

TIME SERIES BASED MACHINE LEARNING PIPELINE TO PREDICT CUSTOMERS' SHIPMENT ORDER TREND

Tharuka Dhananjalee Silva

A dissertation submitted in partial fulfillment of the requirement for
BEng (Hons) Software Engineering Degree in Computer Science

Department of Computing

Informatics Institute of Technology, Sri Lanka

In Collaboration with

University of Westminster, UK

2021

Abstract

Courier shipment industry is widely popular in the international trade due to the fact that anyone could order something over internet and receive the shipment at their doorstep without breaking a sweat. Clients expect their packages to be received at the same day they shipped to the country. Successful courier services who have more than enough delivery resources can fulfill the client's expectations without a problem. Yet, the small scale courier services and startups are more than happy to manage their limited delivery resources to provide for the same customer satisfaction to grow their business. The purpose and the importance of this research is to find an efficient method for the courier services with limited resources on how to manage their delivery resources efficiently for better business.

Machine learning, one of the most significant components in artificial intelligence, enables machines to self-learn from data and perform tasks without human intervention or explicitly programming. Vary in type, machine learning could be employed in many strategies. Time series forecasting is the process of observing the particular time intervals in historical values to extract meaningful statistics. A time series based machine learning model can provide the future predictions by analyzing a dataset. This dissertation is a result of the project to implement a time series based machine learning pipeline to predict the shipment orders received to a courier service and allocate resources for the distribution of the shipments, based on the prediction.

The machine learning pipeline provides time series based machine learning model implemented on LSTM (Long Short-Term Memory networks) algorithm to forecast future predictions. Predicting the shipment trend of shipments a courier service receives in order to allocate the delivery resources and driver allocations are results produced by this dissertation. Compared to the similar systems existing but belongs to different domains, it performs equally or better. The system is available as an API which could be integrated with exiting operative courier systems, and as per the testing purposes, a web graphical interface is facilitated for the user before taking decisions on integration.

Keywords: Machine learning, Machine learning pipeline, Time series forecasting, Algorithm selection, python.