**2020 EECS Graduate Student Research Symposium**

**April, 2020**

|  |  |
| --- | --- |
| Presenter name | Samira Soleimani |

|  |  |
| --- | --- |
| Program  (MS or PhD) | MS |

|  |  |
| --- | --- |
| Number of years in program | 2 |

|  |  |
| --- | --- |
| Department  (ECE or CS) | CS |

|  |  |
| --- | --- |
| Advisor name | Jianhua Chen |

|  |  |
| --- | --- |
| Presenter email | Ssolei1@lsu.edu |

|  |  |
| --- | --- |
| Advisor email | cschen@lsu.edu |

Is this presentation based on your publication?

Yes No

If yes, which of the following items applies:

Accepted journal

Accepted conference

Under review journal

Under review conference

|  |  |
| --- | --- |
| Title | **Mining the Crash Narrative Report of Highway-Rail Grade Crossings: Text Mining Approach** |

|  |  |
| --- | --- |
| Abstract  (max 250 words) | The safety of highway-rail grade crossings is still an issue in the United States (USA). In this project, we performed machine learning techniques to study the safety of the highway-rail grade crossing using the Federal Railroad Administration crash data including unstructured and structured features. The aim was to develop a model to classify the crashes into two categories of “car stuck train” and “train struck car”. This model would help identifying factors causing crashes in different crash types. To avoid losing the information of unstructured crash reports in crash analysis, we employed text mining techniques on the crash unstructured “Narrative” data. Text mining was also used here to compare the similarity of highway rail grade crossing crashes between every two states in the USA. The similarity was visualized in a heatmap matrix based on the crash similarities. The results indicated that states with closer geographical locations have similar crash narrative reports. To dig down deeper into the crash reports a topic modeling technique was applied on data. The Latent Drichlet Allocation (LDA) algorithm identified interesting reasons behind crashes for both crash types when “car struck train” and when “train struck car”. The results showed for “car struck train” crash type, the crash was more likely to be related to drivers’ carelessness, distraction, or loss-of-control over vehicle. However, “train struck car” crashes could be related to both driver’s behavior as well as warning devices potentials problem. |

Selection criteria in the order of importance:

1. Accepted journal
2. Accepted conference
3. Under review journal
4. Under review conference
5. Seniority in the program (PhD students)
6. Seniority in the program (Master students)