Predective Model Report

Classification Summary

Below is a breakdown of how well the model predicted your outcome of interest

Overall Accuracy

80.9%

48,707 out of 60,216 rows tested are classified correctly

See accuracy details

EXCELLENT

Performance for Not Applicable

3.3x better than baseline

Rows predicted to be **Not Applicable** are **3.27x** more likely to be **Not Applicable** than the baseline rate.

To identify the patterns in your data, we trained a custom machine learning model on 80% of your records and validated performance on the remaining 20%. Here's how well the model works.

Predictive Performance

True Positives
2,569/10,778 — **24%**
The model correctly predicted **2,569 of the 10,778** rows in the test set where Main_Cause is Not Applicable

False Positives
1,815/49,438 — **4%**
It incorrectly predicted **1,815 of the remaining 49,438** rows as Not Applicable

True Negatives
47,623/49,438 — **96%**
The model correctly predicted **47,623 of the 49,438** rows in the test set where Main_Cause is not Not Applicable

False Negatives
8,209/10,778 — **76%**
It incorrectly predicted **8,209 of the remaining 10,778** rows as not Not Applicable

Performance Details

Outcome	Accuracy	Precision	Recall	F1 Score	Count
All outcomes	80.9%	-	-	-	60,216
Not Applicable	-	0.59	0.24	0.34	10,778
Human Error	-	0.83	0.96	0.89	47,807
Vehicle Defect	-	0.19	0.01	0.02	1,412
Road Environment...	-	0.00	0.00	0.00	219
BUDDHISTS	-	0.00	0.00	0.00	0

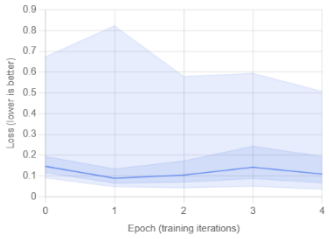
+ See More

Accuracy measures how often a prediction is correct. It is calculated by dividing the number of correct predictions by the total number of predictions.

Advanced Model Details

Different models are compared and the highest quality snapshot is returned at the end of training.

LOSS GRAPH



The blue line shows the average loss over time, and the shaded regions represent the range of loss values during each training step. A lower loss value means that the model is performing better.

MODEL TYPE

Multilayer Neural Network

A baseline neural network that can learn non-linear relationships between inputs and outputs efficiently. This model is a good starting point for most problems.

TRAINED

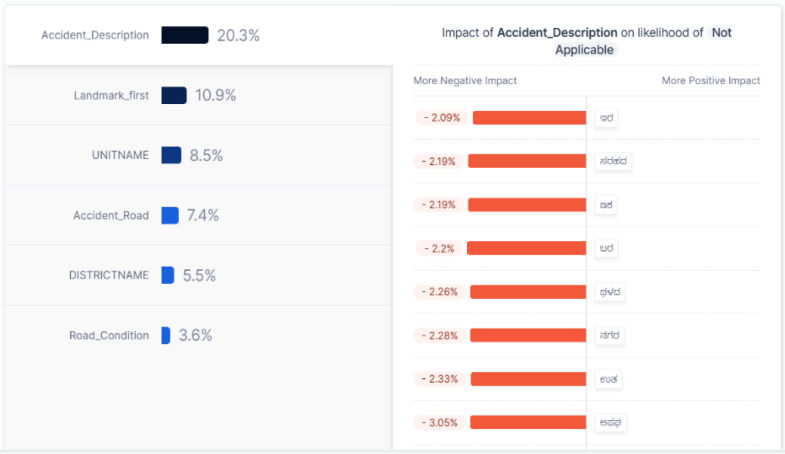
Today At 2:26 PM

OTHER MODELS TESTED

Model Type	Trial	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
Multilayer Neural Network	--	0.007654	82.57%	0.031106	82.59%
Linear Regression	--	0.000349	80.18%	0.000313	83.15%

Top Fields

Fields ranked by their contribution to the prediction results



Top Factors

Individual factors ranked by their contribution to the prediction results

Factor	Rate of Not Applicable	Factor Frequency	Impact on likelihood of Not Applicable
UNITNAME contains Station	83%	2%	+ 13.78%
Surface_Type is Not Applicable	53%	7%	+ 13.59%
UNITNAME contains Police	83%	2%	+ 13%
Accident_Classific... is Not Applicable	47%	7%	+ 10.7%
Road_Condition is No influence on accident	4%	22%	- 9.3%

Example Text:

"Traffic South Police Station "
"Traffic North Police Station "
"Traffic North Police Station "
"Traffic South Police Station "
"Traffic South Police Station "

Rate of Not Applicable

Within this group 83%
3,854 / 4,637 rows
Among all records 18%
53,652 / 301,079 rows

Factor Frequency

7% of Not Applicable records have this factor
3,854 / 53,652 rows
2% of all records have this factor
4,637 / 301,079 rows

+ See more

Segments

Sets of similar records in your dataset grouped by outcome of interest

Likely Not Applicable

Vehicle Defect

This segment is characterized by accidents caused by vehicle defects, often occurring in villages or on national highways, and involving units from central and northern regions of Karnataka.

Not Applicable in this segment — 1.5x vs base rate

Segment size — 20.2% of all records

Uncertain

Road Defect

This segment is characterized by accidents caused by road environment defects, often occurring near landmarks such as houses or poles, and involving units from southern regions of Karnataka.

Not Applicable in this segment — 1.0x vs base rate

Segment size — 53.2% of all records

Unlikely Not Applicable

Human Error

This segment is characterized by accidents caused by human error, often occurring near roads or villages, and involving units from various regions of Karnataka.

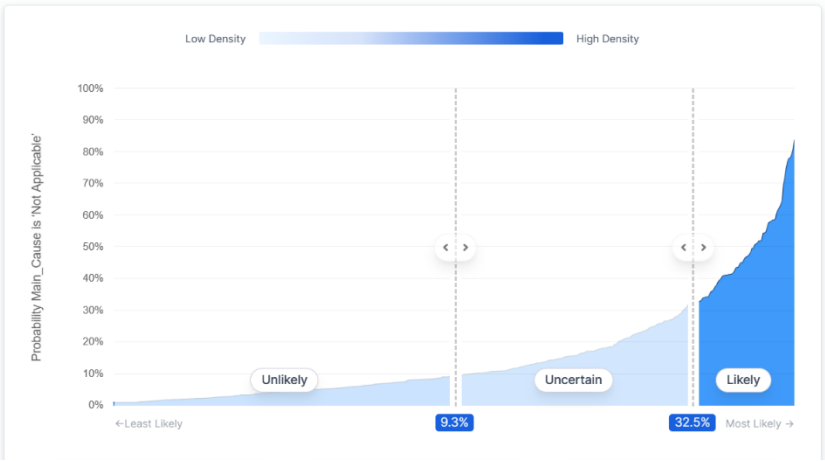
Not Applicable in this segment — 0.7x vs base rate

Segment size — 26.6% of all records

Unlikely Records	Uncertain Records	Likely Records
92.0% not Not Applicable	18.9% Not Applicable	50.7% Not Applicable
vs. base rate 1.1x	vs. base rate 1.0x	vs. base rate 2.8x
Group size 50.0%	Group size 35.0%	Group size 15.0%

Decision Threshold Graph

A subset of your data sorted by outcome of interest probability. Use this to explore decision thresholds.



Sample Rows

Sample rows of your data sorted by outcome of interest probability. Drag the slider to inspect rows with different likelihoods.

Likely Records: Probability of Not Applicable is 80%

← Least Likely — Most Likely →

DISTRICTNAME	UNITNAME	CRIME_NO	YEAR	NOOFVEHICLE_INV.	MAIN_CAUSE
Hassan	Sakaleshpura Rur...	10455075720230...	2023	2	Not Applicable
Mangaluru City	Traffic North Poli...	10980201320210...	2021	3	Not Applicable
Kodagu	Ponnampet PS	10451058320220...	2022	2	Not Applicable
Chamarajanagar	Hanur PS	10467113720220...	2022	2	Not Applicable
Tumakuru	Tumakuru Rural PS	10464108920180...	2018	1	Not Applicable

Download All Sample Rows

8:21:56 pm

PREDICTION

Main_Cause	RI	RoadJunction
Human Error		

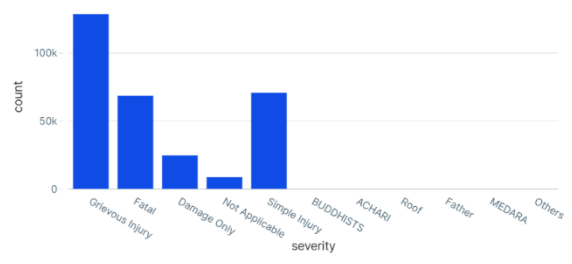
Probability Main_Cause is Hum...	Probability Main_Cause is Not ...	Probability Main_Cause is Vehi...
55.46%	33.76%	10.59%

Probability Main_Cause is Road...	Probability Main_Cause is BUD...	Probability Main_Cause is Acci...
0.16%	0.01%	0.01%

Histogram

Download

This chart reveals the frequency of different levels of accident severity.



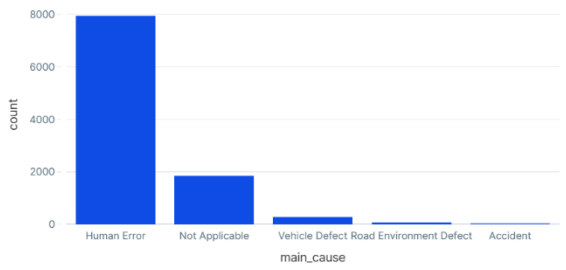
Delete | Hide Summary | Refresh

The histogram of the "severity" column reveals the distribution of different levels of accident severity. This chart provides insights into the frequency of accidents based on their impact, allowing the user to identify vulnerable groups like pedestrians or cyclists. By understanding the severity distribution, the user can analyze the risk factors and develop targeted interventions to improve road safety.

Bar Chart

Download

This chart displays the distribution of main causes of road accidents.



Delete | Hide Summary | Refresh

The chart description of the Bar Chart: (main_cause) provides the distribution of the main causes of road accidents. By analyzing this chart, users can gain insights into the most common reasons for accidents. This information is crucial as it helps prioritize interventions and awareness campaigns to address the identified causes effectively.