

Executive Summary

Sponsor Name: Paychex

Team Members: Shijing Li, Lingyu Ye, Yuan Wang and Yangxin Fan

In this project, we identify the external activities that have significant impacts on Paychex's revenue to build prediction models that will forecast Paychex's future monthly revenue. In the end, we choose the best performed model by using predefined performance metrics and we deliver constructive suggestions for Paychex to increase their future revenue. In terms of methodology, we perform exploratory data analysis, preliminary feature selection, and feature engineering to extract correlated features. Then, we input correlated features into our prediction models starting with SARIMA, build up to hybrid, GAM, and deep learning model --- LSTM. To produce better results, we optimize each model with recursive feature engineering and hyperparameter tuning. We choose our best model based on the lowest MAPE and the best model we get is the FB-Prophet with a MAPE of 4.9% which can be interpreted as highly accurate based on the 5% threshold.