DATA ANOMALY DETECTION IN MARITIME TRAFFIC MANAGEMENT USING DEEP LEARNING

RAMSHA MASHOOD
02-241191-015

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTERS OF SCIENCE (SOFTWARE ENGINEERING)

DEPARTMENT OF SOFTWARE ENGINEERING

BAHRIA UNIVERSITY KARACHI CAMPUS

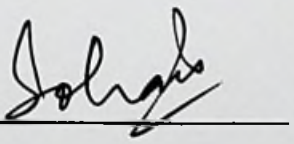
FALL 2021

DECLARATION

Approval for Examination

Scholar's Name: **Engr.Ramsha Mashood**

Registration No. **57901** Program of Study: **MS (Software Engineering)** Thesis Title: **Data Anomaly Detection in Maritime Traffic Management using Deep Learning** It is to certify that the above scholar's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for examination. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index_____% that is within the permissible limit set by the HEC for the MS degree thesis. I have also found the thesis in a format recognized by the BU for the MS thesis.

Principal Supervisor's Signature: 

Date: **3rd February, 2022**

Name: **Dr.Sohaib Ahmed**

Author's Declaration

I, **Engr. Ramsha Mashood** hereby state that my MS thesis titled "**Data Anomaly Detection in Maritime Traffic Management using Deep Learning** is my own work and has not been submitted previously by me for taking any degree from this university **Bahria University Karachi** or anywhere else in the country/world. At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw/cancel my MS degree.

Name of scholar: **Engr. Ramsha Mashood**

Date: **3rd February, 2022**

ACKNOWLEDGMENT

My highest gratefulness is to Allah s.w.t, the one who is great and merciful. Allah s.w.t always opened doors for me which were closed, and I am always in great need of his mercy. Thank my mentor, guide, and supervisor, Dr.Sohaib Ahmed (Principal School of Engineering and Applied Sciences), for his guidance, quick response, and kindness throughout my research. His humble concern, research efforts, and research work are my real motivation. I want to pay special thanks to Dr.Osama Rehman (HOD Software Engineering) for encouraging me to complete my task on time. I am also very thankful to the faculty members, lab Engineers with other non-teaching staff for helping me in different scenarios.

I am also thankful to my colleagues Engr.Ayesha Khan for her moral support and help. While thanks to my other friends, Tooba Zahid, Tooba Mehtab, and Monaliza, for their support in thesis Writing.

My sincere thanks to my parents, who made me what I am today and supported me throughout this research. Thanks to my in-laws, who motivate me during this journey. Last but not least, I would like to express my deep love to my husband and my son, whose genuine love, invaluable concern, and sacrifices surpass all the others. Respect and prayers for my other family members and for all those there to motivate me and encourage me to complete my research.

ABSTRACT

Pandemics and the current world situation affected overall world business. The maritime industry currently depends on different vessel management systems that use Automatic Identification System (AIS) messages to manage Ship activities. AIS messages include both Static and Dynamic data, which contains information related to the position, Heading, and Ship features as AIS data is in massive volume, so the possibility of missing data and anomalous information is present. Different Machine Learning and deep learning algorithms are used to train and detect anomalies to identify those anomalies related to ship classification and ship messages.

CCN, SVM, Random Forest and other deep learning algorithm are used to identify anomalies regarding the position and ship flows. While SQL, PostGre are used for storage purposes. The main goal is to identify which Database is best to store AIS data and which deep learning algorithm performs better to identify anomalies in Ship classification and transceiver class data.

This thesis proposes a solution regarding the storage of AIS data to MongoDB database and identifying anomalies in ship classification and Transceiver class using Deep Learning Algorithm. This approach includes Long Short-Term Memory (LSTM), Convolution Neural Network (CNN), and Artificial Neural Network (ANN) deep Learning algorithm for identifying anomalies in Maritime Traffic Management System.

So, we can conclude that the CNN model provides the highest accuracy, approximately 81%, while other deep learning algorithms like LSTM have 72% and ANN have 79.95% accuracy. The total time consumed to get 1658575 rows takes approximately 0.04sec using MongoDB cloud Database.

TABLE OF CONTENTS