UTM
UNIVERSITI TEKNOLOGI MALAYSIA

PREDICTIVE MODELING OF POLLUTION IN RIVER BASINS USING MACHINE LEARNING TECHNIQUE

NAME
HASLINDA BINTI ABDUL SAHAK

MATRIC NO.
MCS241004

COURSE
MASTER IN DATA SCIENCE

FACULTY
FACULTY OF COMPUTING

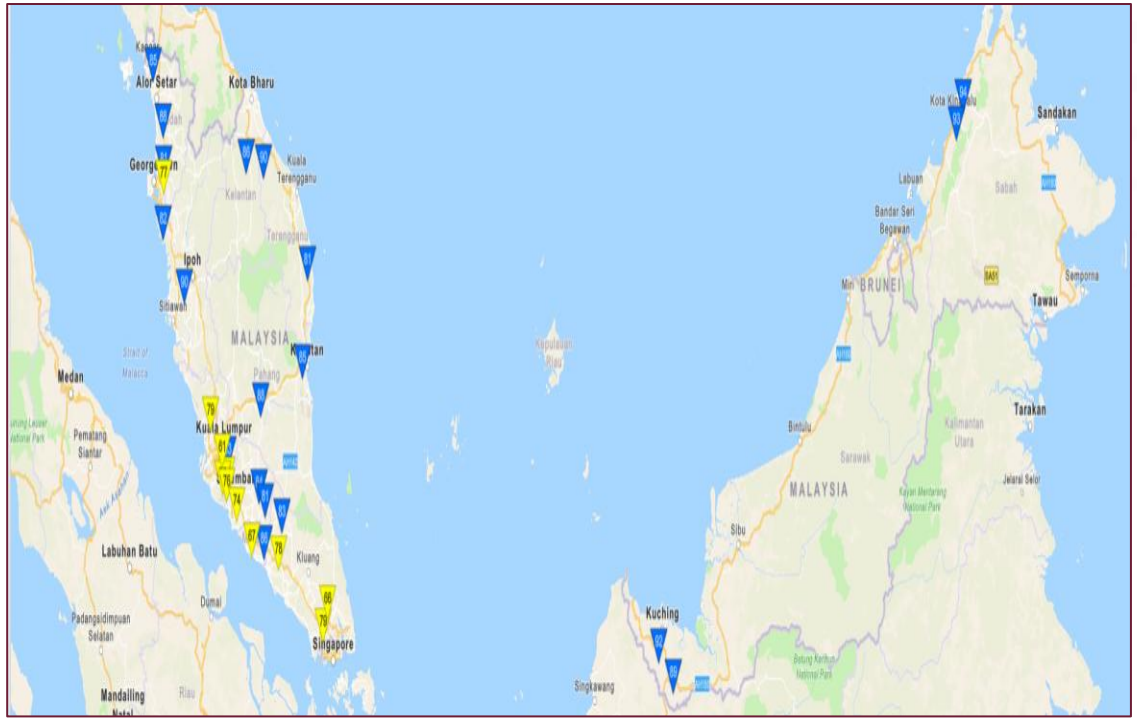
LECTURER
ASSOC. PROF. TS. DR. MOHD SHAHIZAN BIN OTHMAN

Innovating Solutions

PRESENTATION OUTLINE

- Research Overview
- Problem Statement, Research Objective & Research Question
- Scope of Research
- Significant of Research
- Literature Review
- Methodology
- Initial Results
- Conclusion
- Future Works

Research Overview



River Water Quality Index Management System (RQIMS)

Water pollution in river basins is a critical environmental issue that poses significant risks to ecosystems, public health, and economic development.

Effective management of river basins requires a proactive approach to predict and mitigate pollution levels.

This research focuses on leveraging machine learning (ML) techniques to develop predictive models for pollution in river basins, enabling stakeholders to identify and address contamination sources before they reach critical levels.

Research Problem, Research Objective and Research Questions

Research Problem	Research Objective	Research Questions
• River basin pollution is a major environmental challenge, impacting ecosystems, human health, and socio-economic development • Existing methods are time-consuming, labor-intensive, and limited in spatial and temporal coverage • Predicting pollution levels becomes quite impractical.	• To collect historical water quality data • To develop a machine learning models • To evaluate the models according to appropriate dimensions	• What are the effective skills for collecting and pre-processing high-quality, relevant, and consistent water quality data? • Which of the machine learning algorithms Which of the machine learning algorithms • What way does the use of appropriate metrics to evaluate developed models

Scope of Research

A data collection: Sources of historical information for water quality index from Department of Statistics Malaysia (DOSM)

Data Preprocessing: Data cleaning: refers to the preprocessing of these somewhat diverse first data records missing values, outlier information, and inconsistencies.

Model Building and Training: Varying machine learning algorithms and their application for training appropriate models

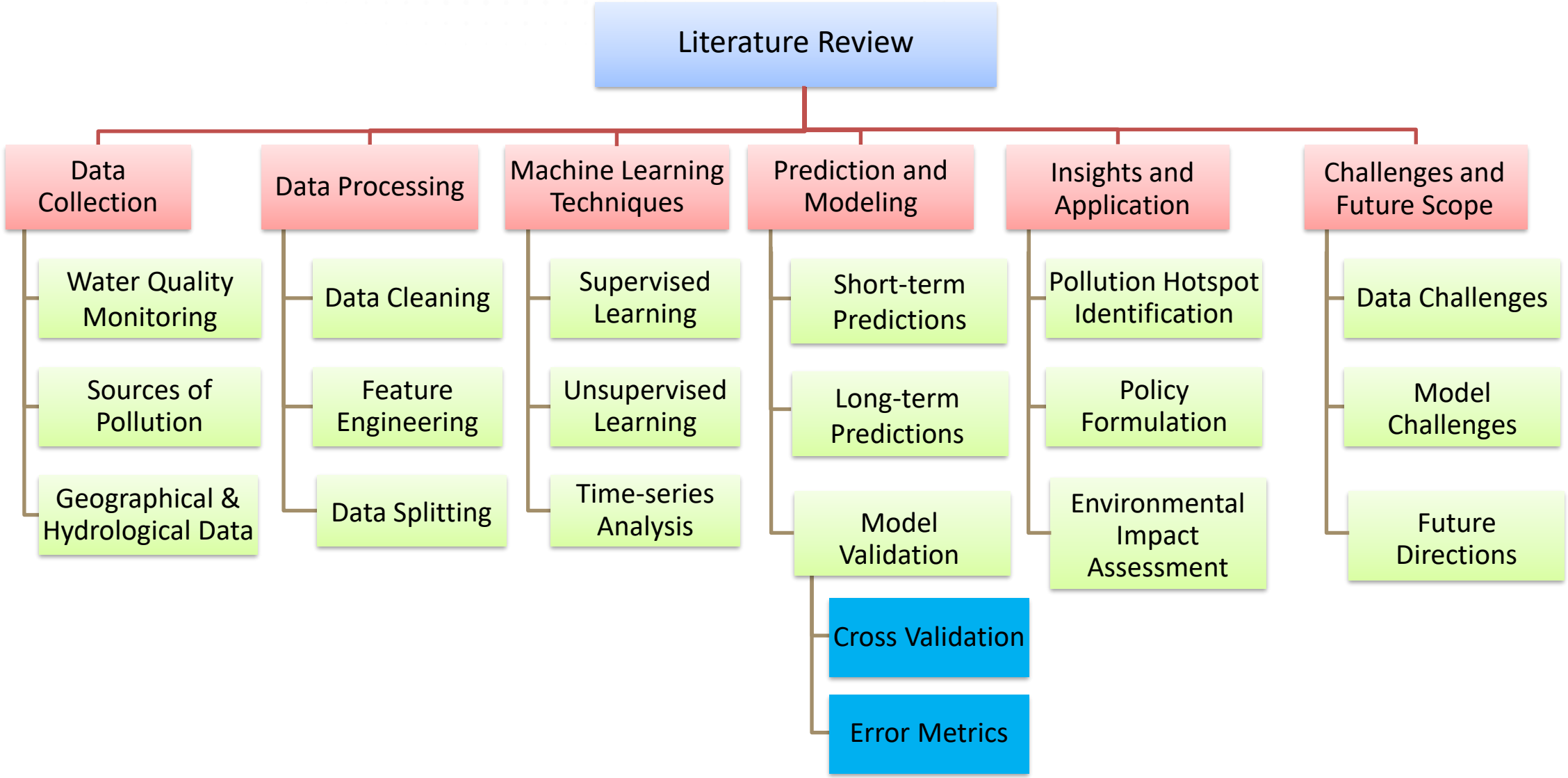
Model Evaluation: Performing evaluations of the model that could be based on metrics

Prediction and Visualization: Future forecasting from the trained model on pollution levels and the drawing of results.

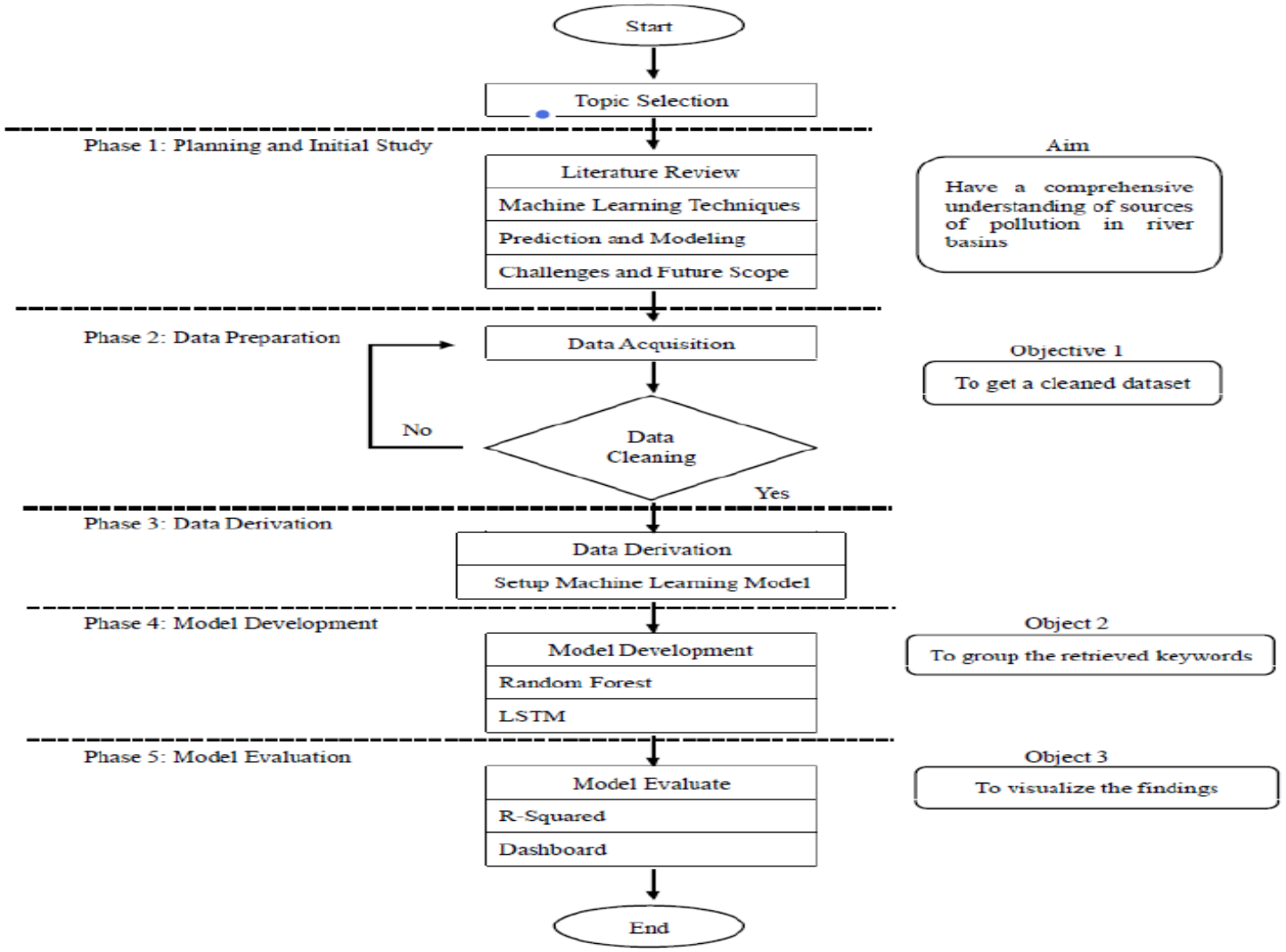
Significant of Research

- Revolutionize water resource management and environmental sustainability.
- Develop advanced models that can accurately predict pollution levels in river basins.
- Enhance water quality monitoring by early detection of pollution events, enabling timely interventions to mitigate environmental degradation
- Contribute to the advancement of machine learning applications, particularly in environmental sciences, opening up new avenues for innovation and problem-solving

Literature Review

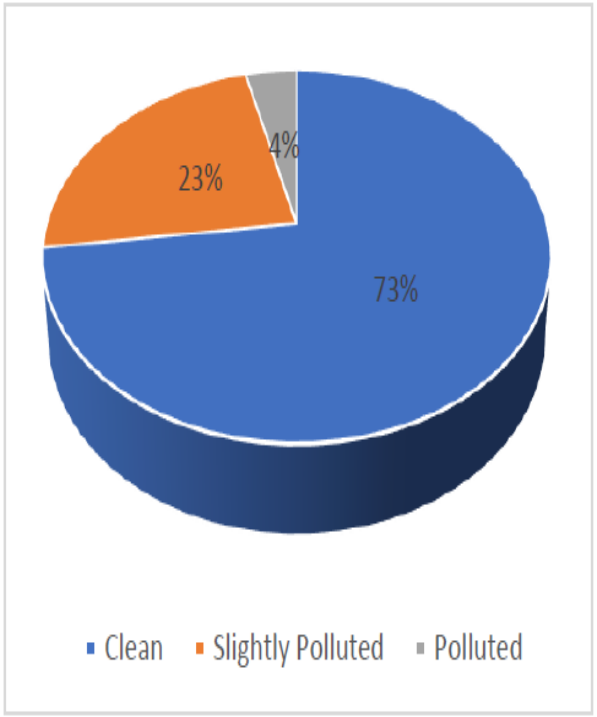


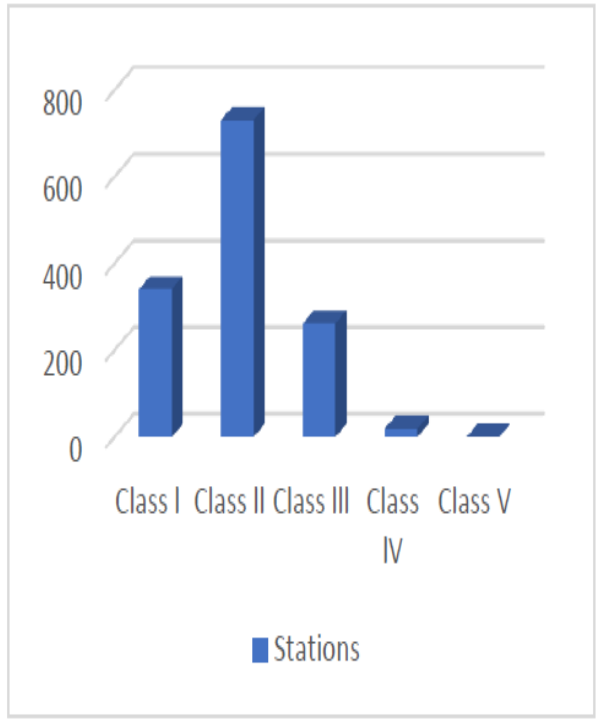
Methodology



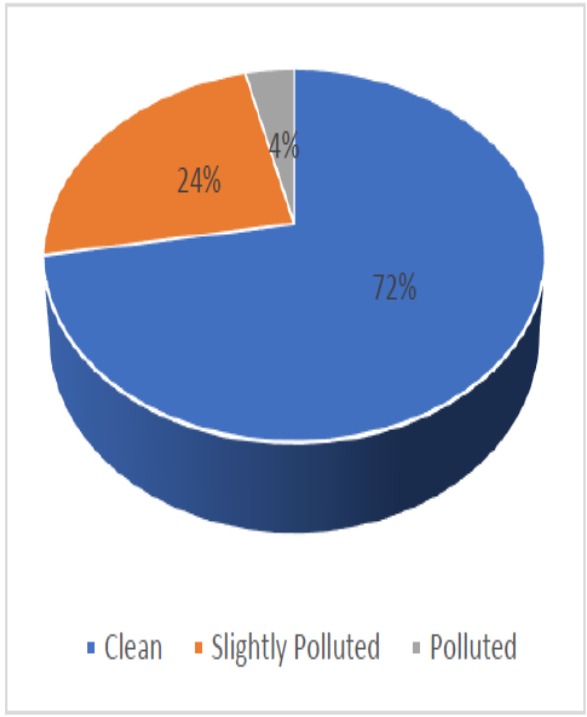
- Phase 1: Planning and Initial Study
- Phase 2: Data Preparation
- Phase 3: Data Derivation
- Phase 4: Model Development
- Phase 5: Model Evaluation

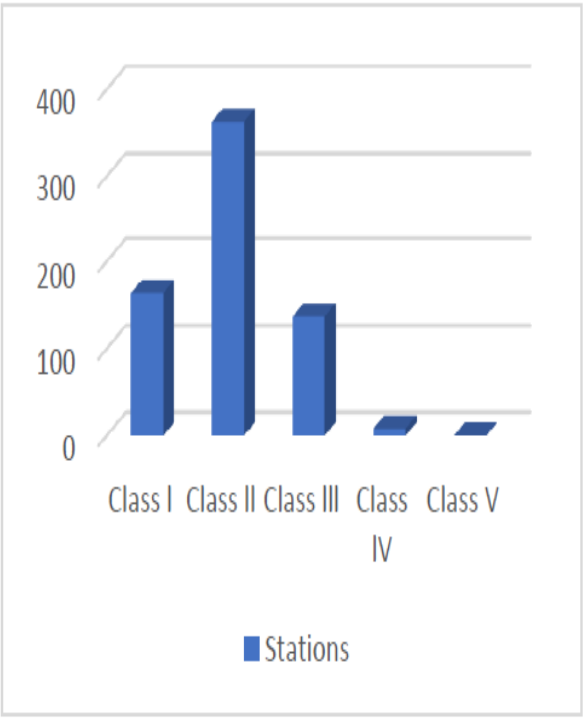
Initial Results





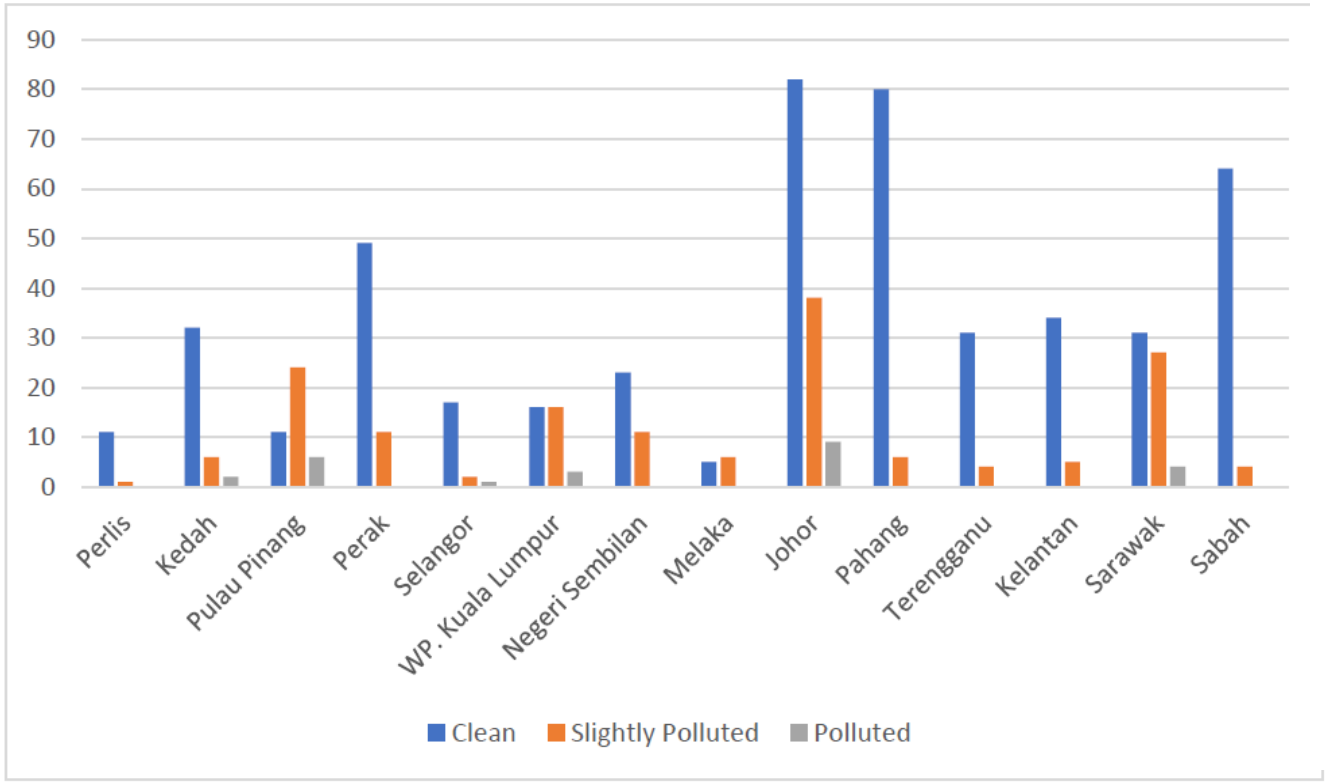
Water quality status by station

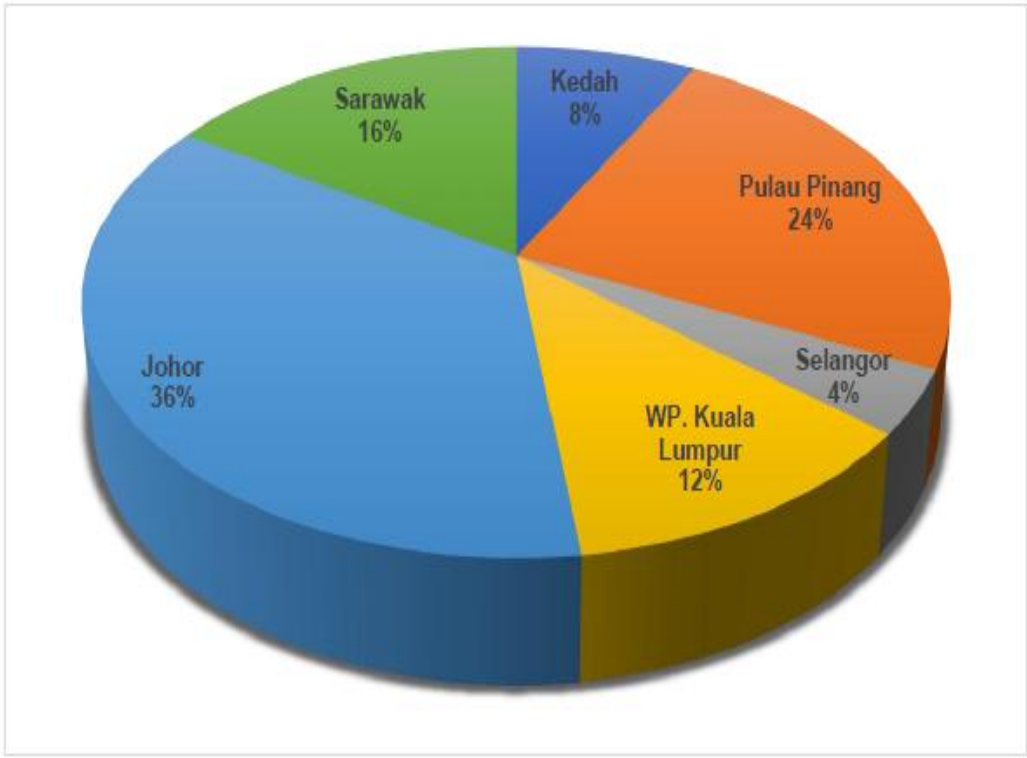




Water quality status by river

Initial Results





Pollution in River Basins by State

Conclusion

- This study demonstrates the potential of machine learning to improve water quality management in Malaysian rivers.
- By analyzing historical water quality data, predictive models were developed to provide accurate and real-time insights into pollution levels.
- Improvements in key indicators, such as Biochemical Oxygen Demand (BOD5) and Suspended Solids (SS), highlight the positive impact of ongoing policies and interventions.
- However, challenges like inconsistent data and system integration must be addressed to enhance model effectiveness.
- Overall, this research provides a practical framework for using machine learning to support sustainable water resource management and pollution reduction in Malaysia.

Future Works

- Future work on predictive modeling of river basin pollution will focus on improving model accuracy, scalability, and real-world application.
- This includes integrating advanced techniques like deep learning and ensemble methods, using real-time data from IoT sensors and satellites, and combining machine learning with GIS for geospatial analysis.
- Models will also be enhanced to simulate scenarios, such as policy changes and climate impacts, and adapted for global application across diverse river basins.
- Collaboration with policymakers, industries, and communities will ensure the models are practical and support effective decision-making for sustainable water management.

THANK YOU

univteknologimalaysia utm.my utmofficial

KERANA TUHAN UNTUK MANUSIA • UNIVERSITI TEKNOLOGI MALAYSIA •

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Innovating Solutions
Menginovasi Penyelesaian