

Three Ideas for the Second Capstone Project

Dovid Burns

1) Build a Toxic Comment Classifier

- a) Toxic comments online are both damaging to people emotionally and prevent productive discussions about sensitive subjects. Many online platforms have difficulties ensuring that conversations are taking place in a healthy way. This project will build a classifier using a dataset that contains comments from Wikipedia's talk page edits in order to classify them as toxic or benign. The model can then be used in many platforms that allow the platforms to automatically detect—and possibly remove—these comments before they offend users or deter users from engaging in communications. The data set is hosted on Kaggle.com. <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>

2) Build a better passenger screening algorithm.

- a) Screening passengers for airport security is an exceedingly difficult task entrusted to TSA. Often, the process delays passengers and has significant impact on their travel experiences. In order to screen people faster, TSA needs better algorithms that will label fewer safe passengers as being potential threats. This project would use a dataset of images collected on the latest generation of scanners to identify the presence of simulated threats under a variety of object types, clothing types, and body types. The data set is hosted on Kaggle.com. <https://www.kaggle.com/c/passenger-screening-algorithm-challenge>

3) Build a credit card fraud detection algorithm

- a) In the United States alone, millions of credit card transactions occur every day. Of these, a very small percent are fraudulent. However, detecting these few unlawful transactions can save credit card companies significant sums of money. This project will use a dataset of almost 300k transactions that took place over the course of two days in September 2013 by European card holders to build an algorithm that can correctly detect the 500 fraudulent transactions. The data set is hosted on Kaggle.com. <https://www.kaggle.com/mlg-ulb/>