

Model Development Phase Template

Date	16 July 2024
Team ID	SWTID1720527361
Project Title	TrafficTellgence: Advanced Traffic Volume Estimation with Machine Learning
Maximum Marks	5 Marks

Feature Selection Report

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
holiday	this feature contains value whether current day is of holiday or not.	No	Rejected due to 99% NULL values, indicating insufficient data for analysis.
temp	This feature contains the temperature in Kelvin unit	No	Excluded because temperature showed minimal impact on efficiency during training and testing so as a process of dimension reduction we rejected this column.

rain	It contains the amount of rain currently raining	Yes	The amount of rain could have impact on traffic movement and volume. Like some cancelling of trips and journey due to rain.
snow	This column contains the amount of snowfall.	No	Not selected because 99% of the data recorded zero snowfall, leading to a large data imbalance.
weather	This column contains current weather like sunny, clear, foggy, etc	Yes	Weather have effect on traffic movement and volume. Like during foggy weather traffic volume might decrease or during clear and sunny weather volume might increase.
date	It has current date in DD-MM-YYYY format	Yes	As we have deleted the holiday column, so date column is important for deciding holiday. So it is selected to help determine holidays and assess the impact of different months and years on traffic volume.
Time	IT has time in 24 hours format in HH:MM: SS	Yes	Time has largest impact on traffic volume. It can be considered crucial as it greatly affects traffic volume, especially during peak office hours and midday lulls.