MULTI-MODAL TRANSPORT DATA INTEGRATION AND ANALYTICS PLATFORM

Haputhanthrige Dilantha Ishara Haputhanthri

(208036G)

Degree of Master of Science

Department of Computer Science & Engineering

University of Moratuwa
Sri Lanka

March 2021

MULTI-MODAL TRANSPORT DATA INTEGRATION AND ANALYTICS PLATFORM

Haputhanthrige Dilantha Ishara Haputhanthri

(208036G)

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science and Engineering

Department of Computer Science & Engineering

University of Moratuwa
Sri Lanka

March 2021

DECLARATION

I, Dilantha Haputhanthri, declare that this is my own work, and this 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 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: 2021-10-06

The above candidate has carried out research for the master's thesis/dissertation under my supervision.

Name of Supervisor: Dr. M.P.A.P. Wijayasiri

06.10.2021

Signature of the Supervisor: Date:

ABSTRACT

Transportation is one of the crucial areas that needs to be optimized by the officials because of the increase in the demand for efficient travel and transportation due to the rapid urbanization. Integration of data from different sources has been explored and introduced as a method to address cross-domain problems like managing assets and resources efficiently. Several data integration methods have been proposed over the years, but the utilization of microservices architecture has been rare, especially in the transportation field. Microservices architecture, supported by container orchestration can be used to realize high availability, scalability, and low-cost operations. In this research, a microservices-based data integration platform was proposed to meet the demand for transportation related data integration. The proposed solution supports data importing, storing, processing, analysis and exporting of several data formats and types. A performance analysis was done to measure the scalability of the platform, accomplished utilizing the container orchestration. A real-world dataset and an experimental setup, hosted on a public cloud were employed for the analysis. The analysis demonstrates that the platform can manage around 500 RPS with a substantially low response time when auto-scaling is enabled. Finally, an approach for transportation mode detection, a use case scenario of the platform is briefly presented.

As an analytics example, another research is done for short-term traffic volume forecasting. Accurate short-term traffic volume forecasting has become an important element in traffic management in intelligent transportation systems. A significant amount of literature can be found on short-term traffic forecasting based on traditional learning approaches, however deep learning based solutions have also produced substantial strides in recent years. In this paper, we propose several long-short term memory (LSTM) based deep learning models to extract inherent temporal and spatial features for traffic volume forecasting. A standard LSTM model, LSTM encoder-decoder model, CNN-LSTM model and a Conv-LSTM model were designed, optimized, and evaluated using a real-world traffic volume dataset for multiple prediction horizons. The experimental results shows that the Conv-LSTM model produced the best performance for the prediction horizon of 15 minutes with a MAPE of 9.03%. At the same time, one of the novelties of the research is the forecasting during the traffic volume anomalies due to the Covid-19.

Keywords: data integration, encoder-decoder, LSTM, microservices, traffic volume forecasting

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Dr. Adeesha Wijayasiri for the continuous support given for my MSc study and research, for his inspiration, and expertise. I am extremely thankful for your advice without which, the present work could not have achieved this point. I could not have imagined a better supervisor and a mentor for my MSc study.

Also, I would like to thank Dr. Dimantha De Silva and Dr. Kutila Gunasekara for their valuable insights and guidance from the early stage of this research. I would like to thank the staff of the Department of Computer Science and Engineering for the support during the research. This research was supported by the Accelerating Higher Education Expansion and Development (AHEAD) Operation of the Ministry of Higher Education, Sri Lanka funded by the World Bank. Also, I would like to acknowledge the support received from the University of Moratuwa Senate Research Grant.

Finally, I would like to thank my family and friends for all the love and support.

Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
API	Application Programming Interface
ARIMA	Auto Regressive Integrated Moving Average
CNN	Convolutional Neural Network
JSON	JavaScript Object Notation
JWT	JSON Web Token
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Square Error
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SDA	Stack Denoise Encoder
SQL	Structured Query Language
SVR	Support Vector Regression