Project Proposal Team 2 Project 4

Nick Mincucci

George Achie

Ian Cook

**Title**

Which Operating System is best?

**Project Description**

The objective of this project is to create a predictive model utilizing the dataset of 700 entries containing information on users’ demographics, device usage patterns, and behavioral classifications. Key features include device models, operating systems, app usage time, battery consumption, data usage, and user demographics (age, gender).

**Research Question**

* Which operating system (Android or iOS) is best for battery life?
* The number of open applications compared to screens on time.

**Data Set to be Used**

Mobile Device Usage and User Behavior Dataset

<https://www.kaggle.com/datasets/valakhorasani/mobile-device-usage-and-user-behavior-dataset?resource>

The dataset to be used contains mobile device usage patterns and user behavior classifications of over 700 users.

**Breakdown of Tasks**

Data Cleaning/Filtering

* Remove duplicates
* Handle missing values
* Standardized data

Data Initialization/Training/Evaluated

* Python script to initialize, train, and evaluate the model
* Utilize SQL or Spark to retrieve the data
* Predictive measure of 75% classification accuracy is met

Data Model Optimization

* Iterative changes utilized to improve the model and create change to the performance
* Documentation of the model and optimization/evaluation process is included

Visualization Development

* Develop appropriate visualizations to show the results of the data and modeling
* Utilize multiple techniques such as Matplotlib and Tableau

Presentation

* Design and implement a 15 minute presentation displaying the visuals and model in an informative fashion. Allowing the audience to both understand the model while seeing appropriate visualizations.

**Ethical Considerations**

Throughout the project, ethical considerations will be considered, particularly concerning data privacy and representation. All personal data will be anonymized, and the visualizations will aim to present data accurately without misleading interpretations.