ANALYSIS OF INCIDENT INJURY SEVERITY ON NJ ROADWAYS

Presenter: Rima Abisaad  
New Jersey Institute of Technology  
NJDOT Annual Research Showcase  
October 17, 2018
OUTLINE

• Introduction
• Background/Objective
• Data Acquisition
• Methodology
• Results
• Conclusions
INTRODUCTION

• In 2015, motor vehicle crashes were the leading cause of death for ages 16 through 23 in the United States

• In 2016, more than 37,000 persons were killed in crashes across the United States

• Beside injury levels, normal flow of traffic is disrupted due to such crashes
INTRODUCTION

Understanding safety problems

Solving them with data-driven decisions

Crash prediction models
BACKGROUND

- Previous research examined crash injury severity using different approaches such as:
  - Poisson Lognormal Regression models
  - Negative Binomial Distribution
  - Zero-inflated Poisson and zero-inflated Negative Binomial models
- The stochastic nature of crashes is poorly described by linear functions
- Other models are explored such as **Artificial Neural Network (ANN)**
ANN (Artificial Neural Networks) CHARACTERISTICS

- Learn from observing data sets
- Mimic brain intelligence
- Generate output from input
- Model complex pattern
CRASH FACTORS

- DRIVER characteristics
- ROADWAY characteristics
- VEHICLE characteristics
- ENVIRONMENTAL CONDITIONS
- TIME characteristics
- TRAFFIC VOLUME & SPEED
The objective of this study is to **develop a ANN model for Crash Injury Severity prediction** that considers:

- **Roadway** characteristics
- **Traffic volume and speed**
- **Environmental** conditions
- **Time** characteristics

Interact with various real-time data

Predict injury severity levels

Support traffic management agencies
* Data for a total of 5,926 crashes was collected on NJ freeways.
METHODOLOGY

• By adopting several forms of ANN, this study aims to identify the best model suitable for predicting crash injury severity.

• The 37 ANNs differ by the following criteria:
  • Type
  • Training algorithm
  • Activation function
  • Number of input neurons, and
  • Input parameters
METHODOLOGY

- Identify potential independent variables
- Select structure of ANN
- Configure network and initialize weights
- Train and test the network
METHODOLOGY

The model is evaluated by calculation of the Root Mean Square Error (RMSE) as follows:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}}
\]

where:

- “\(\hat{y}_i\)” is the observed values for \(i^{th}\) data
- “\(y_i\)” is the estimated ANN values for \(i^{th}\) data
- “\(N\)” is the number of data points
RESULTS

Best performing ANN characteristics:

• Deep-Feed Forward
• 15 input neurons including weighted speed variance
• Resilient Backpropagation training algorithm
• Logistic activation function.
The input layer consists of combinations of the following independent variables:

- Time (Day of the week, Month, Time of the day)
- Alcohol involvement
- Road characteristics
- Median type
- Light conditions
- Environmental conditions
- Posted speed limit
- Pavement condition
- Crash type
- Traffic volume
- Normal speed and crash speed
- Weighted speed variance
ANN computation

Input

Output

- Fatality/Injury/Property damage probability
<table>
<thead>
<tr>
<th>TMC</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4:00</td>
<td>6:00</td>
<td>8:00</td>
<td>10:00</td>
<td>12:00</td>
<td>14:00</td>
<td>16:00</td>
<td>18:00</td>
<td>20:00</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Property damage probability**

**Timestamp**

LOWEST  | HIGHEST
APPLICATION

• Predict potential injury severity level given real-time information.

• Produce injury severity heat maps.

• Show different colors depending on risk

• Identify most dangerous segments prone to fatality and/or injury.
## Sensitivity Analysis

<table>
<thead>
<tr>
<th>TMC</th>
<th>Injury probability/Weekday</th>
<th>Injury probability/Weekend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Timestamp</strong></td>
<td><strong>Timestamp</strong></td>
</tr>
<tr>
<td></td>
<td>4:00</td>
<td>6:00</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SENSITIVITY ANALYSIS

1. Fatalities and injuries are more likely to occur on weekends.
2. Injuries and fatalities happen mostly during the day.
3. Under adverse weather conditions, fatality and injury rates are lower.
SENSITIVITY ANALYSIS

4. Higher posted speed limit, as well as higher travelling speeds more likely lead to fatal crashes.

5. As traffic volume increases, the fatality probability decreases.

6. As speed variance increases, fatality possibility increases.
RESEARCH OPPORTUNITIES

Deploy active measures

Increase safety

Reduce delays and cost
RESEARCH OPPORTUNITIES

• Increase safety service patrol coverage (Maryland CHART, Pennsylvania IF)
• Explore crowdsourced data (Connected Citizens Program)
• Implement stricter speed rules/ Dynamic speed limits
CONCLUSIONS

• The model uses weighted speed variance as an important variable affecting the output.
• The model separates fatality crashes, and the ability to accurately predict fatality probability is not common in the literature and is a substantial contribution of this model.
• The multiplicity and stochastic nature of crash factors make it challenging to predict crash injury severity.
QUESTIONS ?