Machine Learning Specification

C00270395

Ronan Green

Contents

[Introduction 2](#_Toc177559099)

[Ideas 2](#_Toc177559100)

[Summer Olympic Athletes 2](#_Toc177559101)

[Data Application 3](#_Toc177559102)

[Spotify Top Streamed Songs 3](#_Toc177559103)

[Data Application 3](#_Toc177559104)

[Top Rated IMBD TV Shows 3](#_Toc177559105)

[Data Application 3](#_Toc177559106)

[Technologies 4](#_Toc177559107)

[Languages 4](#_Toc177559108)

[Python 4](#_Toc177559109)

[Tools 5](#_Toc177559110)

[Jupyter Notebook: 5](#_Toc177559111)

[Google Colab: 5](#_Toc177559112)

[RapidMiner: 5](#_Toc177559113)

[Weka 3: 5](#_Toc177559114)

[Anacoda: 5](#_Toc177559115)

[MLFlow: 5](#_Toc177559116)

[Frameworks 5](#_Toc177559117)

[Tensorflow: 5](#_Toc177559118)

[Pytorch: 5](#_Toc177559119)

[Appendix 5](#_Toc177559120)

[Dataset Links 5](#_Toc177559121)

[Summer Olympics 5](#_Toc177559122)

[Spotify Streamed Songs 5](#_Toc177559123)

[Top Tv Shows 5](#_Toc177559124)

# Introduction

This specification report includes ideas I propose form my Machine Learning project. Included in this report are the ideas along with various possible technologies I would use and why I would use them. The appendix will have links to sites used as reference for the datasets and information on the technologies.

# Ideas

## Summer Olympic Athletes

This project includes a comprehensive list of athletes who have participated in the summer Olympics since in inaugural year on Athens in 1896 to the 2024 Paris Olympics. The data it includes ranges from the athletes name, sex, country, year city, sport, event and medal won.

### Data Application

* What this data set could be used for is predicting which country will win the most medals.
* Which country produce the most efficient athletes for certain sports.
* Countries that have been doing or worse in recent years and where they could end up in years to come.
* We can also see which country has more proficient males or more proficient females.

## Spotify Top Streamed Songs

This project idea includes the details of the most streamed songs of 2023. It includes details such as artist, song name, bpm of the song, the key the song is mainly in, the number of playlist it is in the day, month and year of release, number of artist on the song and ranking in charts.

### Data Application

* using this data we could see what bpm makes the most popular songs.
* What key listeners find most appealing.
* See what song name format is the most popular
* upcoming and dying artists and possibly the transition of artists between different bpm and keys
* What time of year yields the most successful releases.

## Top Rated IMBD TV Shows

This project includes the data from the top 250 TV shows of all time according to IMBD. It give rankings names, release dates, run time, rating tag and rating.

### Data Application

* This data can be used to see what length of show is the most popular in what years.
* If the runtime of shows is increasing or decreasing due to popularity,
* How many episodes is optimal for popularity.
* What rating tag is the most featured and how these ratings tags correlate to the ratings itself.

# Technologies

## Languages

### Python

Python has become the de facto language for machine learning. This id due to many factors such as its simplicity, versatility and its wide range of libraries and frameworks.

#### Libraries

Python gives access to many useful libraries for machine learning. These libraries include the following:

##### NumPy:

##### Pandas:

##### Matplotlib/Seaborn:

##### Scikit-learn:

##### Keras:

## Tools

### Jupyter Notebook:

### Google Colab:

### RapidMiner:

### Weka 3:

### Anacoda:

### MLFlow:

## Frameworks

### Tensorflow:

### Pytorch:

# Appendix

## Dataset Links

### Summer Olympics

<https://www.kaggle.com/datasets/stefanydeoliveira/summer-olympics-medals-1896-2024>

### Spotify Streamed Songs

<https://www.codecademy.com/resources/blog/machine-learning-engineer-portfolio/>

### Top Tv Shows

<https://www.kaggle.com/datasets/khushipitroda/imdb-top-250-tv-shows>