**The XGBoost Boosting Algorithm and an Application to the Forest Cover Type Dataset**

**Matthias Quinn**

**Matthew Brigham**

**STA 685 - Advanced Data Models**

**Fall 2021**

**Abstract**

This project will focus on the famous XGBoost system and its application on a moderately large dataset. Starting with a history of the system, then exploring the algorithm itself, and finally ending with an application to forest cover types, this project will hopefully provide a framework on which to base further research and applications. The results show that XGBoost outperformed all other models in terms of both accuracy and computation time. These results demonstrate the predictive accuracy and scalability when working with large datasets.