Machine Learning for Click Fraud Prevention

machineLearning

MLND Capstone Proposal
Shan Dou
June 13, 2018

Domain Background

Click fraud has been a “billion dollar” problem facing ***pay-per-click (PPC)*** advertisers. PPC is by far the most widely used compensation model in digital advertising (e.g., Both Google AdWords and Facebook Ads are PPC platforms). As the name “pay-per-click” implies, PPC advertisers pay for every click on their ads. This payment mechanism has clear benefits to advertisers as they don’t need to pay for ads that don’t attract attention, but this same mechanism is also heavily abused by fraudsters via click fraud.

Click fraud is the ill-intentioned clicking of PPC ads by fraudsters to waste and/or to mislead advertisers’ ad spending. According to a recent report by CNBC, click fraud cost advertisers $12.5 billion in 2016 and nearly 20% of total ad spending was wasted. Click fraud not only troubles advertisers, but also hurts revenue streams for ad platforms because fraud has been degrading the appeal of digital advertising. For example, the consumer giant Procter&Gamble slashed its digital ad spending by more than $200 million in 2017 (as reported by The Wall Street Journal). Detecting and preventing click fraud are therefore crucial for the $200 billion market of digital advertising. The goal of this project is to use data mining and machine learning methods to detect click fraud.

Problem Statement

This project is for detecting click fraud specific to ***the advertising of mobile apps***. The expected outcome of the project is a fraud detector that could be used by digital platforms such as TalkingData to protect app developers against click fraud. Although unsupervised learning will be used in explore patterns in the data, this project is framed as a supervised learning problem. Note that for the ground-truths labels needed in supervised learning, this project uses the binary metric of whether or not a click results in an app download instead of whether or not a click is fraudulent. This is because the latter metric is prone to error and biases: fraud labels based on IP and device blacklists are likely to be incomplete, and they are inevitably biased by the procedures used to generate them in the first place (Oentaryo et al., 2014). In short, this project aims to solve a ***supervised learning problem*** that predicts the probability of app download following each ad click. Fraudulent clicks will be identified among clicks that have low download probabilities and are clustered in IP addresses, devices, and/or click time.

Datasets and Inputs

Solution Statement

Benchmark Model

Evaluation Metrics

Project Design