
Analisis Sentimen Pengguna Aplikasi SIMANTAP Pondok Pesantren Salafiyah Syafi'iyah Situbondo

Ach. Zubairi¹, Luckman Ashary²

¹Program Studi Manajemen Bisnis Syariah, Universitas Ibrahimy, Jawa Timur 68374, Indonesia

²Program Studi Akuntansi, Universitas Ibrahimy, Jawa Timur 68374, Indonesia

Abstract

Pondok Pesantren Salafiyah Syafi'iyah Sukorejo Situbondo provides a payment application called SIMANTAP. This application aims to make it easier for guardians of students to make payments for bills given by Islamic boarding schools. Along the way, some obstacles arise; users tend to give their views in a comment. These comments will later provide an overview of the user's response to the SIMANTAP application. This study will analyze comment data using the Naïve Bayes Classifier. The data was taken from September 2021 to October 2021. As a result, users tend to give a positive response of 89%. Based on the test data results using the Naïve Bayes classifier algorithm, the classification can predict quite well; this is shown by the precision results of 82%, 97% recall, and 89% fi-score.

Keywords

Sentiment Analysis; Naïve Bayes Classifier; Islamic boarding school; SIMANTAP

Corresponding Author

Ach. Zubairi

Manajemen Bisnis Syariah, Universitas Ibrahimy, Jawa Timur 68374, Indonesia; dsn.fsei01@gmail.com

PENDAHULUAN

Pondok pesantren adalah sebuah bentuk lembaga pendidikan yang eksistensinya cukup lama di Negara Indonesia dan terbukti memiliki kontribusi besar dalam berbagai aspek kehidupan bangsa (Junaidi, 2017; Mulyo, 2022; Nizar, 2020). Sebagai salah satu bentuk Pendidikan tertua di Indonesia, pengaruh perkembangan teknologi telah memberikan dampak negatif maupun positif bagi kehidupan di pesantren. Salah satu dampak positifnya yaitu teknologi telah membantu pondok pesantren untuk mempermudah wali santri dalam melakukan pembayaran uang tahunan santri Pondok Pesantren Salafiyah Syafi'iyah Sukorejo Situbondo. Yaitu dengan cara menyediakan layanan pembayaran digital menggunakan aplikasi SIMANTAP.

Pembayaran digital merupakan proses penggunaan teknologi untuk melakukan transaksi pembayaran secara elektronik tanpa menggunakan uang tunai (Aulia, 2020; Tarantang et al., 2019). Sehingga meminimalisir adanya kehilangan uang tunai. Mayoritas wali santri menerima dengan baik adanya sistem ini, karena tidak hanya pembayaran uang tahunan santri saja, tetapi pembayaran yang lain juga disediakan, seperti SPP, dana sehat dan DPP. Hal ini tentu sangat mempermudah walisntri yang ingin melakukan pembayaran terutama yang rumahnya jauh dari kota. Cara kerja pembayaran di aplikasi SIMANTAP yaitu dengan membuka aplikasi yang

sudah terpasang di smartphone, lalu memilih layanan pembayaran yang ingin dilakukan lalu mengkonfirmasi layanan pembayaran. Setelah layanan pembayaran dikonfirmasi wali santri akan mendapat nomor rekening dari pengurus pesantren kemudian mentransfer jumlah pembayaran yang telah tertera pada aplikasi.

Seiring perkembangan jasa pembayaran daring pada aplikasi SIMANTAP, masyarakat atau wali santri memiliki opini dan pendapatnya masing-masing. Masyarakat biasanya memberikan opini atau pendapatnya melalui kolom komentar yang telah disediakan pada aplikasi SIMANTAP. Pendapat yang diberikan masyarakat terhadap sistem pembayaran digital SIMANTAP beragam, ada yang merespon baik karena merasa di mudahkan, ada juga yang merespon negatif karena merasa pelayanan yang kurang memuaskan. Apabila diteliti dan dipahami lebih lanjut terhadap pendapat dan opini masyarakat tersebut maka akan didapatkan sebuah sentimen yang apabila dikumpulkan akan mendapatkan kesimpulan apakah sistem pembayaran ini relevan untuk kedepannya atau tidak. Analisis sentimen adalah studi komputasi dari opini-opini, sentimen, serta emosi yang diekspresikan dalam teks. Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau pendapat (Amrullah et al., 2020; Nurhuda et al., 2016). Polaritas mempunyai arti apakah teks yang ada dalam dokumen, kalimat, atau pendapat memiliki aspek positif atau negatif (Amrullah et al., 2020; Nugroho et al., 2016; Wahyudi & Kusumawardhana, 2021).

Teknik pembelajaran mesin atau *machine learning* yang digunakan dalam penerapan metode ini adalah *Naïve Bayes Classifier* (NBC). NBC merupakan salah satu metode *machine learning* yang memanfaatkan perhitungan probabilitas dan statistic (Darwis et al., 2021; Tuhuteru & Iriani, 2018). Analisis sentimen yang dibangun menggunakan NBC memperoleh akurasi sebesar 83% dan rata-rata harmonik sebesar 90,713% (Nugroho et al., 2016). Perbandingan beberapa metode seperti metode *Naïve Bayes*, *K-nearest Neighbor*, dan gabungan *K-means* dan LVQ dalam mengklasifikasikan kategori buku berbahasa Indonesia dengan data yang digunakan berjumlah 200 buku, 150 buku digunakan sebagai data latih, sedangkan 50 buku digunakan sebagai data uji. Dari hasil penelitian yang dilakukan, metode KNN memperoleh akurasi sebesar 96%, kemudian *Naïve Bayes* sebesar 98%, lalu kombinasi *K-Means* dan LVQ menghasilkan akurasi sebesar 92,2%. Metode *Naïve Bayes* mendapatkan hasil akurasi yang tertinggi (Darujati, 2010; Hairani et al., 2018; Wahyuningsih & Utari, 2018). Tingkat keakurasian pengklasifikasian sangat dipengaruhi oleh proses training kompleksitas data juga dapat mempengaruhi hasil akurasi (Chairunisa & Astuti, 2020; Munazilin & Nasta'in, 2023; Nugroho et al., 2016).

METODE PENELITIAN

Pengumpulan Data

Data yang akan digunakan pada penelitian ini bersumber dari ulasan aplikasi SIMANTAP yang kami *scraping* menggunakan *google colab*. Dari data tersebut nantinya akan digunakan untuk data latih dan Sebagian menjadi data uji. Kedua data ini nantinya akan dilakukan *pre-processing* terlebih dahulu untuk diolah menjadi teks yang bersih.

Tahap *Pre-processing*

Pada tahapan *preprocessing* ini, terdiri dari *case folding*, *tokenizing*, *stemming*, dan *filtering*. Berikut ini penjelasan dari tahapan *pre-processing* (Munazilin & Nasta'in, 2023):

1. Pada tahapan *case folding*, teks dilakukan proses perubahan dari huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca pada kalimat.
2. ada tahapan *tokenizing*, setiap kata akan dipisahkan berdasarkan spasi yang ditemukan
3. Pada tahapan *stemming*, yaitu pengubahan kata berimbuhan menjadi kata dasar.
4. Pada tahapan *filtering*, yaitu pembuangan kata-kata tidak penting dari hasil token.

Tahapan Klasifikasi *Naïve Bayes Classifier*

Pada tahap ini dilakukan pengklasifikasian teks bersih yang sudah melalui tahap *pre-processing*. untuk mengetahui kategori dari teks tersebut berdasarkan data latih. Metode *Naïve Bayes* merupakan proses pengklasifikasian atau pengujian suatu teks baru yang belum pernah diketahui kategorinya berdasarkan basis pengetahuan yang dimiliki (Priyono et al., 2016). Pada tahapan ini terdapat empat tahapan.

1. Pembentukan fitur data latih

Bagian ini merupakan tahapan awal melakukan pelatihan data latih. Dalam penelitian ini yang dimaksud dengan fitur adalah kata kunci yang akan menjadi parameter satuan data latih yaitu dokumen(ulasan). Dalam kata lain, fitur adalah kata yang memiliki nilai sentiment. Probabilitas dihitung dengan persamaan sebagai berikut (Setiawan & Nursantika, 2017).

$$p(w_k \setminus v_j) = \frac{nk + 1}{\text{jumlah frekuensi} + \text{jumlah kata}}$$

2. Perhitungan nilai probabilitas data latih

Setelah dibentuk fitur dengan kemunculannya dari data latih selanjutnya menghitung probabilitas dari setiap kelas.

3. Menentukan probabilitas kelas

Setelah dapat probabilitas dari setiap kelas, selanjutnya menghitung probabilitas setiap fitur pada setiap kelas menggunakan persamaan berikut (Priyono et al., 2016).

$$p\left(\frac{wk}{vj}\right) = \frac{f\left(\frac{wk}{c_i}\right)+1}{f(c_i) + |W|}$$

4. Perhitungan Vmap

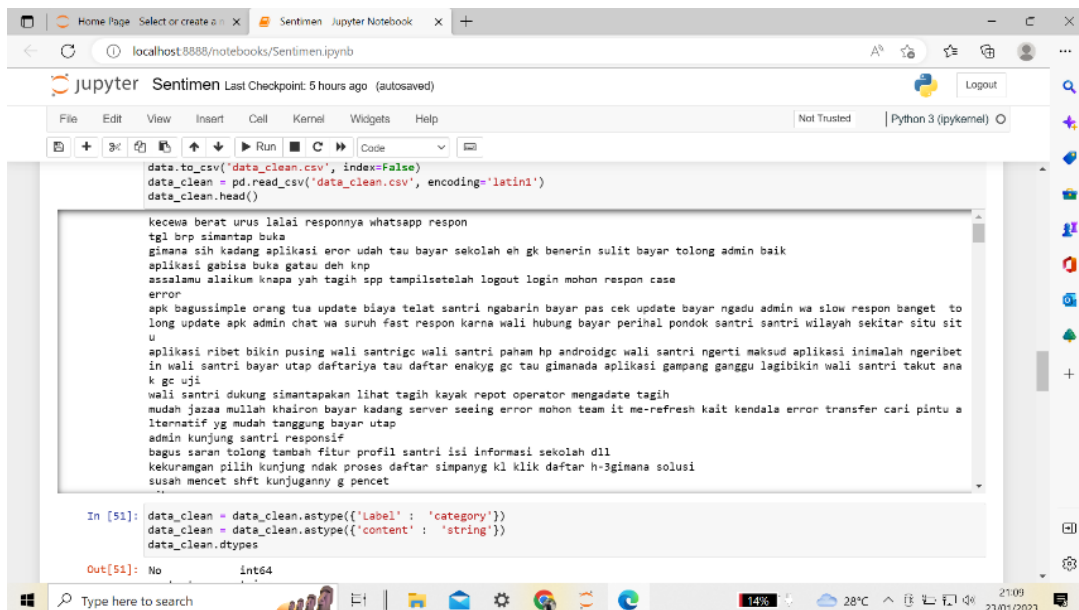
Ditahap ini dilakukan proses perhitungan nilai probabilitas teks berdasarkan data latih. Teks yang akan diuji dilakukan tahap *pre-processing* dan penghitungan Vmap. Vmap adalah perhitungan yang digunakan NBC untuk menentukan probabilitas tertinggi data uji dari masing-masing kelas berdasarkan dari proses learning. Persamaan yang digunakan adalah sebagai berikut (Sentiaji & Bachtiar, 2014).

$$Vmap = argmax\{positif, negatif\}p(wk/c)xp(c)$$

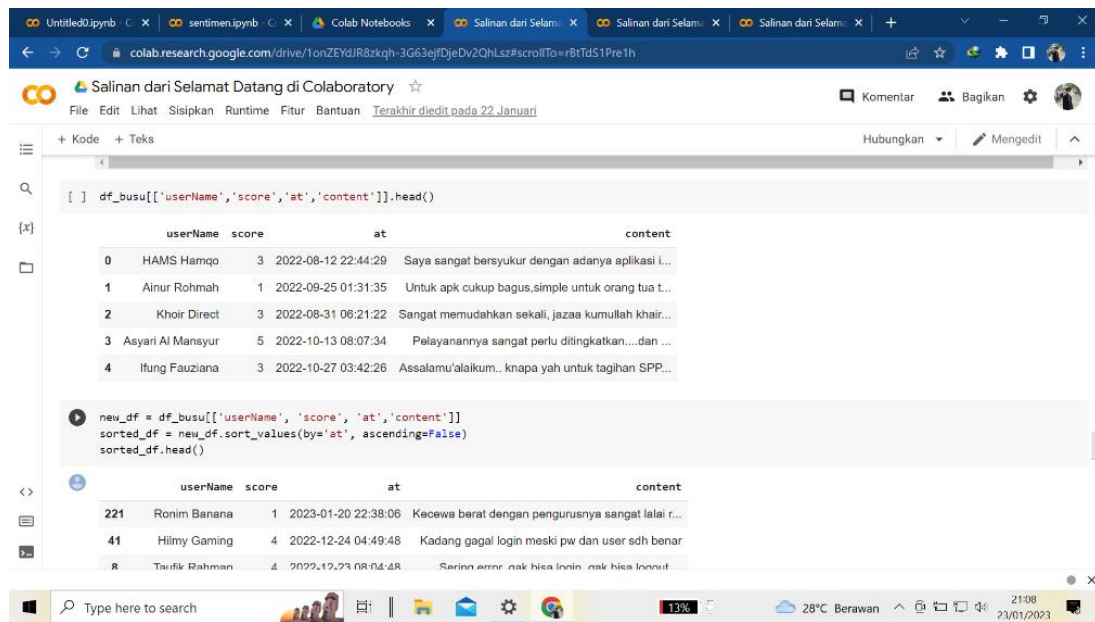
HASIL DAN PEMBAHASAN

Tahapan *Pre-processing*

Sebelum data ulasan digunakan, tahap *pre-processing* dilakukan untuk mendapatkan data bersih. Tahapan yang dilakukan antara lain *case folding*, *tokenizing*, *stemming*, dan *filtering*. Berikut ini adalah gambar data latih 1 sebelum dan setelah dilakukan *pre-processing*.



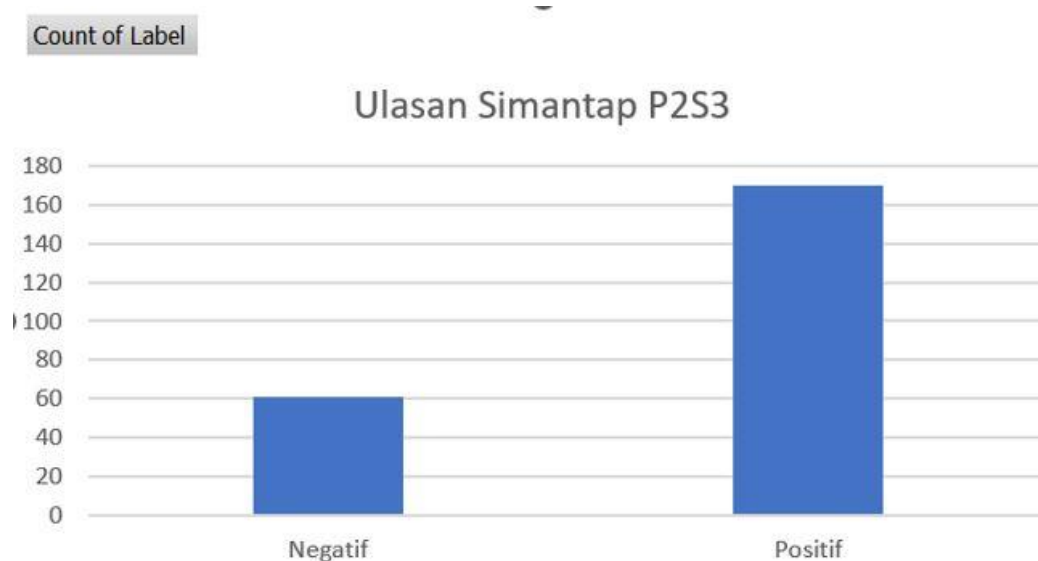
Gambar 1. Data latih sebelum *pre-processing*



Gambar 2. Data latih sesudah *pre-processing*

Penentuan Sentimen

Perbandingan jumlah data ulasan positif dan negatif yang didapatkan setelah melakukan pelabelan secara manual dan penghapusan ulasan yang tidak relevan menunjukkan bahwa dalam rentang waktu September 2021 sampai Oktober 2021, aplikasi SIMANTAP cenderung mendapatkan respon positif dari pengguna. Hal ini dibuktikan dengan sedikitnya data ulasan negatif pada aplikasi SIMANTAP dibandingkan dengan data ulasan positifnya. Hasil tersebut dapat dilihat pada diagram batang di Gambar 3.



Gambar 3. Ulasan perbandingan positif dan negatif aplikasi SIMANTAP

Naïve Bayes Classifier

Klasifikasi menggunakan algoritma *Naïve Bayes* ini dapat menghasilkan sebuah model yang digunakan untuk data uji. Klasifikasi ini menggunakan library dari python yaitu scikit learn. Hasil dari permodelan dalam proses klasifikasi data yang dilatih, nantinya akan digunakan pada data uji coba untuk dijadikan acuan evaluasi terhadap klasifikasi menggunakan algoritma *Naïve Bayes* dalam proses analisis sentiment terhadap aplikasi SIMANTAP. hasil dari data uji terdapat pada gambar 4 berikut.

```

NB Accuracy: 0.8085106382978723
NB Precision: 0.8222222222222222
NB recall Score: 0.9736842105263158
NB f1 Score: 0.891566265060241
confusion matrix:
[[ 1  8]
 [ 1 37]]
=====

```

	precision	recall	f1-score	support
Negatif	0.50	0.11	0.18	9
Positif	0.82	0.97	0.89	38
accuracy			0.81	47
macro avg	0.66	0.54	0.54	47
weighted avg	0.76	0.81	0.76	47

Gambar 4. Hasil dari data uji aplikasi SIMANTAP

Berdasarkan hasil akurasi yang mendekati 100%, yaitu 89% membuktikan cukup tingginya tingkat akurasi analisis sentiment menggunakan klasifikasi *Naïve Bayes*. Hasil precision sebagai rasio data ulasan positif yang diprediksi dengan tepat dari keseluruhan data yang diprediksi juga menunjukkan hasil yang cukup tinggi, yaitu sebesar 89%. Selain itu, hasil recall sebesar 97% yang menunjukkan sensitivitas prediksi data ulasan positif berdasarkan data yang sebenarnya juga cukup tinggi.

KESIMPULAN DAN SARAN

Penelitian analisis sentiment terhadap aplikasi SIMANTAP menggunakan algoritma *Naïve Bayes classifier* menghasilkan kesimpulan bahwa pengguna aplikasi SIMANTAP cenderung

memberikan respon positif terhadap adanya aplikasi karena terdapat 89 % data ulasan positif. Di mana data tersebut lebih banyak dibandingkan data ulasan negatif yang berjumlah 18% data. Dengan demikian hasil ini dapat memperkuat penelitian sebelumnya yang mengatakan bahwa aplikasi sentuh tanahku memberikan kemudahan bagi pengguna. Selain itu, berdasarkan hasil dari data uji menggunakan algoritma *Naïve Bayes classifier*, klasifikasi mampu memprediksi dengan cukup baik, hal ini ditunjukkan dari hasil *precision* sebesar 82 %, *recall* 97%, dan *f1-score* 89 %.

Saran untuk penelitian selanjutnya agar lebih baik adalah menggunakan data ulasan yang lebih banyak dan melakukan penanganan negasi untuk meningkatkan akurasi dari metode yang digunakan.

DAFTAR PUSTAKA

- Amrullah, A. Z., Anas, A. S., & Hidayat, M. A. J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal Bumigora Information Technology (BITE)*, 2(1), 40–44.
- Aulia, S. (2020). Pola perilaku konsumen digital dalam memanfaatkan aplikasi dompet digital. *Jurnal Komunikasi*, 12(2), 311–324.
- Chairunisa, R., & Astuti, W. (2020). Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(5), 805–812.
- Darujati, C. (2010). Perbandingan Klasifikasi Dokumen Teks Menggunakan Metode Naïve Bayes Dengan K-Nearest Neighbor Abstrak. *Univ. Stuttgart*, 13(1), 1–9.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Hairani, H., Nugraha, G. S., Abdillah, M. N., & Innuddin, M. (2018). Komparasi akurasi metode correlated naive Bayes classifier dan naive Bayes classifier untuk diagnosis penyakit diabetes. *InfoTekJar: Jurnal Nasional Informatika Dan Teknologi Jaringan*, 3(1), 6–11.
- Junaidi, K. (2017). Sistem pendidikan pondok pesantren di Indonesia (suatu kajian sistem kurikulum di Pondok Pesantren Lirboyo). *Istawa: Jurnal Pendidikan Islam*, 2(1), 95–100.
- Mulyo, R. P. (2022). Peran Serta Kontribusi Pondok Pesantren Dalam Catatan Sejarah Bangsa Indonesia. *Jurnal Penelitian Agama*, 23(1), 159–174.
- Munazilin, A., & Nasta'in, M. (2023). Analisis Sentimen Pengguna Aplikasi Sistem Pembayaran UTAP Pondok Pesantren Salafiyah Syafi'iyah Situbondo. *Elektriase: Jurnal Sains Dan Teknologi Elektro*, 13(01), 50–55.
- Nizar, M. (2020). Madrasah Diniyah dan Pesantren sebagai Penyeimbang Modernitas. *Sukma: Jurnal Pendidikan*, 4(1), 25–37.
- Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2016). Analisis Sentimen Pada Jasa Ojek

- Online Menggunakan Metode Naive Bayes. *Prosiding Seminar Sains Nasional Dan Teknologi*, 1(1).
- Nurhuda, F., Sihwi, S. W., & Doewes, A. (2016). Analisis sentimen masyarakat terhadap calon Presiden Indonesia 2014 berdasarkan opini dari Twitter menggunakan metode Naive Bayes Classifier. *ITSmart: Jurnal Teknologi Dan Informasi*, 2(2), 35–42.
- Priyono, F., Kanti, S., Dzulfiqar, I., Amirulloh, I., Alvi, A., & Rosiyadi, D. (2016). Analisis Sentimen Media Sosial Opini Ujian Nasional Berbasis Komputer menggunakan Metoda Naive Bayes. *Journal of Electrical And Electronics Engineering*, 1(2), 38–45.
- Sentiaji, A. R., & Bachtiar, A. M. (2014). Analisis sentimen terhadap acara televisi berdasarkan opini publik. *Jurnal Ilmiah Komputer Dan Informatika (KOMPUTA)*, 2(1), 55–60.
- Setiawan, I. A., & Nursantika, D. (2017). Klasifikasi Artikel Berita Menggunakan Metode Text Mining Dan Naive Bayes Classifier. *Prosiding SENIATI*, 3(1), 1–6.
- Tarantang, J., Awwaliyah, A., Astuti, M., & Munawaroh, M. (2019). Perkembangan sistem pembayaran digital pada era revolusi industri 4.0 di indonesia. *Jurnal Al-Qardh*, 4(1), 60–75.
- Tuhuteru, H., & Iriani, A. (2018). Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier. *Jurnal Informatika*, 3(3), 394–401.
- Wahyudi, R., & Kusumawardhana, G. (2021). Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *J. Inform*, 8(2), 8.
- Wahyuningsih, S., & Utari, D. R. (2018). Perbandingan Metode K-Nearest Neighbor, Naive Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit. *Konferensi Nasional Sistem Informasi (KNSI) 2018*, 1(1), 619–623.