

A Machine Learning Approach to Predictive Analytics

John Doe

Dissertation submitted to the University of Sri Jayewardenepura in partial fulfillment of
the requirement for the award of Master in Data Science and Artificial Intelligence.

Declaration

“The work described in this thesis was carried out by me and a report on this has not been submitted in whole or in part to any university or any other institution for another Degree”

John Doe

Contents

Acknowledgements	iii
Abstract	iv
1 Introduction	1
2 Literature Review	1
2.1 Machine Learning in Healthcare	1
2.1.1 Supervised Learning Techniques	1
2.2 Challenges in Predictive Analytics	1
3 Materials and Methods	1
3.1 Research Design	1
3.1.1 Model Selection	2
3.2 Data Collection	2
3.2.1 Data Sources	2
3.3 Data Analysis	2
3.3.1 Statistical Methods	2
4 Results	2
5 Discussion	2
6 Conclusions	2
References	3
Appendices	4

List of Tables

List of Figures

Acknowledgements

I thank my supervisor, Dr. Jane Smith, for her guidance, and my family for their unwavering support.

Abstract**A Machine Learning Approach to Predictive Analytics****John Doe****ABSTRACT**

This dissertation explores the application of machine learning techniques to predictive analytics in healthcare. The study focuses on developing models to predict patient outcomes based on historical data, achieving an accuracy of 87%. Challenges include data preprocessing and model interpretability.

1 Introduction

Predictive analytics has transformed decision-making across industries. This study aims to leverage machine learning for healthcare predictions.

2 Literature Review

Existing research highlights the efficacy of supervised learning in predictive tasks.

2.1 Machine Learning in Healthcare

Studies show high accuracy in disease prediction models.

2.1.1 Supervised Learning Techniques

Techniques like logistic regression and random forests dominate.

I Logistic Regression Simple yet effective for binary outcomes.

II Random Forests Robust for handling complex datasets.

2.2 Challenges in Predictive Analytics

Data quality and interpretability remain key hurdles.

3 Materials and Methods

This section outlines the research methodology.

3.1 Research Design

A quantitative approach was adopted.

3.1.1 Model Selection

Criteria included accuracy and scalability.

I Support Vector Machines Chosen for their performance on small datasets.

II Neural Networks Evaluated for deep learning capabilities.

3.2 Data Collection

3.2.1 Data Sources

Data was sourced from hospital records.

3.3 Data Analysis

3.3.1 Statistical Methods

Regression analysis was applied.

4 Results

The random forest model achieved 87% accuracy.

5 Discussion

Results align with prior studies but highlight preprocessing needs.

6 Conclusions

Machine learning offers significant potential for healthcare predictions.

References

References

- [1] Smith, J. (2023) 'Machine Learning in Healthcare', *Journal of Data Science*, vol. 10, no. 2, pp. 45-60.

Appendices

Appendix A: Dataset Description

The dataset includes 10,000 patient records.