UNIVERSITI TEKNOLOGI MARA

IMPLEMENTING PREDICTIVE ANALYTICS FOR WATER TREATMENT PROCESS OPTIMIZATION AT SEBATIK ISLAND

MOHAMAD SHAHRINAS BIN ASSAN

BACHELOR OF INFORMATION SYSTEMS

(HONS.) BUSINESS COMPUTING

JULY 2026

TABLE OF CONTENT

CONTENT		PAGE
SUPERVISOR APPROVAL		i
STUDENT DECLARATION		ii
ACKNOWLEDGEMENT		iii
ABSTRACT		iv
LIST OF FIGURES		ix
LIST OF TABLES		xii
LIST OF ABBREVIATIONS		xiv
CHAPTER ONE: INTRODUCTION		
1.1	Background of Study	1
1.2	Current Process	3
1.3	Problem Statement	5
1.4	Objectives	7
1.5	Scope	7
1.6	Significance	9
1.7	Project Framework	10
1.8	Gantt Chart	13
1.9	Conclusion	13
CHAPTER TWO: LITERATURE REVIEW		
2.1	Introduction	15
2.2	Water Treatment Process	15
2.1.1	Overview of Water Treatment Processes	16
2.1.2	Operational Challenges in Water Treatment Plants	16
2.1.3	Water Treatment Requirement of Process Optimization	17
2.1.4	AIoT in Water Treatment and Monitoring	17
2.3	Water Level Conditions in Water Treatment Process	18

2.3.1 Definition and types of Water Level Conditions 19
2.3.2 Impacts of Abnormal Water Level Conditions 19
2.3.3 Factors and Challenges Affecting Water Level Conditions 20
2.4 Predictive Analytics 21
2.4.1 Predictive Analytics Techniques Modelling and Algorithms 21
2.4.1.1 Predicting Modelling Approaches in Data-Driven 22
2.4.1.2 Predictive Analytics Algorithms 23
2.4.1.3 Performance Evaluation and Model Validation Techniques 24
2.4.2 Assessing Operational Performance of Predictive System 25
2.4.3 Predictive Analytics Application in Data-Driven Systems 26
2.5 Related Frameworks of Predictive Analysis 27
2.5.1 Cross-Industry Standard Process for Data Mining (CRISP-DM) 27
2.5.2 Team Data Science Process (TDSP) 29
2.5.3 Knowledge Discovery in Databases (KDD) 30
2.5.4 A Modern Data Science Process Model (DASC-PM) 32
2.5.5 Comparative Summary of Predictive Analytics Frameworks 33
2.6 Similar Existing Research 35
2.6.1 Summary of Similar Existing Research 37
2.7 Implications of the Literature Review: Predictive Analytics Implications 39
2.8 Conclusion 41

CHAPTER THREE: METHODOLOGY

3.1 Introduction 42
3.2 CRISP-DM 42
3.3 Project Planning 47
3.4 Business Understanding 48
3.5 Data Understanding 49
3.5.1 Collect Initial Data 49
3.5.2 Explore Data 54
3.5.3 Verify Data Quality 54
3.6 Data Preparation 55
3.7 Modelling 68
3.7.1 Predict Water Level Values for Water Treatment Optimizatin 69
3.8 Evaluation 70
3.9 Deployment 71

ACKNOWLEDGEMENT

Praise and gratitude to Allah SWT for His blessings, mercy and guidance throughout the completion of this research project. I am truly grateful for the strength, patience and perseverance that have been given to me in completing this project successfully within the given time frame.

I would like to express my deepest appreciation to my supervisor, Dr. Norulhidayah Binti Isa, for her continuous guidance, support, and valuable advice throughout the process of conducting this research. Her patience, expertise and constructive feedback have played an important role in improving this project and enhancing my understanding of the research area.

I would also like to express my sincere appreciation to my examiner, Dr. Farah Muna Mohamad Ghazali, for her valuable insights, thoughtful questions and constructive feedback during my final presentation. I truly appreciate her time, evaluation and the opportunity to learn from her valuable comments to further improve this research.

My sincere appreciation also goes to Miss Nor Hasnul Azirah Abdul Hamid, the lecturer for Project Formulation CSP650, for her guidance, encouragement and valuable insights during the early stages of this project. Her advice and support were essential in developing the project proposal and establishing a strong foundation for this research.

I would also like to express my deepest gratitude to Mr. Shahrul Nizam Rahim, CEO of Viucom Sdn. Bhd., for providing me with the opportunity to work on a real industry project and for sharing valuable insights throughout this journey. This experience allowed me to apply the knowledge and skills gained from my academic studies to address a real-world problem, making my Final Year Project a meaningful and rewarding experience.

My deepest gratitude goes to my beloved parents for their endless love, prayers, encouragement, and continuous support throughout my academic journey. Their motivation and sacrifices have been my greatest source of strength in completing this project.

Finally, I would like to thank all my friends and family members for their assistance, cooperation, encouragement, and companionship throughout this research process. Their support has made this journey meaningful and memorable.

ABSTRACT

Water treatment operations at Sebatik Island rely on real-time tank water-level monitoring through the Aeris Artificial Intelligence of Things (AIoT) device and satellite connectivity. However, the existing monitoring process provides limited support for predicting future water-level conditions, analyzing environmental influences and comparing alternative operating-hour scenarios. This project aimed to identify the current water treatment process and challenges at Sebatik Island, implement predictive analytics for water treatment process optimization and visualize the prediction results through a dashboard. The study reviewed related concepts of predictive analytics, AIoT based monitoring, water-level prediction and dashboard decision support to establish the project foundation. The CRISP-DM methodology was applied, involving business understanding, data understanding, data preparation, modelling, evaluation and deployment. Historical water-level, operational, time-based and environmental data were prepared for regression modelling. Linear Regression, Random Forest Regression and Decision Tree Regression were developed and evaluated using 5-fold and 10-fold cross-validation. Model performance was assessed using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R-squared (R^2). Decision Tree Regression using 5-fold cross-validation achieved the best result, with an RMSE of 0.388, MAE of 0.164, and R^2 of 0.948. The selected model was deployed through a Power BI dashboard containing water-level monitoring, prediction and environmental impact analysis, operational scenario comparison, decision support and user guideline pages. The dashboard identified the 20-hour operating scenario as the best overall option because it maintained normal predicted water-level conditions and full time-period coverage while reducing operating duration compared with the 24-hour baseline. Expert evaluation findings showed that the dashboard was functional, learnable, easy to use and satisfactory for operational monitoring. Overall, the project demonstrates that predictive analytics could enhance real-time monitoring by providing scenario-based decision support for water treatment process optimization at Sebatik Island.