

# Linear versus Non-linear Dimensionality reduction

Feature extraction/engineering as preprocessing of  
audio/images/statistical match data

Daniel Runge Petersen, Gustav Svante Graversen, Lars  
Emanuel Hansen, Raymond Kacso, Sebastian Aaholm

Computer Science, cs-22-dat-5-05, 2022-12

Semester Project



Copyright © Aalborg University 2021

Composed and typeset by the authors using the  $\text{\LaTeX}$  Document Preparation System, based on the AAU report template by Daniel Runge Petersen [1].



**Electronics and IT**  
Aalborg University  
<http://www.aau.dk>

## **AALBORG UNIVERSITY**

### STUDENT REPORT

**Title:**

Linear versus Non-linear Dimensionality  
Reduction

**Abstract:**

This is the abstract...
-------------------------

**Theme:**

Theoretical data analysis and modeling

**Project Period:**

Fall Semester 2022

**Project Group:**

cs-22-dat-5-05

**Participant(s):**

Daniel Runge Petersen  
Gustav Svante Graversen  
Lars Emanuel Hansen  
Raymond Kacso  
Sebastian Aaholm

**Supervisor(s):**

Alexander Leguizamon Robayo

**Copies:** 1

**Page Numbers:** 13

**Date of Completion:**

September 27, 2022

*The content of this report is freely available, but publication (with reference) may only be pursued due to agreement with the author.*



# Contents

<b>Preface</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Report outline . . . . .	1
<b>2 Problem Analysis</b>	<b>3</b>
2.1 Problem Statement . . . . .	3
<b>3 Methodology</b>	<b>5</b>
<b>4 Results</b>	<b>7</b>
<b>5 Discussion</b>	<b>9</b>
<b>6 Conclusion</b>	<b>11</b>
<b>Bibliography</b>	<b>13</b>



# Preface

This is the preface. You should put your signatures at the end of the preface.

Aalborg University, September 27, 2022

---

Author 1

<username@student.aau.dk>

---

Author 2

<username@student.aau.dk>

---

Author 3

<username@student.aau.dk>

---

Author 4

<username@student.aau.dk>

---

Author 5

<username@student.aau.dk>

---

Author 6

<username@student.aau.dk>





# Chapter 1

## Introduction

*Universitetsopgave: En faglig undersøgelse af et fagligt problem ved hjælp af faglige teorier og metoder.*

Write about: The initial problem, motivation, and the scope/background of the project/theme. Report outline at the end.

### 1.1 Report outline

The proposed report structure is as follows: \_\_\_\_\_

Is this outline in line with the theoretical theme?

- **Introduction** - This chapter
- **Problem Analysis** - Chapter 2
- **Methodology / Theory and Methods** - Chapter 3
- **Results** - Chapter 4
- **Discussion** - Chapter 5
- **Conclusion** - Chapter 6

The introduction describes the initial problem and the motivation for the project. The Problem Analysis chapter dives into the initial problem and leads to a final problem statement.

The Methodology chapter describes the methods and theory used to explore the problem statement. It also describes the data used and how it was collected/created.

The Results chapter is an evaluation of the results of the project.

The Discussion chapter is a discussion of the results and the project as a whole.

The Conclusion chapter provides a summary of the project and the results. It also provides perspective and reflection on the project and the process.



## Chapter 2

# Problem Analysis

Write about: linear methods, non linear methods, types of data - for example time series, spatial data, etc. - and how to deal with them. Also, write about the different types of problems that can be solved with machine learning? For example, classification, regression, clustering, etc.?

### 2.1 Problem Statement

#### 2.1.1 Audio

For models that predict what music is popular or what genre the music is we would like to see how big of an effect feature engineering has for the model. We would like to investigate which kind of dimensionality reduction works best considering both linear and nonlinear approaches and what they contribute to in the model and when it is a better fit. The performance of these dimensionality reductions is evaluated based on how they affect the performance of the model and their visualisations.

#### 2.1.2 Pokemon

For a model that clasifies Pokemon we would like to see how big of an effect feature engineering has for the model. We would also like to investigate which kind of dimensionality reduction works best and consider both linear and nonlinear approaches and what they each contribute and when theyre correct to use. The performance of these approaches might be evaluated based on their visualisations and how they affect model performance.

#### 2.1.3 Match data

For models that predict the outcome of football matches we would like to see how big the effect of feature engineering has for the model. We would also like to investigate which kind of dimensionality reduction works best considering both linear and nonlinear approaches and what they contribute to in the model and when

it is a better fit. The performance of these dimensionality reductions is evaluated based on how they affect the performance of the model and their visualisations.

## **Chapter 3**

# **Methodology**

Write about theoretical background and methodology of the project.



## **Chapter 4**

# **Results**

Describe the results of the project.





## **Chapter 5**

# **Discussion**

Discuss the results from chapter 4 and compare them to the problem statement in section 2.1. Also, discuss the methodology and the theoretical background in chapter 3. Finally, discuss the project as a whole and the process of the project.



## **Chapter 6**

# **Conclusion**

Based on the discussion in chapter 5, the results from chapter 4 and the problem statement in section 2.1, the following conclusions can be drawn:



# Bibliography

- [1] Daniel Runge Petersen. *AAU-Dat templates*. URL: <https://github.com/AAU-Dat/templates> (visited on 08/17/2022).