



Sistem Analisis Sentimen Untuk Evaluasi Kinerja Dosen dengan Metode Naïve Bayes

Ade Sasmita^{1*}, Gede Aditra Pradnyana², Dewa Gede Hendra Divayana³



^{1,2,3} Progam Studi Pendidikan Teknik Informatika, Universitas Pendidikan Ganesha, Singaraja, Indonesia

ARTICLE INFO

Article history:

Received February 07, 2022

Revised February 12, 2022

Accepted July 19, 2022

Available online October 25, 2022

Kata Kunci:

Evaluasi Kinerja Dosen, Naïve Bayes, Sistem Analisis Sentimen

Keywords:

Lecturer Performance Evaluation, Naïve Bayes, Sentiment Analysis System



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2022 by Author. Published by Universitas Pendidikan Ganesha.

ABSTRAK

Penilaian atau evaluasi dosen dapat dilakukan dengan pengisian angket penilaian oleh mahasiswa. Penilaian yang diberikan mahasiswa dapat memberikan umpan balik dan rekomendasi perbaikan terkait proses kegiatan pembelajaran di universitas. Saat ini data angket tersebut belum diolah dengan menggunakan teknologi pengklasifikasian analisis data. Penelitian ini bertujuan untuk menganalisis sistem analisis sentimen untuk evaluasi kinerja dosen dengan metode naïve bayes. Jenis penelitian ini merupakan penelitian pengembangan perangkat lunak yang menggunakan metode penelitian System Development Life Cycle (SDLC) dengan model waterfall. Subjek yang terlibat dalam penelitian ini terdiri dari admin, dosen dan mahasiswa. Metode pengumpulan data menggunakan observasi dan studi pustaka. Hasil penelitian menunjukkan pengelolaan data tersebut dapat dilakukan dengan menggunakan sistem analisis sentimen dengan metode Naïve Bayes. Data angket akan diklasifikasikan ke dalam 2 (dua) kelas, yaitu positif dan negatif. Data tersebut akan dievaluasi menggunakan k-fold cross validation dan confusion matrix. Hasil evaluasi menunjukkan sistem mampu memprediksi kelas sentimen pada kritik dan saran mahasiswa dengan rata-rata precision sebesar 90.84%, recall sebesar 95.73%, accuracy sebesar 90.6% dan f-measure sebesar 93.22%. Data analisis tersebut dapat dimanfaatkan untuk mendukung hasil evaluasi kinerja dari dosen.

ABSTRACT

Assessment or evaluation of lecturers can be done by filling out an assessment questionnaire by students. Assessments given by students can provide feedback and recommendations for improvements related to the process of learning activities at the university. Currently, the questionnaire data has not been processed using data analysis classification technology. This study aims to analyze the sentiment analysis system for evaluating lecturer performance using the Naïve Bayes method. This type of research is software development research using the System Development Life Cycle (SDLC) research method with the waterfall model. The subjects involved in this study consisted of admins, lecturers and students. Methods of data collection using observation and literature study. The results of the study show that data management can be done using a sentiment analysis system using the Naïve Bayes method. Questionnaire data will be classified into 2 (two) classes, namely positive and negative. The data will be evaluated using k-fold cross validation and confusion matrix. The evaluation results show that the system is able to predict the sentiment class on student criticism and suggestions with an average precision of 90.84%, recall of 95.73%, accuracy of 90.6% and f-measure of 93.22%. The analysis data can be used to support the performance evaluation results of lecturers.

1. PENDAHULUAN

Perguruan tinggi atau universitas pada setiap semesternya melakukan proses penilaian atau evaluasi dosen untuk mengetahui pencapaian tujuan pengajaran (Arsini, 2016; Rahmawati & Trimulyono, 2021). Upaya mencapai pengajaran yang baik tidaklah mudah, tergantung dari dosen yang terlibat langsung dalam proses perkuliahan (Sunardi et al., 2018). Proses perkuliahan penilaian kinerja penting bagi para dosen pengampu mata kuliah untuk melihat hasil proses belajar mengajar (Atmaja & Wijaya, 2019; Syahrir et al., 2021). Evaluasi kinerja kepada para dosen digunakan untuk mengetahui kontribusi yang sudah dilakukan dalam pencapaian tujuan pengajaran. Evaluasi kinerja merupakan sistem penilaian untuk mengukur tingkat keberhasilan aktivitas kerja terhadap tanggung jawab yang diberikan (Rahadi, 2010; Suryani, 2017). Alat evaluasi berupa angket penilaian sudah umum digunakan, angket ini berisi sejumlah pertanyaan yang diberikan kepada responden untuk mendapatkan balikan atau penilaian terhadap objek tertentu (Riany et al., 2016). Salah satu proses penilaian atau evaluasi dosen ini dilakukan dengan pengisian angket penilaian untuk dosen yang diisi oleh mahasiswa. Universitas Pendidikan Ganesha (Undiksha) menyediakan angket dosen yang diberikan oleh mahasiswa melalui SIAK (Sistem

*Corresponding author.

E-mail addresses: ade.sasmita@undiksha.ac.id (Ade Sasmita)

Informasi Akademik), angket ini hanya diisi oleh mahasiswa satu kali setiap akhir semester. Angket atau kuesioner adalah teknik pengumpulan data yang dilakukan dengan memberikan seperangkat pertanyaan kepada responden untuk kemudian dijawab (Risdiyanto, 2008). Angket dosen yang ada di Undiksha merupakan jenis angket terbuka yang isian responden dalam bentuk kalimat saran dan kritik. Pada angket tersebut, mahasiswa dapat memberikan penilaian yang positif maupun negatif terhadap kinerja dosen dalam hal proses perkuliahan yang telah diikutinya, yang pada akhirnya akan menggambarkan tingkat kepuasan mahasiswa terhadap proses perkuliahan di kampus. Selama ini data angket tersebut belum dianalisis dan dimanfaatkan lebih lanjut. Data tersebut jumlahnya mencapai ratusan bahkan ribuan data yang berasal dari seluruh mahasiswa untuk setiap mata kuliah. Data yang ada pada jenis angket terbuka sulit untuk dilakukan proses analisis secara manual dikarenakan pada angket ini isian dari responden berupa kalimat dengan konten beragam. Jika dilakukan analisis satu persatu terhadap isian angket secara manual akan memerlukan waktu dan tenaga, maka diperlukan media yang dapat membantu dalam melakukan proses analisis isian angket terbuka tersebut secara otomatis, sehingga dapat dilakukan lebih efektif dan efisien.

Pengolahan informasi yang berupa komentar atau opini dapat dilakukan dengan mengklasifikasikannya kedalam bentuk komentar positif atau negatif, pengklasifikasian ini sering disebut analisis sentimen. Analisis sentimen adalah bidang ilmu atau metode yang memproses, menganalisa dan mengolah data tekstual berupa pendapat, evaluasi, penilaian, sikap dan emosi publik secara otomatis untuk mendapatkan sebuah sentimen (Sipayung et al., 2016; Sunardi et al., 2018). Tugas dalam analisis sentimen yaitu mengelompokkan aspek positif atau negatif dari teks yang terdapat dalam sebuah dokumen, kalimat atau pendapat (Fitri et al., 2019; Ling et al., 2014). Pemanfaatan analisis sentimen banyak digunakan sebagai bahan untuk mendapatkan umpan balik atau respon terhadap jasa, produk, atau topik yang diberikan agar nantinya dapat digunakan untuk evaluasi. Terdapat dua metode yang umumnya dipakai dalam analisis sentimen yaitu *lexicon-based* dan *learning-based* (Azhar, 2017; Birjali et al., 2021). Pembuatan sebuah analisis sentimen banyak yang harus dipersiapkan terlebih dahulu, salah satunya dengan memilih metode yang akan digunakan. Pada teknik klasifikasi terdapat beberapa algoritma, salah satunya adalah algoritma Naive Bayes yang diterapkan pada penelitian ini. Naive Bayes merupakan proses klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman yang sudah ada di masa lalu sehingga dikenal sebagai Teorema Bayes (Bustami, 2014; Sipayung et al., 2016). Pengklasifikasian dengan metode Naive Bayes dilakukan dengan pendekatan peluang atau probabilitas (Sipayung et al., 2016). Metode ini dapat digunakan untuk menganalisis ulasan atau komentar berbentuk kalimat opini dengan menghasilkan akurasi tinggi untuk klasifikasinya (Gunawan et al., 2018; Gusriani et al., 2016; Sipayung et al., 2016).

Beberapa penelitian sebelumnya tentang analisis sentimen yang berkaitan dengan penelitian ini telah dilakukan dengan kasus yang berbeda-beda. Analisis sentimen terhadap kecenderungan pandangan publik terhadap toko online dengan menggunakan metode Naive Bayes (Gusriani et al., 2016). Sistem analisis sentimen terhadap komentar pelanggan untuk pemesanan kamar dengan tujuan agar dapat membantu pihak hotel dalam mengetahui komentar pelanggan berdasarkan kelompok sentimen dengan metode Naive Bayes (Sipayung et al., 2016). Analisis sentimen untuk menganalisis data review film dan ulasan hotel yang berasal dari website. Penelitian tersebut membandingkan dua algoritma yaitu Naive Bayes dan K-NN (Dey et al., 2016). Analisis sentimen terhadap kuesioner online yang diisi oleh mahasiswa dalam rangka penilaian terhadap dosen menggunakan metode Support Vector Machine (Santoso et al., 2017). Aplikasi penilaian kinerja dosen dengan melakukan analisis sentimen terhadap kuesioner dengan jenis angket tertutup yang diisi oleh mahasiswa (Harison & Faisal, 2017). Semakin banyak data latih yang digunakan maka vocabulary sistem akan menjadi banyak dan hasil klasifikasi akan menjadi lebih akurat (Gunawan et al., 2018). Berdasarkan hal tersebut, penelitian ini bertujuan untuk sistem analisis sentimen untuk evaluasi kinerja dosen dengan metode naive bayes. Adanya sistem ini diharapkan mampu membantu dalam mengolah data angket sehingga menjadi sebuah informasi yang dapat dimanfaatkan untuk mendukung hasil evaluasi kinerja dari dosen.

2. METODE

Penelitian ini merupakan penelitian pengembangan perangkat lunak yang menggunakan metode penelitian *System Development Life Cycle* (SDLC) dengan model *waterfall*. Tahap penelitian dimulai dari *communication* kemudian berkembang ke tahap *planning*, *modeling*, *construction* dan *deployment*. Subjek yang terlibat dalam penelitian ini terdiri dari admin, dosen dan mahasiswa. Admin sebagai subjek yang pengelola sistem, dosen sebagai subjek yang dievaluasi melalui penilaian mahasiswa, dan mahasiswa sebagai subjek yang memberikan penilaian kinerja dosen. Adapun tahapan dalam penelitian ini. Pertama

communication, penelitian diawali dengan melakukan pengumpulan data melalui observasi dan studi pustaka. Pengumpulan data dilakukan dengan melakukan observasi di Universitas Pendidikan Ganesha khususnya di UPT TIK. Dari hasil observasi, didapatkan data kritik dan saran mahasiswa Undiksha dari angket terbuka yang terdapat pada Sistem Informasi Akademik. Pada tahap studi pustaka, dilakukan analisis terhadap penelitian terkait yang sebelumnya sudah pernah dilakukan dengan menggunakan metode yang berkaitan dengan penelitian yang akan dikembangkan. Hasil analisis terhadap beberapa penelitian terkait, menggunakan metode Naïve Bayes sebagai metode klasifikasi. Kedua *planning*, tahap ini akan dianalisis kebutuhan sistem yang akan dikembangkan. Analisis yang dimaksudkan yaitu analisis kebutuhan sistem secara fungsional dan non fungsional. Pada sistem analisis sentimen untuk evaluasi kinerja dosen ini memiliki level pengguna serta fitur didalamnya. Penyusunan kebutuhan sistem dilakukan berdasarkan hasil studi pustaka dan observasi terhadap kebutuhan pengguna. Data yang sudah dikumpulkan pada tahap sebelumnya kemudian difilter dengan cara membuang kritik dan saran yang tidak memenuhi kriteria: (a) kritik dan saran dalam bahasa Inggris; (b) kritik dan saran yang hanya berisi tanda baca; (c) kritik dan saran yang tidak berisi komentar atau kosong; (d) kritik dan saran dengan komentar yang tidak mengandung makna evaluasi kinerja. Proses filtering tersebut data kritik dan saran yang tersisa kemudian dilabeli secara manual berdasarkan jenis sentimennya. Data *preprocessing* merupakan teknik dari data mining yang melibatkan perubahan data mentah menjadi data dengan format yang mudah dimengerti (Gunawan et al., 2018). Tahap *preprocessing* diperlukan untuk mempersiapkan data mentah sebelum dilakukan proses lain atau ke tahap selanjutnya. Tahap ini dilakukan dengan tujuan untuk membersihkan data dari noise, menyeragamkan bentuk kata dan mengurangi volume kata. Terdapat beberapa tahapan pada data *preprocessing*, yaitu *case folding*, *tokenizing*, *stopword removal* atau *filtering*, dan *stemming*. *Case folding* merupakan proses mengubah semua huruf besar pada suatu dokumen atau kalimat menjadi huruf kecil (Gunawan et al., 2018). Huruf yang terdapat pada kalimat harus dirubah karena tidak semua data konsisten menggunakan huruf besar kapital.

Tahap *tokenizing* adalah proses pemisahan suatu rangkaian kata dalam teks komentar berdasarkan karakter spasi, dan dilakukan juga proses menghilangkan tanda baca yang tidak diperlukan (Gunawan et al., 2018). Tahap *stopword removal* atau *filtering* ini merupakan tahap mengambil kata-kata penting dari hasil tokenizing menggunakan algoritma *stopword removal* (membuang kata yang kurang penting). *Stopword* adalah kumpulan dari kata-kata bersifat umum yang tidak memiliki arti dan kemungkinan besar tidak akan memberikan pengaruh prediksi, seperti imbuhan dan pronoun "itu" dan "mereka" (Rachmat & Lukito, 2016). Tahapan setelah *stopword removal* adalah tahapan *stemming*. *Stemming* adalah proses untuk mengubah kata yang terdapat pada kalimat menjadi kata dasar dengan cara memisahkan imbuhan baik awalan (prefiks) maupun akhiran (sufiks) (Wahyudi et al., 2017). Setelah seluruh data sudah dilabeli dan tahap *preprocessing* selesai, maka dilakukan pemisahan menjadi data training dan data testing. Pada data training dilakukan tahap *frequent itemset* yaitu tahap untuk menentukan atribut kata yang akan digunakan pada proses klasifikasi. Setelah data training dan data testing bersih, kemudian dilakukan implementasi algoritma naïve bayes pada data training untuk membangun model probabilitas dari data training.

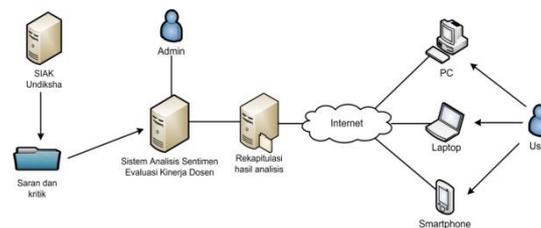
Ketiga perancangan (*modeling*), tahap selanjutnya yaitu tahap perancangan atau modeling. Tahapan ini dilakukan untuk menyederhanakan permasalahan dan ditampilkan melalui tools sehingga memudahkan untuk melihat gambaran sistem yang akan dikembangkan. Tahap desain atau perancangan dari pengembangan sistem meliputi arsitektur sistem, *Flowchart*, *Use Case Diagram*, *Activity Diagram*, ERD (*Entity Relation Diagram*), dan UI (*User Interface*). Dari proses perancangan ini akan menghasilkan rancangan produk atau sistem. Keempat *construction*, setelah rancangan dari sistem dibuat, maka dilanjutkan dengan tahap implementasi dari hasil perancangan menjadi sebuah sistem yang memiliki fungsionalitas sesuai dengan kebutuhan sistem menggunakan bahasa pemrograman. Sistem yang telah selesai dibuat kemudian dilakukan pengujian atau *testing*, terdapat enam jenis pengujian yang dilakukan yaitu *black box testing*, *white box testing*, pengujian kesesuaian, pengujian tingkat keakuratan sistem, pengujian dataset dengan parameter jumlah data training berbeda dan pengujian hasil prediksi sentimen. Pengujian *black box* atau sering disebut dengan pengujian behavioral (pengujian partisi) merupakan pengujian yang terfokus pada spesifikasi fungsional dari perangkat lunak (Novitasari, 2016). Pengujian fungsional ini bertujuan untuk menemukan kesalahan dalam beberapa kategori yaitu fungsi-fungsi yang tidak benar, kesalahan interface, kesalahan kinerja dan kesalahan dalam performa. Pengujian *white box* merupakan pengujian yang menggunakan struktur dan perancangan prosedural untuk memperoleh kasus uji (Mubarak, 2020). Tujuan dilakukan pengujian *white box* yaitu untuk mendeteksi unit dari kode suatu program terdapat kesalahan atau tidak. Pengujian kesesuaian ini dilakukan untuk mengetahui alur proses perhitungan manual dan sistem apakah sudah sama atau tidak. Pengujian pada dataset dengan parameter data training berbeda dilakukan untuk mengetahui apakah jumlah data training memiliki pengaruh terhadap hasil prediksi sistem. Pengujian dilakukan dengan menggunakan jumlah dataset yang tidak sama

pada setiap pengujian. Pengujian hasil prediksi sentimen dilakukan untuk mengetahui apakah susunan kalimat pada suatu komentar memiliki pengaruh terhadap hasil prediksi dari sistem. Parameter pengukuran atau pengujian akurasi digunakan untuk mengevaluasi performansi dalam melakukan klasifikasi dari model yang telah dibuat. Pengujian dilakukan dengan *k-fold cross validation* dan *confusion matrix*. *Confusion matrix* merupakan salah satu alat evaluasi berbentuk tabel yang digunakan untuk memprediksi jumlah ketepatan klasifikasi data uji yang diprediksi benar dan salah oleh model klasifikasi yang kemudian digunakan untuk menentukan kinerja dan akurasi dari model klasifikasi (Sunardi et al., 2018). Pengukuran performa klasifikasi dari sistem dilakukan dengan menghitung *precision*, *recall*, *accuracy* dan *f-measure*. *Precision* merupakan tingkat keberhasilan antara informasi yang diminta dengan jawaban yang diberikan oleh sistem, sehingga kesesuaian data antara hasil klasifikasi sistem dengan kategori sebenarnya dapat diketahui (Riany et al., 2016). *Recall* merupakan presentase tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi, sedangkan *accuracy* merupakan tingkat kedekatan antara kelas hasil prediksi dan kelas sebenarnya (Sunardi et al., 2018). *F-measure* merupakan tingkat keberhasilan dalam temu kembali informasi yang mengkombinasikan nilai *precision* dan nilai *recall* (Riany et al., 2016). Keempat *deployment*, setelah melakukan tahap *communication*, *planning*, *modeling*, dan *construction*, maka dilanjutkan pada tahap akhir yaitu *deployment*. Pada tahap ini menghasilkan sistem yang sudah siap digunakan oleh pengguna. Namun sebelum sistem dapat digunakan dan diimplementasikan secara langsung, diperlukan beberapa proses instalasi agar sistem dapat dijalankan. Maka diperlukan spesifikasi perangkat keras dan spesifikasi perangkat lunak.

3. HASIL DAN PEMBAHASAN

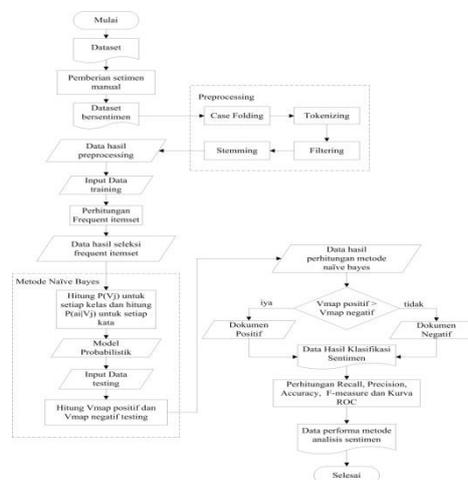
Hasil

Alur kerja sistem analisis sentimen untuk evaluasi kinerja dosen ini dimulai dari pengambilan data saran dan kritik dari SIAK Undiksha berupa file Microsoft Excel. File tersebut kemudian diinputkan oleh administrator pada sistem yang kemudian menghasilkan rekapan dari hasil analisis sentimen. Hasil rekapitulasi dapat diakses oleh user melalui komputer, laptop atau handphone yang memiliki browser. Hasil rancangannya digambarkan dalam bentuk arsitektur komputer yang ditunjukkan pada Gambar 1.



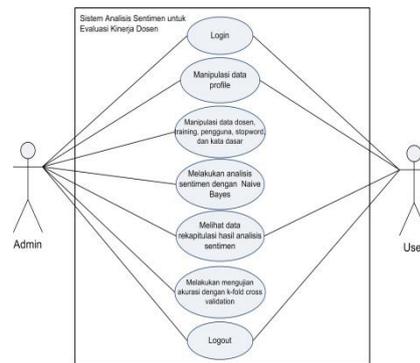
Gambar 1. Arsitektur komputer

Flowchart ini merupakan gambaran suatu alur perjalanan proses satu ke proses selanjutnya, dimana pada penelitian akan dijelaskan alur dari proses analisis sentimen secara manual. Mulai dari tahap persiapan data kemudian menentukan data training dan data testing serta perhitungan dengan naïve bayes. Flowchart metode dapat dilihat pada Gambar 2.



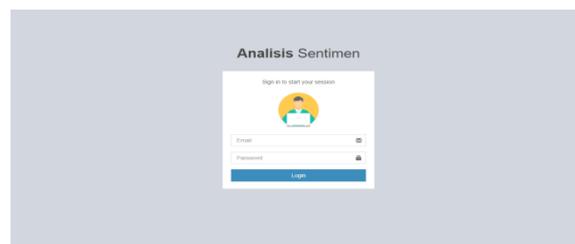
Gambar 2. Flowchart metode

Hasil analisis dan perancangan berupa syarat fungsional yang digambarkan dalam bentuk *use case diagram* seperti ditunjukkan pada **Gambar 3**.



Gambar 3. Use Case Diagram

Use case diagram yang ditunjukkan pada **Gambar 3** tersebut aktor yang berinteraksi dengan sistem dibagi menjadi dua yaitu admin dan user. Admin adalah administrator yang berasal dari UPT TIK Undiksha, sedangkan user adalah kaprodi dari masing-masing program studi. Pada use case diagram tersebut dilihat bahwa pengembangan sistem yang dilakukan menghasilkan beberapa fitur, yakni manipulasi data, menganalisis sentiment dengan Naive Bayes, melihat data rekapitulasi, dan pengujian akurasi dengan *f-fold cross validation*. Manipulasi data merupakan fitur yang digunakan untuk menambah, mengedit dan menghapus data berupa data profile, dosen, training, pengguna sistem, stopword, dan kata dasar. Menganalisis sentimen dengan Naive Bayes merupakan fitur untuk melakukan analisis sentimen terhadap kritik dan saran dari mahasiswa menggunakan metode Naive Bayes. Rekapitulasi merupakan fitur untuk melihat hasil rekapan dari data-data kritik dan saran mahasiswa yang diberikan untuk dosen. Pengujian akurasi merupakan fitur untuk menguji kinerja dari metode yang digunakan. Hasil implementasi tampilan dari sistem ini disesuaikan dengan fungsi yang dapat dilakukan oleh masing-masing pengguna. Adapun implementasi antarmuka perangkat lunak dari sistem analisis sentimen untuk evaluasi dosen. Pertama, halaman login display. Halaman ini akan muncul pertama kali ketika pengguna ingin mengakses sistem. Pengguna harus melakukan login dengan memasukkan email dan password terlebih dahulu. Tampilan sistem dapat dilihat pada **Gambar 4**.



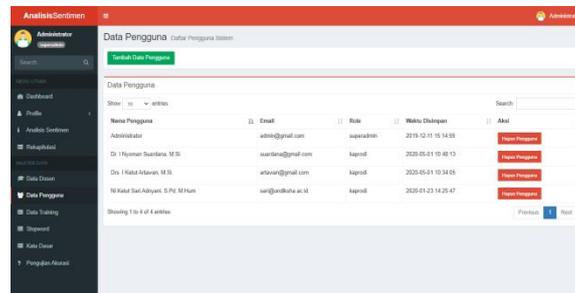
Gambar 4. Tampilan Login Display

Halaman *dashboard* merupakan tampilan yang akan muncul pertama kali ketika telah berhasil melakukan proses login. Pada halaman ini terdapat beberapa menu yang dapat diakses langsung oleh pengguna. Implementasi dari halaman dashboard ditunjukkan pada **Gambar 5**.



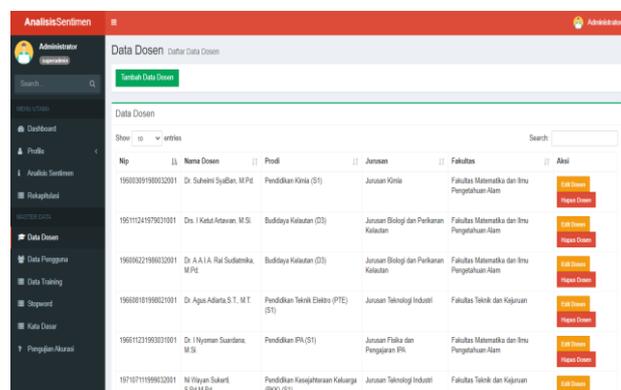
Gambar 5. Halaman Dashboard

Admin bertugas mengelola pengguna sistem yang terdiri dari kaprodi dan admin sendiri. Tampilan untuk mengelola pengguna sistem analisis sentimen untuk evaluasi kinerja dosen ditunjukkan pada Gambar 6.



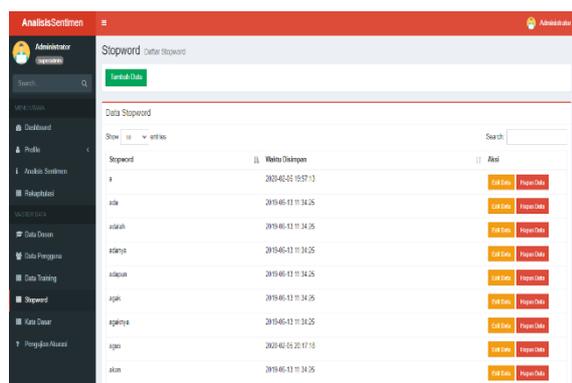
Gambar 6. Halaman Data Pengguna

Admin juga melakukan pengelolaan berupa manipulasi data dosen. Tampilan untuk manipulasi data dosen ditunjukkan pada Gambar 7.

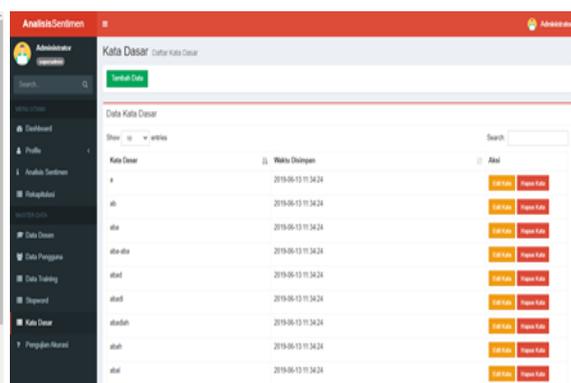


Gambar 7. Halaman Data Dosen

Pengelolaan berupa manipulasi data juga bisa dilakukan oleh administrator terhadap data-data stopword dan kata dasar. Tampilan untuk manipulasi data stopword dan kata dasar ditunjukkan pada Gambar 8 dan 9.

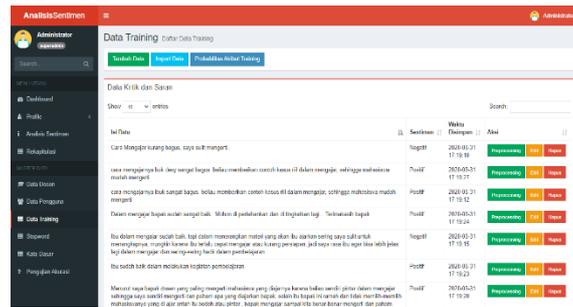


Gambar 8. Halaman Data Stopword



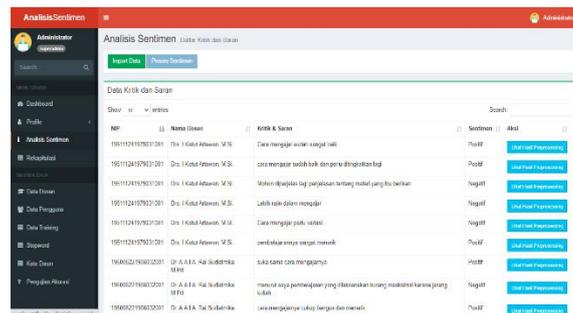
Gambar 9. Halaman Data Kata Dasar

Admin melakukan proses analisis sentimen dengan menambahkan terlebih dahulu data kritik dan saran mahasiswa sebagai data training. Tampilan untuk mengelola data training ditunjukkan pada Gambar 10.



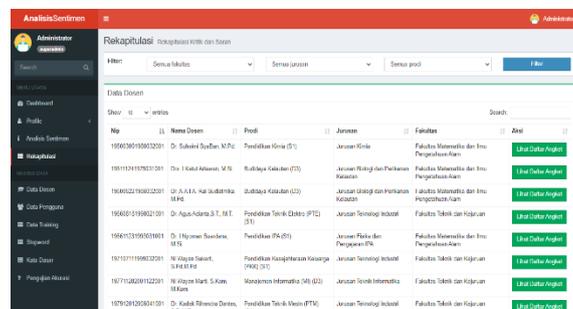
Gambar 10. Halaman Data Training

Admin dapat melakukan proses analisis sentimen dengan mengimport data berupa file Microsoft excel. Berikut tampilan halaman analisis sentimen yang ditunjukkan pada Gambar 11.



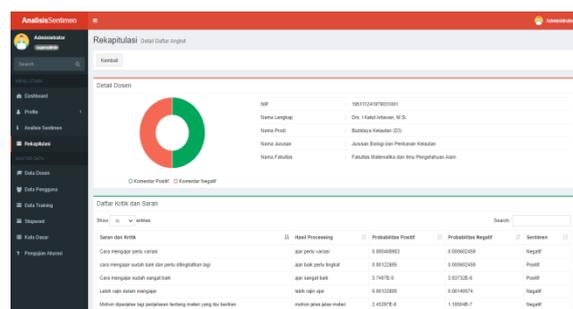
Gambar 11. Halaman Analisis Sentimen

Setelah dilakukan proses analisis sentimen oleh admin, maka sistem akan merekapitulasi hasil analisis untuk setiap dosen. rekapitulasiData rekapitulasi dapat dilihat pada halaman rekapitulasi yang ditunjukkan pada Gambar 12.



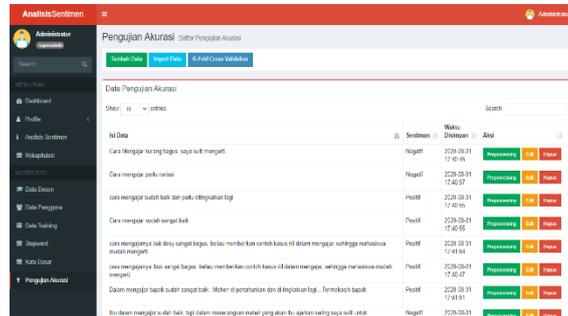
Gambar 12. Halaman Rekapitulasi

Tampilan rekapitulasi pada Gambar 12 tersebut terdapat aksi lihat daftar angket untuk melihat detail dari data kritik dan saran yang diterima oleh dosen. Berikut tampilan detail daftar angket ditunjukkan pada Gambar 13.



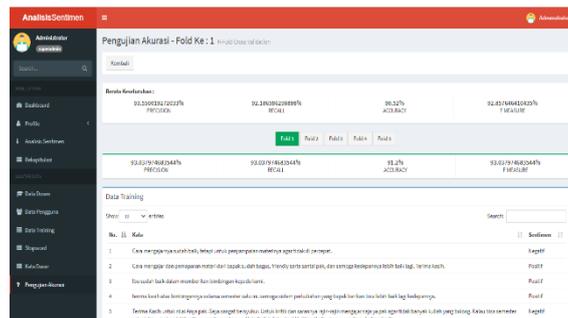
Gambar 13. Halaman Detail Daftar Angket

Tampilan kelola pengujian akurasi ditunjukkan pada Gambar 14. Admin dapat melakukan pengujian dengan *k-fold cross validation* yang bertujuan untuk mengetahui tingkat keakuratan kinerja metode Naïve Bayes.



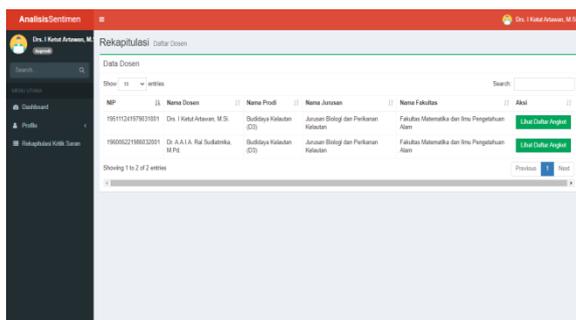
Gambar 14. Halaman Pengujian Akurasi

Pada Halaman Pengujian Akurasi admin juga dapat melihat hasil perhitungan *k-fold cross validation* yang terdapat pada pilihan aksi *k-fold cross validation*. Berikut tampilan halaman *k-fold cross validation* dapat dilihat pada Gambar 15.

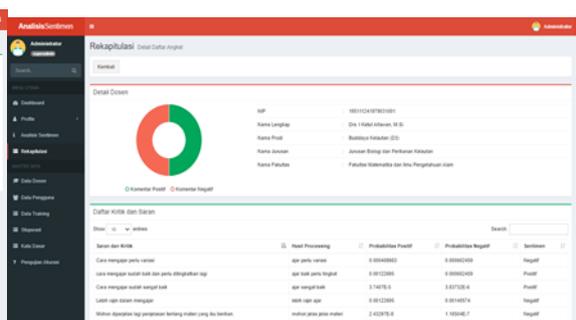


Gambar 15. Halaman k-Fold Cross Validation

User di masing-masing program studi bertugas melihat hasil rekapan kritik dan saran dari masing-masing dosen di program studi tersebut. Tampilan untuk rekapitulasi kritik dan saran dari mahasiswa serta detail rinciannya ditunjukkan pada Gambar 16 dan 17.



Gambar 16. Halaman Rekapitulasi (Kaprodi)



Gambar 17. Halaman Detail Daftar Angket (Kaprodi)

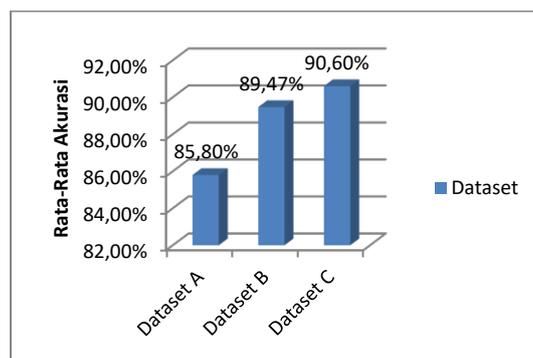
Pengujian perangkat lunak pada penelitian ini menggunakan empat jenis pengujian. Adapun rincian hasil pengujian perangkat lunak. Pertama, *white box testing*. Pengujian *whitebox* dilakukan dengan memperhatikan fungsi-fungsi struktur pada code-code pemrograman yang membangun sistem, dengan tujuan untuk mendeteksi unit dari kode suatu program terdapat kesalahan atau tidak. Pengujian yang telah dilakukan didapatkan hasil bahwa semua fungsi algoritma yang digunakan dalam pembuatan sistem analisis sentimen berjalan dengan baik dan benar sesuai dengan tujuan dari masing-masing algoritma yang diujikan tersebut. Kedua, *black box testing* pada penelitian ini menggunakan teknik *Boundary Value Analysis (BVA)*. *Boundary Value Analysis* merupakan salah satu teknik dalam metode pengujian *black box*

yang terfokus pada proses masukan dengan menguji nilai batas atas dan nilai batas bawah (Jaya, 2018). Hasil pengujian menunjukkan bahwa pengembangan sistem analisis sentimen untuk evaluasi kinerja dosen sudah berfungsi dengan baik dan semua fitur pada sistem ditampilkan sesuai dengan yang diharapkan. Ketiga, uji kesesuaian. Data yang digunakan terlebih dahulu dilakukan proses perhitungan secara manual kemudian dengan data yang sama digunakan pada proses perhitungan di sistem. Hasil pengujian kesesuaian yang telah dilaksanakan didapatkan hasil bahwa tidak ada perbedaan antara perhitungan manual dengan perhitungan pada sistem, hasil akhir yang didapat sama dan sudah sesuai. Pada pengujian akurasi ini digunakan sampel data kritik dan saran sebanyak 2.500 yang terdiri dari 1.690 data positif dan 810 data negatif. Data akan diuji menggunakan *5-fold cross validation*. Berikut hasil pengujian akurasi dari kinerja metode untuk keseluruhan fold dapat dilihat pada Gambar 18.

Tabel 1. Detail Hasil Pengujian Metode

Fold	Data Training	Data Testing	Benar	Salah	Precision (%)	Recall (%)	Accuracy (%)	F-Measure (%)
1	2.000	500	457 (+) 347 (-) 110	43 (+) 27 (-) 16	92.80	95.87	91.60	94.30
2	2.000	500	458 (+) 317 (-) 141	42 (+) 23 (-) 19	90.48	96.13	90.60	93.22
3	2.000	500	458 (+) 302 (-) 156	42 (+) 21 (-) 21	91.94	93.36	91.60	93.62
4	2.000	500	445 (+) 313 (-) 132	55 (+) 35 (-) 20	89.11	95.80	89.40	92.33
5	2.000	500	452 (+) 315 (-) 137	48 (+) 28 (-) 20	89.89	95.52	89.80	92.62

Pada pengujian dataset dengan parameter jumlah data training berbeda, digunakan sampel data sebanyak 2.500 kritik dan saran. Pembagian dataset terdiri dari Dataset A berjumlah 500 data, Dataset B berjumlah 1.500 data dan Dataset C berjumlah 2.500 data. Hasil pengujian rata-rata akurasi dari ketiga dataset yang dapat dilihat pada Gambar 18.



Gambar 18. Grafik Hasil Rata-Rata Akurasi Dataset

Berdasarkan hasil pengujian yang telah dilakukan, maka didapatkan hasil bahwa jumlah data training memiliki pengaruh terhadap hasil prediksi sistem. Hasil pengujian pada dataset A menghasilkan rata-rata akurasi sebesar 85.8%, pada pengujian dataset B menghasilkan rata-rata akurasi sebesar 89.47% dan pada pengujian dataset C menghasilkan rata-rata akurasi sebesar 90.6%. Berdasarkan pengujian yang telah dilakukan mendapatkan hasil bahwa susunan kalimat pada data kritik dan saran mahasiswa tidak mempengaruhi hasil prediksi hasil sentimen pada sistem.

Pembahasan

Hasil Penelitian setelah menerapkan metode Naïve Bayes sebagai metode klasifikasi, pengembangan sistem analisis sentimen untuk evaluasi kinerja dosen memperoleh hasil yang cukup tinggi yaitu dengan rata-rata *precision* sebesar 90.84%, *recall* sebesar 95.73%, *accuracy* sebesar 90.6% dan *f-measure* sebesar 93.22%. Hal ini menunjukkan bahwa pengembangan sistem analisis sentimen dapat berfungsi dengan baik dan telah mampu memprediksi kelas sentimen pada kritik dan saran mahasiswa dengan metode yang digunakan. Pada hasil pengujian menggunakan *5-fold cross validation*, terdapat kesalahan lebih banyak terjadi ketika mengklasifikasikan dokumen bersentimen negatif menjadi positif, sehingga presentase *precision* cenderung lebih kecil dibandingkan dengan *recall*. Hal tersebut terjadi dikarenakan pada data training, kritik dan saran yang bersifat positif lebih banyak dibandingkan dengan negatif sehingga kosakata atau model probabilitas antara positif dan negatif tidak seimbang. Hasil klasifikasi dapat dipengaruhi oleh komposisi jumlah data positif dan negatif yang digunakan, pada penelitian tersebut dikatakan ketika dataset yang memiliki komposisi yang seimbang maka akan memberikan hasil klasifikasi yang lebih baik dikarenakan tidak adanya bias atau pembelokan bobot kata dengan jumlah kemunculan yang tinggi (Novantirani et al., 2015; Suryanputra et al., 2017). Dengan tidak seimbangnyanya dokumen positif dan negatif pada pengujian akurasi menyebabkan terjadi beberapa kekeliruan dalam proses pengklasifikasian.

Pada hasil pengujian dataset dengan parameter jumlah data training yang berbeda, dapat diketahui bahwa jumlah data training dapat mempengaruhi hasil prediksi pada sistem. Dikarenakan semakin banyaknya data training maka sistem memiliki kosakata yang lebih banyak dan beragam, sehingga ketika ada data baru maka sistem bisa lebih tepat dalam memprediksi kelas sentimen. Selain itu, pada penelitian ini untuk proses analisis sentimen dilakukan berdasarkan pembobotan kata yang dihasilkan dari data training. Sehingga ketika terdapat kritik dan saran yang susunan kalimatnya berbeda tetapi memiliki isian kalimat yang sama, hal tersebut tidak mempengaruhi hasil prediksi dari sistem (Fitri et al., 2019; Gunawan et al., 2018). Dikarenakan pengklasifikasian dilakukan berdasarkan pembobotan kata yang dihasilkan dari data training, sehingga dapat menghasilkan presentase *precision*, *recall*, *accuracy* serta *f-measure* yang berbeda-beda disetiap pengujian. Analisis sentimen terhadap kecenderungan pandangan publik terhadap toko online dengan menggunakan metode Naïve Bayes (Gusriani et al., 2016). Sistem analisis sentimen terhadap komentar pelanggan untuk pemesanan kamar dengan tujuan agar dapat membantu pihak hotel dalam mengetahui komentar pelanggan berdasarkan kelompok sentimen dengan metode Naïve Bayes (Sipayung et al., 2016). Analisis sentimen terhadap kuesioner online yang diisi oleh mahasiswa dalam rangka penilaian terhadap dosen menggunakan metode Support Vector Machine (Santoso et al., 2017). Aplikasi penilaian kinerja dosen dengan melakukan analisis sentimen terhadap kuesioner dengan jenis angket tertutup yang diisi oleh mahasiswa (Harison & Faisal, 2017). Terdapat beberapa rekomendasi yang dapat dijadikan bahan pertimbangan untuk pengembangan sistem berikutnya yakni sistem analisis sentimen ini dapat dikombinasikan atau menggunakan metode lain untuk mendapatkan hasil prediksi dengan akurasi yang lebih tinggi. Proses analisis sentimen pada penelitian ini dilakukan hanya berdasarkan pembobotan yang diukur dari angka kemunculan dan mengabaikan makna kata. Pada penelitian selanjutnya diharapkan agar menerapkan juga kedekatan suatu kata dengan kata lainnya (semantic), sehingga dapat membantu sistem dalam menilai kelas sentimen dengan lebih maksimal. Diharapkan pada penelitian selanjutnya untuk algoritma stemming bisa dikembangkan agar kritik dan saran dalam bahasa Inggris dapat diprediksi kelas sentimennya.

4. SIMPULAN

Berdasarkan hasil pengembangan sistem analisis sentimen untuk evaluasi kinerja dosen dengan menggunakan metode Naïve Bayes dapat disimpulkan bahwa, sistem telah mampu memprediksi dan memberikan gambaran hasil sentimen yang terdiri dari 2 (dua) kelas yaitu positif atau negatif pada kritik dan saran mahasiswa. Proses analisis sentimen dilakukan berdasarkan pembobotan kata yang dihasilkan dari data training, sehingga susunan kalimat tidak mempengaruhi hasil prediksi pada sistem. Dalam memprediksi kelas sentimen, jumlah data training memiliki pengaruh terhadap hasil prediksi yang dihasilkan oleh sistem, dikarenakan semakin banyak dan beragamnya kosakata maka sistem akan bisa lebih tepat dalam memprediksi kelas sentimen. Dengan dilakukannya penelitian ini dapat dijadikan solusi dalam pengolahan serta proses analisis terhadap isian angket terbuka secara otomatis sehingga proses analisis dapat dilakukan dengan lebih efektif dan efisien.

5. DAFTAR PUSTAKA

- Arsini, A. (2016). Peningkatan Kompetensi Profesional Dan Inovasi Guru Dalam Mengembangkan Video Pembelajaran Online Melalui Pembuatan Portal "Channel Pembelajaran Sains Berbasis Unity Of Science. *Dimas: Jurnal Pemikiran Agama Untuk Pemberdayaan*, 15(2), 115. <https://doi.org/10.21580/dms.2015.152.749>.
- Atmaja, K. J., & Wijaya, I. N. S. W. (2019). Pengembangan sistem evaluasi kinerja dosen (e-kuesioner) STMIK STIKOM Indonesia. *JST (Jurnal Sains Dan Teknologi)*, 8(1), 55-64. <https://doi.org/10.23887/jstundiksha.v8i1.17290>.
- Azhar, Y. (2017). Metode Lexicon-Learning Based Untuk Identifikasi Tweet Opini Berbahasa Indonesia. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 6(3), 237-243. <https://doi.org/10.23887/janapati.v6i3.11739>.
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134. <https://doi.org/10.1016/j.knosys.2021.107134>.
- Bustami. (2014). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi. *Jurnal Informatika*, 8(1), 884-898.
- Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, 8(4), 54-62. <https://doi.org/10.5815/ijieeb.2016.04.07>.
- Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm. *Procedia Computer Science*, 161, 765-772. <https://doi.org/10.1016/j.procs.2019.11.181>.
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(2), 113-118. <https://doi.org/10.26418/jp.v4i2.27526>.
- Gusriani, S., Wardhani, K. D. K., & Zul, M. I. (2016). Analisis Sentimen Terhadap Toko Online di Sosial Media Menggunakan Metode Klasifikasi Naïve Bayes (Studi Kasus: Facebook Page BerryBenka). *Jurnal Aksara Komputer Terapan*, 5(2).
- Harison, & Faisal, R. (2017). Aplikasi Penilaian Kinerja Dosen pada Proses Belajar Mengajar Berbasis Web: Studi Kasus di Badan Penjamin Mutu Internal Institut Teknologi Padang. *Jurnal Teknologi Dan Sistem Komputer*, 5(2), 89-93. <https://doi.org/10.14710/jtsiskom.5.2.2017.90-94>.
- Jaya, T. S. (2018). Pengujian Aplikasi dengan Metode Blackbox Testing Boundary Value Analysis (Studi Kasus: Kantor Digital Politeknik Negeri Lampung). *Jurnal Informatika Pengembangan IT (JPIT)*, 3(2), 45-48. <https://doi.org/10.30591/jpit.v3i1.647>.
- Ling, J., Kencana, I. P. E. N., & Oka, T. B. (2014). Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square. *E-Jurnal Matematika*, 3(3), 92-99. <https://doi.org/10.24843/MTK.2014.v03.i03.p070>.
- Mubarak, R. (2020). Implementasi Metode White Box Testing Pada Proses Quality Assurance Perangkat Lunak Berbasis Web Dan Mobile Collection System. *Jurnal Teknologi Informasi ESIT*, XV(10), 57-63.
- Novantirani, A., Sabariah, M. K., & Effendy, V. (2015). Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine. *E-Proceeding of Engineering*, 2(1), 1177-1183.
- Novitasari, D. (2016). Perbandingan Algoritma Stemming Porter Dengan Arifin Setiono Untuk Menentukan Tingkat Ketepatan Kata Dasar. *Jurnal String*, 1(2), 120-129. <https://doi.org/10.30998/string.v1i2.1031>.
- Rachmat C, A., & Lukito, Y. (2016). Klasifikasi Sentimen Komentar Politik dari Facebook Page Menggunakan Naive Bayes. *Jurnal Informatika Dan Sistem Informasi*, 02(02).
- Rahadi, D. R. (2010). *Manajemen Kinerja Sumber Daya Manusia*. Tunggal Mandiri Publishing.
- Rahmawati, D. E., & Trimulyono, G. (2021). Validitas Instrumen Penilaian Higher Order Thinking Skills (Hots) pada Materi Keanekaragaman Hayati. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, 11(1), 141. <https://doi.org/10.26740/bioedu.v11n1.p138-147>.
- Riany, J., Fajar, M., & Lukman, M. P. (2016). Penerapan Deep Sentiment Analysis pada Angket Penilaian Terbuka Menggunakan K-Nearest Neighbor. *Jurnal Sisfo*, 06(01), 147-156. <https://doi.org/10.24089/j.sisfo.2016.09.011>.
- Risdianto, E. (2008). Pengembangan Multimedia Interaktif (MPI) pada Praktikum Fisika Dasar I. *Jurnal Exacta*, 6(2), 9-16.
- Santoso, V. I., Virginia, G., & Lukito, Y. (2017). Penerapan Sentiment Analysis pada Hasil Evaluasi Dosen

- dengan Metode Support Vector Machine. *Jurnal Transformatika*, 14(2), 72. <https://doi.org/10.26623/transformatika.v14i2.439>.
- Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi (JSI)*, 8(1), 958–965.
- Sunardi, Fadlil, A., & Suprianto. (2018). Analisis Sentimen Menggunakan Metode Naive Bayes Classifier Pada Angket Mahasiswa. *SAINTEKBU: Jurnal Sains Dan Teknologi*, 10(2), 1–9. <https://doi.org/10.32764/saintekbu.v10i2.190>.
- Suryani, Y. E. (2017). Pemetaan Kualitas Empirik Soal Ujian Akhir Semester Pada Mata Pelajaran Bahasa Indonesia SMA di Kabupaten Klaten. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 21(2), 142–152. <https://doi.org/10.21831/pep.v21i2.10725>.
- Suryanputra, F. O., Yohanes, B. W., & Nugroho, S. (2017). Analisis Sentimen untuk Komentar pada Sistem Pencarian Kost Menggunakan Metode Support Vector Machine (SVM). *TECHNE*, 16(1). <https://doi.org/10.31358/techne.v16i01.157>.
- Syahrir, S., Supriyati, Y., & Fauzi, A. (2021). Evaluasi Dampak Program Pendidikan Jarak Jauh (PJJ) melalui model CIPP pada Kinerja Dosen aspek Pembelajaran pada Masa Pandemi Covid 19. *Jurnal Ilmiah Mandala Education*, 7(1). <https://doi.org/10.36312/jime.v7i1.1716>.
- Wahyudi, D., Susyanto, T., & Nugroho, D. (2017). Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia. *Jurnal Ilmiah SINUS*, 15(2), 49–56. <https://doi.org/10.30646/sinus.v15i2.305>.