Predicting the activity of protein-ligand complexes

Lukas Fallmann

BACHELORARBEIT

eingereicht am

Fachhochschul-Bachelorstudiengang

Medizin- und Bioinformatik

in Hagenberg

im Juni 2023

Advisor:

Micha Johannes Birklbauer, M.Sc.

© Copyright 2023 Lukas Fallmann

This work is published under the conditions of the Creative Commons License *Attribution-NonCommercial-NoDerivatives 4.0 International* (CC BY-NC-ND 4.0)—see https://creativecommons.org/licenses/by-nc-nd/4.0/.

Declaration

I hereby declare and confirm that this thesis is entirely the result of my own original work. Where other sources of information have been used, they have been indicated as such and properly acknowledged. I further declare that this or similar work has not been submitted for credit elsewhere. This printed copy is identical to the submitted electronic version.

Hagenberg, June 27, 2023

Lukas Fallmann

Contents

Declaration iv

Preface vii

Abstract viii

Kurzfassung ix

1 Introduction 1
- 1.1 Overview of the topic . . . 1
- 1.2 Why is it important . . . 1
- 1.3 What are current methods that are prominently used . . . 1
- 1.4 Machine Learning in drug design / activity prediction . . . 1
- 1.5 Interactions . . . 1
- 1.6 Goals . . . 1

2 Methods 2
- 2.1 Data description . . . 2
- 2.2 Data partitioning . . . 2
- 2.3 Machine Learning approaches . . . 2
- 2.4 quality metrics . . . 2
- 2.5 hyperparameter search . . . 2
- 2.6 feature engineering . . . 2

3 Results 3
- 3.1 Results of the models, mostly tables but for best performing models in each category also do some plots like confusion matrices, AUC plots if available, bar plots for comparison between models . . . 3

4 Discussion 4
- 4.1 Performance . . . 4
- 4.2 Improvements . . . 4

A Technical Details 5

B Supplementary Materials 6
- B.1 PDF Files . . . 6

B.2 Media Files 6
B.3 Online Sources (PDF Captures) 6

C Questionnaire **7**

D LaTeX Source Code **8**

References **9**
Online sources 9

Preface

Abstract

This should be a 1-page (maximum) summary of your work in English.

Kurzfassung

An dieser Stelle steht eine Zusammenfassung der Arbeit, Umfang max. 1 Seite. ...

Chapter 1

Introduction

1.1 Overview of the topic

1.2 Why is it important

1.3 What are current methods that are prominently used

1.4 Machine Learning in drug design / activity prediction

1.5 Interactions

1.6 Goals

Chapter 2

Methods

2.1 Data description

2.2 Data partitioning

2.3 Machine Learning approaches

2.4 quality metrics

2.5 hyperparameter search

2.6 feature engineering

Chapter 3

Results

3.1 Results of the models, mostly tables but for best performing models in each category also do some plots like confusion matrices, AUC plots if available, bar plots for comparison between models

Chapter 4

Discussion

4.1 Performance

4.2 Improvements

Appendix A

Technical Details

Appendix B

Supplementary Materials

List of supplementary data submitted to the degree-granting institution for archival storage (in ZIP format).

B.1 PDF Files

Path: /

thesis.pdf . . . . . Master/Bachelor thesis (complete document)

B.2 Media Files

Path: /media

*.ai, *.pdf . . . . . Adobe Illustrator files
*.jpg, *.png . . . . . raster images
*.mp3 . . . . . audio files
*.mp4 . . . . . video files

B.3 Online Sources (PDF Captures)

Path: /online-sources

Reliquienschrein-Wikipedia.pdf [1]

Appendix C

Questionnaire

Appendix D

LaTeX Source Code

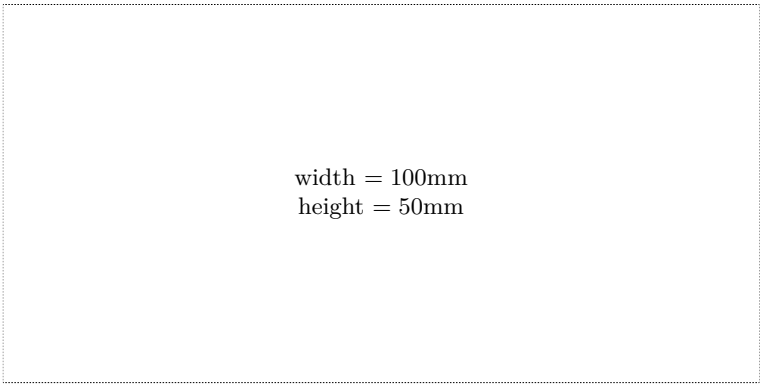
References

Online sources

[1] *Reliquienschrein*. Aug. 29, 2022. URL: https://de.wikipedia.org/wiki/Reliquienschrein (visited on 02/11/2023).

Check Final Print Size

— Check final print size! —



— Remove this page after printing! —