A/B Hypothesis Testing: Ad Campaign performance

Abubakar Alaro Olayemi

Objective of the Study

To design a reliable hypothesis testing algorithm for Brand Impact Optimizer (BIO) service to determine whether a recent advertising campaign resulted in a significant lift in brand awareness.

Methods

- Classical A/B test

classical A/B was used to determine whether to accept or reject the null hypothesis which states that there is no difference in brand awareness between the exposed and control groups in the Ad campaign.

- Machine Learning Model

machine learning models were built to predict the probability of a user clicking the yes button while answering the question: Do you know SmartAd?

These models include:

1. Logistic Regression

2. Decision Tree Model

3. Extreme Gradient Boosting Model

Data

The BIO data for this project is a “Yes” and “No” response of online users to the following question

Q: Do you know the brand SmartAd?

The data used were collected from two experimental groups: control and exposed.

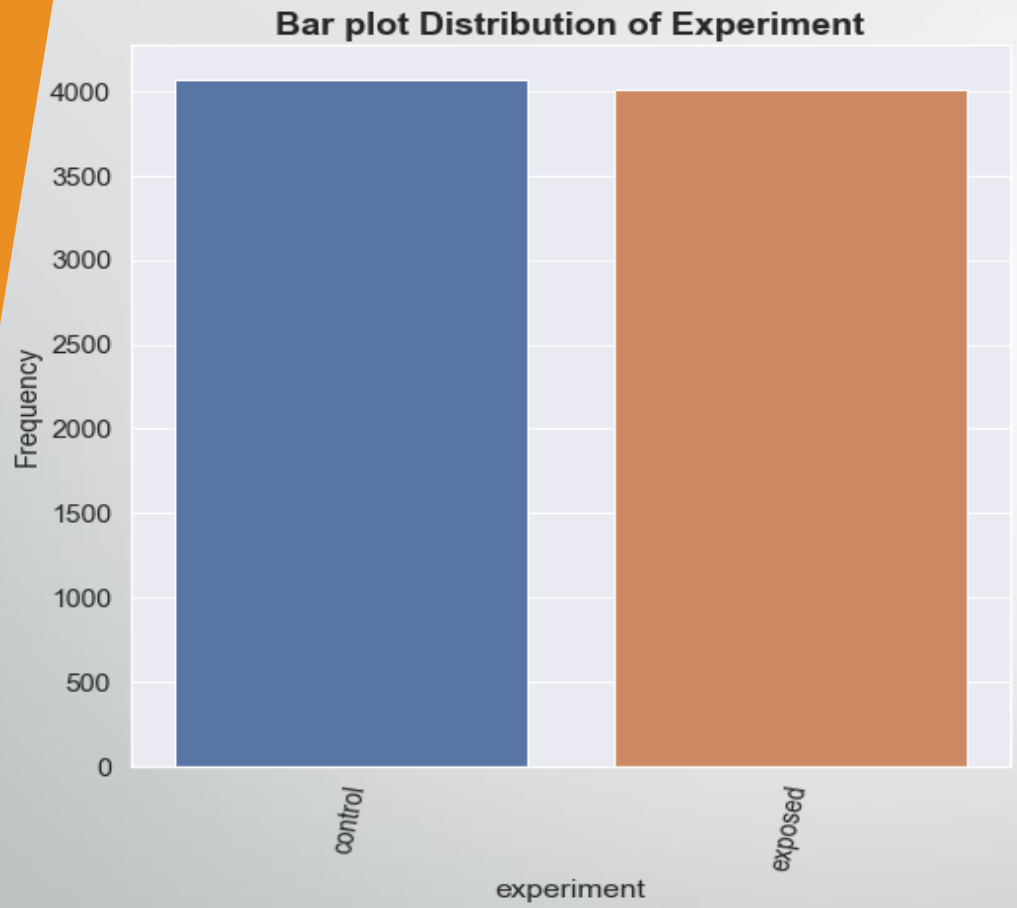
The features of these data include:

- **auction_id**: the unique id of the online user who has been presented the BIO. In standard terminologies this is called an impression id. The user may see the BIO questionnaire but choose not to respond. In that case both the yes and no columns are zero.
- **experiment** : which group the user belongs to - control or exposed.
- **date**: the date in YYYY-MM-DD format
- **hour**: the hour of the day in HH format.
- **device_make** : the name of the type of device the user has e.g. Samsung

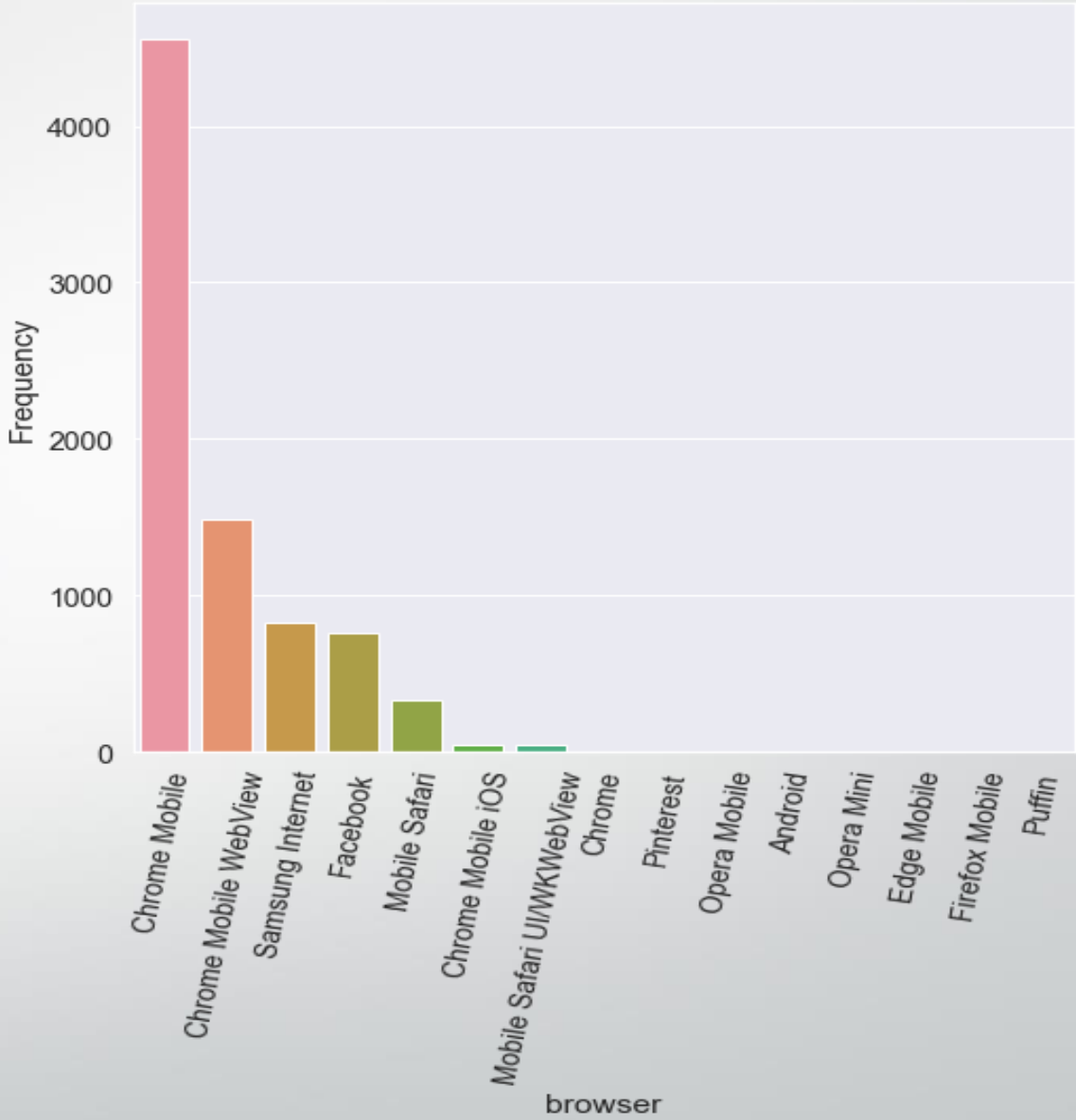
Cont'd

- **platform_os**: the id of the OS the user has.
- **browser**: the name of the browser the user uses to see the BIO questionnaire.
- **yes**: 1 if the user chooses the "Yes" radio button for the BIO questionnaire.
- **no** : 1 if the user chooses the "No" radio button for the BIO questionnaire.

Results: Data Exploration

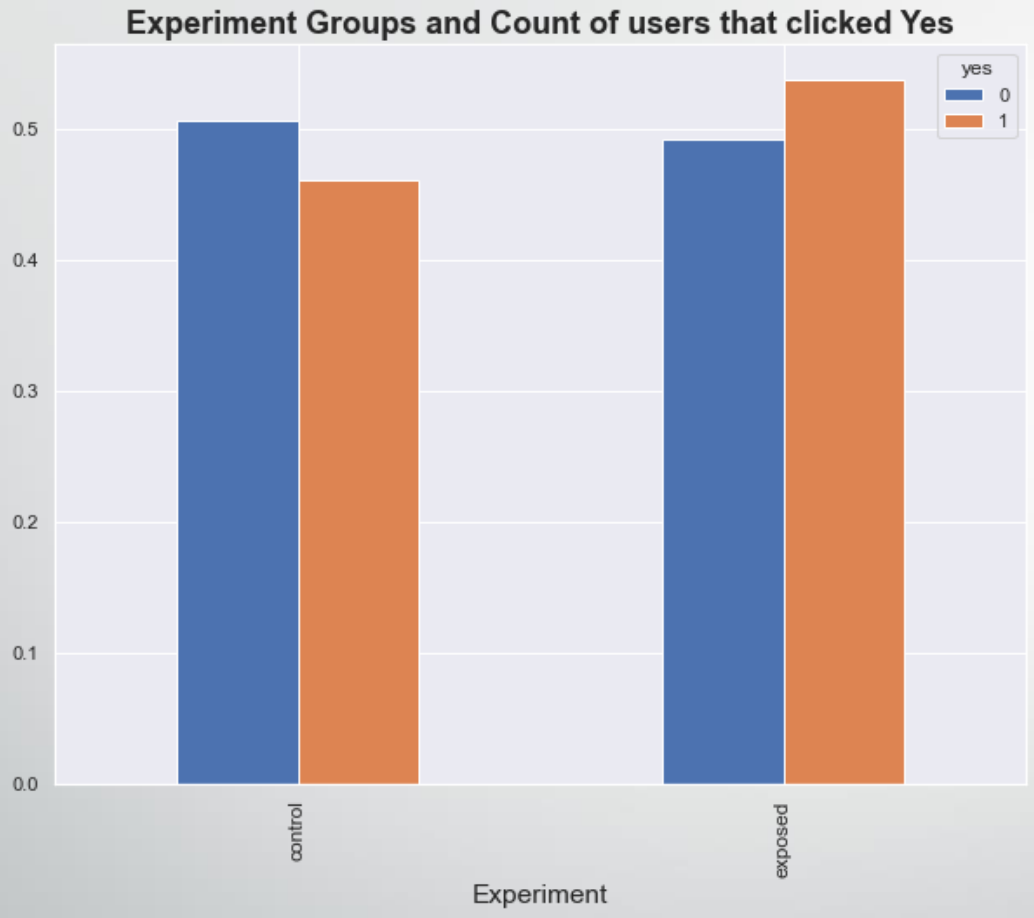


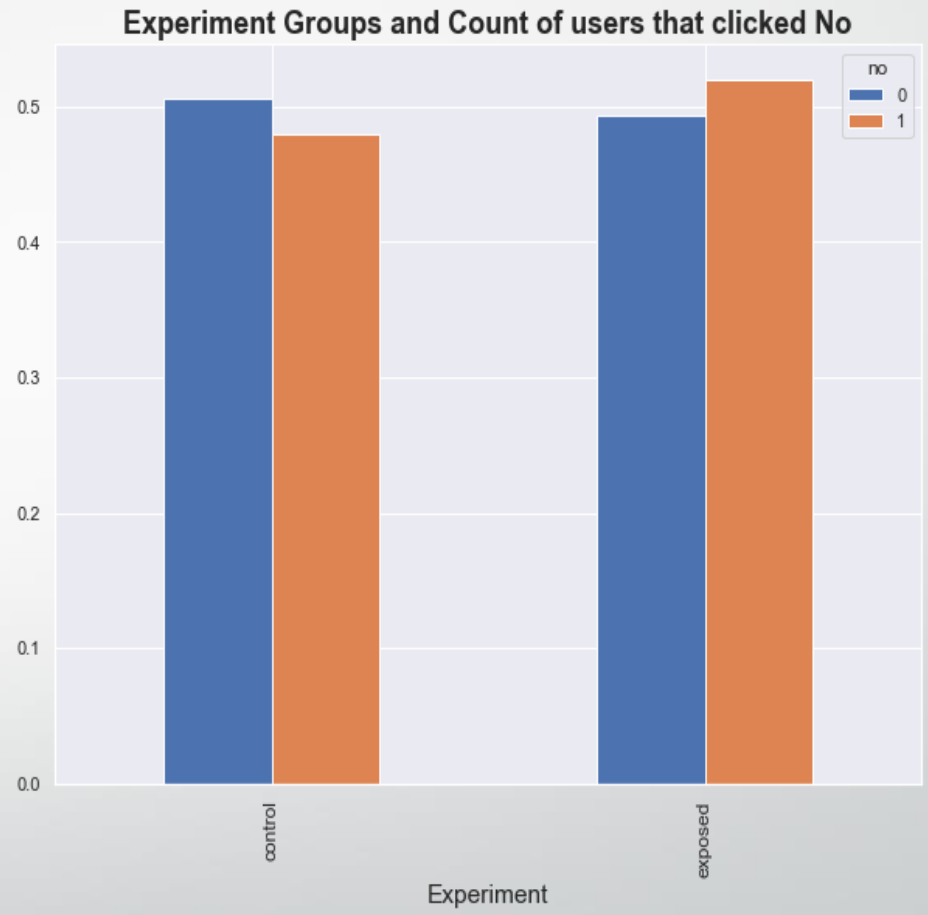
Users were evenly distributed between the groups



Most of those that access the Questionnaire were Chrome Mobile users

Exploration cont'd

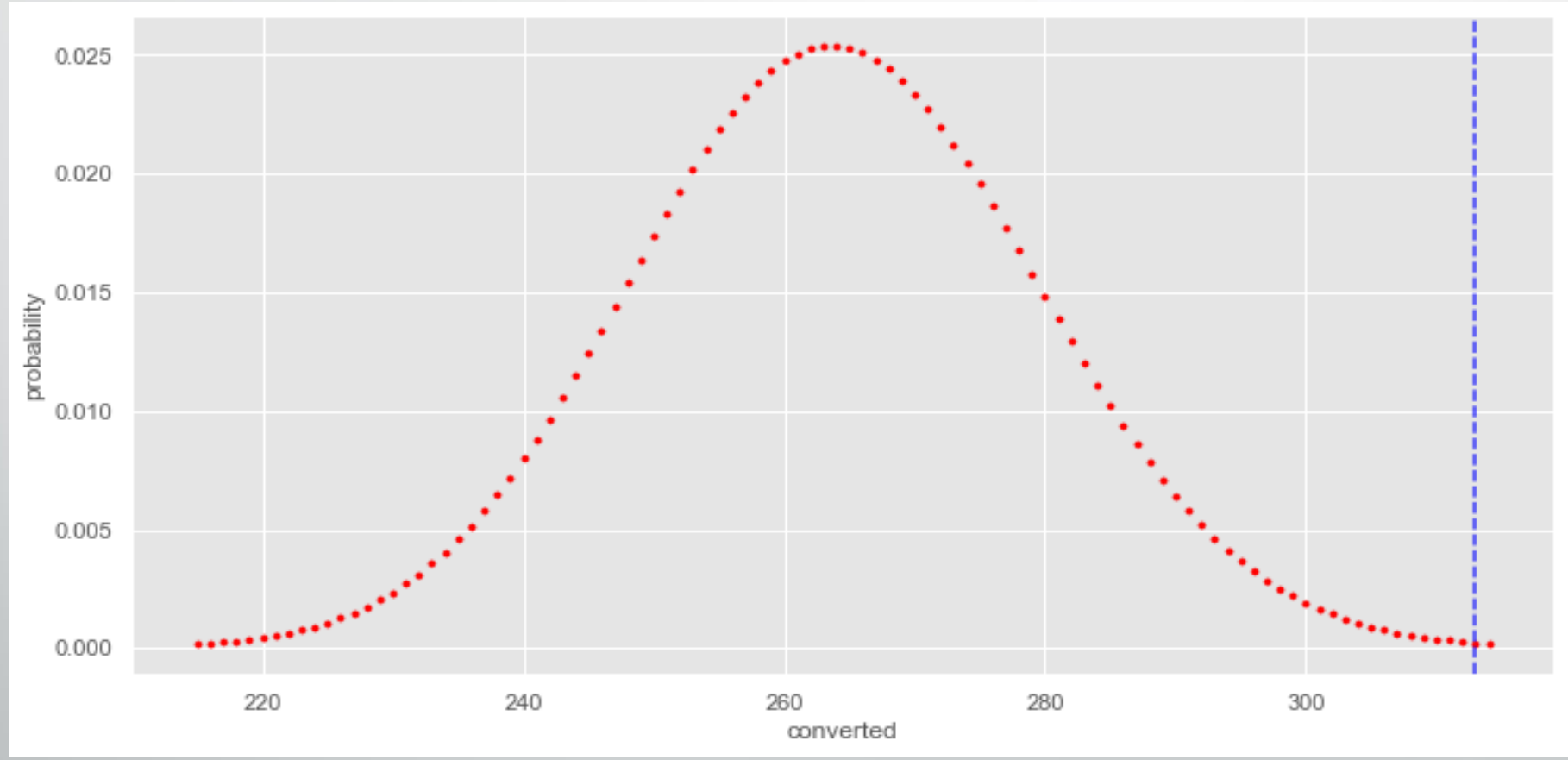




From the charts above, number of users that know about SmartAd in exposed group is more than those that do not.

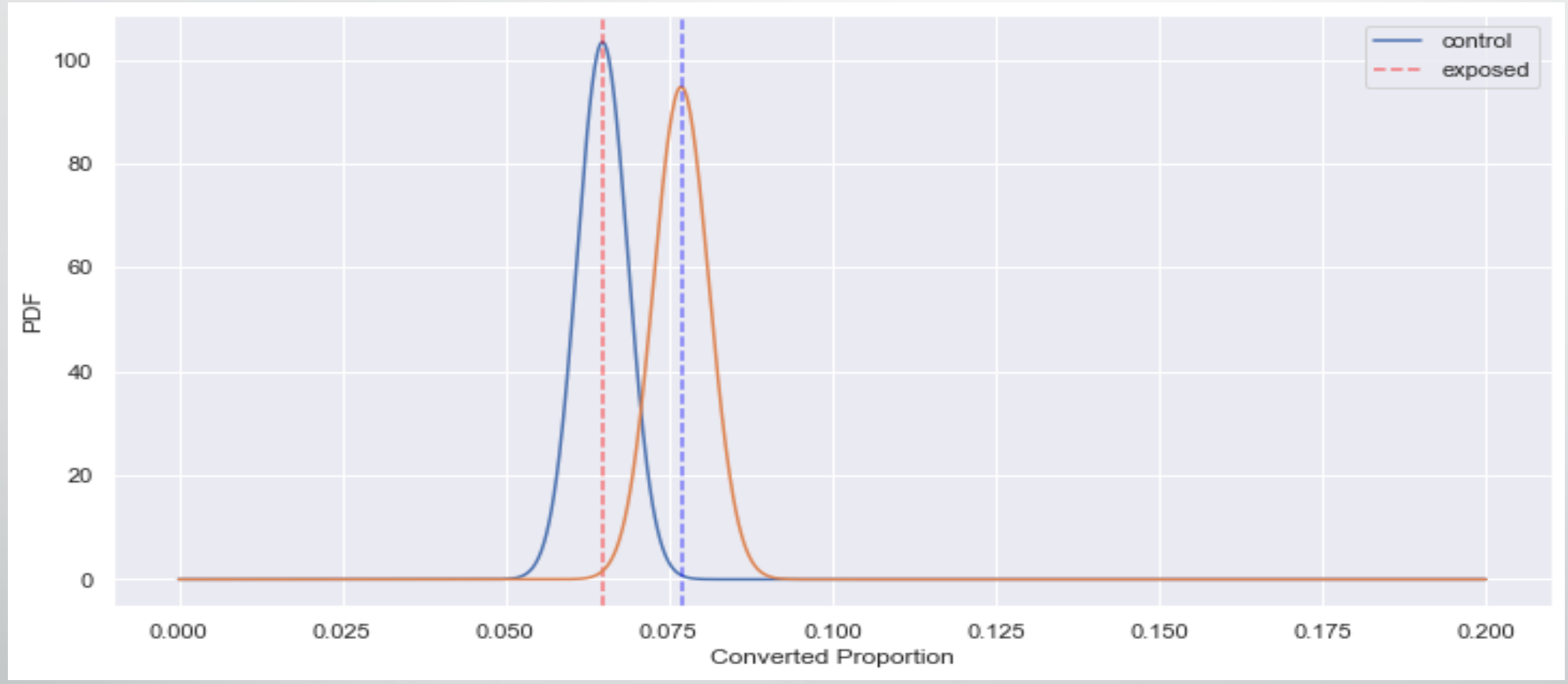
Hypothesis Testing

Null Hypothesis: There is no difference in brand awareness between the two groups
Alternate Hypothesis: There is difference in brand awareness between the two groups



The critical value distribution plot has shown that the null hypothesis is to be rejected as the p-value of approximately 0.03 is less than the critical value of 0.05

Expected Minimum Change between the two groups



The dashed lines represent the mean click rate for both the control and exposed group. The distance between the red dashed line and the blue dashed line is equal to d_hat (Expected Minimum Change).

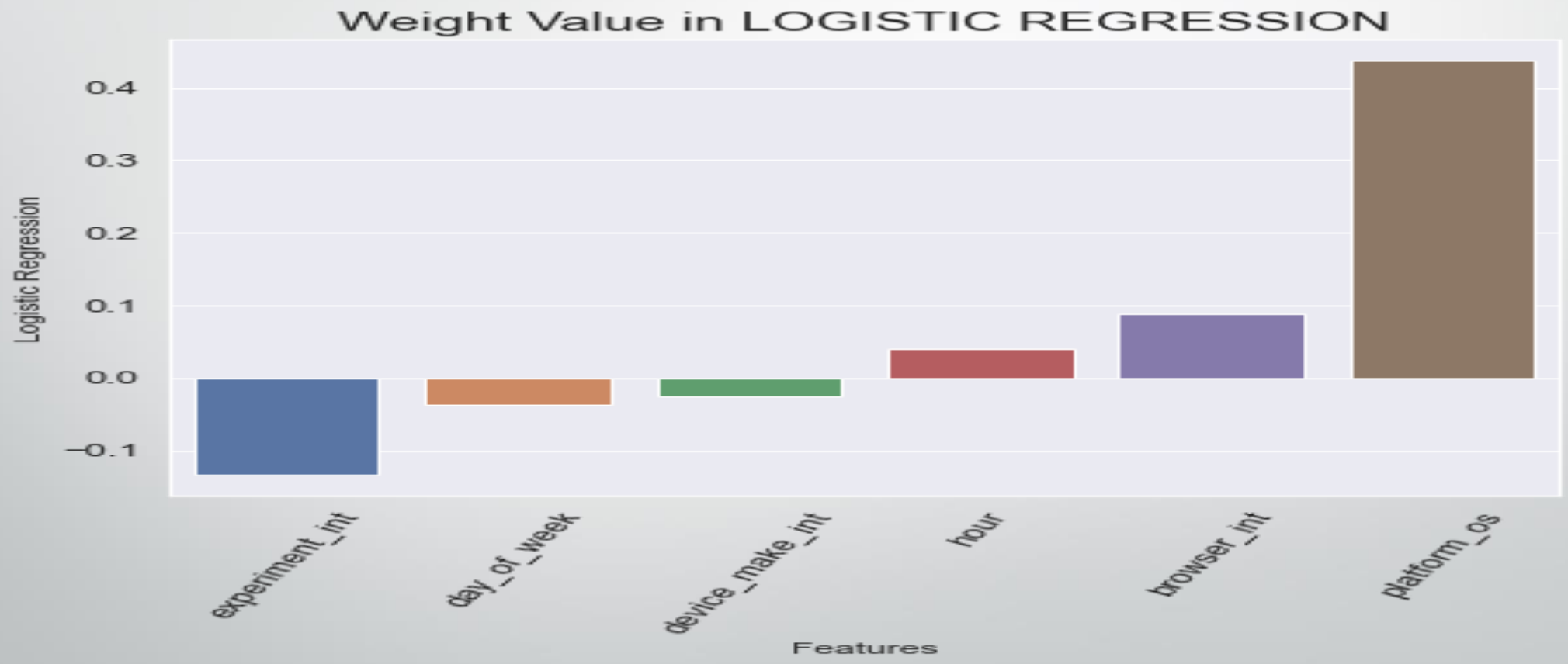
Problem Formulation

Given the data to an experiment ran by SmartAd to verify the likelihood of users to click yes to the Question: Do you Know SmartAd? Build a Machine Learning Model to predict the likelihood of a user to click the yes radio button in answering the question.

- Target Column: yes
- Model Loss Function = LogLoss

Machine Learning Models

Logistic Regression Model

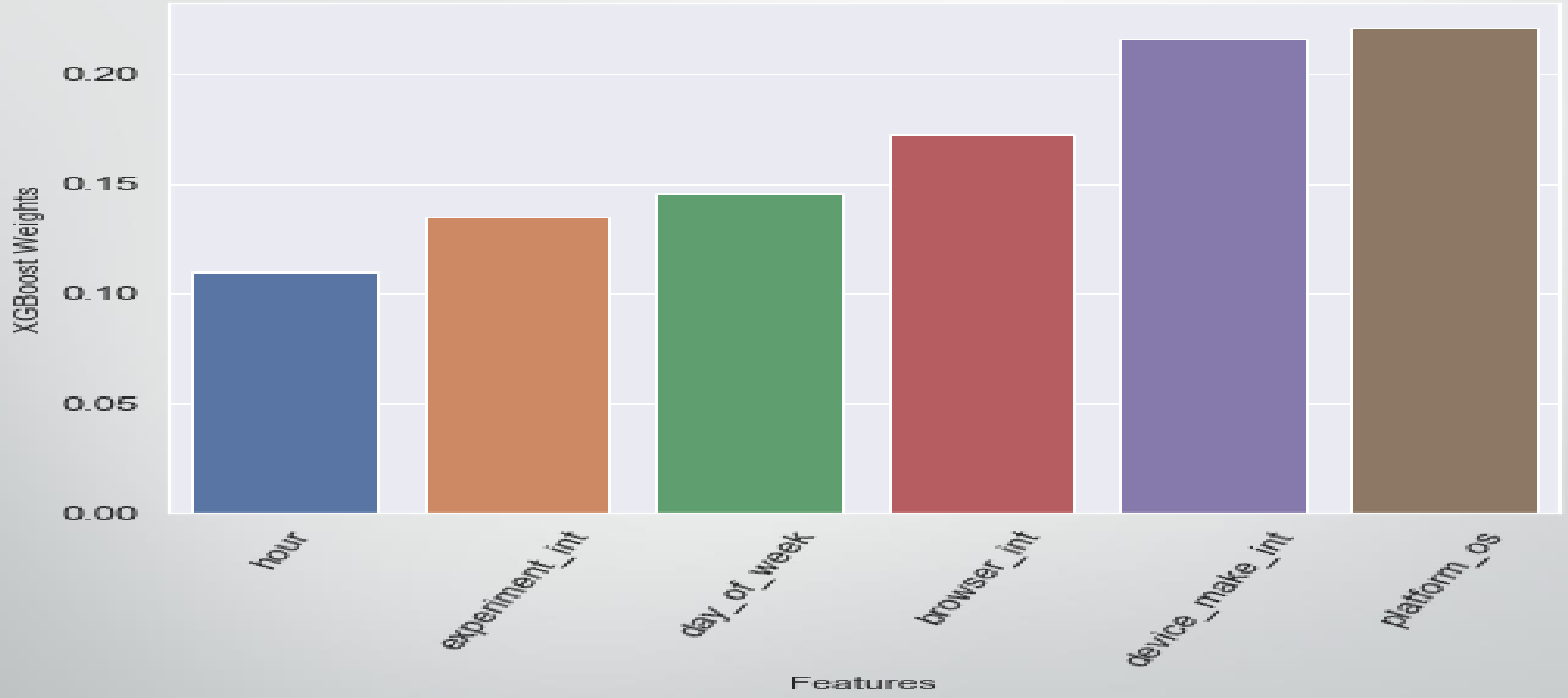


Platform OS is the most important predictor among the features in the data. Experiment has a negative weight value for this model

Decision Tree Classifier model

Hour and Device make are the most important predictors among the features in the data while experiment is the least important

Extreme Gradient Boosting Classifier (XGBoost)



Platform OS and device make are the most important predictors among the features in the data.

Comparison of the two methods

- Classical A/B Hypothesis Testing for Brand Impact Optimizer has revealed that the recent advertising campaign has resulted in a significant brand awareness while the Machine Learning Models have revealed the factors affecting the awareness of SmartAd brand.
- Both classical A/B testing and Machine learning models are based on probability.
- Machine Learning Models provide us with information about magnitudes of each determinant while classical A/B testing provides only test statistics and p-value.

Overall Results

- The Ad that SmartAd runs has resulted into a significant lift in brand awareness as the Null hypothesis was rejected.
- Platform Operating System and Device Make are the most important determinants of brand awareness
- Experiment, Day of week and hour are of little or no significance to the brand awareness.

Recommendations

1. Future Adverts should take into consideration Platform Operating System
2. Different Adverts should be device specific

Limitations

1. Hypothesis Testing is not enough to determine the magnitude of the impact the Ad has on awareness

References

- https://www.kaggle.com/cstorm3000/frequestist-a-b-testing-from-scratch
- https://github.com/mnguyenngo/ab-framework/blob/master/src/plot.py
- https://medium.com/analytics-vidhya/a-b-testing-clearly-explained-56488430156