

THESIS
**MACHINE LEARNING APPROACH DESIGN FOR PREDICTIVE
MAINTENANCE OF THE CENTRIFUGAL PUMPS**

By
Rahmat Pakpahan
22351047

MASTER'S DEGREE
in
INFORMATION TECHNOLOGY
DEPT MASTER OF INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY
The Prominence Tower
Jalan Jalur Sutera Barat No. 15, Alam Sutera
Tangerang, Banten 15143 - Indonesia

March 2026

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.


Rahmat Pakpahan

March 7, 2026

Student

Date

Approved by:


Dr. Ir. Heru Purnomo Ipung, M.Eng

March 7, 2026

Thesis Advisor

Date


Dr. Ir. Eka Budiarto, S.T, M.Sc

March 7, 2026

Thesis Co-Advisor

Date

ABSTRACT

MACHINE LEARNING APPROACH DESIGN FOR PREDICTIVE MAINTENANCE OF THE CENTRIFUGAL PUMPS

By

Rahmat Pakpahan

SWISS GERMAN UNIVERSITY

Centrifugal pumps are widely used in industrial operations, but they are subject to mechanical faults that cause significant downtime and maintenance costs. Traditional maintenance methods often fail to detect early-stage failures; therefore, this thesis presents a machine learning-based predictive analytics approach to improve fault diagnosis and maintenance planning for multistage centrifugal pumps. Adhering to the CRISP-DM methodology, the study demonstrates that integrating a multi-dimensional feature set of flow, suction pressure, and speed enables Tree-based classification models to achieve a very good accuracy of 99.4% and an AUC of 0.998. By prioritizing interpretability and alignment with Hydraulic Institute Standards, the research provides a glass-box diagnostic tool capable of identifying incipient cavitation and deviations from the Best Efficiency Point (BEP). Ultimately, this thesis offers a scalable, physically justifiable solution for modern pump infrastructure, significantly reducing energy waste and the Total Cost of Ownership (TCO) through autonomous, data-driven maintenance.

Keywords: Centrifugal Pumps, Predictive Maintenance, Machine Learning, Fault Diagnosis, Cavitation Detection, Performance Prediction, Tree-based Models, CRISP-DM, Hydraulic Institute Standards.



DEDICATION

I dedicate this thesis to my family who has shaped me to the man I am today. To my mother whose love and wisdom has guided me to where I am. To my brothers and sisters whose perpetual optimism reminds me that there is no great challenge to face.



ACKNOWLEDGEMENTS

First and foremost, I would like to direct our sincerest gratitude towards my advisor: Pak Dr. Ir. Heru Purnomo Ipung, M.Eng and Co-Advisor: Pak Dr. Ir. Eka Budiarto, S.T, M.Sc, whose knowledge and competence guided me through the process of writing this thesis. Secondly, we would like to direct a special appreciation towards my company in Egaturbo, who inspired us with its relentless pursuit of improving service operations while simultaneously creating a better digitalization innovation. Lastly, we would like to direct our gratitude towards the opponents and examiners who helped us refine the work by providing new insights and perspectives.



TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
DEDICATION	5
ACKNOWLEDGEMENTS	6
TABLE OF CONTENTS.....	7
LIST OF ABBREVIATIONS.....	10
LIST OF FIGURES	11
LIST OF TABLES.....	12
CHAPTER 1 - INTRODUCTION.....	13
1.1 Background.....	13
1.2 Research problem	14
1.3 Research question.....	16
1.4 Hypothesis	17
1.5 Thesis structure.....	17
CHAPTER 2 - LITERATURE REVIEW	19
2.1 Centrifugal Pumps	19
2.1.1 Classification of centrifugal pump.....	23
2.1.2 The construction of a centrifugal pump.....	26
2.1.3 Centrifugal pump performance characteristics	27
2.1.4 Centrifugal pump failure modes	28
2.2 Pump maintenance strategies.....	32
2.3 Machine learning	33

2.3.1	Types of machine learning.....	34
2.3.2	Predictive Modeling for Centrifugal Pumps.....	37
2.3.2.1	Classification.....	38
2.3.2.2	Regression.....	42
2.3.3	Machine Learning requirements	44
2.4	Cross-Industry Standard Process for Data Mining (CRISP-DM)	47
2.5	Orange Data Mining	49
2.6	Related Work.....	50
CHAPTER 3 – RESEARCH METHODS		57
3.1	Business Understanding	59
3.2	Data Understanding	60
3.3	Data Preparation	62
3.4	Modeling.....	63
3.5	Evaluation Methods.....	64
3.6	Deployment	66
CHAPTER 4 – MACHINE LEARNING APPROACH DESIGN FOR PUMP CAVITATION DETECTION		67
4.1	Business Understanding	68
4.2	Requirement Specification	71
4.3	Data Understanding	72
4.4	Data preparation	73
4.5	Modeling.....	75
4.6	Evaluation.....	83
4.7	Deployment	85
CHAPTER 5 – MACHINE LEARNING APPROACH DESIGN FOR PUMP PERFORMANCE PREDICTION		86

5.1	Business Understanding	87
5.2	Requirement Specification	87
5.3	Data Understanding	88
5.4	Data preparation	89
5.5	Modeling.....	91
5.6	Evaluation.....	98
5.7	Deployment	99
CHAPTER 6 – RESULTS AND DISCUSSIONS.....		100
6.1	Results overview	100
6.2	Discussions	103
6.3	General guidelines machine learning approach design	105
CHAPTER 7 – CONCLUSIONS AND RECOMMENDATIONS.....		109
7.1	Conclusions overview	109
7.2	General Recommendations.....	111
7.3	Recommendations for the company	112
7.4	Future research	113
GLOSSARY		115
REFERENCES		116
CURRICULUM VITAE.....		118

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
AUC	Area Under the Curve
BEP	Best efficiency point
CRISP-DM	Cross Industry Standard Process for Data Mining
CNN	Convolutional Neural Network
DE	Drive end
GPU	Graphics Processing Unit
ISO	International standard organization
LSTM	Long Short-Term Memory
MCSA	Motor current signature analysis
ML	Machine learning
MLP	Multi-Layer Perceptrons
NDE	Non drive end
NPSHA	Net positive suction head available
NPSHR	Net positive suction head required
OEM	Original equipment manufacturer
OPEX	Operational expenditures
ROI	Return on Investment
RUL	Remaining useful life
SVR	Support Vector Regression
TCO	Total Cost of Ownership

LIST OF FIGURES

Figures	Page
Figure 1. 1 Fish bone diagram for research problem	16
Figure 2. 1 Types of pumps (Ifidon et al., 2025)	20
Figure 2. 2 Types of centrifugal pumps (Karassik and McGuire, 2012)	21
Figure 2. 3 A simple figure of a centrifugal pump (Fausing Olesen and Shaker, 2020)	22
Figure 2. 4 Cross-section of a centrifugal pump (Karassik and McGuire, 2012)	23
Figure 2. 5 Radial flow centrifugal pump (Karassik and McGuire, 2012)	24
Figure 2. 6 Diagram of axial flow pump (Karassik and McGuire, 2012)	25
Figure 2. 7 Diagram of a mixed flow pump (Karassik and McGuire, 2012)	26
Figure 2. 8 Main components of a centrifugal pump (Karassik and McGuire, 2012) ..	26
Figure 2. 9 Data driven predictive maintenance (Fausing Olesen and Shaker, 2020) ..	33
Figure 2. 11 Types of machine learning (Razzaq and Shah, 2025)	34
Figure 2. 13 CRISP-DM Process Model for Data Mining (Wirth and Hipp, 2000)	47
Figure 2. 13 Orange data mining view	49
Figure 3. 2 Equipment photo	69
Figure 3. 3 Performance test procedures	70

LIST OF TABLES

Tables	Page
Tabel 2. 1 Related work	53
Tabel 4. 1 The basic statistics of dataset	73
Tabel 4. 2 The pump cavitation rank	74
Tabel 4. 3 Supervised ML result	76
Tabel 4. 4 Tree-based models with Ps	79
Tabel 4. 5 Tree-based models with speed	80
Tabel 4. 6 Tree-based models with flow	82
Tabel 4. 6 Tree-based models with Ps, speed & flow	83
Tabel 5. 1 The basic statistics of dataset	89
Tabel 5. 2 The pump operation rank	90
Tabel 5. 4 Tree-based models with the high-gain data	95
Tabel 5. 4 Tree-based models with full hydraulic context	96
Tabel 5. 4 Tree-based models with suction-side integrity	97