INTELLIGENT DATA-DRIVEN METHODS FOR DEMAND AND PRICE PREDICTION IN THE SHIPPING INDUSTRY

Ayesha Ubaid

School of Computer Science

Faculty of Engineering and Information Technology

University of Technology Sydney, NSW, Australia

A Thesis submitted in fulfilment

Of the requirements of the degree of

Doctor of Philosophy

May 10th 2021

Certificate of Original Authorship

I, ***Ayesha Ubaid*** declare that this thesis, is submitted in fulfilment of the requirements for the award of ***Doctor of Philosophy*****, in the *****FEIT***** at the University of Technology Sydney.**

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 10th May 2021

SYNOPSIS

Machine Learning has found its applications in many industrial and commercial domains. However, there are few ignorant industries that still lack digitization and require fusion of AI into their processes. Once such industry is the container-shipping industry. The global supply chain is complex, with cargo volumes that are highly seasonal driven by events, consumer related, Christmas and Chinese New Year, agriculture harvests further impacted by extreme weather occurrences and changes to the Geo, Political regulatory environment affecting trade. In contrast, supply, the shipping capacity is fixed in the short run; this results in periods of mismatched supply and demand and therefore shipping price volatility. The Shipping lines are trapped by the current spot market pricing practice of using date validity and not the vessel voyage. This causes a disconnect between price, demand and supply, a problem that is compounded by the shipping line's enterprise systems. Entrenched operational silos within the lines are resulting in missed revenue opportunities through a lack of real time visibility into the availability of equipment and vessel space which are key inputs to price decisions. Forty percent of all shipping containers moved around the world are purchased on the short-term spot market; their commercial terms are set manually with emails and phone calls. Spot market price should be a product of supply and demand. However, the industry as a whole has little visibility into the state of the market in real time, and carriers are making sub optimal pricing decisions and their customer procurement decisions because of this.

Australian Linear Shipping industry is lacking digitization hence the visibility into the industry statistics is missing. Without real time visibility of the market as a whole, future spot pricing decisions are based mainly on a carrier's internal assessment of current and future booking build up and net contribution targets. This can, and usually does, present a different and misleading picture of market conditions. With the limited market wide information and inability to use vessel voyage specific price as a lever to steer the cargo opportunity to the optimum vessel sailing, opportunity cost materializes.

In this research, we want to empower Australian Container Shipping Industry with machine learning capabilities to provide a market wide view of current and future demand (both for short-term and long-term), optimal spot pricing model and machine learning based prices prediction model that can predict spot pricing based on current demand and available capacity (supply).

To do so, the first step is to explore and model the relationship between demand, supply, and price. This can provide the current market scenario in which industry is operating. Based on the results inferred, a model is designed to set optimal spot pricing. In order to price effective spot pricing model, statistical analysis is done to discover the relationship between demand, capacity and price. Based on inferred results, historic data and spinning companies' limitations for price quotation, a novel mathematical model is designed that can calculate optimal pricing based on the mentioned factors. Finally, for making contract pricing more effective, the demand forecast performed earlier is used with the available capacity to predict optimal container pricing based on demand and supply using a regression based multivariate machine learning model. This predicted price will shorten the gap between contract and spot pricing in the container shipping industry. Hence, this research provides visibility into future demand and will allow shipping lines and their customers to make better-informed pricing and procurement decisions.

ACKNOWLEDGEMENTS

I would like to acknowledge the blessing of Almighty Allah, the most gracious and beneficent for giving me this opportunity to complete my Doctoral Dissertation under their supervision of Associate Professor Farookh Khadeer Hussain and Associate Professor Asif Qamar Gill.

Additionally, I would like to acknowledge the efforts of my family specially my husband Ubaid and mother, Shaheen without whose support and prayers I would have never been the person I am today. A special thanks to my kids, Suleman and Mustafa who despite their quite young age of 5 and 3 were able to understand that their mum has to study. Their positive and helping attitude made things easy for me. To all of my family, I hope this thesis completion will bring positive changes in our lives.

Moreover, I would like to thank my supervisors whose trust and faith in me made me achieve this goal quite comfortably. The never-ending support, compassion, encouragement and superb guidance are the key to my success. This thesis is as much of their work as mine. Finally, I would like to thank the Centre of Artificial Intelligence (CAI) for supporting me throughout this whole three-year tenure.

I would like to dedicate this thesis to my supervisor Dr. Farookh Khadeer Hussain and my beloved husband Ubaid-UR-Rehman who are my real mentors.

LIST OF PUBLICATIONS

JOURNAL PAPER

- Modeling Shipment Spot Pricing in the Australian Container Shipping Industry: Case of ASIA-OCEANIA trade lane Research Question 1

- Container Shipment Demand Forecasting in the Australian Shipping Industry: Case Study of Asia-Oceania Trade Lane Research question 2 and 3 (under review)

CONFERENCE PAPERS

- Framework for feature selection in health assessment systems. Research Question 1.

- Machine Learning-Based Regression Models for Price Prediction in the Australian Container Shipping Industry: Case Study of Asia-Oceania Trade Lane. Research Question 4.

TABLE OF CONTENTS

Table of Figures

Table of Tables

LIST OF ACRONYMS

ML	Machine learning
UI	User interface
DS	Dataset
FS	Feature selection
LR	Linear R
NNs	Neural networks
AR	Autoregressive
SVR	Support vector regressor
SVM	Support vector machines
GBR	Gradient boosted regression
RFR	Random forecast regressor
RMSE	Root mean squared error
R2 SCORE	R square score
EDA	Exploratory data analysis
MAPE	Mean absolute percentage error
LSTM	Long-term Short-term memory
ARIMA	Autoregressive moving average
SARIMA	Seasonal Autoregressive moving average
PACF	Partial auto correlation function