|  |  |
| --- | --- |
|  | |
| Machine Learning Assignment  TU060 : H&M Personalised Fashion Recommendations  Kaggle Competition | |
|  | |
| Ciaran Finnegan – Part Time – First Year 2021/2022  MSc in Computer Science (Data Science)  Student No : D21124026  8/5/2022 |  |
|  |  |

Table of Contents

1 Introduction 3

1.1 Purpose of Report 3

1.2 Project Strategy/Approach 3

1.3 Python Coding Best Practice 3

2 Data Preparation + Project Implementation 4

2.1 Data Import + Analysis 4

2.2 Data Preparation + Enrichment 4

2.3 Creating Training and Test Datasets 4

2.4 Scaling Datasets 4

2.5 Implementation of Nearest Neighbour Model 4

3 Model Evaluation Strategy 5

3.1 Find Predicted Products for Each Customer 5

3.2 Find Actual Products for Each Customer (Local Test Data) 5

3.3 Local Evaluation Metric 5

3.4 Local Evaluation Results 5

4 Observations + Conclusions 6

4.1 Establishing a ‘Baseline’ 6

4.2 Optimisation Steps 6

4.3 Project Observations/Future Suggestions 6

5 Kaggle Submission 7

6 References 8

# Introduction

## Purpose of Report

The..

## Project Strategy/Approach

The..

## Python Coding Best Practice

The..

# Data Preparation + Project Implementation

## Data Import + Analysis

<Diagram>

Figure 1: Data Flow Diagram

## Data Preparation + Enrichment

<Diagram>

Figure 1: Data Flow Diagram

## Creating Training and Test Datasets

<Diagram>

Figure 1: Data Flow Diagram

## Scaling Datasets

<Diagram>

Figure 1: Data Flow Diagram

## Implementation of Nearest Neighbour Model

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

# Model Evaluation Strategy

## Find Predicted Products for Each Customer

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

## Find Actual Products for Each Customer (Local Test Data)

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

## Local Evaluation Metric

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

## Local Evaluation Results

<Diagram>

Small Sample

Low Medium Sample

Higher Volume Sample

# Observations + Conclusions

## Establishing a ‘Baseline’

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

.

## Optimisation Steps

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

.

## Project Observations/Future Suggestions

<Diagram>

Including features used, prediction algorithm(s) used, parameters, etc.;

.

# Kaggle Submission

One could ipso lorem…

.

# References

[1] Polkovnikov, I. (2016). Unified Control and Data Flow Diagrams Applied to Software Engineering and other Systems. Retrieved 15 April 2022, from https://doi.org/10.48550/arXiv.1610.02374

[2]

[3]