Quantum Machine Learning

by

Kristian Wold

Thesis

for the degree of

Master of Science

Department of Physics

Faculty of Mathematics and Natural Sciences

University of Oslo

August 2021

This master's thesis is submitted under the master's program *Computational Science*, with program option *Physics*, at the Department of Informatics, University of Oslo. The scope of the thesis is 60 credits.

www.duo.uio.no

Print production: Reprosentralen, University of Oslo

Abstract

Acknowledgements

Kristian Wold
Oslo, August 2021

Abbreviations

AURKA	Aurora Kinase A
AURKB	Aurora Kinase B
AURKC	Aurora Kinase C
CDK	Cyclin-Dependent Kinase
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CML	Chronic Myelogenous Leukemia
CPC	Chromosomal Passenger Complex
DOF	Degrees of Freedom
EGFR	Epidermal Growth Factor Receptor
GROMACS	GROningen MAchine for Chemical Simulations
HDX	Hydrogen-Deuterium Exchange
INCENP	Inner Centromere Protein
MD	Molecular Dynamics
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
PBC	Periodic Boundary Conditions
PCA	Principal Component Analysis
PK	Protein Kinase
PKA	Protein Kinase A
RMSD	Root-Mean-Square Deviation
RMSF	Root-Mean-Square Fluctuation
VMD	Visual Molecular Dynamics

Glossary

Dipole Blockade	Phenomenon in which the simultaneous excitation of two atoms is inhibited by their dipolar interaction.
Cavity Induced Transparency	Phenomenon in which a cavity containing two atoms excited with light at a frequency halfway between the atomic frequencies contains the number of photons an empty cavity would contain.

Nomenclature

c	Speed of light ($2.997\ 924\ 58 \times 10^8$ ms^{-1})
$\hbar$	Planck constant ($1.054\ 572\ 66 \times 10^{-34}$ Js)
k_B	Boltzmann constant ($1.380\ 658 \times 10^{-23}$ JK^{-1})
Z_0	Impedance of free space ($376.730\ 313\ 461\ \Omega$)
μ_0	Permeability of free-space ($4\pi \times 10^{-7}$ Hm^{-1})

Contents

1

Introduction and Objective of the Study

1.1 Introduction

1.1.1 Parameter Identification

1.2 Objective of the Study

1.3 How This Thesis is Organized

How This Book is Organized (ISL)

Our view is that one must understand simple methods before trying to grasp more complex ones. Hence, after giving an overview of the supervising learning problem in Chapter 2, we discuss linear methods for regression and classification in Chapters 3 and 4. In Chapter 5 we describe splines, wavelets and regularization/penalization methods for a single predictor, while Chapter 6 covers kernel methods and local regression. Both of these sets of methods are important building blocks for high-dimensional learning techniques. Model assessment and selection is the topic of Chapter 7, covering the concepts of bias and variance, overfitting and methods such as cross-validation for choosing models. Chapter 8 discusses model inference and averaging, including an overview of maximum likelihood, Bayesian inference and the bootstrap, the EM algorithm, Gibbs sampling and bagging, A related procedure called boosting is the focus of Chapter 10. In Chapters 9–13 we describe a series of structured methods for supervised learning, with Chapters 9 and 11 covering regression and Chapters 12 and 13 focusing on classification. Chapter 14 describes methods for unsupervised learning. Two recently proposed techniques, random forests and ensemble learning, are discussed in Chapters 15 and 16. We describe undirected graphical models in Chapter 17 and finally we study high- dimensional problems in Chapter 18. At the end of each chapter we discuss computational considerations important for data mining applications, including how the computations scale with the number of observations and predictors. Each chapter ends with Bibliographic Notes giving background references for the material.

Part I

Theoretical Background

2

Theoretical Background

2.1 Theory

2.1.1 Project Theory 1

This is subsection 2.1.1.

Citation is done with BibTeX [**Sakurai**].

Cross-reference equations such as

$$E = mc^2 \tag{2.1}$$

with Equation (2.1).

Figure 2.1.1 brings the noise from the **figures folder**.

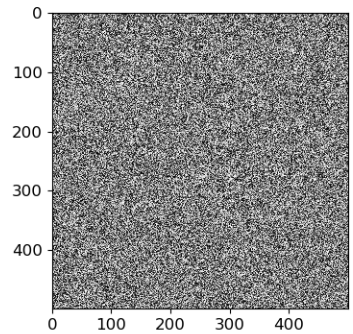

Figure 2.1.1: *Make some noise.*

Figure 2.1.2 shows a happy animal found in the **Images folder**.

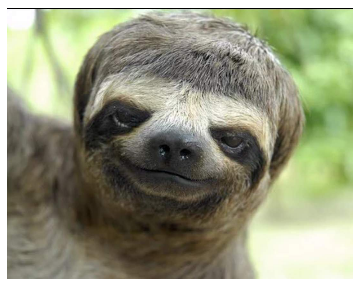

Figure 2.1.2: *Sloths are arboreal mammals noted for slowness of movement and for spending most of their lives hanging upside down in the trees.*

Source: Insert image source here

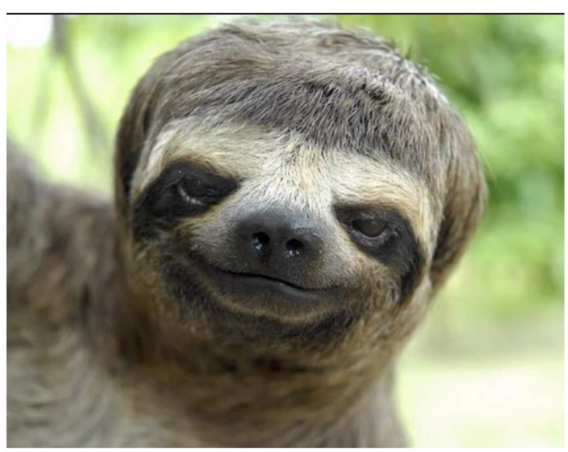

Figure 2.1.3: *Sloths are arboreal mammals noted for slowness of movement and for spending most of their lives hanging upside down in the trees.*

Source: Insert image source here

Table 2.1.1 is from the **tables folder**.

Table 2.1.1: *From pandas to latex.*

x	x^2	x^3
0.250	0.062	0.016
0.500	0.250	0.125
0.750	0.562	0.422

Table 2.1.2 tabulates some values with alternating row colors.

Table 2.1.2: *Alternating background color for rows.*

α	β	γ
0.1	0.2	0.3
0.4	0.5	0.6
0.7	0.8	0.9

Table with nice rulers

Table 2.1.3: *Generic table with different sized rulers.*

header1	header2	header3	header4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4

Table with nice rulers and alternating rows

Table 2.1.4: *Generic table with alternating rows and different sized rulers.*

header1	header2	header3	header4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4

Table with nice rulers and alternating rows 2

Given

$$f : \mathbb{R} \to \mathbb{R},$$

Table 2.1.5: *Generic table with alternating rows and different sized rulers.*

header1	header2	header3	header4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4
col1	col2	col3	col4

magic happens

$$\int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x$$

This is

Rewrite this!

2.2 Figures and Tables

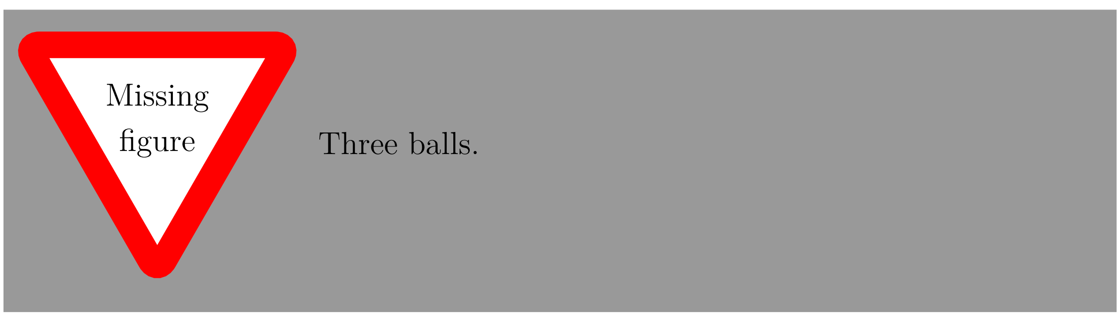

Figure 2.2.1: *Three balls.*

Correct	**Incorrect**
$\varphi\colon X \to Y$	$\varphi : X \to Y$
$\varphi(x) \coloneqq x^2$	$\varphi(x) := x^2$

Table 2.2.1: *Proper colon usage.*

It is now easy to tell that Birch and Swinnerton-Dyer are two people.

Correct	**Incorrect**
$A \implies B$	$A \Rightarrow B$
$A \impliedby B$	$A \Leftarrow B$
$A \iff B$	$A \Leftrightarrow B$

Table 2.2.2: *Proper arrow usage.*

Correct	**Incorrect**
-1	-1
1–10	1-10
Birch–Swinnerton-Dyer conjecture	Birch-Swinnerton-Dyer conjecture
The ball — which is blue — is round. The ball—which is blue—is round.	The ball - which is blue - is round.

Table 2.2.3: *Proper dash usage.*

Correct	**Incorrect**
“This is an ‘inner quote’ inside an outer quote”	"This is an ’inner quote’ inside an outer quote"

Table 2.2.4: *Proper quotation mark usage. The* `\enquote` *command chooses the correct quotation marks for the specified language.*

3

Bayesian Inference

3.1 The Bayesian Paradigm

Part II

Methodology & Computational Approach

4
Computational Approach

4.1 Method

4.1.1 Project Method 1

Part III

Results & Discussion

5

Results and Discussion

5.1 Results

5.1.1 Project Results 1

Part IV

Conclusion & Future Research

6
Conclusion & Future Research

6.1 Conclusion

6.2 Future Research

Appendices

A

Appendix A

A.1 Appendix.append(stuff)

Appendix A

B
Abbreviations and Units

B.1 List of Abbreviations

AURKA	Aurora Kinase A
AURKB	Aurora Kinase B
AURKC	Aurora Kinase C
CDK	Cyclin-Dependent Kinase
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CML	Chronic Myelogenous Leukemia
CPC	Chromosomal Passenger Complex
DOF	Degrees of Freedom
EGFR	Epidermal Growth Factor Receptor
GROMACS	GROningen MAchine for Chemical Simulations
HDX	Hydrogen-Deuterium Exchange
INCENP	Inner Centromere Protein
MD	Molecular Dynamics
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance

PBC Periodic Boundary Conditions

PCA Principal Component Analysis

PK Protein Kinase

PKA Protein Kinase A

RMSD Root-Mean-Square Deviation

RMSF Root-Mean-Square Fluctuation

VMD Visual Molecular Dynamics

B.2 Units

M	molar (1 mol/L)
μs	microsecond (10^{-6} s)
ns	nanosecond (10^{-9} s)
ps	picosecond (10^{-12} s)
Å	Ångström (10^{-10} m)