



# Implementasi Algoritma *Naïve Bayes Classifier* Dalam Klasifikasi *Review Program Career Preparation Eduwork*

Adinda Alifatul Jannah<sup>1</sup>, Achmad Fauzan<sup>2</sup>,

<sup>1, 2</sup> Prodi Statistika, FMIPA Universitas Islam Indonesia

Corresponding author email: [achmadfauzan@uii.ac.id](mailto:achmadfauzan@uii.ac.id)

**Abstract:** The development of digital technology in the industrial era 5.0 is directly proportional to the increasingly strict job selection qualifications. Degrees and education do not guarantee someone to get a job if they are not equipped with critical skills and global skills so that careful career preparation is needed. Eduwork as a provider of skills improvisation, mentoring, and job distribution courses launched the Career Preparation Program as a form of guidance service for improvising skills of job seekers. This study aims to identify feedback and ratings from users of the newly launched Career Preparation Program by Eduwork using *Naïve Bayes Classifier* algorithm implemented by sentiment analysis. Based on the research, the performance produced from the test results carried out on a ratio of 70:30 between train data compared to test data resulted in an accuracy rate of 92.5%, with a precision rate is 1 While in the Recall Class a value of 96.72% was produced (positive class: positive). The majority of Career Preparation Program users provide positive feedback and are satisfied with Eduwork services.

**Keywords:** *Naïve Bayes Classifier*, Sentiment analysis, Review, Career preparation program

**Abstrak:** Perkembangan teknologi digital di era industri 5.0 berbanding lurus dengan kualifikasi seleksi kerja yang semakin ketat. Gelar dan pendidikan tidak menjamin seseorang memperoleh pekerjaan jika tidak dibekali dengan *critical skill* dan *global skill* sehingga diperlukan persiapan karir yang matang. Eduwork sebagai penyedia kursus improvisasi *skill*, mentoring, dan penyaluran kerja meluncurkan *Career Preparation Program* sebagai bentuk layanan bimbingan untuk improvisasi *skill* para pencari pekerjaan. Penelitian ini bertujuan untuk mengidentifikasi *feedback* serta penilaian dari pengguna *Career Preparation Program* yang baru diluncurkan oleh Eduwork menggunakan analisis sentimen yang diimplementasikan dalam algoritma *Naïve Bayes Classifier*. Berdasarkan penelitian, diperoleh performa yang di hasilkan dari hasil pengujian yang dilakukan terhadap perbandingan 70:30 antar data *train* berbanding data *test* dihasilkan nilai akurasi sebesar 92.5%, dengan nilai presisi 1 Sementara pada *Class Recall* dihasilkan nilai 96,72% (*positive class: positive*). Para pengguna *Career Preparation Program* mayoritas memberikan *feedback* positif dan puas dengan pelayanan Eduwork.

**Kata kunci:** *Career preparation program*, *Naïve Bayes Classifier*, Analisis sentimen, Review

## I. PENDAHULUAN

Persaingan dunia kerja di era digitalisasi industri 5.0 berkembang semakin ketat. Hal tersebut sejalan dengan perkembangan teknologi digital yang menuntut *skill* dan profesionalisme pekerja. Banyak perusahaan besar yang telah menggunakan perkembangan teknologi kecerdasan buatan (*artificial intelligence*) dengan menggunakan robot atau mesin dalam proses industri. Hal tersebut membawa pengaruh besar dalam penurunan angkatan kerja. Menurut data BPS, di Bulan Agustus 2022 terdapat 8.425.931 orang atau 5.86% jumlah pengangguran terbuka di Indonesia. Proporsi pengangguran terbuka mayoritas alumni sekolah kejuruan sebanyak 9.42%, lulusan SMA 8.57%, lulusan SMP 5.95%, lulusan universitas dan diploma 9.39%, dan sisanya lulusan sekolah dasar [1]. Angka tersebut mengalami penurunan pada Februari 2023 menjadi 7.99 juta orang dengan proporsi tertinggi berasal dari Provinsi Banten [2].

Gelar dan pendidikan tidak menjamin seseorang memperoleh pekerjaan jika tidak dibekali dengan *critical skill* dan *global skill* (terutama kemampuan komunikasi dan bahasa asing). Angka pengangguran terbuka yang lebih besar dari lapangan kerja yang ada ternyata tidak optimal untuk menemukan sumber daya manusia yang sesuai kualifikasi dari perusahaan. Perusahaan akan memberi rangkaian seleksi yang banyak untuk menentukan kualifikasi sumber daya manusia yang dapat bersaing dan memiliki profesionalitas [3]. Berdasarkan survei *Talent Development* pada *Engineering Career Center* (ECC) mengenai permasalahan para pencari kerja saat *apply* pekerjaan di perusahaan,



diperoleh beberapa permasalahan, seperti minimnya kemampuan *public speaking*, kepercayaan diri saat wawancara, kebingungan mencari informasi di dunia kerja, dan kurangnya pengetahuan dalam pembuatan *Curriculum Vitae* [4]. Dari survei diatas diperoleh informasi bahwa kesulitan kerja tidak hanya dialami oleh pekerja berpengalaman yang telah diterima disuatu perusahaan, namun juga oleh para pencari kerja.

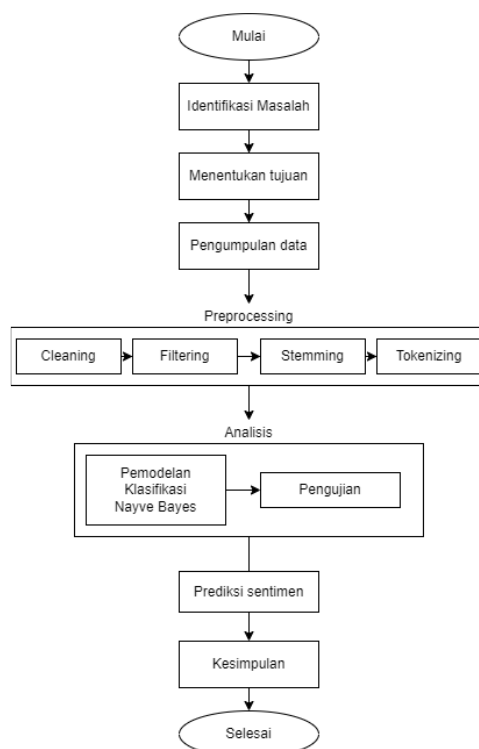
Diperlukan persiapan karir atau *career development* menuju rangkaian seleksi perusahaan dan pengembangan skill agar lebih unggul dari pencari kerja yang lain sangat dibutuhkan. Karena marak digunakan dalam meningkatkan kualitas SDM, banyak program-program *career development* yang diadakan oleh lembaga-lembaga kursus dan penyedia kegiatan mentoring, salah satunya adalah Eduwork. Eduwork merupakan *platform e-learning* yang membantu pengembangan kemampuan dan menyediakan layanan *mentoring* serta penyaluran kerja. Eduwork memiliki banyak program unggulan yang dapat dipilih dalam menelusuran minat, improvisasi kemampuan, bahkan *career development* yang baru-baru ini diluncurkan.

*Career Preparation Program* merupakan program berbentuk *career development* sebagai upaya mempersiapkan dan memberikan layanan bimbingan kepada pada pencari kerja untuk menjadi sumber daya yang unggul dari para pencari kerja lainnya. Program ini menjawab permasalahan para pencari kerja dalam mempelajari persiapan seleksi pekerjaan. Para pencari kerja akan dibimbing langsung dengan praktisi profesional secara privat. Program ini baru diperkenalkan ke masyarakat pada tahun 2023 sehingga dibutuhkan banyak masukan dan kecenderungan masyarakat dalam menerima adanya program tersebut. Dalam proses analisis data, digunakan *text mining* jenis analisis sentimen untuk mengidentifikasi *review* pengguna *Career Preparation Program*. Analisis sentimen mengelompokkan *review* dalam ungkapan positif, negatif, bahkan netral. Beberapa pendekatan sentimen yang marak diimplementasikan antara lain *Naïve Bayes Classifier*, SVM, *K Nearest Neighbor*, *Lexicon*, dan masih banyak pendekatan lainnya [5]. Pendekatan sentimen yang ada memiliki beberapa perbedaan pada proses pengklasifikasian, seperti *Naïve Bayes Classifier* yang melibatkan sedikit kuantitas data *train* dengan proses pengklasifikasian jauh lebih cepat dari yang lainnya dalam tiga kelompok, yaitu *positive*, *negative*, atau netral [6] Hal tersebut menjadi faktor penggunaan analisis sentimen pada penelitian ini dengan memperhatikan kuantitas data yang sedikit.

Berdasarkan penelitian Permadi [7], diperoleh informasi performa akurasi klasifikasi dan proporsi kategori hampir 90% pada *review* restoran di Singapura. Penelitian yang dilakukan Hayuningtyas dan Retnosari [8] menyatakan nilai akurasi *review* pendapat publik akan penggunaan bahasa yang baik dalam Taman Mini Indonesia Indah dengan analisis *Naïve Bayes* sebesar 70%. Dari beberapa hasil analisis terdahulu, dalam penelitian ini peneliti ingin meneliti *review* program untuk mengetahui umpan balik serta penilaian dari pengguna *Career Preparation Program* yang baru diluncurkan oleh Eduwork.

## II. METODE PENELITIAN

Data pada penelitian ini berbentuk primer berdasarkan wawancara pada 10 pengguna pertama *Career Preparation Program* Eduwork. Data dihimpun pada 10-16 Februari 2023. Data berbentuk teks dengan konsep *text mining* dan dianalisis menggunakan analisis sentimen dengan pendekatan *Naïve Bayes Classifier*. Hanya terdapat satu variabel *review* dalam proses analisis ini.

**Gambar 1.** Tahapan Penelitian

### 3.1 *Text Mining*

Penggalian teks merupakan proses dengan tujuan mengklasifikasi segala bentuk informasi yang berbentuk teks untuk menentukan pola kata yang bervariasi [9]. *Text Mining* biasanya berupa *preprocessing* data dalam penelitian untuk mentransformasikan data mentah menjadi data yang rapi dan tertata. *Preprocessing* data dalam *text mining* terdiri atas *cleaning*, *filtering*, *stemming*, dan *tokenizing*. Pada penelitian ini tidak dilakukan *case folding* karena seluruh data tidak menggunakan huruf kapital [10].

1. *Cleaning*: Membersihkan data dari karakter atau simbol-simbol yang berpengaruh pada proses analisis.
2. *Filtering*: Menghapus teks atau kata tidak penting berdasarkan *stopword*.
3. *Stemming*: Pengecekan kata untuk diselaraskan sehingga mengurangi daftar kata.
4. *Tokenizing*: Pemisahan teks menggunakan spasi sebagai entitas yang siap di analisis.

### 3.2 Analisis Sentimen

Analisis sentimen merupakan bentuk *text mining* untuk mengidentifikasi, ekstraksi, dan mengolah opini atau *review* untuk memperoleh klasifikasi dan umpan balik dari pandangan masyarakat terhadap sesuatu yang dapat dinilai. Sentimen mengacu pada permasalahan yang telah ditentukan yang biasanya didasarkan pada subjek penelitian [11].

### 3.3 Algoritma *Naïve Bayes Classifier*

Algoritma ini menghasilkan pengklasifikasian bersifat *supervised learning* yang menggunakan pendekatan peluang atau probabilitas beberapa tahun yang akan datang dari informasi masa lalu. Model yang digunakan bersumber pada teorema Bayes seperti persamaan (1).

$$P(A|X) = \frac{P(A) \cdot P(X|A)}{P(X)} \quad (1)$$

Persamaan (1) memuat nilai  $X$  sebagai atribut,  $A$  menyatakan kelas,  $P(A|X)$  sebagai probabilitas  $A$  jika  $X$  telah terjadi,  $P(X|A)$  merupakan probabilitas  $X$  jika  $A$  telah terjadi, probabilitas kejadian  $A$  dinyatakan dalam  $P(A)$ , dan  $P(X)$  adalah probabilitas kejadian  $X$  [12].

Algoritma ini juga dapat menentukan kategori setiap dokumen menggunakan rumus:

$$C^* = \arg \max_{C \in \mathcal{C}} p_{\mu} p(W_k | C_i) \quad (2)$$

Dengan nilai  $A^*$  dalam persamaan (2) merupakan kategori yang mempunyai peluang kemunculan dan seluruh nilai di sisi kiri adalah nilai data *training* yang dimaksimalkan [13].

### III. HASIL DAN PEMBAHASAN

Peneliti melihat terlebih dahulu data yang digunakan dalam penelitian dalam sebuah *wordcloud*. *Wordcloud* merupakan salah satu bentuk visualisasi yang digunakan untuk menggambarkan kata dalam sebuah matriks agar lebih mudah dipahami dan menarik [14]. *Wordcloud* data review produk *Career Preparation Program* Eduwork seperti pada gambar berikut.



**Gambar 2** *Wordcloud* Data Penelitian

*Wordcloud* diatas menggambarkan beberapa kata yang paling banyak disebutkan dalam *review* adalah *mentor*, *interview*, *baik*, *seru*, *materi*, *simulasi*, dan *portofolio*. Kata ‘*mentor*’ menggambarkan bahwa kegiatan ini dilayani oleh *mentor* profesional dibidangnya. Adanya *mentor* menjadi poin utama dalam *Career Preparation Program* ini. Keberadaan *mentor* sangat penting dalam program ini. Kata dengan kuantitas terbanyak kedua ialah ‘*interview*’. Hal ini menjelaskan bahwa salah satu kendala yang dialami pengguna dan dapat diselesaikan oleh pembinaan dari *Career Preparatin Program* adalah *interview* atau wawancara kerja. *Materi interview* menjadi materi terpenting dalam program ini. Kata berikutnya adalah ‘*baik*’ dan ‘*seru*’ yang menandakan bahwa *mentor* yang membersamai pengguna bersifat baik dan menyampaikan materi pembelajaran dengan suasana kelas yang seru. Selanjutnya terdapat kata ‘*simulasi*’ yang menandakan keunggulan materi *interview* disertai dengan simulasi agar para pengguna lebih mendalami materi dan merasakan atmosfer wawancara yang sebenarnya. Kata terakhir yang memiliki kuantitas besar adalah kata ‘*portofolio*’ yang menandakan kesulitan tertinggi kedua yang dialami pengguna sehingga memutuskan mendaftarkan dirinya ke *Career Preparation Program* ini. Pembuatan *portofolio* sangat penting agar unggul dalam persaingan seleksi kerja. Pembuatan *portofolio* menjadi program unggulan lain setelah adanya simulasi *interview*. Terdapat banyak kata lain dengan kuantitas yang rendah. Kata-kata yang banyak disebutkan memiliki konotasi positif sehingga mayoritas pengguna memberikan *feedback* positif jika diperhatikan dari *wordcloud* yang ada.

#### 1. *Text Preprocessing*

Tahap ini mempersiapkan data dengan mentransformasikan teks yang tidak terstruktur agar terstruktur dan baku tanpa tambahan apapun. Tahap ini mempersiapkan proses berikutnya yaitu *text mining* dengan analisis sentimen untuk memperoleh informasi berdasarkan *feedback* yang ingin dicapai. *Text preprocessing* yang dilakukan adalah *cleaning*, *stemming*, *filtering*, dan *tokenizing*. Jika data yang dimiliki tidak dipersiapkan menggunakan *text preprocessing*, maka akan memperlambat analisis dan memberi informasi yang tidak optimal.

## 2. *Cleansing*

Proses *cleansing* dilakukan untuk menghilangkan karakter, tanda baca, ataupun simbol mengganggu (tanpa berdampak) pada proses analisis sentimen, salah satunya menghilangkan titik (.), tanda koma (,), atau karakter yang bukan berupa huruf dan angka. Proses ini mempersingkat waktu analisis dan mempermudah algoritma dalam klasifikasi [15].

**Tabel 1** Proses *Cleansing*

Sebelum	Sesudah
memberi byk informasi belajar, mulai dari pembuatan cv, wawancara, smp dgn seluruh rangkaian pencarian pekerjaan.	memberi byk informasi belajar mulai dari pembuatan cv wawancara smp dgn seluruh rangkaian pencarian pekerjaan

## 3. *Stemming*

Pada tahapan *stemming*, data diperbaiki agar tidak terdapat kesalahan pengedaan atau kata singkatan sebagai upaya penyeragaman kata dan pengecekan kata agar lebih ringkas berdasarkan kata baku yang mempunyai kesamaan arti namun berbeda bentuk penulisan, termasuk imbuhan yang dihilangkan [16].

**Tabel 2** Proses *Stemming*

Sebelum	Sesudah
memberi byk informasi belajar mulai dari pembuatan cv wawancara smp dgn seluruh rangkaian pencarian pekerjaan	beri banyak informasi belajar mulai dari buat cv wawancara sampai dengan seluruh rangkaian cari kerja

## 4. *Filtering*

Tahapan berikutnya adalah filterisasi sebagai upaya untuk menghilangkan teks atau kata yang tidak penting dalam proses sentimen. Dalam penelitian ini, *stopword* yang dijadikan acuan berasal dari kamus *stopwords*. Nantinya kata-kata yang telah diseragamkan akan diklasifikasikan untuk mengikuti proses analisis atau dihilangkan [17].

**Tabel 3** Proses *Filtering*

Sebelum	Sesudah
beri banyak informasi belajar mulai dari buat cv wawancara sampai dengan seluruh rangkaian cari kerja	beri informasi belajar buat cv wawancara rangkaian cari kerja

## 5. *Tokenizing*

Pada tahap *preprocessing* yang terakhir, dilakukan pemisahan kata menjadi beberapa potongan kata atau token. Karakter yang kurang dari dua dan bukan huruf atau angka akan dieliminasi. Pola ini digunakan untuk melihat pola *review* yang bernilai dalam *text matrix* yang akan diolah. Tahap ini telah disesuaikan dengan tahap-tahap sebelumnya dan hanya menyisakan kata-kata yang menunjang analisis [18].

**Tabel 4** Proses *Tokenizing*

Sebelum	Sesudah
beri informasi belajar buat cv wawancara rangkaian cari kerja	“beri” “informasi” “belajar” “buat” “cv” “wawancara” “rangkaian” “cari” “kerja”

## 6. *Skenario Pengujian*

Dengan 10 *review* masa percobaan peluncuran *career preparation program*, dilakukan skenario perbandinga data *training* dengan data *testing* yang merupakan sebuah tahap yang tidak dapat terlewat dalam mengembangkan model *machine learning*. Pembagian data ini bertujuan untuk menguji sejauh mana kinerja model *machine learning* dalam memprediksi hasil yang akurat pada *new dataset* yang akan ada di masa mendatang. Data *training* adalah bagian data yang berguna sebagai



pelatih model *machine learning*. Data *training* merupakan *dataset* yang sudah diketahui labelnya atau target *outputnya*, sehingga model *machine learning* dapat mempelajari pola pada data tersebut dan menghasilkan model yang dapat melakukan prediksi dengan akurat [18-19]. Contoh penggunaan perbandingan ini dapat diperhatikan dalam tabel 5 dan 6.

**Tabel 5** Contoh data *training*

No	Class	Keywords
1	Positif	Baik (4), Seru (4), Asik (2)
2	Negatif	Kaku (2), Dadakan (2)
3	Netral	Interview (8), Mentor (7)

**Tabel 6** Contoh data *testing*

Kaku	Interaksi	Seru
------	-----------	------

### 7. Analisis Algoritma *Naïve Bayes Classifier*

Pembagian ini menggunakan perbandingan 70:30 dan siap dianalisis menggunakan algoritma *Naïve Bayes Classifier*. Data *training* membentuk peluang dan diujikan pada tabel peluang oleh data *testing*. Tabel 7 menyatakan proporsi pembagian data yang menguji *machine learning*.

**Tabel 7** Proporsi Perbandingan Data

No	Class	Data Training	Data Testing
1	Positif	703	107
2	Negatif	12	107
3	Netral	81	107

Dari penerapan *Naïve Bayes Classifier*, diperoleh asumsi probabilitas untuk setiap  $P(X)$ . Asumsi yang dimiliki adalah *review* positif dan negatif dari pengguna *Career Preparation Program*.

**Tabel 8** Probabilitas Klasifikasi

Positif		Negatif	
Positif	703	Negatif	12
P (positif)	0.983	P (negatif)	0.017

Contoh penggunaan pada kata mentor yang terdiri atas 7 kata data *training*, diperoleh nilai  $P(\text{mentor}|\text{Positif}) = 0.875$  dan  $P(\text{mentor}|\text{negatif}) = 0.0765$  menggunakan rumus analisis pada persamaan (1).

### 8. Analisis Sentimen dengan *Confusion Matrix*

Setelah dibuat proporsi data, dilakukan klasifikasi berdasarkan *output actual* dan *prediction* data. Penggunaan *confusion matrix* mengidentifikasi banyaknya prediksi benar dan salah pada setiap kelas negative dan kelas positif.

**Tabel 9** *Confusion Matrix*

	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	TP	FP
<i>Negative</i>	FN	TN

Keterangan [7]:

- TP berarti *True Positive* yang menyatakan kuantitas data bernilai *actual* positif diklasifikasi dalam positif.
- TN berarti *True Negative*, yang menyatakan kuantitas data bernilai *actual* negatif diklasifikasi dalam negatif.
- FN berarti *False Negative*, yang menyatakan kuantitas data bernilai *actual* negatif tetapi diklasifikasi salah.



- FP berarti *False Positive*, yang menyatakan kuantitas data bernilai *actual* positif tetapi diklasifikasi salah.

Jika telah diperoleh angka yang menyusun *confusion matrix*, maka nilai tersebut digunakan untuk mengukur nilai *accuracy*, *precision*, dan *recall*. Berdasarkan analisis yang telah dilakukan, *review* program dengan *class 0* terklasifikasi *True* berjumlah 328 kata, disamping itu *class 1* dengan klasifikasi benar sebanyak 106 kata. Akurasi pada klasifikasi benar adalah 93.75%. Sedangkan klasifikasi salah pada *class 0* dan *class 1* tidak lebih dari 30 kata dengan akurasi 92%. Berdasarkan informasi berikut, diperoleh performa yang di hasilkan dari hasil pengujian yang dilakukan terhadap 70% data *train* dengan perbandingan 30% data *test* menghasilkan *accuracy* senilai 92.5%, dengan nilai presisi 1 Sementara pada *Class Recall* dihasilkan nilai 96,72% (*positive class: positive*).

#### IV. KESIMPULAN

1. Berdasarkan *sentiment analysis* menggunakan pendekatan *Naïve Bayes Classifier* yang telah dilakukan, diperoleh probabilitas klasifikasi positif sebesar 0.983 dan klasifikasi negatif sebesar 0.017 dengan nilai akurasi pada pengujian proporsi antara data *train* berbanding data *test* 70:30 sebesar 92.5%. Peluang *review* positif sebesar 98.3% menyatakan banyaknya *review* positif yang diberikan oleh pengguna *Career Preparation Program*. Hal tersebut bernilai cukup besar dan menyatakan bahwa model yang digunakan dapat menganalisis performa *review* pengguna *Career Preparation Program*.
2. Emosi dasar yang dihasilkan paling banyak adalah positif dan kepercayaan terhadap produk *Career Preparation Program Eduwork*

#### UCAPAN TERIMA KASIH

Penulis berterima kasih pada Eduwork (PT. Talenta Sinergi Group) yang telah memberikan izin penelitian serta seluruh responden penelitian yang merupakan pengguna *Career Preparation Program Eduwork*. Rasa terima kasih juga penulis ucapkan pada panitia Seminar Nasional Sains Data (SENADA) 2023 yang telah mengadakan seminar nasional dan melakukan publikasi pada makalah yang penulis buat. Tidak lupa rasa terima kasih pada segenap rekan yang sudah membantu proses penelitian sehingga makalah ini dapat diselesaikan dengan baik.

#### REFERENSI

1. BPS, "Pengangguran Terbuka Menurut Pendidikan Tertinggi yang Ditamatkan 1986 - 2022," Badan Pusat Statistik, 9 Januari 2023. [Online]. Available: <https://www.bps.go.id/statictable/2009/04/16/972/pengangguran-terbuka-menurut-pendidikan-tertinggi-yang-ditamatkan-1986---2022.html>. [Accessed 20 Februari 2023].
2. A. Ahdiat, "Jumlah dan Tingkat Pengangguran Terbuka di Indonesia (Februari 2019-Februari 2023)," Katadata Media Network, 5 Mei 2023. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2023/05/05/awal-2023-ada-79-juta-pengangguran-di-indonesia>. [Accessed 7 Mei 2023].
3. Akhlanudin, "Gema Uhamka," Universitas Muhammadiyah Prof. Dr. Hamka, 16 Juni 2016. [Online]. Available: <https://gema.uhamka.ac.id/2016/06/16/kenapa-lulusan-perguruan-tinggi-makin-susah-mendapat-pekerjaan/#:~:text=Susah%20terserapnya%20lulusan%20perguruan%20tinggi,manajemen%20perusahaan%2C%E2%80%9D%20kata%20Lilis..>
4. R. A. Ariestanie, "Training "Upgrading Self Value" untuk Mempersiapkan Mahasiswa dan Alumni dalam Menghadapi Proses Seleksi," in *Proceeding Indonesia Career Center Network Summit IV*, Samarinda, 2019.
5. P. A. Permatasari, Linawati and L. Jasa, "Survei Tentang Analisis Sentimen Pada Media



- Sosial," *Majalah Ilmiah Teknologi Elektro*, vol. XX, no. 2, pp. 177-186, 2021.
6. N. T. Romadhoni, I. Santoso and S. Budilaksono, "Perbandingan Metode Naive Bayes, KNN, dan Decision Tree Terhadap Analisis Sentimen Transportasi KRL Commuter Line," *Jurnal IKRA-ITH Informatika*, vol. 3, no. 2, pp. 1-9, 2019.
  7. V. A. Permadi, "Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura," *Jurnal Buana Informatika*, vol. 11, no. 2, pp. 141-151, 2020.
  8. R. Y. Hayuningtyas and R. Sari, "Analisis Sentimen Opini Publik Bahasa Indonesia Terhadap Wisata TMII Menggunakan Naive Bayes dan PSO," *Jurnal Techno Nusa Mandiri*, vol. 16, no. 1, pp. 37-42, 2019.
  9. A. V. Sudiantoro and E. Zuliarso, "Analisis Sentimen Menggunakan Text Mining Dengan Algoritma Naive Bayes Classifier," in *Prosiding SINTAK*, Semarang, 2018.
  10. B. Gunawan, H. S. Pratiwi and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *Jurnal Edukasi dan Penelitian Informatika*, vol. 4, no. 2, pp. 113-118, 2018.
  11. E. Sipayung, H. Maharani and I. Zefanya, "Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem Informasi*, vol. 8, no. 1, pp. 958-965, 2016.
  12. S. M. Salsabila, A. A. Murtopo and N. Fadhillah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naive Bayes Classifier," *Jurnal Minfo Polgan*, vol. 11, no. 2, pp. 30-35, 2022.
  13. D. Darwis, N. Siskawati and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, pp. 131-145, 2021.
  14. M. Pradana, "Penggunaan Fitur Wordcloud dan document term matrix dalam Text Mining," *Jurnal Ilmiah Informatika*, vol. 8, no. 01, pp. 38-43, 2020.
  15. A. F. Rozi and A. S. Purnomo, "Analisis Sentimen Untuk Respon Masyarakat Terhadap Universitas (Studi Kasus : Universitas Mercu Buana Yogyakarta)," *Jurnal Information System & Artificial Intelligence*, vol. 1, no. 2, pp. 53-60, 2021.
  16. N. Saputra, "Analisis Sentimen dengan Preprocessing Kata," *Jurnal Dinamika Informasi*, vol. 7, no. 1, pp. 45-57, 2018.
  17. L. Ardiani, H. Sujaini and Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *Jurnal Sistem dan Teknologi Informasi*, vol. 8, no. 2, pp. 183-190, 2020.
  18. R. Julianto, E. D. Bintari and Indrianti, "Analisis Sentimen Layanan Provider Telepon Seluler pada Twitter menggunakan Metode Naive Bayesian Classification," *Journal of Big Data Analytic and Artificial Intelligence*, vol. 3, no. 1, pp. 23-30, 2017.
  19. Y. Astuti, I. R. Wulandari, A. R. Putra and N. Kharomadhona, "Naive Bayes untuk Prediksi Tingkat Pemahaman Kuliah Online Terhadap Mata Kuliah Algoritma Struktur Data," *Jurnal Edukasi dan Penelitian Informatika*, vol. 8, no. 1, pp. 28-32, 2022.





20. D. Untari, "Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree dan Algoritma C4.5," *Jurnal Informasi & Teknologi*, vol. 2, no. 1, pp. 23-29, 2020.