

Analisis Penggunaan *Fingerprint* Dengan Metode *K-Nearest Neighbors* Bagi Pegawai dan Dosen (Studi Kasus: Universitas Stella Maris Sumba)

¹Fransiska Asi, ²Gergorius Kopong Pati, ³Alexander Adis

^{1,2,3}Universitas Stella Maris Sumba

Alamat Surat

Email: [1Fransiska Asi @gmail.com](mailto:1Fransiska.Asi@gmail.com)*, 2gregkopong80@gmail.com,

Article History:

Diajukan: 23 Januari 2025; Direvisi: 18 Februari 2025; Accepted: 8 April 2025

ABSTRAK

Penelitian ini mengukur keakuratan dan efektivitas *K-Nearest Neighbors* dalam menganalisis pemakaian *fingerprint*. Dalam penelitian ini, penulis mengumpulkan data terbaru dengan teknik web scraping. Data tersebut kemudian dibersihkan dan diberi label, menghasilkan ulasan positif dan ulasan negatif. Proses preprocessing dilakukan, termasuk case folding, tokenisasi, filtering, dan stemming. Setelah tahap pengolahan data, penulis menerapkan algoritma K-Nearest Neighbor (KNN). Data tersebut diuji untuk menghasilkan ulasan negatif dan ulasan positif. Hasil penelitian menunjukkan tingkat akurasi KNN. Penelitian ini diharapkan memberikan pemahaman tentang penggunaan KNN dalam menganalisis sentimen pengguna *fingerprint*. Dataset melibatkan parameter berat badan dan faktor-faktor lain yang berpengaruh. Model dikembangkan dan diuji menggunakan teknik cross-validation untuk memastikan konsistensi kinerja. Temuan penelitian menunjukkan bahwa penerapan Algoritma K-NN dengan Wrapper preprocessing dapat meningkatkan akurasi penentuan *Fingerprint* Dengan Metode *K-Nearest Neighbors* Bagi Pegawai Dan Dosen. Penerapan metode K-Nearest Neighbor dan K-Nearest Neighbor dengan Wrapper sebagai tahap preprocessing dalam menentukan keterangan berat manusia mendapatkan hasil nilai akurasi yang sama yaitu sebesar 90,48%.

Kata kunci : Analysis. *Fingerprint*, *K-Nearest Neighbors*

ABSTRACT

This research measures the accuracy and effectiveness of K-Nearest Neighbors in analyzing fingerprint usage. In this research, the author collected the latest data using web scraping techniques. The data is then cleaned and labeled, resulting in positive reviews and negative reviews. Preprocessing processes are carried out, including case folding, tokenization, filtering, and stemming. After the data processing stage, the author applies the K-Nearest Neighbor (KNN) algorithm. The data was tested to produce negative reviews and positive reviews. The research results show the level of accuracy of KNN. This research is expected to provide an understanding of the use of KNN in analyzing fingerprint user sentiment. The dataset involves body weight parameters and other influencing factors. The model was developed and tested using cross-validation techniques to ensure consistency of performance. Research findings show that the application of the K-NN algorithm with Wrapper preprocessing can increase the accuracy of determining fingerprints using the K-Nearest Neighbors method for employees and lecturers. Applying the K-Nearest Neighbor and K-Nearest Neighbor methods with Wrapper as a preprocessing stage in determining human weight information results in the same accuracy value of 90.48%.

Keywords: Analysis. *Fingerprint*, *K-Nearest Neighbors*

1. PENDAHULUAN

Dengan perkembangan teknologi yang semakin maju, dapat menyesuaikan dengan segala perkembangan dan kemajuan teknologi karena seiring dengan perkembangan dan kemajuan teknologi kadang tidak menyadari bahwa teknologi sangat erat kaitannya baik dalam keseharian juga dalam dunia kerja. Sistem pengidentifikasian sidik jari (fingerprint) hanya digunakan di kalangan aparat keamanan untuk menemukan jati diri korban atau tersangka kejahatan. Kini kegunaannya telah bergeser hingga ke perusahaan-perusahaan komersil maupun lembaga pendidikan. Sidik jari manusia merupakan bukti materi yang sangat penting. Tidak ada sidik jari yang identik di dunia ini sekalipun di antara dua saudara kembar (Jusuf et al., 2013).

Penerapan absensi berbasis teknologi bertujuan untuk meminimalisir atau menghilangkan suatu bentuk kecurangan pegawai dalam proses absensi dengan cara titip-menitip absen. Pegawai yang berperilaku seperti itu secara administratif mereka disiplin hadir masuk kantor tapi kenyataannya pegawai tersebut sebenarnya tidak masuk kerja. Sistem absensi ini juga memiliki data rekaman waktu sehingga pegawai yang datang terlambat ataupun pulang mendahului akan terdeteksi. Langkah ini dilakukan untuk melakukan pengawasan terhadap kinerja pegawai supaya tidak ada manipulasi data dan aksi titip-menitip absen.

Universitas Stella Maris Sumba dalam meningkatkan disiplin pegawai dan Dosen penerapan absensi fingerprint (sidik jari) sudah berjalan dengan efektif, dikarenakan selama diterapkannya absensi fingerprint sangat membantu untuk meningkatkan disiplin terhadap kehadiran pegawai dan Dosen namun dikarenakan adanya kerusakan pada fingerprint sehingga absensi berjalan secara manual. Algoritma K-Nearest Neighbors merupakan algoritma populer yang termasuk dalam grup instance-based learning. Metode ini digunakan untuk klasifikasi data yang jaraknya dekat, ada pendapat lain juga mengatakan bahwa K-Nearest Neighbors digunakan untuk menganalisis dan menambang main data. Metode K-Nearest Neighbors merupakan teknik lazy-learning yang melakukan proses klasifikasi berdasarkan pada data yang jaraknya dekat (Widodo, 2021).

Terdapat penelitian mengenai analisis penggunaan metode K-Nearest Neighbor (K-NN) sebelumnya menggunakan algoritma K-Nearest Neighbor (K-NN). Penelitian sebelumnya yaitu penerapan algoritma K-Nearest Neighbor pada sentimen review agen travel (Ernawati and Wati 2018). Analisis sentimen kurikulum 2013 pada Twitter menggunakan Ensemble Feature dan metode K-Nearest Neighbor (Mentari, Fauzi, and Muflikhah 2018). Algoritma K-Nearest Neighbor (K-NN) merupakan metode lazy learning dimana tidak ada model yang dipelajari dari data testing, jadi hanya belajar dari contoh uji harus diklasifikasikan. Algoritma K-Nearest Neighbor (K-NN) bertujuan untuk mengklasifikasikan objek berdasarkan atribut dan data training (Salam, Zeniarja, and Khasanah 2018). Sentimen analisis tweet berbahasa Indonesia dengan Deep Belief Network (DBN) dan hasil DBN dibandingkan dengan Naïve Bayes dan Support Vector Machine (SVM). Tujuan penelitian tersebut adalah untuk mengetahui hasil sentimen terhadap tweet berbahasa Indonesia di Twitter dan hasil akurasi DBN sebesar 93.31%, Naïve Bayes sebesar 79.10%, dan SVM sebesar 92.18%. Sentimen Analisis Tweet Pornografi Kaum Homoseksual Indonesia Di Twitter dengan Naive Bayes. Tujuan penelitian ini adalah mengetahui hasil sentimen positif dan negatif terhadap data uji tweet dan berdasarkan hasil pengujian tersebut dapat disampaikan kepada pengguna Twitter secara luas untuk menggunakan Twitter secara tepat. Hasil sentimen analisis terhadap 500 data uji menunjukkan bahwa nilai sentimen negatif sangat tinggi yaitu 68.4%. Sedangkan hasil perbandingan akurasi kedua metode adalah metode Naïve Bayes sebesar 87.48% dan k-NN 85.40% dimana metode Naïve Bayes lebih akurasi dibanding metode kNN.

Dari uraian diatas maka penulis tertarik untuk melakukan penelitian dengan judul “Analisis Penggunaan Fingerprint dengan metode K-Nearest Neighbor (K-NN) bagi Pegawai dan Dosen di Universitas Stella Maris Sumba”.

2. METODE

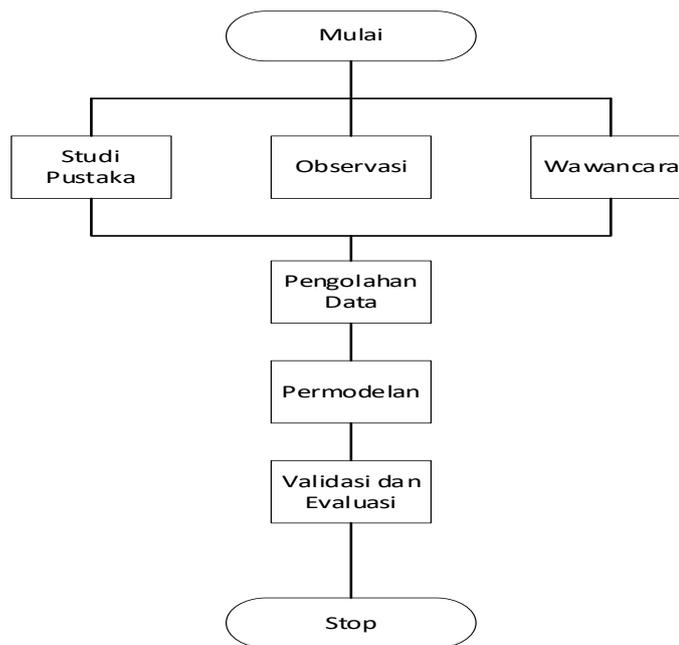
Penelitian ini akan menggunakan metode kualitatif yang dimana pengumpulan datanya merupakan kegiatan dengan menggunakan panca indra, bisa penglihatan, penciuman, pendengaran atau lebih dikenal dengan penelitian observasi. Dimana Untuk memperoleh data yang dibutuhkan untuk menjawab segala permasalahan dalam penelitian. Dalam memilih masalah penelitian terdapat beberapa pertimbangan salah satunya adalah ketersediaan sumber data. Penelitian kuantitatif lebih bersifat *explanationn* (menerangkan, menjelaskan), karena itu bersifat *to learn about the people* (masyarakat objek), sedangkan penelitian kualitatif lebih bersifat *understanding* (memahami) terhadap fenomena atau gejala sosial, karena bersifat *learn about the people* (masyarakat sebagai objek).

Menurut V. Wiratna Sujarweni (2018) sumber data adalah subjek dari data penelitian itu di peroleh. Dalam Kamus Besar Bahasa Indonesia data di artikan sebagai kenyataan yang ada dan berfungsi sebagai bahan sumber untuk menyusun suatu pendapat, keterangan yang benar dan keterangan atau bahan yang di pakai untuk penalaran dan penyelidikan. Kesimpulan nya sumber data adalah subjek dari data penelitian di mana mejadi pendatang sebuah informasi yang menjadi data.

Untuk memperoleh data sebagai bahan pembuatan skripsi, maka dilakukan beberapa cara teknik pengumpulan data yaitu:

1. Observasi Yaitu pengumpulan data dengan cara melakukan pengamatan secara langsung di Universitas Stella Maris Sumba guna memperoleh data-data yang dapat membantu dalam proses perancangan sistem yang meliputi prosedut pemberian koesioner.
2. Wawancara
Metode wawancara ini dilakukan dengan Wawancara secara langsung kepada Pimpinan untuk memperoleh data-data yang dibutuhkan dalam penelitian.
3. Studi Pustaka
4. Studi pustaka dilakukan dengan membaca buku, jurnal, referensi yang menguraikan teori-teori mempunyai studi kasus yang berkaitan dengan penggunaan *fingerprint* dan hal-hal terkait lainnya.

Berikut pada gambar 1 merupakan alur dari tahapan penelitian.



Gambar 1. Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

Proses data pada penelitian ini membagi data-data yang telah dilakukan proses preprocessing menjadi dua bagian yaitu data latih dan data uji. Pembuatan data latih ini bertujuan sebagai pedoman dalam menentukan hasil data uji. Data komentar yang disimpan dalam excel bisa dilihat pada gambar dibawah ini:

| | TEXT | SENTIMEN |
|----|--|----------|
| 1 | | |
| 2 | Guru selalu memeberikan motivasi dalam belajar | positif |
| 3 | Guru selalu membimbing apa yang baik dan salah | positif |
| 4 | Guru mengajar sangat baik dan mudah dipahami | positif |
| 5 | Materi yang disampaikan sangat bagus | positif |
| 6 | Guru selalu mengajarkan tata krama yang baik dan santun | positif |
| 7 | Materi yang diajarkan menjadi mudah dimengerti karena pembahasan tidak terlalu cepat | positif |
| 8 | Materi yang diajarkan selalu diberi contoh sehingga mudah dipahami | positif |
| 9 | Guru sangat tegas sehingga murid menjadi rajin belajar dan mengerjakan tugas | positif |
| 10 | Guru selalu sabar dalam mengajar di sekolah | positif |
| 11 | saya puas dengan relasi yang terbangun antara sesama guru di sekolah | positif |
| 12 | Penyampaian materi sangat mudah dipahami | positif |
| 13 | Guru mengarajar sangat detail | positif |
| 14 | Guru mengajar secara bertahap dan lembut | positif |
| 15 | Guru menyampaikan materi sangat menyenangkan | positif |
| 16 | Guru selalu mengingatkan yang benar dan salah | positif |
| 17 | Guru selalu mengajar dan membimbing dengan baik meski sulit untuk diajari | positif |
| 18 | Guru membuat mata pelajaran lebih mudah dimengerti | positif |
| 19 | Guru memberikan ilmu yang belum saya miliki dan berguna untuk masa depan | positif |

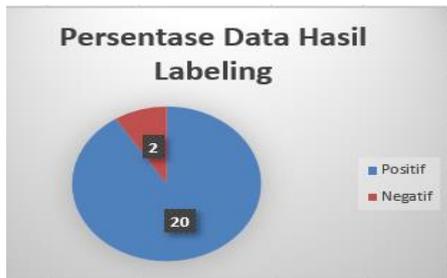
Gambar 4.1 Contoh Data Latih

Persentase Data Hasil Labelling

Pada proses pengambilan data yang dilakukan berdasarkan koesioner yang dibagikan kepengunjung dengan menghasilkan data dengan jumlah 22 komentar.

Setelah seluruh komentar berhasil disimpan selanjutnya dilakukan proses pelabelan. Dalam proses pelabelan yang dilakukan seharusnya untuk menentukan komentar yang didapatkan apakah bernilai positif, negative. amun pada proses pelabelan pada penelitian ini terdapat beberapa komentar yang tidak sesuai dengan sentimen komentar yang sebenarnya.

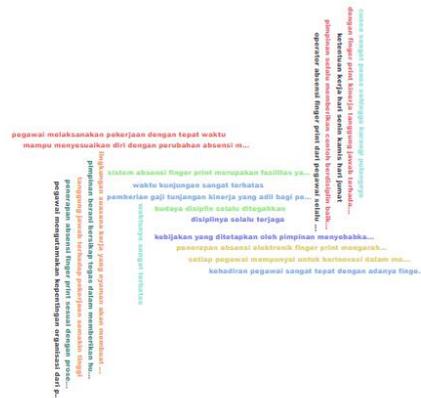
Dan hasil akhir dari proses pelabelan data menghasilkan data sebanyak 20 komentar bernilai positif, 2 komentar bernilai negatif. Persentase data bernilai positif, negatif dan positif negatif dan dapat dilihat pada gambar 4.8



Gambar 4. 2 Persentase Data Hasil Labeling

Ekstraksi Fitur

Setelah terbentuknya file yang akan dijadikan dataset, maka selanjutnya data tersebut akan dibentuk menjadi sebuah model klasifikasi. Namun sebelum membentuk model, ada beberapa tahapan yang harus dilakukan agar terbentuknya suatu model yang baik. Yang pertama dilakukan adalah membaca file xlsx dan kemudian dilakukan tokenisasi terhadap seluruh. Dokumen dalam file tersebut. Berdasarkan hasil tokenisasi yang dilakukan, maka penulis juga ingin mengetahui frekuensi kata yang banyak diperbincangkan oleh konsumen, untuk itu penulis memvisualisasikannya dalam bentuk wordcloud pada gambar 4.18.



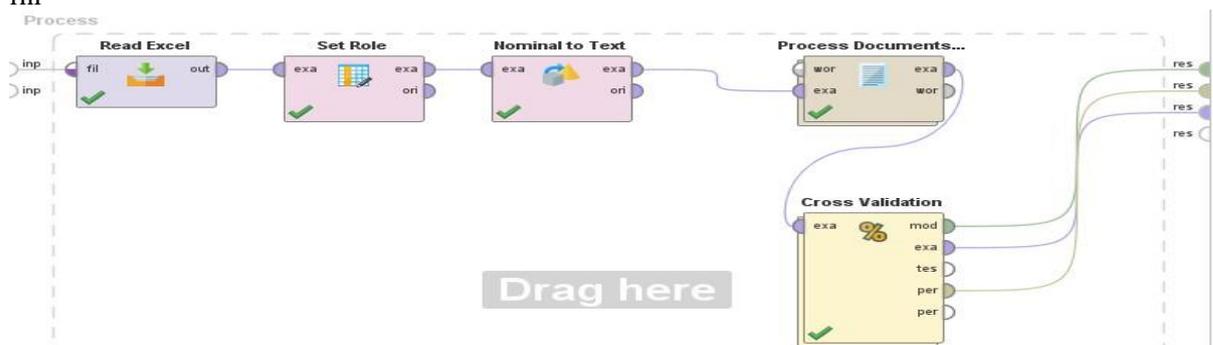
Gambar 4. 3 Proses Document From Data

Perhitungan Akurasi Dengan RapidMiner

Penelitian ini menggunakan tools RapidMiner versi 10.1. perhitungan akurasinya akan proses dengan menggunakan metode Naive Bayes Classifier dan Support Vektor Machine. Untuk mengetahui tingkat akurasi dari kedua metode tersebut.

Proses Dokument dengan metode Naive Bayes Classifier

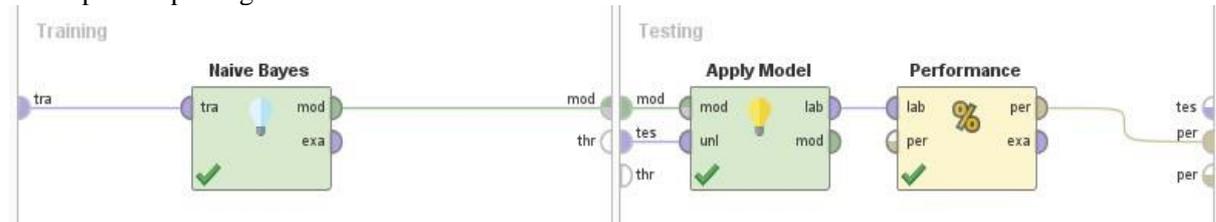
Pada proses ini diuji dengan data latih dari excel. Seperti yang terlihat pada gambar dibawah ini



Gambar 4.4 Proses Dokumen Data Latih

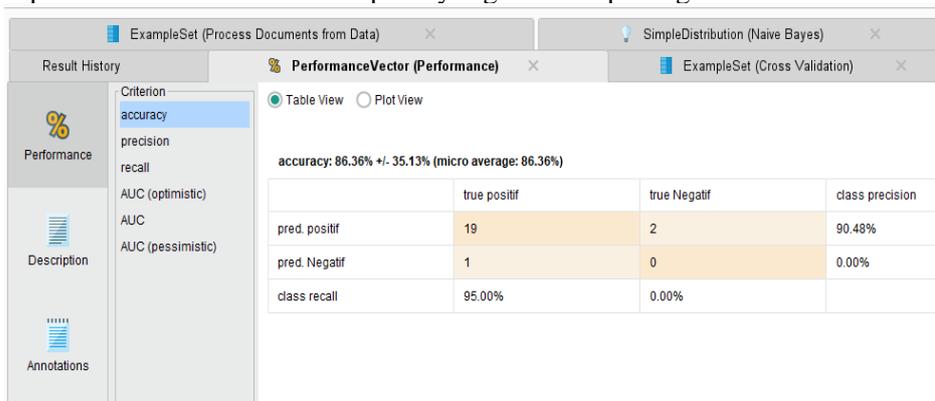
Perhitungan Akurasi dengan Naive Bayes Classifier pada RapidMiner

Proses *Training* dan *Testing* dengan metode Naive Bayes Classifier pada RapidMiner ditampilkan pada gambar dibawah ini:



Gambar 4.5 Proses *Training* dan *Testing* dengan Naive Bayes Classifier pada RapidMiner

Hasil perhitungan klasifikasi menggunakan metode Metode Naive Bayes Classifier diperoleh sebesar 76.92% seperti yang terlihat pada gambar dibawah.



Gambar 4.6 Akurasi dengan Metode Naive Bayes Classifier pada RapidMiner

Hasil Pengujian Naive Bayes

Akurasi yang diuji pada tools RapidMiner dengan menggunakan metode Naive Bayes Classifier. Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier sebesar 90,48%. oleh karena itu tingkat akurasi yang diperoleh dengan menggunakan metode yang diuji pada tools RapidMiner.

Tabel 3. Hasil Perhitungan Akurasi

| Metode | Tingkat Predikat | True Positif | True Negatif |
|--------------------------------------|------------------|--------------|--------------|
| Naive Bayes Classifier | Pred. Positif | 19 | 2 |
| | Pred. Negatif | 1 | 0 |
| Akurasi pada tools RapidMiner | | | |
| Naive Bayes Classifier | | 90,48% | |

4. SIMPULAN DAN SARAN

Berdasarkan Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier adalah: tingkat akurasi dengan metode Naive Bayes Classifier sebesar 90,48%. memiliki hasil akurasi yang cukup baik untuk menganalisis sentimen komentar atau tweet dari pengunjung. Hasil tingkat akurasi juga menunjukkan adanya kepuasan pegawai dan Dosen

5. DAFTAR PUSTAKA

- Ernawati, Siti, and Risa Wati. 2018. "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel." *Jurnal Khatulistiwa Informatika* 6(1): 64–69.
- Hendri, Rifwan. 2018. "No Title." <https://travel.tempo.co/read/1149739/enam-manfaat-traveling-dikehidupan-yang-semakin-sibuk>.
- Mentari, Nurul Dyah, M Ali Fauzi, and Lailil Muflikhah. 2018. "Analisis Sentimen Kurikulum 2013 Pada Sosial Media Twitter Menggunakan Metode K-Nearest Neighbor Dan Feature Selection Query Expansion Ranking." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya* 2(8): 2739–43.
- Salam, Abu, Junta Zeniarja, and Rima Septiyan Uswatun Khasanah. 2018. "Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekpress Indonesia)." *Prosiding SINTAK*: 480–86.
- Wilianto, L., Pudjiantoro, T. H., & Umbara, F. R. 2017. Analisis Sentimen Terhadap Tempat Wisata Dari Komentar Pengunjung Dengan Menggunakan Metode Naive Bayes Classifier Studi Kasus Jawa Barat. *Jurnal Prosiding Sntatif Vol. 4*.
- Liu, B. 2012. *Sentiment analysis and opinion mining. Synthesis lectures on human language technologies*. California: Morgan & Claypool Publishers.
- Kim, S. B., Han, K. S., Rim, H. C., & Myaeng, S. H. 2006. "Some effective techniques for naive bayes text classification." *IEEE transactions on knowledge and data engineering*, 18(11), 1457-1466.