UNIVERSTIY OF HAMBURG

MASTER THESIS

Predicting Protein Crystallization Conditions using Machine Learning

Author:
Michael HÜPPE

Supervisors:
Prof. Dr. Fabian Michael Kern

A thesis submitted in fulfillment of the requirements for the degree of Master of Science

in the

Machine Learning in Bio Informatics
Department of Computer Science

September 30, 2025

Declaration of Authorship

I, Michael HÜPPE, declare that this thesis titled, "Predicting Protein Crystallization Conditions using Machine Learning" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"This one is for the boys with the booming system"

Nicki Minaj

UNIVERSTIY OF HAMBURG

Abstract

Faculty Name
Department of Computer Science

Master of Science

Predicting Protein Crystallization Conditions using Machine Learning

by Michael HÜPPE

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgments and the people to thank go here, don't forget to include your project advisor...

Contents

List of Figures

List of Tables

List of Abbreviations

LAH	**L**ist **A**bbreviations **H**ere
WSF	**W**hat (it) **S**tands **F**or

Physical Constants

Speed of Light $c_0 = 2.997\,924\,58 \times 10^8\,\mathrm{m\,s^{-1}}$ (exact)

List of Symbols

a	distance	m
P	power	W ($\mathrm{J\,s^{-1}}$)
ω	angular frequency	rad

For/Dedicated to/To my…

Chapter 1

Decoding the Crystal Recipe: Predicting Protein Crystallization Conditions via Machine Learning

1.1 Introduction

Chapter 2

Theory

2.1 Proteins

2.1.1 Protein Crystallization

2.2 Trees

Chapter 3

Data

3.1 The Protein database

3.1.1 Acquisition

3.2 Protein Properties

3.3 Crystallization Conditions

3.4 Relationships

Chapter 4

Results

Chapter 5

Discussion

Appendix A

Data Appendix