Ad Tracking Fraud Detection

Lixuan Mao, netID: lm769
Ang Li, netID: al2386

Proposal:

This is a competition from Kaggle.

Fraud risk is everywhere, but for companies that advertise online, click fraud can happen at an overwhelming volume, resulting in misleading click data and wasted money. With over 1 billion smart mobile devices in active use every month, China is the largest mobile market in the world and therefore suffers from huge volumes of fraudulent traffic.

TalkingData, China's largest independent big data service platform, covers over 70% of active mobile devices nationwide. They handle 3 billion clicks per day, of which 90% are potentially fraudulent. And in this project, we are challenged to build an algorithm that predicts whether a user will download an app after clicking a mobile app ad. The dataset covering approximately 200 million clicks over 4 days.

Alternate Proposals:

1. StarCraft 2 AI Building

We are both fans of game StarCraft II, as recently Blizard and DeepMind published a paper about the brand new API available in SC2. We are very interested in building an AI controlled unit against some other

enemy units controlled by SC2 original robot.
API Documentation
PaperLink

2. Yelp Food Photo Classification

Yelp posts its challenge of food photo classification based on its own dataset. In this challenge, we are asked to build models that have the ability to identify different kinds of food in photos.
We are very interested in building some deep neural network models that are powerful enough to detect and classify food in photos.

Reference Paper:
Application of artificial neural network in food classification