## Initial plan

Predicting online shoppers' purchasing intention using machine learning

Author: Akash Deoraj Supervisor: Yuhua Li

## **Description:**

Retail shopping is continuing to shift to E-commerce shopping and as a result the dynamics of shopping is changing around the world. E-commerce has already become a major form of retail market. Online customers often browse pages of e-commerce sites before they place orders or abandon their browsing without purchase. This information can help businesses to better cater to customer preferences and help both the business and customers mutually by recommending products specific to each customer and therefore increasing sales for the businesses. However, most of the time customers visiting these online websites may not make any purchase at all. This could be for various reasons i.e., Price of product or window shopping. It is important to predict customers' purchasing intention so that retention measures like e.g., recommending suitable products can be taken to convert potential customers into purchasers. This project aims to use the information customers may leave in the form of the trace of browsing history data or user information when they visit an online shopping site. With the help of this information, the project aims to predict online shoppers' purchasing intention by using clickstream and session information data. The project aims to create a machine learning model based on this information in order to predict customer's purchasing intension.

Data will be used from this repository to build Machine Learning model:

https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset.

## **Objective:**

- **Background Reading:** Background Reading consists of reviewing Machine learning concepts and studying its application in the context of this project.
- Initial Data Set and Exploratory Data Analysis: Analysing data in order to gain a preliminary understanding and acquaintance with the dataset. Performing Descriptive statistics for mean, medium, mode, standard deviation and Data shaping for Pivoting data, grouping data, filtering data, etc.
- Data Modelling and Pre-processing: data is subjected to various checks and scrutiny in order to remedy issues of missing values, spelling errors, normalizing/standardizing values such that they are comparable, transforming data. This is to improve the quality of data and therefore the accuracy of the model.
- Implementation: Building the Model. This includes application the correct Machine Learning Algorithm i.e., Supervised learning, Unsupervised learning and Reinforcement learning. Hyperparameter optimization- tweaking parameters of the machine learning algorithm that directly impacts the learning process and prediction performance and feature selection- the process of selecting a subset of features from an initially large volume of features. Aside from achieving highly accurate models, one of the most important aspect of machine learning model building is to obtain actionable insights and in order to achieve that it is important to be able to select a subset of important features from the vast number.
- **Model performance:** Evaluation of the performance of regression models are performed to assess the degree at which a fitted model can accurately predict the values of input data.
- **Final Reporting:** Writing the final report and reviewing the work.

## Gantt chart

	Weeks are based on second Semester calender dates												
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	_
													_
Background Reading													_
													_
Intial Dataset and Exploratory Data Analysis													_
													_
Data Modelling and Pre-processing													+
													+
Implemenation													+
Madel motormance & Adjustments													+
Model performance & Adjuctments													+
Poviow (Open a work)													
Review (Office a week)													
Final Reporting													+