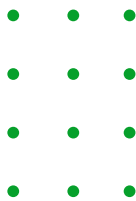


Analisis Prediktif Faktor Keselamatan pada Kecelakaan Lalu Lintas

RYAN BADAI ALAMSYAH - IBM Skillsbuild



Table of contents



01

Dataset

02

Overview

03

Proses Analisis

04

Insight

05

Kesimpulan dan
Rekomendasi

06

AI Support
Explanation



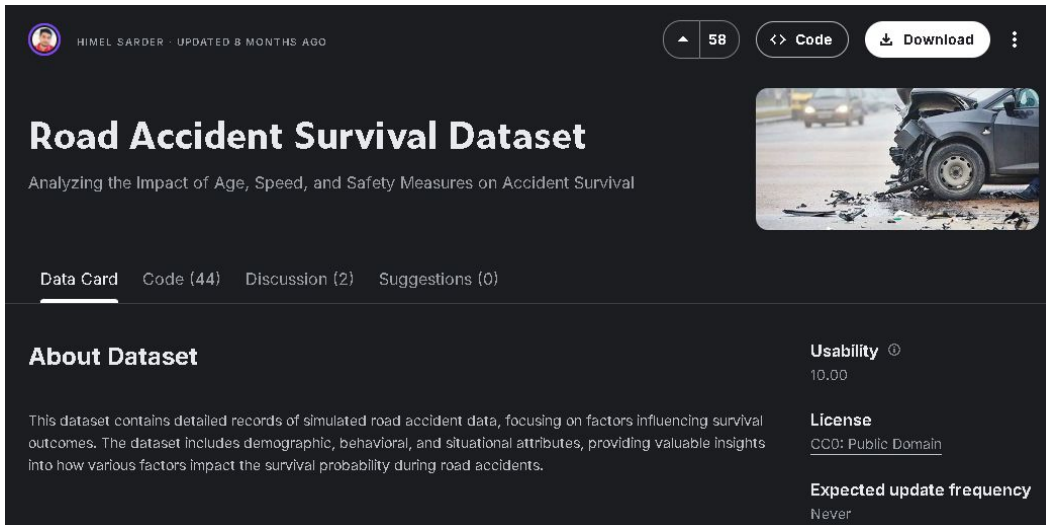
DATASET

Fitur

- Age
- Gender
- Speed_of_Impact
- Helmet_Used
- Seatbelt_Used
- Survived

Link

<https://www.kaggle.com/datasets/himelsarder/road-accident-survival-dataset/data>



The screenshot shows the Kaggle dataset page for 'Road Accident Survival Dataset' by Himel Sarder, updated 8 months ago. The page has a dark theme. At the top, there's a header with the user's profile, name, and update date. On the right, there are buttons for 'Code' (58), 'Download', and a menu icon. The main title 'Road Accident Survival Dataset' is prominently displayed, followed by a subtitle 'Analyzing the Impact of Age, Speed, and Safety Measures on Accident Survival'. Below this, there are tabs for 'Data Card' (selected), 'Code (44)', 'Discussion (2)', and 'Suggestions (0)'. The 'About Dataset' section describes the data as simulated road accident records focusing on survival factors. On the right side, there are three metrics: 'Usability' (10.00), 'License' (CC0: Public Domain), and 'Expected update frequency' (Never). A small image of a car accident is shown in the top right corner.

HIMEL SARDER · UPDATED 8 MONTHS AGO

58 Code Download

Road Accident Survival Dataset

Analyzing the Impact of Age, Speed, and Safety Measures on Accident Survival

Data Card Code (44) Discussion (2) Suggestions (0)

About Dataset

This dataset contains detailed records of simulated road accident data, focusing on factors influencing survival outcomes. The dataset includes demographic, behavioral, and situational attributes, providing valuable insights into how various factors impact the survival probability during road accidents.


Usability ⑩
10.00

License
[CC0: Public Domain](#)

Expected update frequency
Never



OVERVIEW

- Kecelakaan lalu lintas merupakan isu serius. Memahami faktor-faktor yang memengaruhi keselamatan korban dapat membantu dalam perumusan kebijakan dan edukasi publik.
 - Menganalisis dataset kecelakaan lalu lintas untuk mengidentifikasi faktor-faktor yang paling berpengaruh terhadap keselamatan korban selamat atau tidak.
 - Membangun dan membandingkan beberapa model machine learning untuk memprediksi apakah seorang korban selamat atau tidak berdasarkan data yang ada.
 - Memberikan rekomendasi berdasarkan temuan analisis.
- 

PROSES ANALISIS

- Pengumpulan dan Pembersihan Data
- Exploratory Data Analysis (EDA)
- Persiapan Model
- Pelatihan Model

INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

Berdasarkan LLM Granite, tiga faktor yang memengaruhi kelangsungan hidup pada kecelakaan lalu lintas adalah:

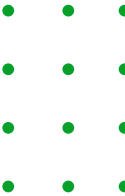
- Penggunaan Sabuk Pengaman
- Kecepatan kendaraan
- Penggunaan Helm

```
7     {df}
8     ---
9
10     Based on the data provided, Identify and briefly explain the top 3 factors
11     """
12
13 response = output.invoke(prompt)
14 print(response)
```

As a data analyst, to identify the top factors influencing road accident survival

1. **Seatbelt Usage (Seatbelt_Used):** This appears to be a strong indicator of survival.
2. **Speed of Impact (Speed_of_Impact):** The speed at which an impact occurs is a critical factor.
3. **Helmet Usage (Helmet_Used):** For those involved in accidents where head injury is a factor, helmet usage is significant.

It's important to note that while these factors appear significant based on a hypothesis test, further analysis is needed to confirm their impact. This interpretation assumes that the dataset contains a sufficient sample size.

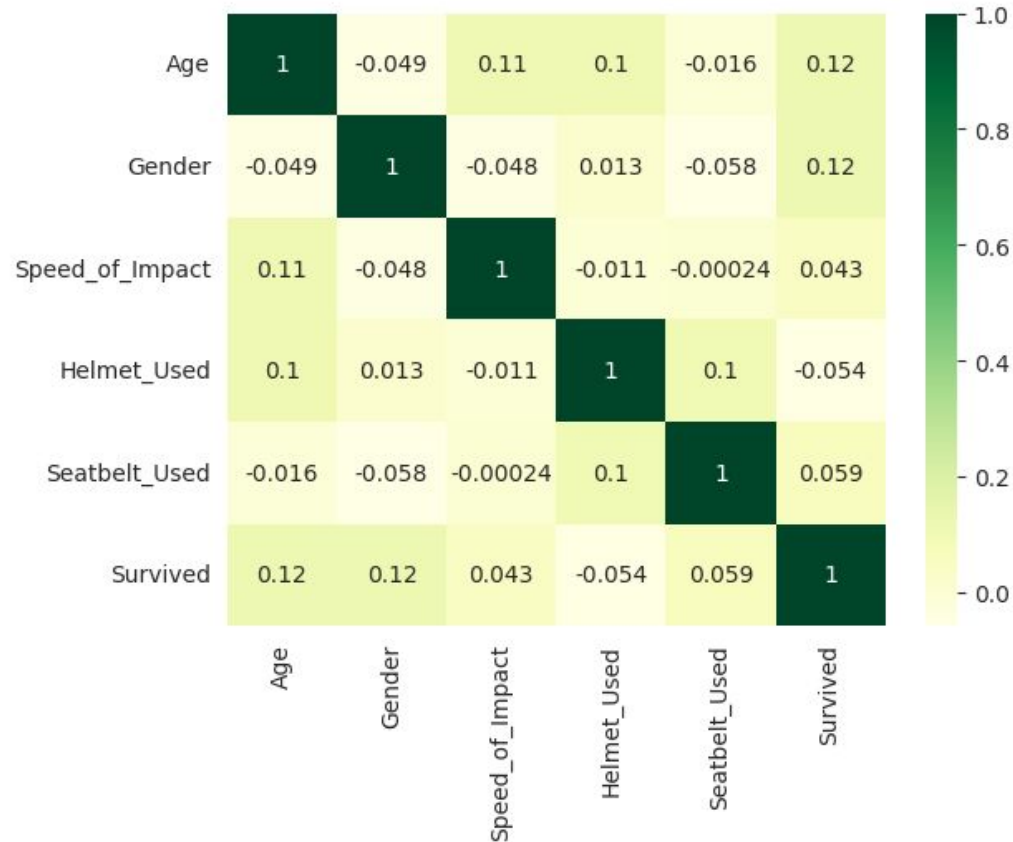


INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

Apabila dari confusion matrix di samping, diketahui faktor-faktor yang memengaruhi keselamatan korban adalah sebagai berikut:

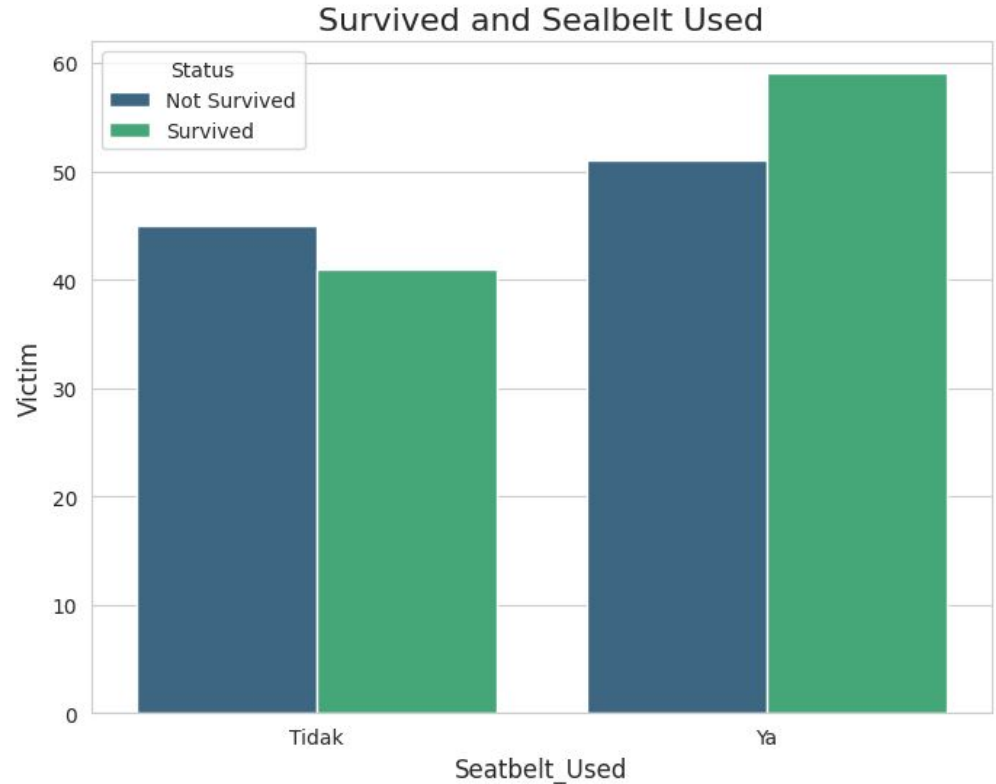
- Penggunaan Sabuk Pengaman
- Kecepatan kendaraan
- Jenis kelamin dan Umur



INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

- • Penggunaan sabuk pengaman secara konsisten mengurangi risiko fatalitas pada korban kecelakaan.

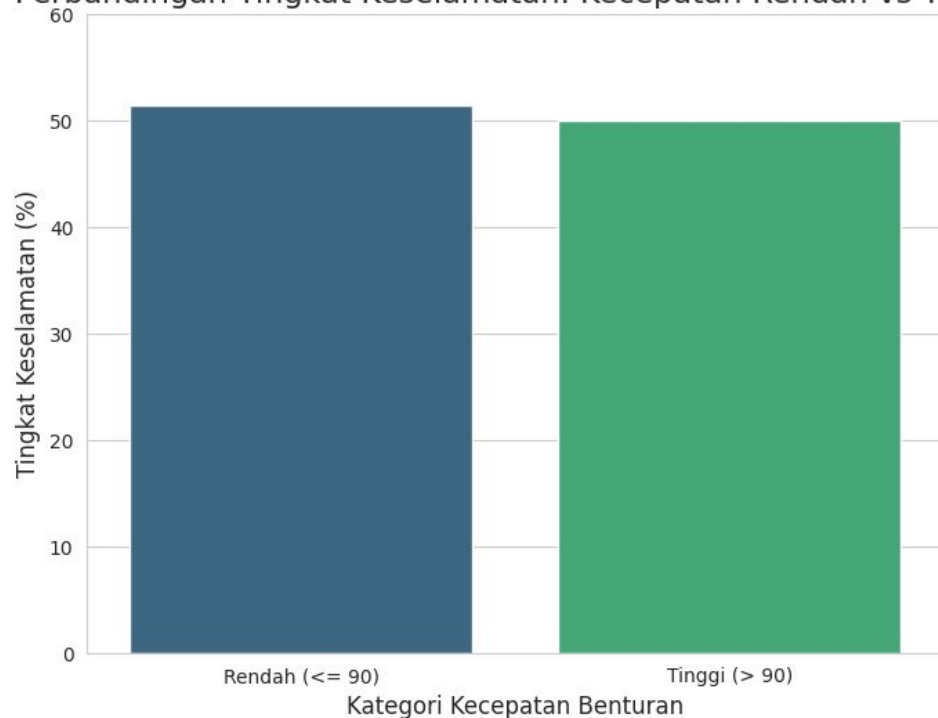


INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

- Faktor kecepatan juga salah satu faktor dalam keselamatan korban kecelakaan.
- Visualisasi di samping menunjukkan kecepatan rendah lebih tinggi tingkat keselamatannya daripada kecepatan tinggi.

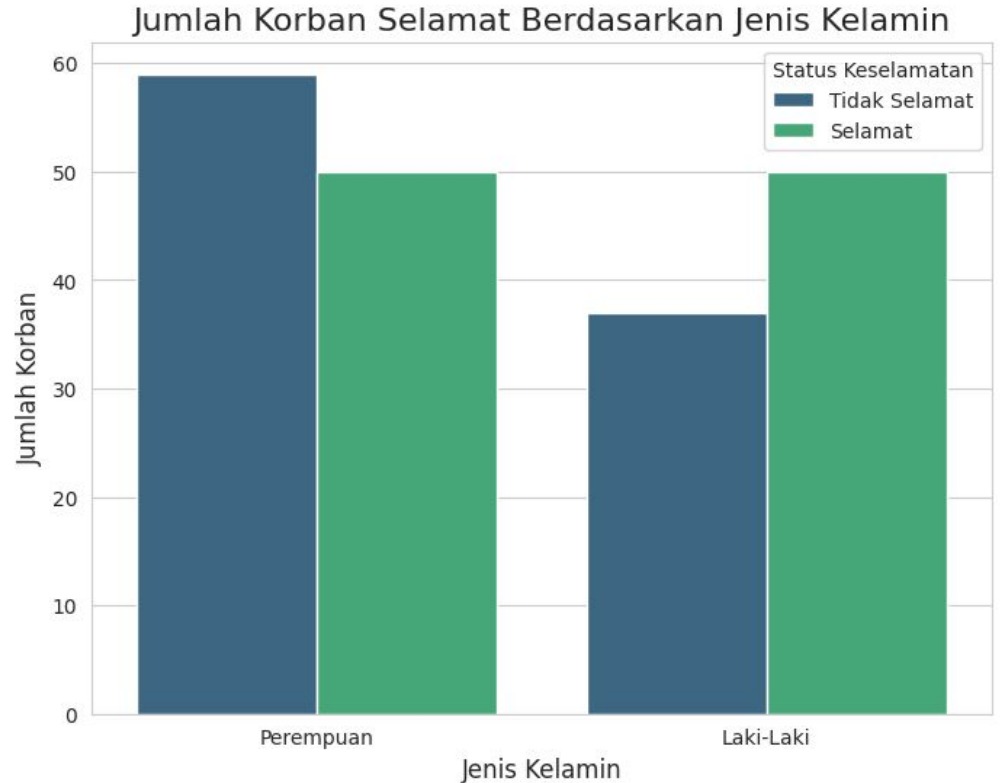
Perbandingan Tingkat Keselamatan: Kecepatan Rendah vs Tinggi



INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

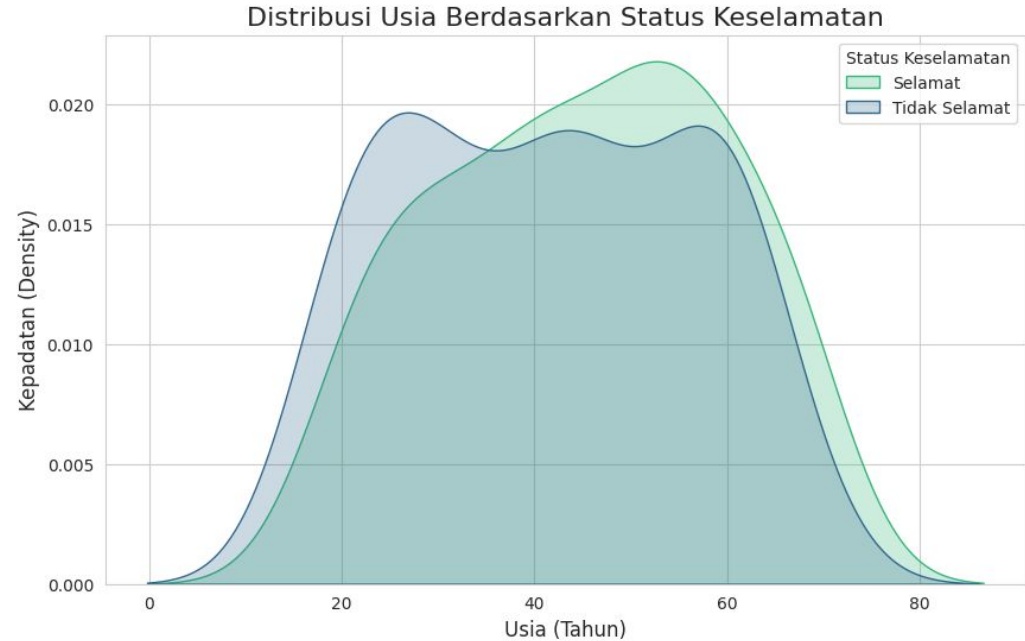
- Jenis kelamin juga faktor yang memengaruhi keselamatan korban kecelakaan.
- Dari visualisasi di samping menunjukkan perempuan lebih banyak menjadi korban tidak selamat di kecelakaan.



INSIGHT

Faktor Kelangsungan Hidup pada Kecelakaan Lalu lintas

- Berdasarkan faktor usia, ditemukan bahwa jumlah korban jiwa tertinggi berada pada kelompok usia muda, khususnya pada rentang usia 20-30 tahun.



INSIGHT

Faktor pada Kecelakaan Lalu lintas

Berdasarkan LLM dari Granite, ada tiga faktor utama yang menjadi penyebab kecelakaan lalu lintas, yaitu:

- Jenis Kelamin
- Umur
- Kecepatan Berkendara

****1. Gender:****

To determine the gender most frequently involved in accidents, we can count the occurrences of each gender.

- Female: 116 accidents
- Male: 79 accidents

From the count, females are more frequently involved in accidents than males (116 vs. 79).

****2. Age:****

To identify the most common age range, we would typically group the ages into ranges (e.g., 20-29, 30-39, etc).

****3. Speed of Impact:****

To find the most frequent range of speed at which accidents occur, we would again need the 'Speed_of_Impact'

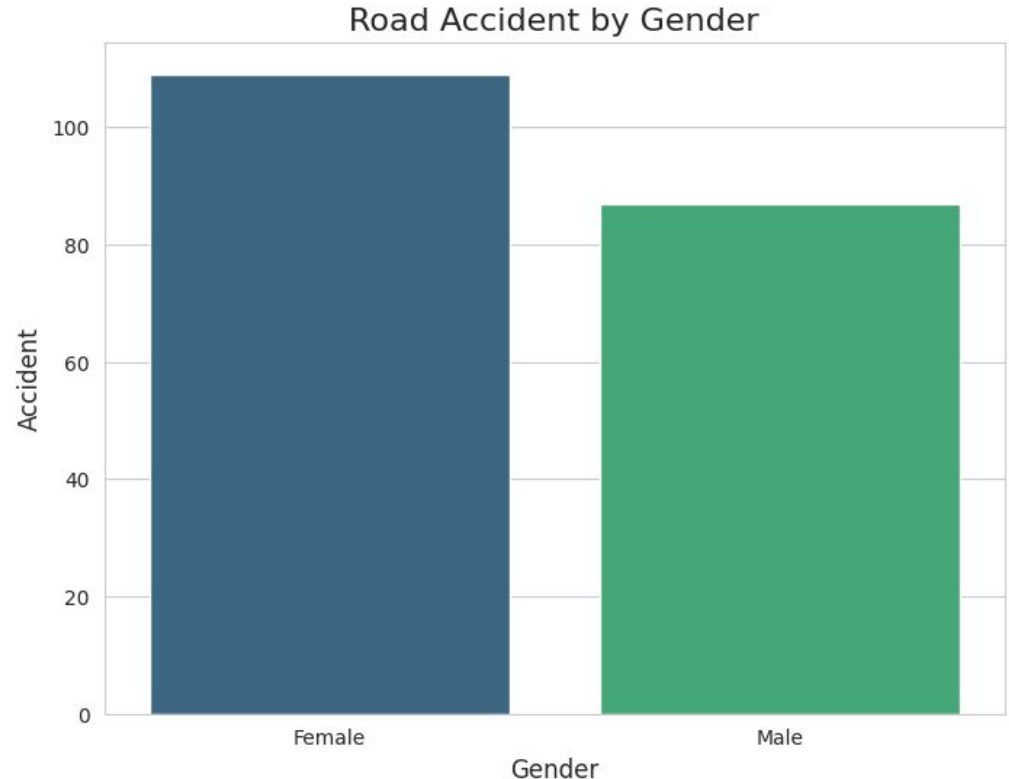
****Concise Summary Profile:****

- ****Gender Distribution:**** Females are more frequently involved in accidents (116 occurrences) compared to males (79).
- ****Age Range Insight:**** Without pre-defined age groupings, we cannot definitively state the most common age range.
- ****Speed of Impact Insight:**** Similarly, without categorized speed ranges, we cannot identify a most frequent speed range.

INSIGHT

Faktor pada Kecelakaan Lalu lintas

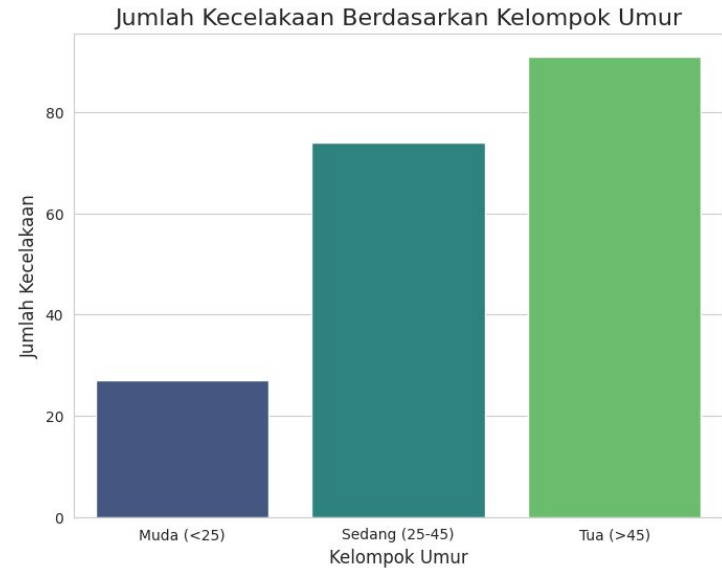
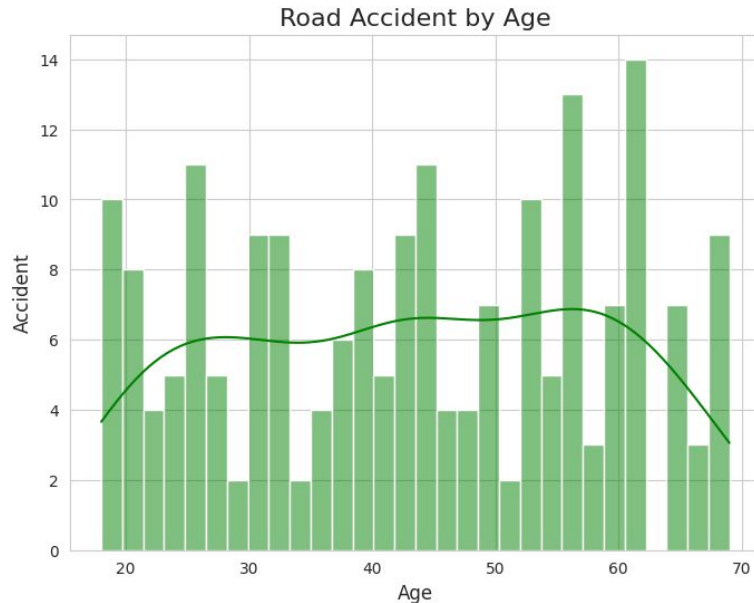
- Berdasarkan faktor jenis kelamin, data menunjukkan bahwa jumlah kecelakaan yang melibatkan perempuan lebih tinggi daripada yang melibatkan laki-laki.



INSIGHT

Faktor pada Kecelakaan Lalu lintas

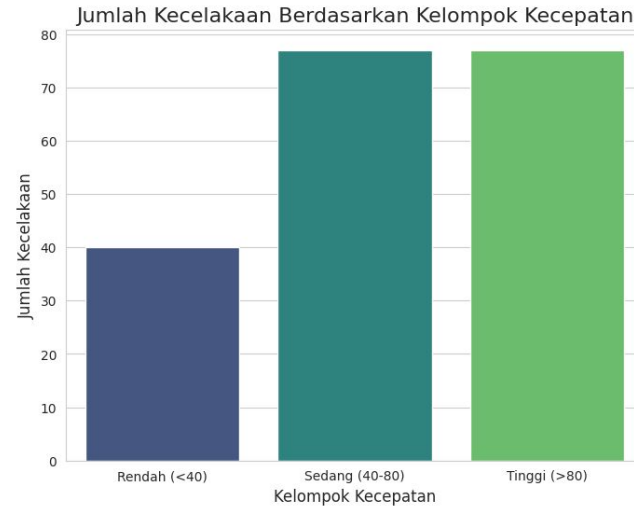
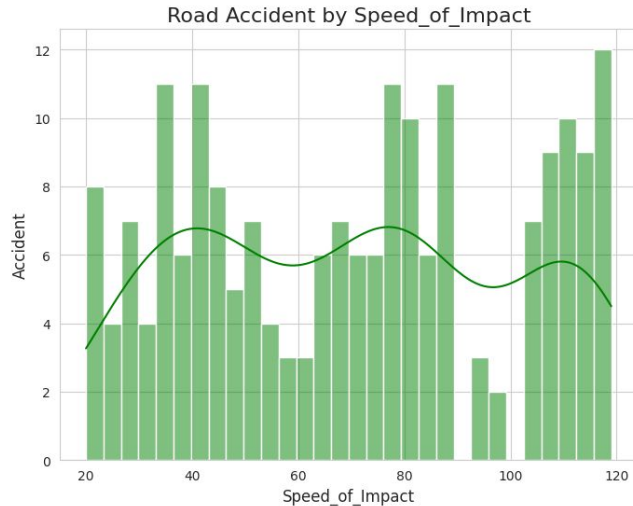
Berdasarkan faktor usia, data menunjukkan tingginya angka kecelakaan lalu lintas yang melibatkan kelompok usia tua.



INSIGHT

Faktor pada Kecelakaan Lalu lintas

Berdasarkan faktor kecepatan, dapat disimpulkan bahwa terdapat korelasi positif antara kecepatan berkendara yang tinggi dengan meningkatnya risiko kecelakaan lalu lintas.



INSIGHT

Model Klasifikasi Prediksi Kelangsungan Hidup pada Kecelakaan Lalu Lintas Menggunakan Machine Learning

Model Name	Accuracy Score	F1 Score	Precision Score	Recall Score
Logistic Regression	0.525	0.512821	0.47619	0.555556
K-Nearest Neighbors	0.55	0.571429	0.5	0.666667
Support Vector Machine	0.6	0.636364	0.538462	0.777778
Decision Tree	0.525	0.55814	0.48	0.666667
Random Forest	0.475	0.511628	0.44	0.611111
Extreme Gradient Boosting	0.425	0.378378	0.368421	0.388889

INSIGHT

Model Klasifikasi Prediksi Kelangsungan Hidup pada Kecelakaan Lalu Lintas Menggunakan LLM Granite

Based on the patterns observed from the provided examples, it appears that individuals

Applying these patterns to the new cases:

- New Case 4: Given the high speed of impact (117.0) despite using a seatbelt, the clas
- New Case 5: An older female using both safety measures (helmet and seatbelt) at a mod
- New Case 6: A young male with high speed of impact (49.0) and using safety measures,
- New Case 7: Similar to Case 6, this male with a very high speed of impact (116.0) and
- New Case 8: A middle-aged male with a high speed of impact (47.0) and using safety ge

Classifications:

- New Case 4: Did Not Survive
- New Case 5: Survived
- New Case 6: Survived
- New Case 7: Did Not Survive
- New Case 8: Did Not Survive

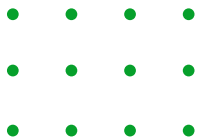
Please note that this classification system is based on a very limited set of examples



KESIMPULAN & REKOMENDASI

Kesimpulan

- Analisis menggunakan LLM Granite dan evaluasi Confusion Matrix menunjukkan bahwa faktor utama yang memengaruhi keselamatan korban adalah penggunaan sabuk pengaman dan kecepatan kendaraan.
- Adapun faktor yang sering ditemukan pada profil kecelakaan dalam dataset ini adalah jenis kelamin, usia korban, dan kecepatan saat berkendara.
- Dari beberapa model machine learning yang diuji, Support Vector Machine (SVM) terbukti menjadi model paling akurat dengan tingkat akurasi 0.6, meskipun dilatih menggunakan dataset yang terbatas. Selain itu, klasifikasi menggunakan LLM Granite juga menunjukkan kinerja yang baik.





KESIMPULAN & REKOMENDASI

Rekomendasi

- **Memprioritaskan Kampanye Keselamatan Publik:** Mengingat kecepatan dan penggunaan sabuk pengaman adalah faktor yang paling dapat dikontrol, kampanye keselamatan jalan raya perlu memfokuskan pesan pada dua hal ini. Edukasi mengenai bahaya fatal akibat kecepatan tinggi dan pentingnya mengenakan sabuk pengaman dapat menjadi intervensi paling efektif untuk menekan angka fatalitas.
- **Memberikan Pelatihan Khusus untuk Kelompok Berisiko:** Direkomendasikan untuk menyediakan pelatihan khusus bagi perempuan dan kelompok usia lanjut, karena data menunjukkan tingginya keterlibatan kedua kelompok ini dalam kecelakaan lalu lintas.
- **Mengembangkan Model dengan Data Tambahan:** Model yang ada saat ini perlu ditingkatkan lebih lanjut untuk memperkuat akurasi. Disarankan untuk memperkaya dataset dengan variabel kontekstual lain seperti jenis kendaraan, kondisi cuaca, tipe jalan, dan waktu kejadian.

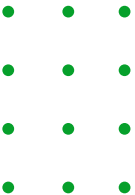


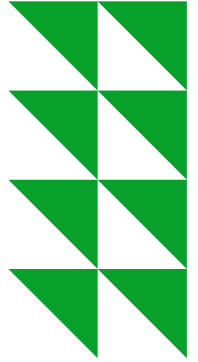


AI SUPPORT

LLM IBM Granite:

- • - ibm-granite/granite-3.3-8b-instruct
- • - Membantu dalam menganalisis data untuk mengidentifikasi variabel-variabel kunci yang memengaruhi keselamatan korban, sekaligus memahami profil umum dari kecelakaan yang terjadi.
- • - Membuat model klasifikasi untuk memprediksi keselamatan korban, sebagai pendekatan alternatif selain machine learning konvensional.





Terima Kasih

