Texas A&M Transportation Institute

Single and Multiple Vehicle Crashes due to Hydroplaning: A Case Study of Interpretable Machine Learning

Subasish Das (TTI)
Anandi Dutta (UTSA)
Kakan Dey (WVU)
Mohammad Jalayer (RU)
Abhisek Mudgal (IIT BHU)

Wednesday, January 15, 2020

Overview

Synopsis

Study Design and Analysis

Key Findings

Synopsis

Synopsis

- Hydroplaning is a specific condition of tire skidding on wet pavement in which the automobile tires float on a layer of water at a high speed, causing the driver to lose control of the vehicle.
- Hydroplaning is a dangerous phenomenon that can cause roadway crashes and result in severe injuries or fatalities.
- Although hydroplaning is a major contributor to roadway crashes, it is not typically reported by police in crash databases.
- A framework to classify various crash attributes from police reports and identify hydroplane crashes is strongly needed.

Study Design and Analysis

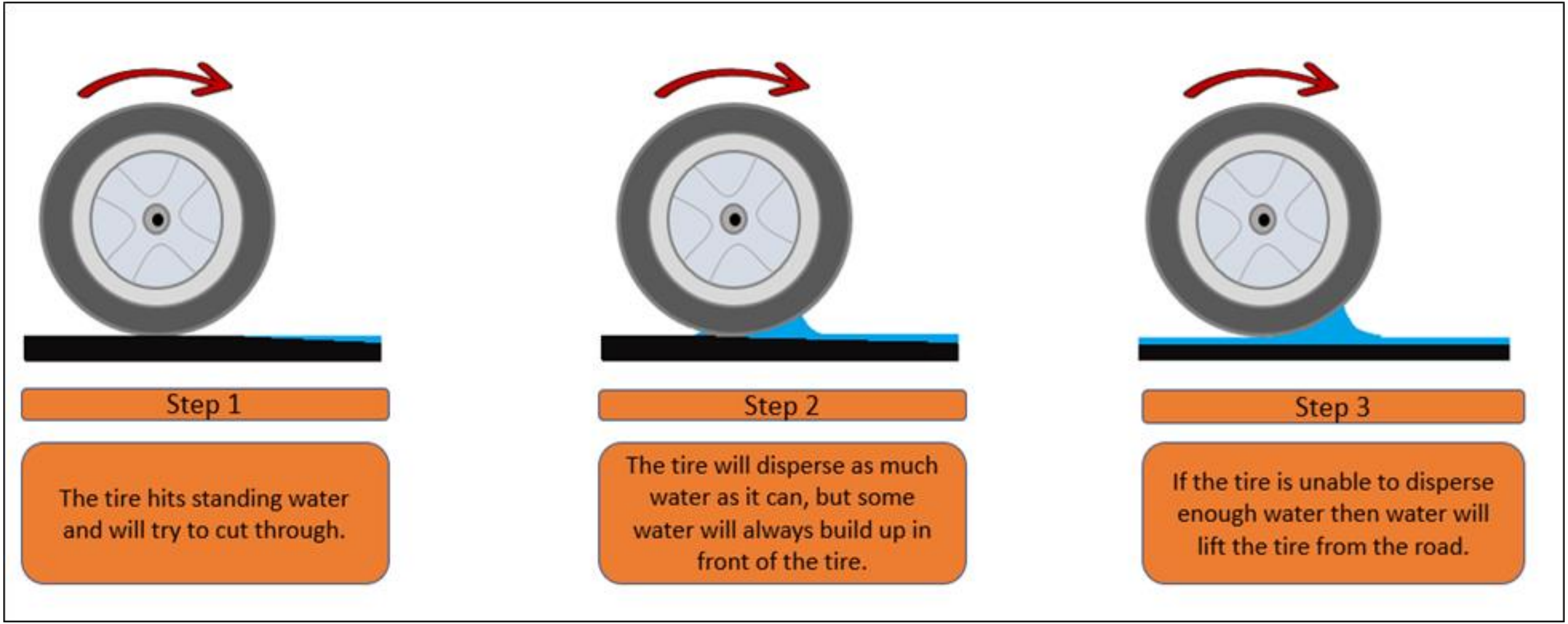
Objectives

- The two major objectives:
  - to develop a framework for applying machine learning models to unstructured textual content for classifying the number of vehicle involvements in a crash, and
  - to apply an interpretable machine learning framework that evaluates the effectiveness of keywords in determining the classification.

Data

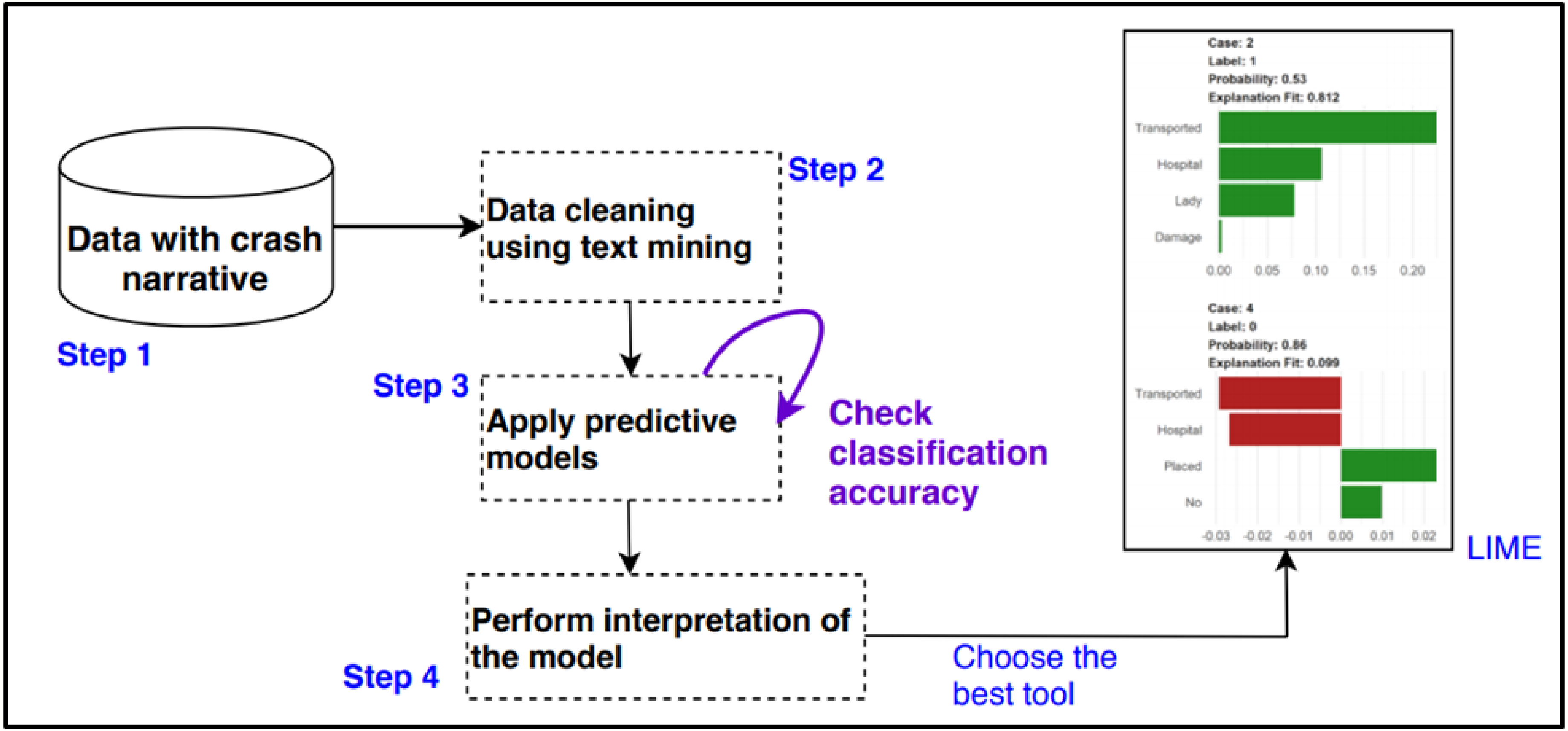
- Crash narrative reports in text format that were collected from police-reported crashes in Louisiana from 2010 to 2016.
- No filtering option to identify hydroplaning related crashes.
- Used a text searching algorithm to identify crashes associated with the keywords: ‘hydroplane’ and ‘hydroplaning.’
- Identified 703 crash reports that might be related to hydroplaning.
- Manual inspection was involved in removing reports that were out of scope of hydroplaning. This resulted in a total of 652 remaining crash reports.

Mechanism of Hydroplaning



Source: Yeager, R. W. Tire Hydroplaning: Testing, Analysis, and Design. The Physics of Tire Traction: Theory and Experiment (D. F. Hays and A. L. Browne, eds.), Springer US, Boston, MA, 1974, pp. 25–63.

Framework



Data with crash narrative
Step 1
Data cleaning using text mining
Step 2
Step 3
Apply predictive models
Check classification accuracy
Step 4
Perform interpretation of the model
Choose the best tool
LIME
Case: 2
Label: 1
Probability: 0.53
Explanation Fit: 0.812
Transported
Hospital
Lady
Damage
0.00
0.05
0.10
0.15
0.20
Case: 4
Label: 0
Probability: 0.86
Explanation Fit: 0.099
Transported
Hospital
Placed
No
-0.03
-0.02
-0.01
0.00
0.01
0.02

Methodology

- Step 1: Data collection.
- Step 2: Data cleaning.
- Step 3: Application of predictive modeling.
- Step 4: Application of interpretable machine learning model.

Stop word Removal from Crash Narratives

driver vehicle 1 advised she was traveling north bound on i-310 between us 90 and la 3127 when her vehicle began to hydroplane due to the heavy rain, at which time she lost control. driver vehicle 1 stated her vehicle exited the roadway beyond the shoulder to the right and into the grassy area when the front of her vehicle struck a tree. driver vehicle 1 advised after striking the tree her vehicle rolled over onto the roof of the vehicle and becoming partially submerged in the water. driver vehicle 1 stated she was able to get her seat belt off and exit the vehicle and walk to the shoulder of the roadway. driver vehicle 1 advised she was not injured during the crash and refused medical attention at the scene of the crash deputy gregory observed moderate to severe damage to the front bumper area and an unknown extent of water damage to the interior of vehicle 1 due to being partially submerged in the water. deputy gregory found driver vehicle 1 at fault in the crash for driving too fast for the weather conditions however no citation was issued.

After removing stopwords →

driver vehicle 1 traveling between vehicle began hydroplane heavy rain lost control. driver vehicle 1 vehicle exited the roadway beyond shoulder right into grassy area front her vehicle struck tree. driver vehicle 1 advised after striking tree vehicle rolled over onto roof vehicle and becoming partially submerged in water. driver vehicle 1 able get seat belt off and exit vehicle and walk shoulder roadway. driver vehicle 1 not injured during crash and refused medical attention scene crash deputy observed moderate severe damage front bumper area unknown extent of water damage interior vehicle 1 being partially submerged water deputy driver vehicle 1 at fault crash for driving too fast weather conditions no citation issued

Study Design

Model	Class (Observed)	Train (438 crashes)		Test (218 crashes)	
		Single (Predicted)	Multiple (Predicted)	Single (Predicted)	Multiple (Predicted)
SVM	Single	200	103	87	53
	Multiple	59	72	38	40
RF	Single	195	108	81	59
	Multiple	56	75	33	45
XGBoost	Single	215	88	92	48
	Multiple	57	74	26	52

“Why Should I Trust You?” Explaining the Predictions of Any Classifier

Marco Tulio Ribeiro
University of Washington
Seattle, WA 98105, USA
marcotcr@cs.uw.edu

Sameer Singh
University of Washington
Seattle, WA 98105, USA
sameer@cs.uw.edu

Carlos Guestrin
University of Washington
Seattle, WA 98105, USA
guestrin@cs.uw.edu

ABSTRACT

Despite widespread adoption, machine learning models remain mostly black boxes. Understanding the reasons behind predictions is, however, quite important in assessing *trust*, which is fundamental if one plans to take action based on a prediction, or when choosing whether to deploy a new model. Such understanding also provides insights into the model, which can be used to transform an untrustworthy model or prediction into a trustworthy one.

In this work, we propose LIME, a novel explanation technique that explains the predictions of *any* classifier in an interpretable and faithful manner, by learning an interpretable model locally around the prediction. We also propose a method to explain models by presenting representative indi-

how much the human understands a model’s behaviour, as opposed to seeing it as a black box.

Determining trust in individual predictions is an important problem when the model is used for decision making. When using machine learning for medical diagnosis [6] or terrorism detection, for example, predictions cannot be acted upon on blind faith, as the consequences may be catastrophic.

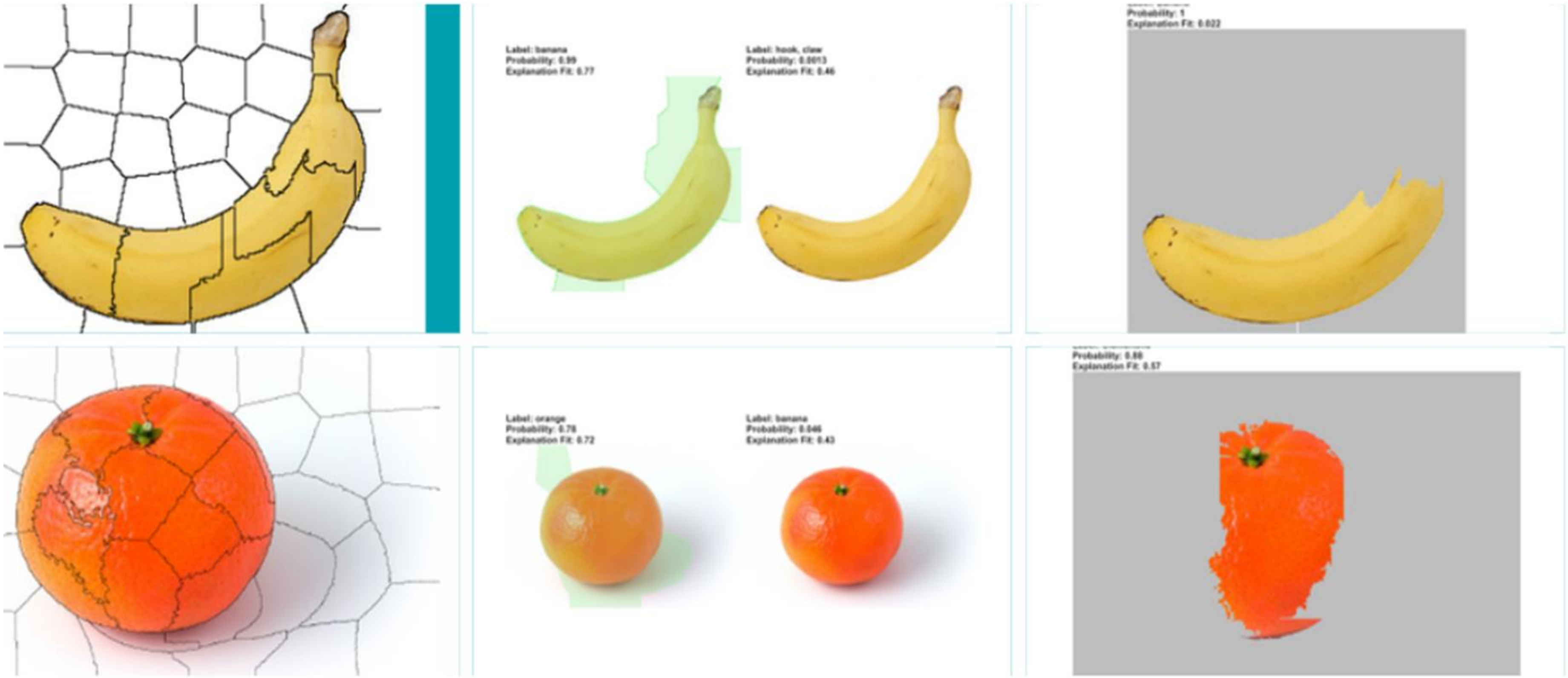
Apart from trusting individual predictions, there is also a need to evaluate the model as a whole before deploying it “in the wild”. To make this decision, users need to be confident that the model will perform well on real-world data, according to the metrics of interest. Currently, models are evaluated using accuracy metrics on an available validation dataset. However, real-world data is often significantly different, and further, the evaluation metric may not be indicative of the

Resource: *https://github.com/subasish/Awesome-Interpretable-Papers*

Why do We Need Interpretation?

- The black-box approach makes ML difficult to be being presented in front of the practitioners.
- There are many algorithms that can make precise predictions with almost zero interpretation. The good performance of the predictive model is enough for many practical problems.
- Similarly, interpretability is a crucial part of solving the problem.

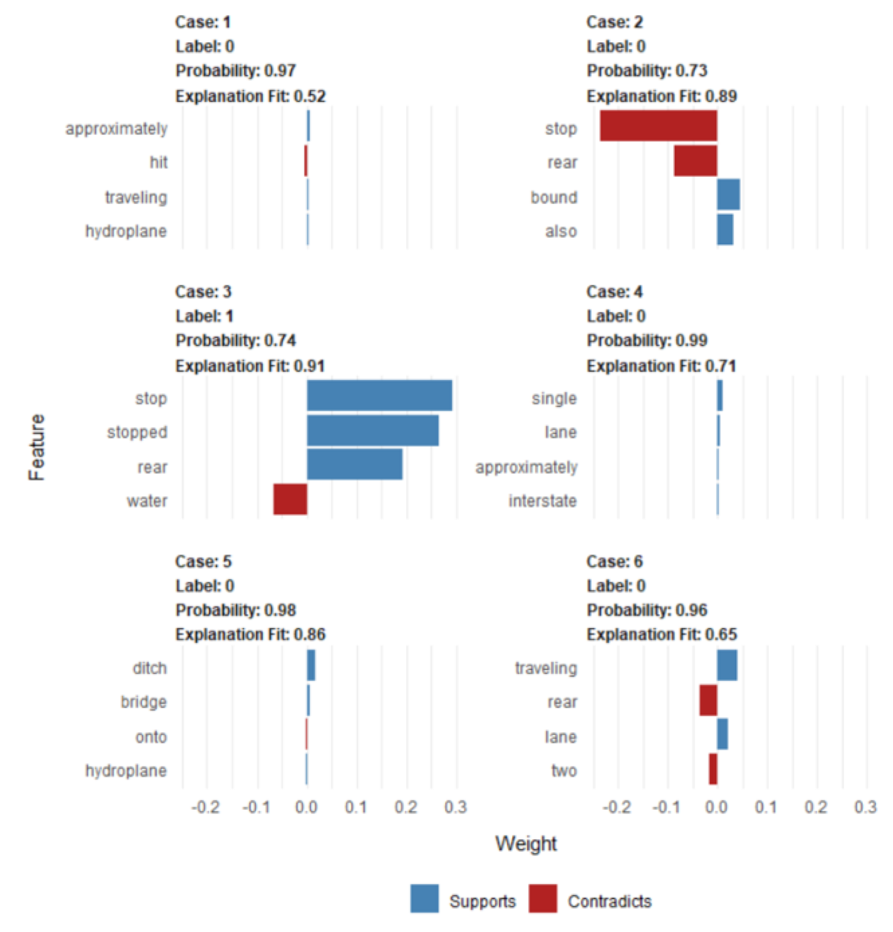
Explanation from Image Models



https://shirinsplayground.netlify.com/2018/06/keras_fruits_lime

Model Performance

Model	Data	Sensitivity	Specificity	Accuracy	Balanced Accuracy	Precision	F-score
SVM	Train	0.6601	0.5496	0.6267	0.6048	0.7722	0.7117
	Test	0.6214	0.5128	0.5826	0.5671	0.6960	0.6566
RF	Train	0.6436	0.5725	0.6221	0.6080	0.7769	0.7040
	Test	0.5786	0.5769	0.5780	0.5777	0.7105	0.6378
XGBoost	Train	0.7096	0.5649	0.6659	0.6372	0.7904	0.7478
	Test	0.6571	0.6667	0.6606	0.6619	0.7797	0.7132



(a) Randomly selected six reports (Level 1: multiple vehicles, level 0: single vehicle)

(b) Text explanation of first 3 reports in the left-hand side (Level 1: multiple vehicles, level 0: single vehicle)

Key Findings

Key Findings

- The framework developed by the research team can be replicated to other crash-related classifications (e.g., collision type) from the crash narratives.
- The XGBoost classifiers demonstrated a high prediction power; the three machine learning models predicted the severity of injury outcomes for the data with an accuracy of 66%.
- This proves that underlying trends can be revealed and analyzed through machine learning models.

Questions?

Texas A&M Transportation Institute

Subasish Das

s-das@tti.tamu.edu

979-317-2153