

This thesis deals with comparing two approaches to modelling and predicting time series: a traditional one (the ARIMAX model) and a modern one (gradiently boosted decision trees within the framework of the XGBoost library). In the first part of the thesis we introduce the theoretical framework of supervised learning, the ARIMAX model and gradient boosting in the context of decision trees. In the second part we fit the ARIMAX and XGBoost models which both predict a specific time series, the daily volume of the S&P 500 index, which is a crucial task in many branches. After that we compare the results of the two approaches, we describe the advantages of the XGBoost model, which presumably lead to its better results in this specific simulation study and we show the importance of hyperparameter optimization. Afterwards, we compare the practicality of the methods, especially in regards to their computational demands. In the last part of the thesis, a hybrid model theory is derived and algorithms to get the optimal hybrid model are proposed. These algorithms are then used for the mentioned prediction problem. The optimal hybrid model combines ARIMAX and XGBoost models and performs better than each of the individual models on its own.