Ray Xia

8 August 2023

CSS 497 B

Professor Retik

Machine Learning Model for Price Markup Tool

At PACCAR, my capstone internship project was focused on creating a machine learning solution to estimate a dealer’s markup in price when selling a truck to a customer. A PACCAR Financial Corporation (PFC) salesperson currently has to take time to manually assess many factors such as buyer history, truck model, credit score, truck condition, and more to make an estimation of the dealer markup when selling a truck to a dealer. To maximize profits, PFC must understand a dealer’s markup to know how much to sell trucks to the dealer. My project automates this process with machine learning and simplifies a salesperson’s job.

I utilized many tools for my project. I programmed in Python, using different tools such as Jupyter Notebook, scikit-learn, pandas, and ydata-profiling for my data exploration and cleaning. After cleaning my data, I selected important variables to make my RandomForestRegressor model and performed hyper tuning to increase accuracy of the model.

After my model was complete, it was time for deployment. I used DVC (Data Version Control) for data source control and used GitHub Actions to automate my training, evaluation, and deployment of the machine learning model. I used FastAPI to make an API that accepts parameters to feed to my machine learning model and output an estimated dealer markup. Using GitHub Actions workflows, I was able to automatically create a Docker image of my application and store it on Google Cloud Platform’s (GCP) Artifact Registry, where GCP Cloud Run would use that image to have an accessible web API for PFC.

My project is still under development, but when my model is ready for production, it will be able to support PFC when conducting sales to dealers. My API will save time and also assist in price negotiations with dealers, with a likeliness in increasing profit margins,