



Penerapan Metode Naive Bayes Dalam Prediksi Penyebab Kecelakaan Kerja CV. Deka Utama

Monica Putri Rahayu¹, Yusti Farlina²

¹Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
Jl. Cemerlang No 8 Sukakarya Kecamatan Warudoyong 0266-6251992

²Sistem Informasi Akuntansi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
Jl. Cemerlang No 8 Sukakarya Kecamatan Warudoyong 0266-6251992

monicaputri0792@gmail.com

Abstrak-- Perkembangan teknologi yang sangat pesat dan teknologi merupakan suatu kebutuhan bagi penggunanya. Sejahtera ini teknologi telah banyak diterapkan pada beberapa bidang, salah satu contoh adalah kehadiran teknologi di sector konstruksi yang memudahkan para penggunanya untuk melakukan proses prediksi terhadap faktor penyebab kecelakaan kerja konstruksi. Kecelakaan dalam bekerja merupakan suatu kejadian yang tidak diharapkan dan tidak direncanakan. Kecelakaan kerja dapat menghambat jalannya pekerjaan dan akan mempengaruhi hasil dan lamanya waktu pekerjaan itu sendiri. Salah satu cara agar dapat meminimalisir terjadinya kecelakaan kerja adalah dengan memprediksi faktor apa saja yang dapat menyebabkan kecelakaan kerja. *Data mining* merupakan salah satu cara untuk mendapatkan informasi yang tersimpan pada database yang berjumlah besar. Data kecelakaan kerja yang terdapat pada suatu perusahaan konstruksi hanya digunakan sebagai laporan perusahaan saja. Dalam kenyataannya, data tersebut dapat memberikan informasi yang lebih dari sekedar laporan. Salah satu informasi yang dapat diambil dari data kecelakaan kerja perusahaan adalah informasi tentang prediksi penyebab kecelakaan kerja. Metode untuk memprediksi yang menghasilkan data akurat adalah metode *Naive Bayes*. Dalam penelitian ini menggunakan 78 data sampel yang diolah sehingga menghasilkan nilai akurasi sebesar 96,15%.

Kata kunci: *Naive Bayes, Kecelakaan Kerja, Akurasi*

Abstract - Progress of technology and technology is a necessity for its users. The technology has been widely applied in several fields, one example is the presence of technology in the construction sector that makes it easy for its users to make predictions on the factors that cause construction work accidents. An accident at work is an unexpected and unplanned event. Work accidents can hamper the work and will affect the results and the length of time the work itself. One way to minimize workplace accident's to predict what factors can cause work accidents. Data mining is one way to get information stored in a large number of databases. Work accident data contained in a construction company is only used as a company report. In reality, the data can provide information that is more than just a report. One of the information that can be taken from the company's occupational accident data is information about the prediction of the causes of work accidents. The method for predicting that produces accurate data is the Naive Bayes method. In this research, 78 sample data were processed to produce an accuracy value of 96.15%.

Keywords: *Naive Bayes, Work Accident, Accuracy*

I. PENDAHULUAN

Pendahuluan Jasa konstruksi adalah salah satu sektor industri yang memiliki resiko kecelakaan kerja yang cukup tinggi dibanding sektor industri lainnya [1]. Kecelakaan kerja adalah kejadian yang tak diharapkan dan tidak terduga. Tidak diharapkan karena kecelakaan adalah peristiwa yang menyebabkan penderitaan mulai yang paling ringan, sedang hingga sampai yang paling berat dan disertai kerugian material. Tidak terduga, karena peristiwa itu tidak dalam bentuk perencanaan atau tidak mengandung unsur kesengajaan [2]. Secara global menurut ILO (*International Labour Organization*) atau Organisasi Perburuhan Internasional terdapat lebih dari 250 juta kasus

kecelakaan ditempat kejadian sekitar 1,2 juta pekerja yang meninggal dan sakit akibat kecelakaan ditempat kerja. Sedangkan di Indonesia sendiri sepanjang tahun 2019 BPJS Ketenagakerjaan melaporkan terdapat 77.925 kasus kecelakaan kerja [3].

CV. Deka Utama adalah salah satu perusahaan konstruksi pembangunan bersekala kecil dan menengah yang sering menangani proyek pembangunan dari pemerintah Kota Sukabumi. Salah satu proyek yang pernah dijalankan oleh CV. Deka Utama adalah pembuatan dan pemasangan besi separator sepanjang jalan R.E Martadinata Kota Sukabumi yang tidak lepas dari resiko

* Korepondensi.

Alamat E-mail : jurnal.larik@bsi.ac.id.

Diterima 30 July 2021; Direvisi 20 January 2021; Diterima 21 January 2021

© 2021 Jurnal Larik.

kecelakaan pada pekerjaanya. Berdasarkan data yang diperoleh dari pihak CV. Deka Utama terdapat beberapa kasus kecelakaan kerja baik yang ringan hingga sedang selama pelaksanaan proyek berlangsung dalam kurun waktu 3 bulan. Mulai dari pekerja yang terkena serpihan las besi, terluka karena goresan benda tajam, tertancap paku yang disebabkan karena peralatan kerja yang tidak digunakan dengan baik hingga peralatan yang disimpan sembarangan.

Data mining adalah sebuah proses yang menggunakan teknik matematika, statistik, *machine learning* dan kecerdasan buatan untuk mengidentifikasi dan menguraikan informasi dan pengetahuan yang terikat dari berbagai database [4]. Data mining sering pula disebut sebagai Knowledge Discovery in Database (KDD) yaitu kegiatan pengumpulan, pemakaian data, historis menemukan keteraturan, dan hubungan atau pola dalam suatu data set yang berukuran besar [5]. Data mining terbagi menjadi dua kategori utama yaitu deskriptif dan prediktif. Deskriptif bertujuan untuk menurunkan atau mengekstraksi pola-pola (trend, korelasi, teritori, anomali, dan cluster) yang meringkas hubungan pokok dalam data. Sedangkan prediktif bertujuan untuk memprediksi atribut tertentu dan nilai berdasarkan nilai dari atribut-atribut lain. Data mining memiliki beberapa macam metode untuk klasifikasi, asosiasi, dan juga prediksi. Metode yang dapat digunakan untuk prediksi diantaranya adalah metode C 4.5 dan metode Naive Bayes.

Metode C 4.5 atau sering disebut pula dengan pohon keputusan merupakan pengembangan dari algoritma ID3 [6] Algoritma C4.5 memiliki kelebihan seperti fleksibel, mudah dipahami, dan dapat divisualisasikan dalam bentuk gambar atau pohon keputusan. Namun metode C4.5 ini memiliki beberapa kekurangan, yaitu sulit untuk mendesain pohon keputusan yang optimal, kualitas dari hasil keputusan yang didapat sangat tergantung dari bagaimana pohon keputusan tersebut didesain dan kerap terjadi overlap.

Metode Naive Bayes adalah sebuah teknik yang berbasis probabilitic sederhana yang berdasar pada penerapan aturan Bayes atau Teorema Bayes dengan asumsi ketidak ketergantungan atau indenpendensi yang kuat [7]. Algoritma Naive Bayes memiliki kelebihan seperti dapat dipakai untuk data kuantitatif dan kualitatif, proses penghitungannya cepat dan efisien, mudah dipahami dan mudah untuk dibuat. Selain kelebihan, metode naive bayes juga memiliki beberapa kekurangan yaitu jika probabilitas kondisionalnya bernilai nol maka probabilitas prediksinya juga akan bernilai sama yaitu nol, dan metode naive bayes ini hanya dirancang untuk mendeteksi kata kata saja dan tidak dapat berupa gambar.

Dalam penelitian ini penulis menggunakan metode naive bayes dalam memprediksi faktor-faktor penyebab kecelakaan kerja untuk mengetahui apa saja

yang dapat menyebabkan kecelakaan kerja konstruksi di CV. Deka Utama terus terjadi. Usaha dalam menemukan faktor-faktor penyebab kecelakaan konstruksi perlu adanya pemahaman yang lengkap tentang mengapa dan bagaimana kecelakaan kerja sering terjadi. Karena masih tingginya angka kecelakaan kerja khususnya pada CV atau perusahaan konstruksi kecil. Tentunya hasil dari penelitian ini dapat digunakan sebagai panduan untuk mengurangi resiko terjadinya kecelakaan kerja dibidang konstruksi sehingga dapat meningkatkan kualitas dari suatu perusahaan konstruksi dan dapat mengurangi kerugian waktu, materil dan jiwa. Tujuan dari penelitian ini untuk mengetahui apa saja faktor yang menyebabkan kecelakaan kerja konstruksi selalu terjadi. Menerapkan bagaimana metode naive bayes dalam memprediksi faktor-faktor penyebab kecelakaan kerja konstruksi pada CV. Deka Utama. Membantu pemilik perusahaan konstruksi berskala kecil untuk lebih memahami dan mengetahui tentang standar keamanan dan keselamatan dan keamanan proyek.

II. METODE PENELITIAN

2.1. Tahapan Penelitian

Tahapan penelitian dalam skripsi ini mengenai prediksi faktor penyebab kecelakaan CV. Deka Utama dapat digambarkan sebagai berikut



Gambar 1. Tahapan Penelitian

Berikut merupakan penjelasan mengenai tahapan dari penelitian yaitu:

1. Studi Literatur
Melakukan pengumpulan bahan informasi dan literatur yang berkaitan dengan judul penelitian.
2. Pengumpulan Data
Dalam pengumpulan data pada penelitian ini menggunakan data sekunder yang diperoleh dari pihak CV. Deka Utama.
3. Pengolahan Data Dengan *Microsoft Excel*
Mengolah data sekunder dengan memasukkan data kedalam *Microsoft Excel* dan melakukan perhitungan.
4. Mengolah Data Menggunakan Aplikasi *Rapid Miner*
Data yang telah diolah kedalam *Microsoft Excel* kemudian dimasukkan pada aplikasi *Rapid Miner* untuk menguji data testing.
5. Mengolah Data Kedalam Metode *Naive Bayes*
Pada tahap ini, data yang telah dimasukkan kedalam aplikasi *Rapid Miner* kemudian diolah menggunakan metode *Naive Bayes* dimana hasil dari pengolahan data tersebut akan menghasilkan validasi data dan akurasi yang sesuai dengan data yang telah diuji dalam aplikasi *Rapid Miner*.
6. Menentukan Hasil Akurasi Data
Data yang telah diolah menggunakan aplikasi *Rapid Miner* dan metode *Naive Bayes* akan menghasilkan akurasi data hasil dari kriteria faktor penyebab kecelakaan kerja pada CV. Deka Utama.
7. Kesimpulan
Penarikan kesimpulan diambil berdasarkan analisa data dimana data yang diolah sudah sesuai dengan tujuan penelitian.

2.2. Instrumen Penelitian

Instrumen penelitian merupakan suatu alat yang digunakan untuk mendukung fenomena alam maupun sosial. Instrumen penelitian berperan penting dalam menentukan kualitas suatu penelitian.

Beberapa hal penting yang menjadi instrumen pada penelitian ini adalah sebagai berikut :

1. Penelitian ini menggunakan data riset, data yang akan dimasukkan dan diolah yaitu data yang akan menjadi sampel digunakan sebagai instrumen untuk memperoleh data dalam proses klasifikasi prediksi dalam faktor penyebab kecelakaan kerja.

Percobaan yang dilakukan dalam memprediksi faktor penyebab kecelakaan kerja dalam penelitian ini dengan menghitung probabilitas prior serta probabilitas posterior dengan menggunakan data sebanyak 78 record data. Hasil probabilitas prior seperti pada tabel 1 terdapat dua kelas yaitu Kecelakaan dan Tidak Terjadi Kecelakaan.

Tabel 1. Tabel Probabilitas Prior

Atribut		Jumlah Kasus J	Terjadi Kecelakaan(S1)	Tidak (S2)	P(X) Ci	
					Terjadi Kecelakaan	Tidak
Total		39	19	20		

2. Data yang sudah tersedia dalam bentuk *Microsoft Excel* dengan variabel sebanyak 78 data yang diambil dari data kecelakaan kerja CV. Deka Utama selama tahun 2018-2019.
3. Perangkat lunak yang digunakan dalam pengolahan data adalah *Microsoft Excel* dan *Rapid Miner*.
4. Algoritma yang digunakan dalam penelitian ini menggunakan metode *Naive Bayes*.

2.3. Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh pekerja bangunan CV. Deka Utama di dengan data angka kecelakaan kerja selama tahun 2018 hingga 2019 sebanyak 78 kasus. Jumlah tersebut sekaligus menjadi sampel dalam penelitian ini sehingga disebut penelitian survey sensus atau penelitian jenuh.

2.4. Metode Analisis Data

Metode analisis data dalam penelitian ini menggunakan klasifikasi *Data Mining* dengan metode *Naive Bayes*. Untuk menentukan klasifikasi yang digunakan pada suatu masalah diperlukan cara sistematis untuk mengevaluasi bagaimana metode yang bekerja dan membandingkan dengan yang lain

III. HASIL DAN PEMBAHASAN

3.1. Hasil Penelitian

Hasil penelitian yang telah dilakukan dalam memprediksi faktor penyebab kecelakaan kerja pada CV. Deka Utama dengan menggunakan metode *Naive Bayes*. Dari penelitian ini diperoleh sampel dengan jumlah data sebanyak 78 data dari jumlah keseluruhan laporan kecelakaan kerja sepanjang tahun 2018 hingga 2019 pada CV. Deka Utama

3.2. Perhitungan Manual

3.2.1. Menghitung Probabilitas Prior

Rentang Usia						
	< 40 tahun	21	13	8	1,461538	2,5
	> 40 tahun	18	6	12	3,166667	1,666667
Alat Penunjang Keselamatan						
	Ya	14	2	12	9,5	1,666667
	Tidak	25	17	8	1,117647	2,5
Personal Adjustment						
	Menguasai	23	7	16	2,714286	1,25
	Tidak	16	12	4	1,583333	5
Kondisi Mesin						
	Baik	26	9	17	2,111111	1,176471
	Tidak	13	10	3	1,9	6,666667
Disiplin						
	Ya	22	3	19	6,333333	1,052632
	Tidak	17	16	1	1,1875	20

3.2.2. Menghitung Probabilitas Posterior

Setelah selesai perhitungan probabilitas prior, langkah selanjutnya adalah menghitung probabilitas posterior terhadap kasus baru.

Tabel 2. Tabel Probabilitas Posterior

Data X		P(X Ci)	
Atribut	Nilai	Terjadi	Tidak
Rentang Usia	> 40 tahun	2,642857	1,863636
Alat Penunjang Keselamatan	Tidak	1,121212	2,733333
Personal Adjustment	Tidak Menguasai Semua Bidang	1,761905	5,125
Kondisi Mesin dan Alat	Baik	2,176471	1,171429
Disiplin	Tidak	1,15625	20,5

Setelah nilai atribut dihitung, kemudian langkah selanjutnya adalah mengkalikan semua hasil sesuai dengan data X yang dicari nilai *class* nya. Berikut adalah penjelasan perhitungannya :

$$P(X | \text{Kecelakaan} = \text{Terjadi}) = P(\text{Rentang Usia} > 40 \text{ tahun} | \text{Kecelakaan} = \text{Terjadi}) * P(\text{Alat Penunjang Keselamatan} = \text{Tidak} | \text{Kecelakaan} = \text{Terjadi}) * P(\text{Personal Adjustment} = \text{Tidak Menguasai Semua Bidang} | \text{Kecelakaan} = \text{Terjadi}) *$$

$$P(\text{Kondisi Mesin dan Alat} = \text{Baik} | \text{Kecelakaan} = \text{Terjadi}) * P(\text{Disiplin} = \text{Tidak} | \text{Kecelakaan} = \text{Terjadi}) = 2,642857 * 1,121212 * 1,761905 *$$

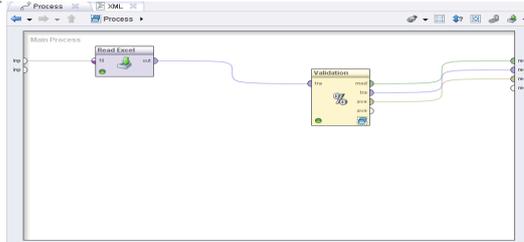
$$2,176471 * 1,15625 = 13,13858062$$

$$P(X | \text{Kecelakaan} = \text{Tidak Terjadi}) = P(\text{Rentang Usia} > 40 \text{ tahun} | \text{Kecelakaan} = \text{Tidak Terjadi}) * P(\text{Alat Penunjang Keselamatan} = \text{Tidak} | \text{Kecelakaan} = \text{Tidak Terjadi}) * P(\text{Personal Adjustment} = \text{Tidak Menguasai Semua Bidang} | \text{Kecelakaan} = \text{Tidak Terjadi}) * P(\text{Kondisi Mesin dan Alat} = \text{Baik} | \text{Kecelakaan} = \text{Tidak Terjadi}) * P(\text{Disiplin} = \text{Tidak} | \text{Kecelakaan} = \text{Tidak Terjadi}) = 1,863636 * 2,733333 * 5,125 * 1,171429 * 20,5 = \mathbf{626,9274944}$$

3.2. Validasi

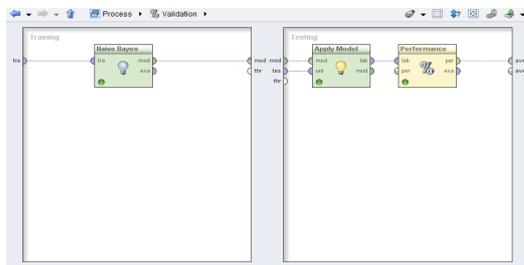
Untuk melakukan validasi data yaitu dengan menganalisis berbagai model dan memilih model dengan kinerja yang baik, setelah melakukan pembacaan file dan *import data* di *split read excel* kemudian *split read excel* dihubungkan dengan *split cross validation*.

Berikut adalah pengujian validasi menggunakan *Rapid Miner*:



Gambar 2. Main Process

Proses selanjutnya adalah pada *split process cross validation* dilakukan pengeklikan sebanyak dua kali hingga muncul tampilan dua bagian, yaitu bagian *testing* dan bagian *training*. Untuk data *training* menggunakan metode *Naive Bayes* sedangkan di *testing* menggunakan *apply model* dan *performace clasification* untuk mengetahui nilai *accuracy*, nilai *recall*, nilai *precision*, *sample distribution*, *AUC*, *AUC (optimistic)* dan *AUC (pessimistic)*. Berikut adalah proses dari validasi:



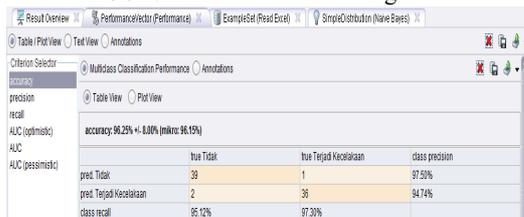
Gambar 3. Proses X-Validasi

3.3. Hasil Performance Vector

Hasil dari proses akan menghasilkan *performance vector* yang memperoleh nilai *accuracy*, *recall*, *precision*, *simple distribution*, *AUC*, *AUC (Optimistic)* dan *AUC (Pessimistic)* sebagai berikut :

1. Accuracy perhitungan Rapid Miner

Dari hasil jumlah data yang telah diklasifikasikan maka dapat diketahui hasilakurasi dari perhitungan Rapid Miner sebesar 96.15% dari hasil data testing.

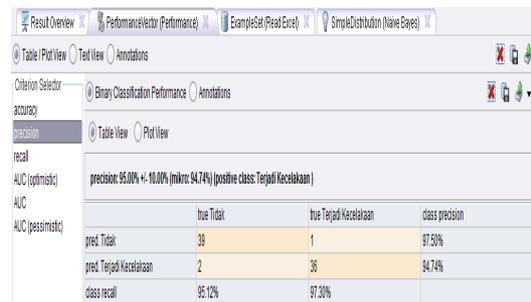


Gambar 4. Accuracy

Jumlah *True Positif (TP)* adalah 39 record dan diklasifikasikan sebagai tidak terjadi kecelakaan dan *False Positif 2 record* diklasifikasikan sebagai terjadi kecelakaan yang ternyata tidak terjadi. *False Negatif (FN)* 1 record diklasifikasikan sebagai terjadi kecelakaan tapi tidak terjadi. Kemudian *True Negative (TN)* sebanyak 36 record diklasifikasikan sebagai terjadi kecelakaan.

2. Precision perhitungan Rapid Miner

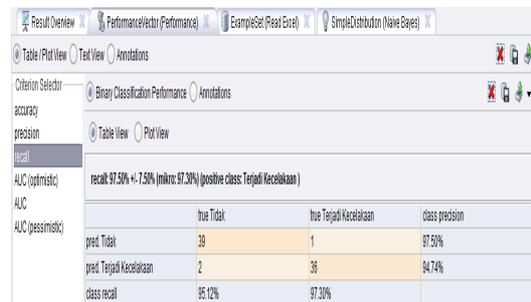
Hasil dari nilai *precision* ini adalah 97.30% untuk *class* terjadi kecelakaan 95.12% untuk tidak terjadi.



Gambar 5. Precision

3. Recall perhitungan Rapid Miner

Hasil dari perhitungan nilai *recall* pada *Rapid Miner* sebesar 95.12% untuk *class* tidak terjadi dan 97.30% untuk *class* terjadi kecelakaan.



Gambar 6. Recall

4. AUC (Area Under Curve)

Dapat diperhatikan dari gambar dibawah ini bahwa *AUC* dari algoritma *Naive Bayes* pada penelitian ini adalah 0.987%.



Gambar 7. AUC

5. AUC (Optimistic)

Pada *AUC (Optimistic)* dapat menghasilkan nilai 0.987% dengan *positif class*= terjadi kecelakaan.



Gambar 8. *AUC (Optimistic)*

6. *AUC (Pessimistic)*

AUC (Pessimistic) menghasilkan nilai 0.987% dengan *positif class* = tidak terjadi kecelakaan.



Gambar 9. *AUC (Pessimistic)*

7. Simple Distribution

Hasil dari penghitungan klasifikasi dengan metode Naive Bayes ini menghasilkan dua class yang terdapat dalam simple distribution yaitu class terjadi kecelakaan dan tidak terjadi kecelakaan. Hasil dari pengolahan data dalam penelitian ini menunjukkan nilai class terjadi kecelakaan sebesar 0.474% dan class tidak terjadi kecelakaan sebesar 0.526%.



Gambar 10. *Simple Distribution*

IV. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Dari hasil penelitian penerapan metode *Naive Bayes* pada prediksi faktor penyebab kecelakaan kerja CV. Deka Utama maka dapat diambil kesimpulan sebagaiberikut:

1. Pengolahan *Data Mining* menggunakan metode *Naive Bayes* dapat diterapkan pada bidang jasa konstruksi dan bangunan dalam memprediksi faktor penyebab kecelakaan kerja.

2. Hasil dari pengujian menggunakan metode *Naive Bayes* pada penelitian ini dapat ditarik kesimpulan bahwa metode *Naive Bayes* dapat menghasilkan hasil yang akurat dalam memprediksi faktor penyebab kecelakaan kerja.
3. Dari hasil penelitian ini dapat diketahui faktor apa saja yang sangat berpeluang dalam menyebabkan kecelakaan kerja konstruksi.
4. Hasil dari penelitian ini dapat digunakan untuk menganalisa faktor penyebab kecelakaan kerja CV. Deka Utama dan tentu dapat membantu untuk menekan angka kecelakaan kerja.

4.2. Saran

Saran-saran yang diajukan untuk penelitian selanjutnya yaitu:

1. Penelitian ini dapat dikembangkan dengan membandingkan atau menggabungkan algoritma lain untuk mendapatkan hasil yang lebih akurat.
2. Untuk pengembangan dan perbaikan model perlu dilakukan guna memperoleh hasil yang lebih akurat dalam memprediksi faktor penyebab kecelakaan kerja.

V. REFERENSI

- [1] Elisa. (2017). Pembangunan Jembatan Dengan Jasa Kontruksi. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 36-40. <https://doi.org/10.30865/komik.v3i1.1664C>
- [2] asadei D, Serra G, Tani K. Implementation of a Direct Control Algorithm for Induction Motors Based on Discrete Space Vector Modulation. *IEEE Transactions on Power Electronics*. 2007; 15(4): 769-777. (pada contoh ini Vol.15, Issues 4, and halaman 769-777)
- [3] Angraini, A. (2018). Resiko Adanya Kecelakaan Di Tempat Kerja. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 02(01), 34-44.
- [4] BPJS Ketenagakerjaa. (2019). Kasus Kecelakaan Selama tahun 2019. <https://www.bpjsketenagakerjaan.go.id/tentang-kami.html>
- [5] Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data
- [5] Susanto, H., & Sudiyatno, S. (2014). Data mining untuk memprediksi prestasi siswa berdasarkan sosial ekonomi, motivasi, kedisiplinan dan prestasi masa lalu. *Jurnal*

- Pendidikan Vokasi*, 4(2), 222–231.
<https://doi.org/10.21831/jpv.v4i2.2547>.
- [6] Rizki, M., Devrika, D., Umam, I. H., & Lubis, F. S. (2020). Aplikasi Data Mining dalam penentuan layout swalayan dengan menggunakan metode MBA. *Jurnal Teknik Industri: Jurnal Hasil Penelitian Dan Karya Ilmiah Dalam Bidang Teknik Industri*, 5(2), 130.
<https://doi.org/10.24014/jti.v5i2.8958>.
- [7] Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, 4(2), 151.
<https://doi.org/10.24076/citec.2017v4i2.106>