

## LIST OF CONTENTS

CHAPTER I INTRODUCTION .....	1
1.1 Background .....	1
1.2 Research Problem.....	2
1.3 Research Scope .....	2
1.4 Research Objectives .....	3
1.5 Research Benefits.....	3
CHAPTER II LITERATURE REVIEW .....	4
CHAPTER III THEORETICAL BASIS .....	10
3.1 Rotary Equipment.....	10
3.1.1 Pump .....	10
3.2 Sensor for Rotary Equipment .....	11
3.2.1 Vibration Sensor/Transducer.....	11
3.2.2 Temperature Sensor .....	16
3.3 Classification in Machine Learning.....	18
3.3.1 Random Forest (RF) Classifier .....	19
3.3.2 Gradient Boosting .....	20
3.4 Hyperparameter Tuning.....	22
3.4.1 Grid Search Cross-Validation .....	23
3.4.2 Randomized Search Cross-Validation .....	23
3.4.3 Bayesian Optimization.....	24
3.5 Evaluation of Classification Models .....	25
3.5.1 Confusion Matrix and Evaluation Metrics.....	25
3.5.2 Cross-Validation.....	28
CHAPTER IV RESEARCH METHODOLOGY .....	30
4.1 Tools and Materials .....	30
4.2 Data Collection.....	31
4.2.1 Dataset Description.....	31
4.2.2 Similarity-Based Modelling (SBM) to Calculate ‘Estimate’ Feature... 35	
4.2.3 Labelling Mechanism for Advisory and Diagnostic Advisory Indications .....	37
4.2.4 Data Collection Architecture .....	37

4.3 System Design .....	38
4.3.1 Data Acquisition .....	41
4.3.2 Data Preprocessing .....	41
4.4 Training Machine Learning Models .....	45
4.4.1 Random Forest .....	45
4.4.2 XGBoost .....	47
4.4.3 CatBoost .....	48
4.4.4 Baseline Model Training .....	49
4.4.5 Hyperparameter Tuning .....	49
4.4.6 Model Training with Optimized Hyperparameters .....	50
4.5 Model Evaluation .....	50
4.5.1 Classification Report .....	50
4.5.2 Confusion Matrix .....	51
4.5.3 Cross-Validation (Stratified K-Fold) .....	52
4.6 Model Saving and Serialization .....	52
4.7 Deployment to Dashboard .....	53
CHAPTER V IMPLEMENTATION .....	55
5.1 Data Loading and Exploratory Data Analysis (EDA) .....	55
5.1.1 Data Loading .....	55
5.1.2 Data Structure and Label Relationships .....	56
5.1.3 Column and Value Inspection .....	57
5.1.4 Time-Series Visualization and Label Mapping .....	59
5.2 Data Preprocessing Implementation .....	61
5.2.1 Label Correction and Completion .....	61
5.2.2 Label Renaming and Standardization .....	64
5.2.3 Column Filtering .....	65
5.2.4 Empty Sensor Columns Removal .....	66
5.2.5 Duplicate Removal .....	66
5.2.6 Outlier Removal .....	67
5.3 Model Training and Tuning .....	68
5.3.1 Feature Preparation and Data Splitting .....	68
5.3.2 Class Weight Handling .....	69

5.3.3 Initial Model Training .....	70
5.3.4 Hyperparameter Tuning .....	71
5.3.5 Model Training using Optimized Parameters .....	75
5.3.6 Model Saving .....	77
5.4 Model Evaluation .....	78
5.4.1 Evaluation Metrics .....	78
5.4.2 Cross-Validation Accuracy .....	79
5.4.3 Confusion Matrix Visualization .....	80
5.5 Streamlit Dashboard Implementation .....	81
5.5.1 System Architecture .....	81
5.5.2 Model and Sheet Selection via Sidebar: .....	82
5.5.3 Upload Dataset .....	84
5.5.4 Data Preprocessing .....	85
5.5.5 Quick Summary and Sensor Behavior Visualization: .....	86
5.5.6 Advisory/Diagnostic Class Distribution and Prediction Result Table: .	89
5.5.7 Manual Prediction Interface .....	91
5.5.8 Deployment and Usability .....	92
CHAPTER VI RESULTS AND DISCUSSION .....	93
6.1 Baseline Models Performance .....	93
6.1.1 Test Set Performance Summary .....	94
6.1.2 Cross-Validation Accuracy Summary .....	95
6.1.3 Confusion Matrix Visualizations (Baseline Models) .....	96
6.1.4 Test Set Classification Report Summary Table and Discussion .....	98
6.2 Tuned Model Performance .....	103
6.2.1 Optimized Hyperparameter Summary .....	104
6.2.2 Test Accuracy Summary .....	106
6.2.3 Cross-Validation Accuracy Summary .....	107
6.2.4 Confusion Matrix Visualization .....	108
6.2.5 Post-Tuning Classification Report Summary and Key Observations..	110
6.3 Tuned Models Performance Analysis .....	114
6.3.1 Accuracy Improvements Overview .....	115
6.3.2 Classification Report Comparison .....	117

6.4 Model Performance Across All Sheets .....	119
6.5 Sensor Sheet Characteristics.....	120
6.5.1 Temperature Sensor Sheets .....	121
6.5.2 Vibration Sensor Sheets .....	121
6.6 Streamlit Dashboard Results Analysis .....	122
CHAPTER VII CONCLUSION .....	123
7.1 Conclusion.....	123
7.2 Recommendations .....	124
BIBLIOGRAPHY .....	126
APPENDIX.....	129