

Applied Data Science Capstone

Car Accident Severity Analysis: Seattle, Washington

The project aims to understand the factors which play a role in the severity of accidents using Machine Learning Models.

Submitted by: Kishan Kumar S

Contents

1. Introduction	3
1.1 Background	3
1.2 Problem	3
1.3 Stakeholders	3
2. Understanding Data	4
2.1 Data Cleaning	4
2.2 Feature Selection	5
3. Methodology	6
3.1 Data Collection	6
3.2 Exploratory Analysis	6
3.3 Machine Learning Model Selection	8
4. Results	8
4.1 Decision Tree Analysis	8
4.1.1 Classification Report	8
4.1.2 Confusion Matrix	9
4.2 Logistic Regression	9
4.2.1 Classification Report	9
4.2.2 Confusion Matrix	10
4.3 k-Nearest Neighbor	10
4.3.1 Best kNN value	11
4.3.2 Classification Report	11
5. Discussion	11
5.1 Average f1-Score	12
5.2 Precision	12
5.3 Recall	13
6. Conclusion	13
7. Recommendations	14
7.1 Public Development Authority of Seattle (PDAS)	15
7.2 Car Drivers	16
7. References	16

1. Introduction

1.1 Background

Seattle, also known as the Emerald city, is Washington State's largest city, with home to a large tech industry with Microsoft and Amazon headquartered in its metropolitan area. As of 2020, it has a total metro area population of 3.4 million (www.macrotrends.net). The total number of personal vehicles in Seattle in the year 2016 hit a new high of nearly 444,000 vehicles. In one South Lake Union census tract, the car population has more than doubled since 2010 (www.seattletimes.com). The increase in car ownership rates can lead to higher numbers of accidents on the road because of a simple probability. Worldwide, approximately 1.35 million people die in road crashes each year, on average 3,700 people lose their lives every day on the roads and an additional 20-50 million suffer non-fatal injuries, often resulting in long-term disabilities.

1.2 Problem

The world as a whole suffers due to car accidents, including the USA. National Highway Traffic Safety Administration of the USA suggests that the economical and societal harm from car accidents can cost up to \$871 billion in a single year. According to 2017 WSDOT data, a car accident occurs every 4 minutes and a person dies due to a car crash every 20 hours in the state of Washington while Fatal crashes went from 508 in 2016 to 525 in 2017, resulting in the death of 555 people. The project aims to predict how severity of accidents can be reduced based on a few factors.

1.3 Stakeholders

The reduction in severity of accidents can be beneficial to the Public Development Authority of Seattle which works towards improving those road factors and the car drivers themselves who may take precaution to reduce the severity of accidents.