PREDICTING AD CLICKS USING MACHINE LEARNING

Julia Kang, Sam Kahr, Ryan Koch

Who are we?

Julia Kang

Philadelphia, PA
Cornell '19, Operations Research
FinTech, Cooking, Running

Sam Kahr

San Francisco, CA
Xoogler
Making music, Russia, Ryan

Ryan Koch

Boston, MA
M.S. Human Factors Eng.
Making useful things, yoga, Sam

What is our project goal?

GENERAL GOAL

Predict mobile ad clicks given advertising data

TECH GOALS

Learn & apply ML models, use Google Cloud, and create/deploy websites
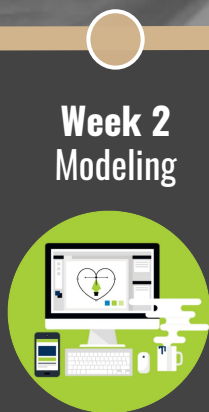
DELIVERABLES

Build generalizable ML web application for click prediction

WHAT WAS THE TIMELINE?

Week 1
Brainstorming

- Obtain Dataset
- Setup Google Cloud & Github

Week 2
Modeling

- Apply different models

Week 3
Merging

- Compare models
- Design Front End
- Google Next

Week 4
Polishing

- Implement & deploy Front End

Week 5
Delivery

- Polish Front End
- Present to Team

What data did we use?

SOURCE

Kaggle : Avazu's Click-Through Rate Prediction (2015)

kaggle

DATA TYPES

Categorical | Numeric
Anonymized Hashed Values

SIZE
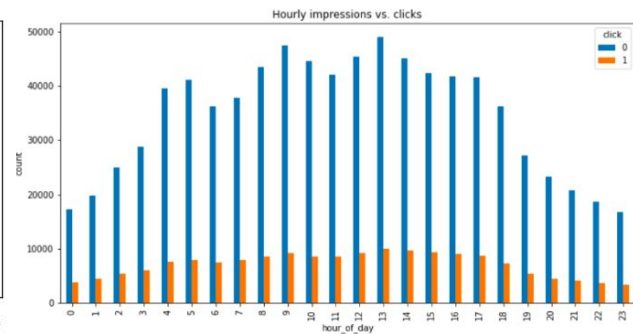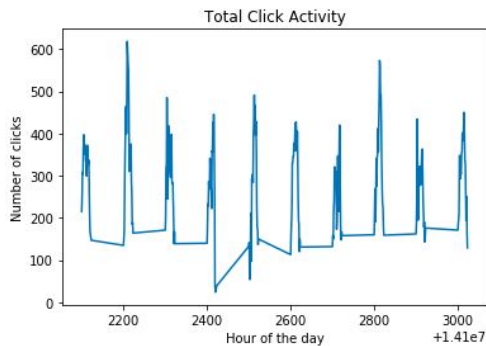
40,428,968 rows x 24 Columns

TARGET VARIABLE

'Click' : Binary
(1 = Click, 0 = No Click)

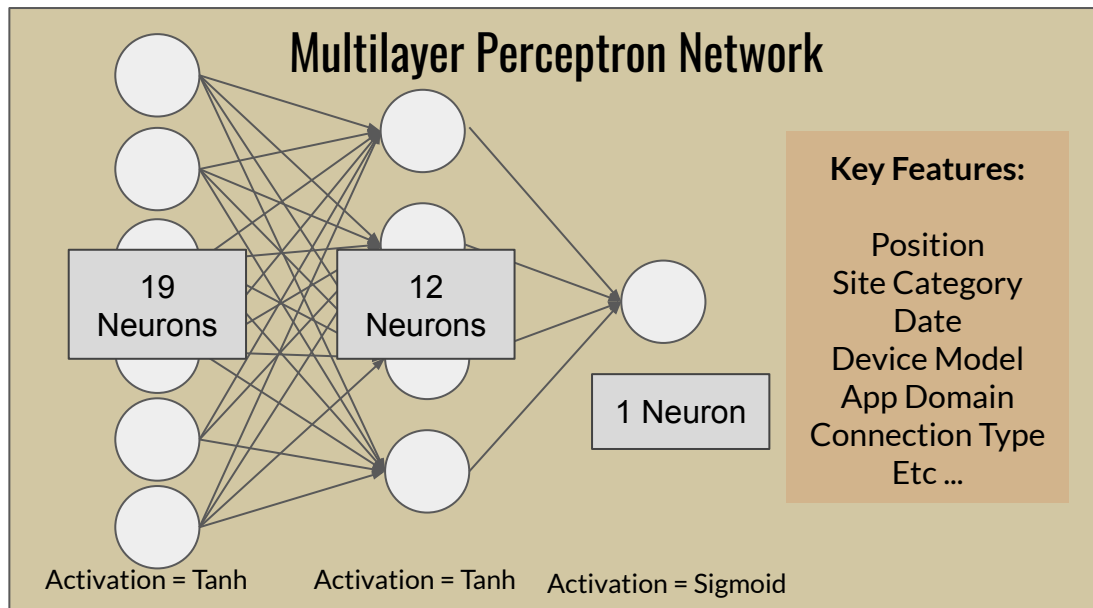# What did we do with the data?

DATA VISUALIZATIONS

Understanding Relationships

Click vs. _____

What did we do with the data?

MODELING

Multilayer Perceptron Network

19 Neurons

12 Neurons

1 Neuron

Activation = Tanh   Activation = Tanh   Activation = Sigmoid

Key Features:

Position
Site Category
Date
Device Model
App Domain
Connection Type
Etc ...

What did we do with the data?

FEATURE ENGINEERING

Creating Relationships

Feature Bagging

user_id	app_id
John	uber
Rachel	uber
Conrad	uber
John	lyft
Rachel	lyft

user_id	app_id	user_bagged	bag_codes
John	uber	Uber, lyft	01
Rachel	uber	Uber, lyft	01
Conrad	uber	Uber	02
John	lyft	Uber, lyft	01
Rachel	lyft	Uber, lyft	01

What did we do with the data?

WEB DASHBOARD

Presenting Relationships
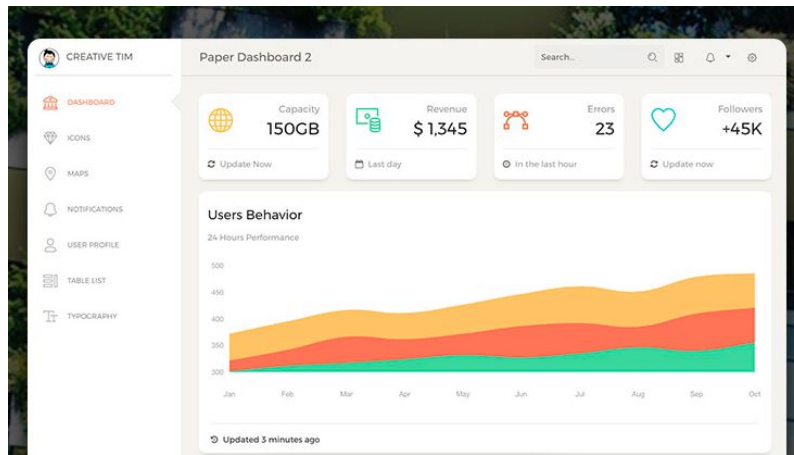
What is our site architecture ?
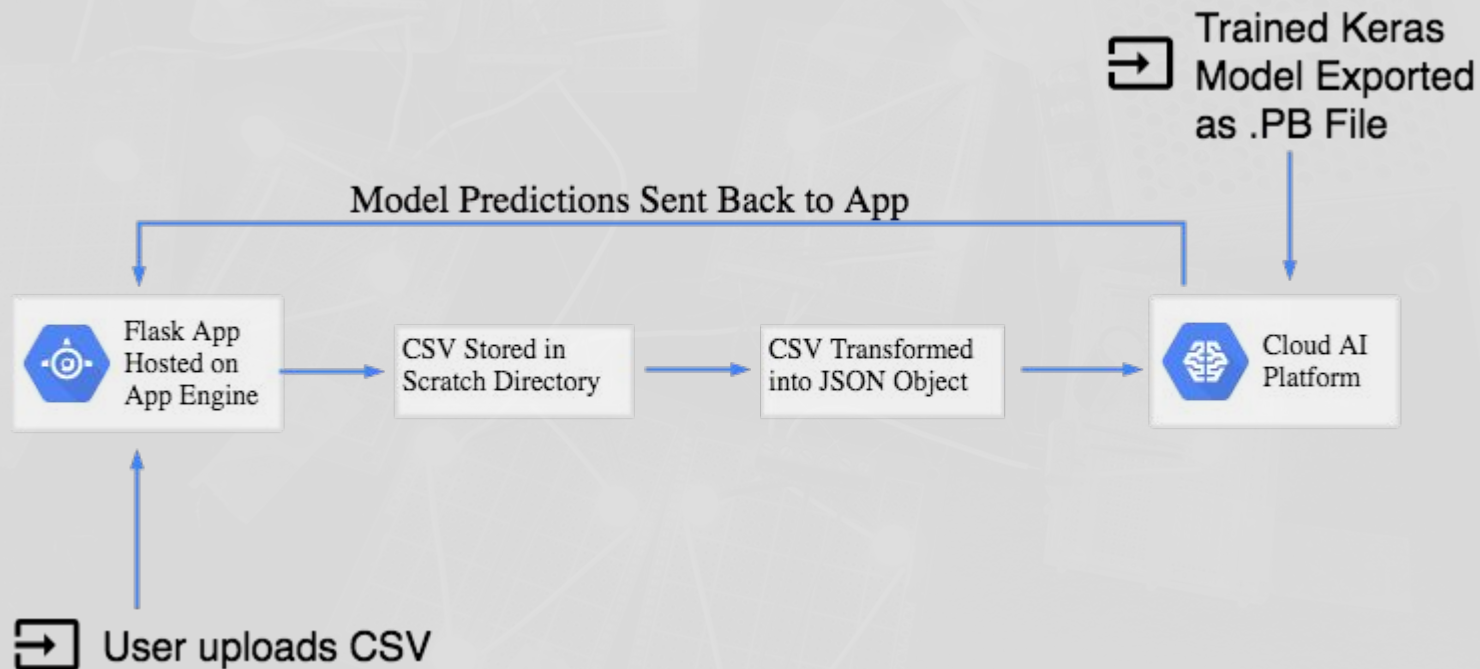
- DEMO -

https://kanalyzers.appspot.com/

What did we learn?

- Predictive modeling, classification
- Data science
  - Feature engineering
  - Bagging
  - Visualization & exploration
- Cloud architecture
- Web app. development
- Team work

What are our next steps?

Feature Engineering
➜ Control Dimensionality
➜ Normalize Data

Better Model Performance
➜ Ensembling

Configure Cloud
➜ Larger Uploads
➜ Faster Prediction

THANK YOU