StackPros

Coding Test

- **Programming Language**: Python
- **Submission file format**: Jupyter Notebook

1. Data Description

A mega-online shopping mall provides different baby brands to customers. In order to optimize online advertisements throughout websites, they want to predict the view-to-click probability given the information below.

The data consists of:

1. **Action**: the "interaction" that users completed when online ad was exposed. There are two types of actions in this dataset, **View** (a.k.a Impression) or **Click.**
   (*Reference: https://en.wikipedia.org/wiki/Impression_(online_media)*)
2. **ID**: unique user IDs
3. **Action Time**: the time when action (View/Click) happened, in microseconds format
4. **Website**: the website where the ad was published
5. **Banner Size**: the banner size of ad
6. **Brand**: the brand name that the ad was promoting
7. **Colour**: the colour of product showcased in the ad
8. **Interaction Time**: user's interaction time with each ad (sec)

2. Task

- **Goal: Predict the Likelihood of Click**
- **What we would like to see in your Jupyter Notebook:**
  - EDA Process (univariate/multivariate analysis, filling missing data etc.)
  - Feature Selection (select existing features and explanation)
  - Feature Engineering (create new features and explanation)
  - Select and Create Model
  - Hyper parameter Tuning if you feel it's needed
  - Model Validation process
  - ...

During this test, you may not be able to finish the entire modelling process – *that's ok!*

Apply your knowledge and thought leadership as much as possible.

Thank you for putting in the hard work, and StackPros will contact you soon.