AdTracking Fraud Detection Challenge

JungHyun Kim
Student ID : 20175052
woori4829@unist.ac.kr

SeongBeom Park
Student ID : 20165112
amita90@unist.ac.kr

Team Alias: OMG (Oh, ML God)

1 Introduction

Billions of mobile devices are being active every month. In other words, most of people are revealed on on-line market place. By utilizing it, i.e. mobile advertising is key for all companies to inform their products, specifically mobile apps. However, mobile device is usually controlled by *touch*, so many users are easy to mis-click the advertisement. But, in the company's point of view, advertisement channels can drive up costs by simply clicking on the ad at a large scale, so such fraud click can results an overwhelming volume of misleading click data: waste of money to handle the data. Therefore, detecting and controlling such fraud click is an important skill to save money. The goal of this project is determining whether a person will download the application after clicking the advertisement.

2 Data Analysis

2.1 Source

Dataset is obtained from kaggle project of TalkingData [1]. The dataset includes 200 million clicks over 4 days. Each data point indicating one click consists of the mobile device information, advertisement information, whether the application is downloaded after clicking the advertisement. Detailed information of label and features are in the next sub-section.

2.2 Detailed Description

Label

- **is_attributed:** Indicating the app was downloaded or not

Features

- **ip:** IP address of the device which click the advertisement
- **app:** application ID for marketing
- **device:** type of the mobile device
- **os:** version of operating system used by the mobile device
- **channel:** channel ID of mobile advertisement publisher
- **click_time:** timestamp of clicking the advertisement in UTC
- **attributed_time:** timestamp of download the application after clicking the advertisement

---

[1]TalkingData, AdTracking Fraud Detection Challenge, https://www.kaggle.com/c/talkingdata-adtracking-fraud-detection/data