF dan SVM memiliki beberapa kelebihan. Pertama, metode ini relatif mudah diimplementasikan menggunakan berbagai perangkat lunak *machine learning*. Kedua, kebutuhan sumber daya komputasi relatif rendah dibandingkan model berbasis *deep learning*. Ketiga, metode ini mampu memberikan hasil klasifikasi yang stabil pada berbagai jenis data teks.

Di sisi lain, pendekatan ini juga memiliki beberapa keterbatasan. TF-IDF belum mampu menangkap hubungan semantik antar kata sehingga representasi fitur masih bergantung pada frekuensi kemunculan kata. Selain itu, performa SVM dipengaruhi oleh kualitas data, distribusi kelas, dan parameter model yang digunakan. Oleh karena itu, beberapa penelitian mulai mengombinasikan TF-IDF dengan metode representasi fitur yang lebih modern untuk meningkatkan kualitas klasifikasi.

Implikasi terhadap Pengembangan Aplikasi Pembelajaran

Berdasarkan hasil kajian, analisis sentimen memberikan manfaat yang signifikan bagi pengembang aplikasi pembelajaran. Informasi yang diperoleh dari ulasan pengguna dapat dimanfaatkan untuk mengidentifikasi aspek yang paling banyak mendapatkan apresiasi maupun keluhan. Hasil tersebut dapat digunakan sebagai dasar dalam menentukan prioritas pengembangan fitur, memperbaiki kualitas layanan, serta meningkatkan kepuasan pengguna.

Selain bagi pengembang aplikasi, hasil analisis sentimen juga memberikan manfaat bagi institusi pendidikan dan peneliti. Informasi mengenai persepsi pengguna dapat dijadikan bahan evaluasi terhadap efektivitas penggunaan aplikasi pembelajaran serta menjadi dasar dalam pengembangan penelitian selanjutnya pada bidang pendidikan digital.

Peluang Penelitian Selanjutnya

Hasil kajian menunjukkan bahwa penelitian mengenai analisis sentimen masih memiliki peluang yang sangat besar untuk dikembangkan. Penelitian selanjutnya dapat mengeksplorasi penggunaan metode representasi fitur berbasis *word embedding* maupun model *transformer* untuk meningkatkan kemampuan model dalam memahami konteks bahasa.

Selain itu, penelitian mendatang juga dapat melakukan perbandingan berbagai algoritma klasifikasi pada dataset yang sama sehingga diperoleh gambaran yang lebih komprehensif mengenai performa masing-masing metode. Penggunaan dataset yang lebih besar, proses pelabelan yang lebih baik, serta evaluasi menggunakan berbagai metrik juga menjadi peluang untuk meningkatkan kualitas penelitian di masa mendatang.

KESIMPULAN

Berdasarkan hasil kajian literatur yang telah dilakukan, dapat disimpulkan bahwa analisis sentimen merupakan salah satu pendekatan yang memiliki peran penting dalam mengolah data ulasan pengguna aplikasi pembelajaran menjadi informasi yang bermanfaat bagi pengembang maupun institusi pendidikan. Meningkatnya jumlah ulasan pengguna pada berbagai platform distribusi aplikasi telah menghasilkan data tekstual dalam jumlah besar yang sulit dianalisis secara manual. Oleh karena itu, pemanfaatan teknik Natural Language Processing (NLP) dan *machine learning* menjadi solusi yang efektif untuk mengidentifikasi kecenderungan opini pengguna secara otomatis, cepat, dan lebih objektif. Berbagai penelitian menunjukkan bahwa hasil analisis sentimen dapat dimanfaatkan untuk mengevaluasi kualitas layanan, meningkatkan pengalaman pengguna (*user experience*), serta mendukung proses pengambilan keputusan dalam pengembangan aplikasi pembelajaran.

Hasil sintesis terhadap berbagai penelitian terdahulu menunjukkan bahwa metode Term Frequency–Inverse Document Frequency (TF-IDF) masih menjadi salah satu teknik ekstraksi fitur yang paling banyak digunakan dalam analisis sentimen berbasis teks. Metode ini dinilai memiliki keunggulan berupa proses

komputasi yang relatif sederhana, efisien, mudah diimplementasikan, serta mampu menghasilkan representasi fitur yang baik pada berbagai jenis data teks. Selain itu, TF-IDF dapat dikombinasikan dengan berbagai algoritma klasifikasi sehingga tetap menjadi pendekatan yang relevan, terutama pada penelitian yang menggunakan metode *machine learning* klasik. Meskipun demikian, beberapa penelitian juga menunjukkan bahwa TF-IDF memiliki keterbatasan dalam memahami hubungan semantik antar kata sehingga kualitas hasil klasifikasi sangat dipengaruhi oleh tahapan *text preprocessing* yang diterapkan sebelum proses ekstraksi fitur dilakukan.

Kajian ini juga menunjukkan bahwa Support Vector Machine (SVM) merupakan salah satu algoritma klasifikasi yang paling banyak diterapkan dalam penelitian analisis sentimen. Popularitas algoritma ini dipengaruhi oleh kemampuannya dalam menangani data berdimensi tinggi, menghasilkan proses klasifikasi yang stabil, serta memiliki kemampuan generalisasi yang baik pada berbagai karakteristik dataset. Berbagai penelitian pada bidang pendidikan maupun bidang lainnya melaporkan bahwa kombinasi TF-IDF dan SVM mampu memberikan performa klasifikasi yang kompetitif dibandingkan beberapa algoritma *machine learning* lainnya, terutama ketika didukung oleh proses *text preprocessing* yang baik. Oleh karena itu, kombinasi kedua metode tersebut masih menjadi salah satu pendekatan yang paling banyak digunakan sebagai dasar dalam pengembangan sistem analisis sentimen.

Selain mengidentifikasi keunggulan metode, hasil kajian juga menunjukkan bahwa keberhasilan proses analisis sentimen tidak hanya ditentukan oleh pemilihan algoritma klasifikasi, tetapi dipengaruhi oleh berbagai faktor lain, seperti kualitas dataset, teknik pelabelan data, tahapan *case folding*, *cleaning*, *tokenization*, *stopword removal*, *stemming*, serta keseimbangan distribusi kelas sentimen. Perbedaan karakteristik dataset dan strategi *preprocessing* yang digunakan pada setiap penelitian menyebabkan variasi hasil yang diperoleh, sehingga proses evaluasi model perlu mempertimbangkan konteks penelitian dan karakteristik data yang dianalisis. Hal tersebut menunjukkan bahwa tidak terdapat satu metode yang selalu memberikan hasil terbaik pada seluruh kondisi, melainkan efektivitas metode sangat dipengaruhi oleh kualitas data dan proses pengolahannya.

Perkembangan teknologi kecerdasan buatan juga menunjukkan bahwa model berbasis *deep learning* dan *transformer*, seperti BERT dan IndoBERT, mulai banyak digunakan dalam penelitian analisis sentimen karena memiliki kemampuan yang lebih baik dalam memahami konteks bahasa. Meskipun demikian, hasil kajian menunjukkan bahwa pendekatan berbasis TF-IDF dan SVM masih memiliki posisi yang penting sebagai metode dasar (*baseline*) karena lebih sederhana, efisien, mudah direproduksi, dan tidak memerlukan sumber daya komputasi yang tinggi. Oleh sebab itu, pendekatan ini masih sangat relevan untuk diterapkan pada institusi pendidikan maupun penelitian yang memiliki keterbatasan infrastruktur komputasi.

Secara akademik, kajian ini memberikan kontribusi berupa sintesis terhadap perkembangan penelitian mengenai penerapan TF-IDF dan SVM dalam analisis sentimen ulasan aplikasi pembelajaran. Hasil sintesis tersebut diharapkan dapat menjadi referensi bagi peneliti dalam memahami perkembangan metode, mengidentifikasi kelebihan dan keterbatasan pendekatan yang telah digunakan, serta menemukan peluang penelitian baru yang masih terbuka. Selain itu, bagi pengembang aplikasi pembelajaran, hasil kajian ini dapat dijadikan dasar dalam menentukan pendekatan analisis sentimen yang sesuai untuk mengevaluasi kualitas layanan berdasarkan umpan balik pengguna secara berkelanjutan.

Sebagai rekomendasi, penelitian selanjutnya disarankan untuk memperluas cakupan kajian dengan melibatkan lebih banyak artikel dari jurnal internasional bereputasi, membandingkan metode *machine learning* klasik dengan pendekatan berbasis *deep learning* maupun *transformer*, serta mengkaji penerapan teknik ekstraksi fitur yang lebih modern. Selain itu, penelitian lanjutan juga dapat mengembangkan kajian pada berbagai jenis aplikasi pendidikan, seperti *Learning Management System* (LMS), aplikasi pembelajaran berbasis kecerdasan buatan, maupun platform pembelajaran adaptif. Dengan demikian, hasil penelitian diharapkan mampu memberikan kontribusi yang lebih komprehensif terhadap pengembangan ilmu

pengetahuan, khususnya pada bidang *Natural Language Processing*, *text mining*, dan analisis sentimen dalam ekosistem pendidikan digital.

DAFTAR PUSTAKA

- Aggarwal, C. C. (2023). *Machine learning for text*. Springer.
- Fahima, M. M. F., & Sumathipala, P. (2025). Sentiment analysis and its application in the education sector: A comprehensive review. *Elementaria: Journal of Educational Research*, 3(1), 105–131. <https://doi.org/10.61166/elm.v3i1.89>
- Fikri, M., & Sarno, R. (2019). A comparative study of sentiment analysis using SVM and SentiWordNet. *Indonesian Journal of Electrical Engineering and Computer Science*, 13(3), 902–909.
- Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing* (3rd ed., draft). Pearson.
- Kurniawan, M. B., Hanafi, H., Hikmianto, R., & Muslihah, I. (2025). Hyperparameter optimization of TF-IDF and SVM via grid search for sentiment analysis of Traveloka customer reviews. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*.
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions* (2nd ed.). Cambridge University Press.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.
- Permata Sari, A. A. I. I., & Suhartana, I. K. G. (2025). Penggunaan metode SVM dan Naive Bayes pada analisis sentiment ulasan aplikasi Edlink. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(3), 575–584. <https://doi.org/10.24843/JNATIA.2025.v03.i03.p11>
- Prasetyo, S. A., Yusuf, D., & Syamsu, M. (2026). Penerapan Natural Language Processing untuk analisis sentimen ulasan aplikasi Allo Bank menggunakan algoritma SVM. *Jurnal Sistem Informasi (JUSIN)*, 7(1), 65–75.
- Putra, J. A., Dharmawan, A., & Gondohanindijo, J. (2024). Sentimen analisis aplikasi Digitalent Mobile menggunakan Naïve Bayes dan SVM dengan ekstraksi fitur TF-IDF. *INTECOMS: Journal of Information Technology and Computer Science*.
- Rokhman, A., Berlilana, B., & Arsi, P. (2025). Comparison of text representation methods for sentiment analysis using Support Vector Machine. *Journal of Advances in Information and Industrial Technology*.
- Sa'ad, M. I., Muflikhah, L., & Bachtiar, F. A. (2025). Perbandingan SVM dan IndoBERT untuk analisis sentimen layanan akademik mahasiswa. *Bulletin of Information Technology*, 7(2). <https://doi.org/10.47065/bit.v7i2.2913>
- Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). Sentiment analysis and opinion mining on educational data: A survey. *arXiv*. <https://arxiv.org/abs/2302.04359>
- Syabina, N. A., Sulistyorini, P., & Sulistiyaningsih, R. (2025). Analisis sentimen ulasan pengguna aplikasi Stimuler menggunakan Support Vector Machine (SVM) dan TF-IDF. *Smart Comp: Jurnalnya Orang Pintar Komputer*. <https://doi.org/10.30591/smartcomp.v15i2.10091>
- Zebua, D. B. M., & Dwidasmara, I. B. G. (2025). Analisis sentimen ulasan aplikasi Citilink menggunakan metode Support Vector Machine dengan TF-IDF. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(2), 439–446. <https://doi.org/10.24843/JNATIA.2025.v03.i02.p24>

- Algipari, R. Z., & Tresnawati, S. (2025). Sentiment analysis of reading difficulties in Grade 7 secondary school students using the Support Vector Machine algorithm. *Journal of Artificial Intelligence and Software Engineering*.
- Madjid, M. F., Ratnawati, D. E., & Rahayudi, B. (2025). Sentiment analysis on app reviews using Support Vector Machine and Naïve Bayes classification. *Sinkron: Jurnal dan Penelitian Teknik Informatika*. <https://doi.org/10.33395/sinkron.v8i1.12161>
- Nabilah, A., Sari, N. I., Afrina, M., & Ibrahim, A. (2026). The sentiment analysis of Indonesian startup application reviews using TF-IDF + SVM and FastText: A comparative study. *JITeCS (Journal of Information Technology and Computer Science)*, 10(3). <https://doi.org/10.25126/jitecs.2025103807>
- Ramadhan, H. S., Akbar, A. S., Sinaga, K. Y., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). Sentiment analysis of AI adoption in Indonesian higher education using machine learning and transformer-based models. *arXiv*. <https://arxiv.org/abs/2604.27439>
- Weko, H. Y., & Suparwito, H. (2025). SVM and ensemble majority voting algorithm on sentiment analysis of using ChatGPT in education. *International Journal of Applied Sciences and Smart Technologies*.
- Zahra, N. Z., Farhanatussaidah, S., Afifah, N. N., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). Benchmarking Logistic Regression, SVM, Naive Bayes, and IndoBERT fine-tuning for sentiment analysis on Indonesian product reviews. *arXiv*. <https://arxiv.org/abs/2605.03439>