MACHINE LEARNING BASED MODELING APPROACH TO PREDICT DELIVERY PROMISE TIME FOR E-COMMERCE

Irippuge Supun Sharmal Fernando

219332D

Master of Science in Computer Science

Department of Computer Science Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka
July 2023

MACHINE LEARNING BASED MODELING APPROACH TO PREDICT DELIVERY PROMISE TIME FOR E-COMMERCE

Irippuge Supun Sharmal Fernando

219332D

Thesis/Dissertation submitted in partial fulfilment of the requirements for the degree Master of Science in Computer Science

Department of Computer Science Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka
July 2023

DECLARATION

"I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 24/07/2023

The supervisor/s should certify the thesis/dissertation with the following declaration.

The above candidate has carried out research for the Masters/MPhil/PhD thesis/ Dissertation under my supervision.

Name of the supervisor: Dr. Sapumal Ahangama

Signature of the supervisor: Date: 24/07/2023

ACKNOWLEDGEMENT

I would like to express my appreciation to Dr. Sapumal Ahangama, who served as my research supervisor, for providing valuable guidance and unwavering support throughout this research endeavour.

Additionally, I am grateful to the faculty members of the Department of Computer Science Engineering for their encouragement, support, and guidance that have enabled me to fully utilize my knowledge and capabilities.

I owe a debt of gratitude to my parents who provided me with the necessary space and facilities to work hard on this research project. Additionally, I am grateful to my family for their assistance whenever I needed it.

Lastly, I would like to extend a special thanks to all of my fellow batchmates for their supportive camaraderie that has played an integral role in the successful completion of this research.

ABSTRACT

To smoothly function an online retail business, provision of reliable and fast delivery service is essential. A customer's buying decision on an ecommerce platform is discouraged due to erroneous promise timelines defined at the point of order placement. The ecommerce industry is under increasing pressure to create a data-driven framework that can accurately forecast the distribution of delivery time and set precise promise delivery times. By exploiting the relationships between delivery time and relevant predictor variables, developed tree based model, quantile regression, quantile regression forest model, LightGBM trained with quantile regression model, and LSTM deep learning model to generate conditional distribution forecasts. The models are evaluated using two methods of scoring rules namely pinball loss function and continuous ranked probability score (CRPS). LSTM model incorporated with quantile regression generated best results out of all developed, trained and tested from real world dataset obtained from ABC.com - Sri Lanka's largest ecommerce platform, where multi-origin marketplace seller fulfillment was considered as the novelty contribution. Furthermore, it was revealed that hour of order placement and origin to destination combination impacted most on delivery time and highlighted the importance of integrating the forecasting model to decision making structure. As a limitation of this study and eligible future work, forecasting the delivery time with most recent order distribution restricting to the ongoing logistics capacity utilizations can be stated.And also consideration on incorporating real-time data and cost-sensitive approaches to improve the practicality and effectiveness of forecasting in e-commerce businesses.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES