Milestone 1

DSC550-T301 Data Mining

Joshua Greenert

10/8/2022

**Introduction**

In a data-driven world, consumerism reigns supreme. Users, nowadays, can order products through several different applications from the safety of their home, or wherever they may be. For companies, this surge in eCommerce provides an opportunity for significant growth given the right direction and appropriate action. However, to provide the proper guidance and direction for businesses to take they must understand when to advertise to achieve the maximum amount of revenue from their customers. When is the best time to advertise to customers? Further, when is the best time to advertise to online customers to get the most conversions? This paper will address that concern by utilizing data from 12,330 unique user sessions to create graphical visuals that explain online behavior

.

**Methodology**

To discover insights into the eCommerce world, we will be reviewing data provided by the UCI Machine Learning Repository. The data will be cleaned to remove unnecessary columns, convert strings to Boolean binary integers, and many other standard preparation practices. Afterwards, the columns within the dataset will be checked for statistical significance to determine the best features to train the machine learning model upon, and what data can be considered the target goal. From the information gleaned from this process, a model will be created to suggest to businesses when the best time — days/months/timeline — would be the most effective for online sale conversions using advertisements. After several iterations of this model building process, a final direction will be provided to stakeholders to give insight on how to properly develop their eCommerce strategies and allocate the appropriate funding to support it.