

KLASIFIKASI LINEARITAS SERAPAN LULUSAN SEKOLAH MENENGAH KEJURUAN MENGGUNAKAN ALGORITMA NAIVE BAYES

Hirga Ertama Putra ^{*1)}, Cahyo Crys dian ²⁾, M. Imammudin ³⁾

1. S2 Teknik Informatika, Teknologi, UIN Maulana Malik Ibrahim Malang, Indonesia
2. S2 Teknik Informatika, Teknologi, UIN Maulana Malik Ibrahim Malang, Indonesia
3. S2 Teknik Informatika, Teknologi, UIN Maulana Malik Ibrahim Malang, Indonesia

Article Info

Kata Kunci: Algoritma Naive Bayes; Klasifikasi; Linearitas; Serapan Lulusan

Keywords: Naive Bayes algorithm; Classification; Linear-bag; Graduate Uptake

Article history:

Received 26 March 2024

Revised 9 April 2024

Accepted 23 April 2024

Available online 1 June 2024

DOI :

<https://doi.org/10.29100/jipi.v9i2.5721>

* Corresponding author.

Corresponding Author

E-mail address:

zegapoetra@gmail.com

ABSTRAK

Pendidikan merupakan instrumen dasar dalam peningkatan kualitas suatu bangsa dan negara. Melalui Sumber Daya Manusia (SDM) yang handal akan menambah kesejahteraan dan peningkatan kualitas bangsa. Secara yuridis SMK dibentuk untuk memberikan alumni yang siap bekerja. Berdasarkan CNN Indonesia menyampaikan bahwa TPT (Tingkat Pengangguran Terbuka) tamatan SMK masih merupakan yang paling tinggi dibandingkan tamatan jenjang pendidikan lainnya. Berbeda dengan kondisi yang ada di sekolah, banyak sekolah yang sudah bisa menyalurkan alumninya dengan cara mengadakan kegiatan jobfair. Dengan menggunakan *Machine Learning* diharapkan dapat membantu untuk mengklasifikasikan linearitas serapan lulusan siswa SMK. Algoritma *Naive Bayes* dipilih untuk mengolah *dataset* serapan lulusan SMK untuk mendapatkan klasifikasi linearitasnya. Berdasarkan hasil penelitian ini didapatkan 623 alumni yang terserap secara linier dan 59 alumni yang terserap tidak linier dari total data sebanyak 682 baris data.

ABSTRACT

Education is a basic instrument in improving the quality of a nation and state. Through reliable Human Resources (HR), we will improve the welfare and improve the quality of the nation. Legally, Vocational Schools were formed to provide alumni who are ready to work. According to CNN Indonesia, the TPT (Open Unemployment Rate) for vocational school graduates is still the highest compared to graduates of other levels of education. In contrast to the conditions that exist in schools, many schools are able to channel their alumni by holding job fair activities. By using Machine Learning it is hoped that it can help to classify the linearity of absorption of vocational school student graduates. The Naive Bayes algorithm was chosen to process the vocational school graduate uptake dataset to obtain a linearity classification. Based on the results of this research, it was found that 623 alumni were absorbed linearly and 59 alumni were absorbed non-linearly from a total of 682 rows of data.

I. PENDAHULUAN

Pendidikan merupakan instrumen dasar dalam peningkatan kualitas suatu bangsa dan negara. Melalui Sumber Daya Manusia (SDM) yang handal akan menambah kesejahteraan dan peningkatan kualitas bangsa. Tercatat dalam sebuah data dalam Badan Pusat Statistik (BPS) bahwa tingkat pengangguran pada Februari 2023 sebesar 5,45 persen dari jumlah penduduk Indonesia [1]. Lulusan Sekolah Menengah Kejuruan (SMK) masih menjadi penyumbang pengangguran paling banyak, berdasarkan CNN Indonesia menyampaikan bahwa TPT (Tingkat Pengangguran Terbuka) tamatan SMK masih merupakan yang paling tinggi dibandingkan tamatan jenjang pendidikan lainnya yaitu sebesar 9,6 persen [2]. Selain dari CNN Indonesia, dituliskan juga pada penelitian sebelumnya yang sudah meneliti tentang indikator Sekolah Menengah Kejuruan (SMK) adalah penyumbang pengangguran terbesar di Provinsi Jawa Timur [3]. Dinamika permasalahan ketenagakerjaan dan pengangguran di Indonesia juga menjadi fokus oleh beberapa peneliti sebelumnya [4]. Dalam penelitian tersebut menyebutkan juga masalah pengangguran yang sangat kompleks dan proses penyelesaiannya butuh banyak pihak yang harus bekerja sama. Dituliskan bahwa aspek kelembagaan adalah elemen kunci yang perlu diperbaiki dalam upaya menurunkan pengangguran.

Penelitian yang dilakukan oleh Amalia [5] menyampaikan bahwa dengan memanfaatkan data kelulusan siswa SMK sebagai sumber datanya, diharapkan dapat menghasilkan informasi tentang pola tingkat kelulusan siswa SMK melalui teknik *datamining*. Salah satunya dengan cara mengklasifikasi data lulusan dengan metode *Naive Bayes*. *Naive Bayes* adalah metode klasifikasi dengan menggunakan metode probabilitas dan statistik yang digunakan untuk memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya yang dikenal sebagai *Teorema Bayes* [6]. Pada penelitian yang dilakukan oleh Purnama,dkk [7] menyimpulkan bahwa klasifikasi merupakan bagian dalam *machine learning* dimana klasifikasi merupakan sebuah proses pembelajaran dari sebuah fungsi target yang mampu memetakan tiap himpunan atribut ke dalam salah satu label *class dependen* yang sudah didefinisikan sebelumnya. Tahapan dalam klasifikasi antara lain adalah proses *training* dan proses *testing*. Proses *training* merupakan proses pembelajaran menggunakan data *training*. Proses *testing* merupakan proses melakukan pengujian terhadap model yang diperoleh pada proses *training* dengan menggunakan data *testing*. Dengan adanya penelitian sebelumnya dan sumber yang lain, dapat disimpulkan bahwa metode *Naive Bayes* adalah salah satu algoritma yang digunakan untuk mengklasifikasikan sekelompok data. Selain itu, jika ditinjau dari hasil nilai *accuracy* dari setiap penelitian sebelumnya metode *Naive Bayes* tergolong algoritma yang mempunyai kecepatan dan ketepatan yang paling baik untuk mengklasifikasi serapan lulusan.

Dari latar belakang tersebut penulis melakukan sebuah penelitian terkait linearitas serapan lulusan SMK, untuk menjawab isu yang ada bahwa lulusan SMK bukan penyumbang pengganguran paling banyak. Dengan tema penelitian mengklasifikasikan serapan lulusan SMK menggunakan metode *Naive Bayes*. Penelitian ini akan menggunakan *feature extraction* jurusan, tahun lulus, nilai UKK, nilai KOMJARDAR, dan nilai PROGDAS. Dengan *feature extraction* tersebut yang akan dijadikan sebuah parameter untuk mendapatkan hasil dari linearitas serapan lulusan SMK.

II. METODE PENELITIAN

A. Metode Naive Bayes

Metode *Naive Bayes* adalah metode yang memberikan cara yang berprinsip untuk memasukkan informasi eksternal ini ke dalam proses analisis data [8]. Proses ini dimulai dengan distribusi probabilitas yang sudah diberikan untuk kumpulan data yang dianalisis. Karena distribusi ini diberikan sebelum data apa pun dipertimbangkan, itu disebut distribusi sebelumnya. Kumpulan data baru memperbarui distribusi sebelumnya menjadi distribusi *posterior*. Alat dasar untuk pemutakhiran ini adalah *Teorema Bayes*. *Teorema Bayes* mewakili latar belakang teoritis untuk pendekatan statistik untuk masalah klasifikasi inferensi induktif.

Algoritma Naive Bayesian menyediakan cara membangun yang probabilistik dalam menyelesaikan sebuah kasus [9]. Pendekatan ini menghitung probabilitas untuk setiap nilai variabel kelas untuk nilai variabel masukan tertentu. Dengan bantuan probabilitas bersyarat, untuk catatan tertentu yang tidak diketahui, modelnya menghitung hasil dari semua nilai kelas target dan menghasilkan pemenang yang diklasifikasikan.

Persamaan $P(c|x)$ yang digunakan oleh metode *Naive Bayes* dapat disajikan sebagai berikut :

$$\text{Posterior} = \frac{\text{Likelihood} \cdot \text{Prior}}{\text{Evidance}}$$

$$P(c|x) = \frac{P(x|c) P(c)}{P(x)}$$

Keterangan:

$P(c|x)$ = Probabilitas hipotesis berdasarkan kondisi posteriori probabilitas

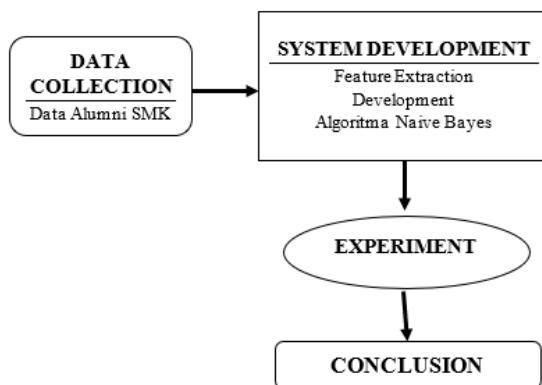
$P(x|c)$ = Probabilitas berdasarkan kondisi c (likelihood)

$P(c)$ = Probabilitas hipotesis prior probabilitas (prior)

$P(x)$ = Probabilitas x (evidance)

B. Prosedur Penelitian

Metode yang digunakan pada penelitian ini, akan dijelaskan pada Gambar 1 tentang prosedur penelitian.



Gambar 1 Prosedur Penelitian

C. Data collection

Pengumpulan data adalah tahapan yang penting dalam sebuah penelitian. Ketersediaan data menentukan dalam proses pengolahan dan analisa selanjutnya. Pengumpulan data harus dilakukan dengan cara yang tepat dan menjamin bahwa data yang diperoleh itu benar dan dapat dipertanggungjawabkan. Cara pengumpulan data yang bersifat teoritis dilakukan dengan mempelajari beberapa literatur, jurnal-jurnal penelitian dan sumber-sumber informasi lain yang ada hubungannya dengan permasalahan yang akan dibahas.

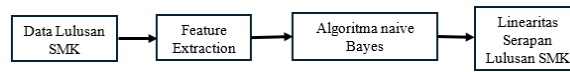
Penelitian ini menggunakan data primer yang didapatkan dari data alumni 3 tahun terakhir (2021, 2022, dan 2023) pada SMK Telkom Malang. Pengambilan data dilakukan dengan menggunakan aplikasi alumni tracert yang dimiliki oleh sekolah dan data nilai diperoleh dari pihak kurikulum sekolah. Jumlah data alumni pada SMK Telkom Malang yang terkumpul sebanyak 682 alumni. Parameter yang akan digunakan dapat dilihat pada Tabel 1 tentang kriteria untuk parameter.

TABEL 1
KRITERIA PARAMETER

<i>Parameter</i>	<i>Kriteria</i>
<i>Jurusan</i>	RPL TKJ
<i>Tahun Lulusan</i>	2021 2022 2023
<i>Bidang</i>	Teknologi Tidak Teknologi
<i>Nilai UKK</i>	Nilai UKK > 90 Nilai UKK = 80 – 89 Nilai UKK = 70 – 79 Nilai UKK < 69
<i>Nilai KOMJARDAS</i>	Nilai KOMJARDAS > 90 Nilai KOMJARDAS = 80 – 89 Nilai KOMJARDAS = 70 – 79 Nilai KOMJARDAS = 60 – 69 Nilai KOMJARDAS < 59
<i>Nilai PROGDAS</i>	Nilai PROGDAS > 90 Nilai PROGDAS = 80 – 89 Nilai PROGDAS = 70 – 79 Nilai PROGDAS = 60 – 69 Nilai PROGDAS < 59

D. System Development

System development yang digunakan pada penelitian ini, akan dijelaskan pada Gambar 2 tentang desain sistem.



Gambar 2 Desain Sistem

Desain *system development* yang ditampilkan pada Gambar 2 menjelaskan bahwa akan dilakukan proses pengumpulan data lulusan SMK. Selanjutnya akan ditambahkan dengan feature extraction yang berisi Jurusan, Tahun Lulus, Bidang, Nilai Uji Kompetensi Kejuruan (UKK), Nilai Komputer Jaringan Dasar (KOMJARDAS), Nilai Pemrograman Dasar (PROGDAS). Kemudian akan diproses dengan algoritma *Naive Bayes* yang digunakan untuk mendapatkan hasil yaitu linearitas serapan lulusan SMK.

E. Feature Extraction

Tahap feature extraction menentukan class klasifikasi diantaranya *class* linier dan *class* tidak linier. Setelah menentukan *class*, tahapan berikutnya adalah menentukan tingkatan jawaban dari kriteria pada masing-masing parameter tersebut dan terlihat pada Tabel 2 tentang tingkatan kriteria.

TABEL 2
TINGKATAN KRITERIA

Parameter	Kriteria	Tingkatan
Jurusan	RPL	RPL 1
		RPL 2
		RPL 3
	TKJ	TKJ 1
		TKJ 2
		TKJ 3
Tahun Lulus	2021	Lulusan 2021
	2022	Lulusan 2022
	2023	Lulusan 2023
Bidang	Teknologi	Kategori Teknologi
	Tidak Teknologi	Kategori Tidak Teknologi
Nilai UKK	Nilai = 90 – 100	Sangat Kompeten
	Nilai = 80 – 89	Kompeten
	Nilai = 70 – 79	Cukup Kompeten
	Nilai < 69	Tidak Kompeten
Nilai KOMJARDAS	Nilai = 90 – 100	Predikat A
	Nilai = 80 – 89	Predikat B
	Nilai = 70 – 79	Predikat C
	Nilai = 60 – 69	Predikat D
	Nilai < 59	Predikat E
Nilai PROGDAS	Nilai = 90 – 100	Predikat A
	Nilai = 80 – 89	Predikat B
	Nilai = 70 – 79	Predikat C
	Nilai = 60 – 69	Predikat D
	Nilai < 59	Predikat E

Tahap berikutnya adalah membentuk basis kasus yang akan menggunakan 682 basis kasus dari data alumni. Kemudian dilakukan pencocokan dengan algoritma *Naive Bayes*, untuk menentukan hasil probabilitasnya. Berikut rician dataset yang akan digunakan untuk proses klasifikasi ditunjukkan pada Tabel 3 tentang data alumni.

TABEL 3
DATA ALUMNI

NO	Nama	Jurusan	Tahun Lulus	UKK	KOM JAR DAS	PROG DAS	Industri	Bidang
1	ACHMAD AZHAR ALDIANSYAH	RPL	2023	85	77	81	PT TELKOM AKSES KEDIRI	Teknologi
2	ADINDA KHARISMA PUTRI TANDAWIJAYA	RPL	2023	76	87	84	CV SINARA SUKSES BERSAMA	Teknologi
3	AIZAR RAHIMA SUPRAYITNO	RPL	2023	95	91	91	BINA NUSANTARA UNIVERSITY	Sistem Informasi
4	AMELIA	RPL	2023	85	86	84	PT FINNET	Teknologi

	INDRIANA						INDONESIA	
5	ARDHIAN CALWA NUGRAHA	RPL	2023	98	86	86	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATIKA
...
680	SYFANADYA WENING ADI	TKJ	2021	81	85	94	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATIKA
681	YUDISTIRA ANDRA FIRMANSYAH	TKJ	2021	71	88	92	PLN UDIKLAT PANDAAN	Teknologi
682	YUWANDI RULI BAYHAQQI	TKJ	2021	87	85	93	TELKOM UNIVERSITY BANDUNG	Sistem Informasi

Penentuan data alumni pada tabel 3 tentang data alumni didapat dari data alumni tracet dan digabungkan dengan data nilai dari alumni SMK Telkom Malang.

III. HASIL PENELITIAN DAN PEMBAHASAN

Berdasarkan data alumni dan parameter yang sudah ditentukan, maka muncul sebuah dataset yang disajikan pada Tabel 4 tentang dataset alumni. Pada Tabel 4 tentang dataset alumni, semua parameter yang sudah ditentukan akan disesuaikan.

TABEL 4
DATASET ALUMNI

NO	Nama	Jurusan	Tahun Lulus	UKK	KOM JAR DAS	PROG DAS	Industri	Bidang
1	ACHMAD AZHAR ALDIANSYAH	RPL 1	2023	B	C	B	PT TELKOM AKSES KEDIRI	Teknologi
2	ADINDA KHARISMA PUTRI TANDAWIJAYA	RPL 1	2023	C	B	B	CV SINARA SUKSES BERSAMA	Teknologi
3	AIZAR RAHIMA SUPRAYITNO	RPL 1	2023	A	A	A	BINA NUSANTARA UNIVERSITY	Sistem Informasi
4	AMELIA INDRIANA	RPL 1	2023	B	B	B	PT FINNET INDONESIA	Teknologi
5	ARDHIAN CALWA NUGRAHA	RPL 1	2023	A	B	B	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATIKA
...
680	SYFANADYA WENING ADI	TKJ 3	2021	B	B	A	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATIKA
681	YUDISTIRA ANDRA FIRMANSYAH	TKJ 3	2021	C	B	A	PLN UDIKLAT PANDAAN	Teknologi
682	YUWANDI RULI BAYHAQQI	TKJ 3	2021	B	B	A	TELKOM UNIVERSITY BANDUNG	Sistem Informasi

Pengujian terhadap metode klasifikasi *Naive Bayes* dilakukan dengan cara teknik *split validation* dan menghasilkan *confusion matrix*, dimana teknik *split validation* adalah dataset yang digunakan dibagi kedalam dua bagian yakni 75% (511 baris) dari dataset akan dijadikan sebagai *data training* dan 25% (171 baris) dari dataset dijadikan sebagai *data testing*. *Confusion matrix* adalah performa dari sebuah algoritma yang dihasilkan dari proses klasifikasi dan dikomparasi dengan data aktual. Berikut data training yang ditampilkan pada Tabel 5 tentang *data training*.

TABEL 5
DATA TRAINING

NO	Nama	Jurusan	Tahun Lulus	UKK	KOM JAR DAS	PROG DAS	Industri	Bidang	Aktual
1	ACHMAD AZHAR ALDIANSYAH	RPL 1	2023	B	C	B	PT TELKOM AKSES KEDIRI	Teknologi	Linier
2	ADINDA KHARISMA PUTRI TANDAWIJAYA	RPL 1	2023	C	B	B	CV SINARA SUKSES BERSAMA	Teknologi	Linier
3	AIZAR RAHIMA SUPRAYITNO	RPL 1	2023	A	A	A	BINA NUSANTARA UNIVERSITY	Sistem Informasi	Linier

4	AMELIA INDRIANA	RPL 1	2023	B	B	B	PT FINNET INDONESIA	Teknologi	Linier
5	ARDHIAN CALWA NUGRAHA	RPL 1	2023	A	B	B	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATI KA	Linier
...
509	MUHAMMAD FACHRI AFIF	RPL 2	2021	C	B	B	SEVEN INC JOGJA (DOUBLE)	Teknologi	Linier
510	MUHAMMAD RAFI PUTRA SATRIA	RPL 2	2021	B	B	B	PT JENDELA DIGITAL INDONESIA (INDONESIA IT)	Teknologi	Linier
511	MUHAMMAD ZIDANE IQBAL	RPL 2	2021	A	B	B	PT. Digital Netwerk Venture Indonesia	Teknologi	Linier

Selanjutnya adalah membuat sebuah data testing, dimana data testing didapat dari 25% dari dataset alumni. Data testing berisi 171 baris data yang ditampilkan pada Tabel 6 *data testing*.

TABEL 6
DATA TESTING

NO	Nama	Jurusan	Tahun Lulus	UKK	KOM JAR DAS	PROG DAS	Industri	Bidang	Aktual
1	Mukhammad Haidar	Nuril RPL 2	2021	C	B	A	Politeknik Malang	Negeri TEKNIK INFORMATI KA	Linier
2	NABILA JENNANDA RIDWAN	RPL 2	2021	C	B	C	PDAM NGAWI	Teknologi	Linier
3	NAFARRAS AYU PATRYCHIA	RPL 2	2021	C	B	B	BINA NUSANTARA UNIVERSITY	BUSINESS MANAGEM ENT	Tidak Linier
4	NISRINA IZDIHAR ARDANA PUTRI	RPL 2	2021	C	A	B	BINA NUSANTARA UNIVERSITY	Sistem Informasi	Linier
5	RAFI RAIHAN FIRDAUS	RPL 2	2021	C	B	B	TELKOM UNIVERSITY BANDUNG	Teknik Telekomunik asi	Linier
...
169	SYFANADYA WENING ADI	TKJ 3	2021	B	B	A	TELKOM UNIVERSITY BANDUNG	TEKNIK INFORMATI KA	Linier
170	YUDISTIRA ANDRA FIRMANSYAH	TKJ 3	2021	C	B	A	PLN UDIKLAT PANDAAN	Teknologi	Linier
171	YUWANDI RULI BAYHAQQI	TKJ 3	2021	B	B	A	TELKOM UNIVERSITY BANDUNG	Sistem Informasi	Linier

Berdasarkan tabel training yang sudah ditampilkan pada tabel 5 tentang *data training*, maka untuk perhitungan *prior* adalah sebagai berikut :

$$P_{(Linier)} = \frac{623}{682} = 0,91$$

$$P_{(TidakLinier)} = \frac{59}{682} = 0,086$$

- Nilai *Evidence* adalah sebagai berikut :

$$P_{(rpl)} = \frac{354}{682} = 0,51$$

$$P_{(tkj)} = \frac{328}{682} = 0,48$$

$$P_{(2021)} = \frac{233}{682} = 0,341$$

$$P_{(2022)} = \frac{232}{682} = 0,340$$

$$P_{(2023)} = \frac{217}{682} = 0,318$$

- Nilai *Likelihood* adalah sebagai berikut :

$$P_{(rpl|Linier)} = \frac{317}{682} = 0,46$$

$$P_{(rpl|TidakLinier)} = \frac{37}{682} = 0,054$$

$$P_{(tkj|Linier)} = \frac{306}{682} = 0,44$$

$$P_{(tkj|TidakLinier)} = \frac{22}{682} = 0,032$$

$$P_{(2021|Linier)} = \frac{208}{682} = 0,30$$

$$P_{(2021|TidakLinier)} = \frac{25}{682} = 0,036$$

$$P_{(2022|Linier)} = \frac{222}{682} = 0,32$$

$$P_{(2022|TidakLinier)} = \frac{10}{682} = 0,014$$

$$P_{(2023|Linier)} = \frac{193}{682} = 0,28$$

$$P_{(2023|TidakLinier)} = \frac{24}{682} = 0,035$$

- Jadi hasil nilai *Posterior* adalah sebagai berikut :

$$Posterior = \frac{Likelihood \cdot Prior}{Evidance}$$

$$P(linier|rpl) = \frac{P(rpl|linier) * P(linier)}{P(rpl)} = \frac{0,46 * 0,91}{0,51} = \frac{0,41}{0,51} = 0,80$$

$$P(linier|tkj) = \frac{P(tkj|linier) * P(linier)}{P(tkj)} = \frac{0,44 * 0,91}{0,48} = \frac{0,40}{0,48} = 0,83$$

Dari perhitungan tersebut, untuk parameter jurusan RPL mendapat nilai 0,80 dan jurusan TKJ mendapat nilai 0,83. Dapat disimpulkan bahwa untuk jurusan TKJ lebih banyak alumni yang terserap secara linier daripada jurusan RPL.

$$P(linier|2021) = \frac{P(2021|linier) * P(linier)}{P(2021)} = \frac{0,30 * 0,91}{0,341} = \frac{0,273}{0,341} = 0,80$$

$$P(linier|2022) = \frac{P(2022|linier) * P(linier)}{P(2022)} = \frac{0,32 * 0,91}{0,340} = \frac{0,291}{0,340} = 0,85$$

$$P(linier|2023) = \frac{P(2023|linier) * P(linier)}{P(2023)} = \frac{0,28 * 0,91}{0,318} = \frac{0,254}{0,318} = 0,80$$

Dari perhitungan tersebut, untuk parameter tahun lulus 2022 mendapat nilai 0,85 dan tahun lulus 2021 dan 2022 mendapat nilai 0,80. Dapat disimpulkan bahwa untuk tahun lulus 2022 lebih banyak alumni yang terserap secara linier daripada lulusan tahun 2021 dan lulusan tahun 2023.

Penelitian yang ditulis oleh Amalia [5] sebelumnya, menggunakan atribut nilai Bahasa Indonesia, Matematika, IPA dan Jumlah Nilai dengan jumlah data sebanyak 100 baris data. Pada penelitian ini menggunakan atribut jurusan, tahun lulus, nilai UKK, nilai KOMJARDAS, nilai PROGDAS dan bidang keahlian dengan jumlah data sebanyak 682 baris data. Selain dari jumlah data dan atribut, pembaharuan yang dilakukan pada penelitian ini adalah dengan menggunakan bahasa *Python* yang sebelumnya dengan menggunakan *Rapid Miner*.

IV. KESIMPULAN

Berdasarkan dari hasil pembahasan, dapat disimpulkan bahwa dari data alumni sebanyak 682 alumni yang dijadikan dataset dengan parameter jurusan, tahun lulus, nilai Uji Kompetensi Kejuruan (UKK), nilai Komputer Jaringan Dasar (KOMJARDAS), nilai Pemrograman Dasar (PROGDAS), dan bidang. Data alumni menjadi data primer yang akan digunakan untuk mengklasifikasikan serapan lulusan sekolah. Dari data primer tersebut dibuatlah dataset yang akan digunakan untuk proses uji coba metode dengan menyesuaikan parameter yang sudah dibuat. Dataset dibagi menjadi 2 bagian yaitu 75% dari *dataset* sebanyak 511 baris menjadi *data training* dan 25% sebanyak 171 baris menjadi *data testing*. Setelah dilakukan pengujian dengan metode *Naive Bayes* menghasilkan klasifikasi serapan lulusan sebesar 623 alumni terserap sesuai dengan jurusannya (linier) dan 59 siswa terserap

tidak sesuai dengan jurusannya (tidak linier). Hasil dari pengujian tersebut adalah untuk jurusan RPL mendapat nilai 0,80 sedangkan untuk jurusan TKJ mendapat nilai 0,83. Berdasarkan dari tahun lulus, alumni lulusan tahun 2021 diklasifikasikan dengan nilai 0,80, untuk tahun lulus 2022 dengan nilai 0,85 dan untuk tahun lulus 2023 mendapat nilai 0,80.

Berdasarkan proses klasifikasi dengan metode *Naive Bayes* didapatkan kesimpulan bahwa lulusan SMK Telkom Malang dengan jumlah 682 alumni, dimana 623 alumni terserap linier sesuai dengan bidangnya dan 59 alumni yang tidak terserap secara linier. Dengan rincian pada tahun lulus 2022, alumni banyak terserap sesuai dengan bidangnya daripada tahun 2021 dan 2023 serta jurusan TKJ yang paling banyak terserap sesuai bidangnya.

DAFTAR PUSTAKA

- [1] Said, A. (2023). Berita Resmi Statistik. Retrieved from <https://webapi.bps.go.id/download.php?f=2yo9/GXVDpdcQ2YjOsugqZU8/czFtlw2hegol0umTtlh727beuilKARfDmzFhje-qlND6OJ9qMAd+6whUUIgk+XSyQJ3Bv6aLsWY+8kMxHRzWix6SHOJ4uLZKfFm8ux7tEkxDLORByv6bjy+sqxd+zueNIWXJ+iFKreuvTbG7HyVd7j/4uP5sVTpgtOe-oJmplpU0GCZpGFf4UcTG26GJy0pTUYv8LheX016CSqz791KJ/8W5hUPe54vflVnm6rk+9EgsGSVnCVJ59peSNuEjX2UnDr8rW/2LIFpzJG0gDUI=>
- [2] CNN Indonesia. (2023, May 5). Pengangguran di RI Terbanyak Lulusan SMK. Retrieved 02/29/2024 from <https://www.cnnindonesia.com/ekonomi/20230505130917-92-945695/pengangguran-di-ri-terbanyak-lulusan-smk>
- [3] Mukhlason, A., Winanti, T., & Yundra, E. (2020). Analisa Indikator SMK Penyumbang Pengangguran di Provinsi Jawa Timur.
- [4] Suhandi, Wiguna, W., & Quraysin, I. (2021). Dinamika Permasalahan Ketenagakerjaan dan Pengangguran di Indonesia. <https://doi.org/10.46306/vls.v1i1>
- [5] Amalia, R. (2020). Penerapan Data Mining untuk Memprediksi Hasil Kelulusan Siswa Menggunakan Metode *Naive Bayes*. JUI SI, 06(01).
- [6] Wiwik Sri Rahayu Ginantra, N. L., Nur Arifah, F., Hadi Wijaya, A., Septarini, R. S., Ahmad Dewa Putu Yudhi Ardiana, N., Effendy, F., ... Surya Negara, E. (2021). Data Mining dan Penerapan Algoritma.
- [7] Indra Purnama, D., Lathifah Islami, R., Sari, L., & Robinson Sihombing, P. (2021). Analisis Klasifikasi Data Tracer Study Dengan Support Vector Machine Dan Neural Network. Jurnal Sistem Komputer Dan Kecerdasan Buatan.
- [8] Kantardzic, M. (2020). DATA MINING Concepts, Models, Methods, and Algorithms Third Edition.
- [9] Kotu, V., & Deshpande, B. (2015). Predictive analytics and data mining : concepts and practice with RapidMiner.
- [10] Hutapea, E., & Enggar Harususilo, Y. (2019, August 29). SMK Bisa, Siswa SMK Indonesia Raih 15 Medali pada WSC 2019 Rusia. Retrieved 02/29/2024 from <https://edukasi.kompas.com/read/2019/08/29/21175261/smk-bisa-siswa-smk-indonesia-raih-15-medali-pada-wsc-2019-rusia>
- [11] Kementerian Pendidikan dan Kebudayaan. (2018, February 1). Terobosan Model Pembelajaran di SMK. Retrieved 02/29/2024 from <https://www.kemdikbud.go.id/main/blog/2018/02/terobosan-model-pembelajaran-di-smk#:~:text=Tujuan%20pendidikan%20di%20SMK%20adalah,percepatan%20dan%20peningkatan%20kompetensi%20siswa.>
- [12] IDN Times. (2023, October 20). Potret Karya Siswa SMK di Penjuru Negeri. Retrieved 02/29/2024 from <https://bali.idntimes.com/life/education/irma/potret-karya-siswa-smk-di-penjuru-negeri?page=all>
- [13] Lukman, O. : (2018). Kinerja Bursa Kerja Khusus (BKK) SMK Dalam Penyiapan Lulusannya Memasuki Dunia Kerja. Prodi Pendidikan Teknik Elektro (Vol. 8). Retrieved from <http://journal.student.uny.ac.id/ojs/index.php/elektro>
- [14] MALANG POSCO MEDIA. (2023, October 31). Gandeng 35 Perusahaan, SMK PGRI 3 Malang Gelar Job Fair 2023; Fasilitas Lulusan Masuk Dunia Kerja. Retrieved 02/29/2024 from <https://malangposcomedia.id/gandeng-35-perusahaan-smk-pgri-3-malang-gelar-job-fair-2023-fasilitas-lulusan-masuk-dunia-kerja/>
- [15] Menteri Ketenagakerjaan Republik Indonesia. Peraturan Menteri Ketenagakerjaan Republik Indonesia nomor 39 Tahun 2016 Tentang Penempatan Tenaga Kerja (2016).
- [16] Janu, S., Tyas, S., Febianah, M., Solikhah, F., Kamil, A. L., & Arifin, W. A. (2021). Analisis Perbandingan Algoritma *Naive Bayes* dan C.45 dalam Klasifikasi Data Mining untuk Memprediksi Kelulusan, 8(1).
- [17] Prasajo, M., Sriyanto, & Triwidiyanti, J. (2021). Prediksi Prestasi Siswa SMK Masuk Pasar Kerja Menggunakan Teknik Data Mining (Studi Kasus SMKN 1 Kota Agung Timur Tanggamus, Lampung).
- [18] Presiden Republik Indonesia. Undang-Undang Republik Indonesia nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional (2003). Available at: <https://peraturan.bpk.go.id/Details/43920/uu-no-20-tahun-2003>
- [19] Ramadhan, L. A. (2022, June 11). Siswa SMK Telkom Malang Disiapkan Jadi Pekerja Digital Profesional. Beritajatim.Com. Retrieved 02/20/2024 from <https://beritajatim.com/siswa-smk-telkom-malang-disiapkan-jadi-pekerja-digital-profesional>
- [20] Rizmayanti, A. I., Hidayati, N., Nugraha, F. S., Gata, W., & Mandiri, S. N. (2021). Penerapan Data Mining Untuk Memprediksi Kompetensi Siswa Menggunakan Metode Decision Tree (Studi Kasus SMK Multicom Depok). JURNAL SWABUMI, 9(1), 2021.
- [21] Farid Rifai, M., Jatnika, H., & Valentino, B. (2019). Penerapan Algoritma *Naive Bayes* Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS), 12(2).
- [22] Shihab, M. Q. (2002). Tafsir Al-Mishbah : pesan, kesan dan keserasian Al-Qur'an / M. Quraish Shihab (Vol. 15).
- [23] Sinaga, K., & Handoko, K. (2021). Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa Dengan Metode *Naive Bayes*. JURNAL COMASIE (Vol. 04).
- [24] Fortino, A. (2023). Data Mining and Predictive Analytics for Business Decisions.
- [25] Widyawati, S. (2023, October 4). SMKN 11 Kota Malang Gelar Job Fair, Diikuti 22 Industri Kreatif & Umum, Serap Lulusan SMK Siap Kerja. Retrieved 02/29/2024 from <https://jatim.tribunnews.com/2023/10/04/smkn-11-kota-malang-gelar-job-fair-diikuti-22-industri-kreatif-umum-serap-lulusan-smk-siap-kerja>
- [26] Wydiastuty Kusuma, L. (2019). Prediksi Kemampuan Lulusan SMK untuk Dapat Bersaing Di Dunia Kerja dengan Menggunakan *Naive Bayes*: Studi Kasus SMK Buddhi Tangerang. JURNAL ALGOR (Vol. 1). Retrieved from <https://jurnal.buddhidharma.ac.id/index.php/algord/index>