

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 \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: 5

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

Preface	vii
1 Introduction	1
2 Problem Statement	3
2.1 Audio	3
2.2 Pokemon	3
2.3 Match data	3
Bibliography	5

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

Write about: The initial problem, motivation, and the scope/background of the project/theme. Add an outline of the report.

Chapter 2

Problem Statement

2.1 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.

2.2 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 aproaches might be evaluated based on their visualisations and how they affect model performance.

2.3 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.

Bibliography

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