



ARTIFICIAL INTELLIGENCE GROUP ASSIGNMENT

BY:

Audrey Njenga

Edward Obala

Kelvin Wachira

Ziithe Hiwa

January 2021

1 Overview

The main objective of this assignment was to show our understanding of the various tools and libraries used in AI. This document serves as clear documentation of our process and understanding with extracts and examples of the methods used. To achieve this, we chose a Covid 19 Dataset, with the goal of studying and monitoring the growth of Covid 19 Cases using the various AI tools

2 About the Dataset

The chosen dataset (insert name of dataset) was a fairly large dataset, including data on confirmed, recovered, deceased cases of Covid19 worldwide. To match the set criteria, we filtered out the dataset to only include data within the African continent, using one of the stated libraries.

3 Tools

Tools used were

- Anaconda
- Python
- Jupyter Notebook
- TensorFlow
- Keras
- NumPy
- SciPy
- Matplotlib
- Pandas
- Scikit-Learn

3.1 Pandas

Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.

We used panda to read data on the jupyter notebook using the following code

```

import csv

#to read the data file on covid

import pandas as pd

#read file

covidAfrica = pd.read_csv("covid_19_clean_complete.csv")

covidAfrica.head(10)

#To use only African countries we create a filter
covidAfrica.loc[covidAfrica["WHO Region"]=="Africa"]

```

```

covidAfrica.loc[covidAfrica["WHO Region"]=="Africa"]
covidAfrica

```

The above code was used to filter and clean out the data from out side of Africa

3.2 NumPy

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

```

import pandas as pd
import numpy as np
import tensorflow as tf

##Monitoring Recovered Cases Using Numpy

#load the dataset
covid_data = pd.read_csv("covid_19_clean_complete.csv")

#filter for Africa
covidAfrica = covid_data.loc[covid_data["WHO Region"]=="Africa"]
covid_array = np.array(covidAfrica[1:])
covid_array
#slice the array to extract specific row
# covid_array[1,:]

#slice the array to extract recovered cases column

```

```

cases = covid_array[:,7]
cases

cases.astype(int)

#Calculate total recovered cases in Africa -sum up column
total = np.sum(cases)
print(total)
# changes recovered cases array into a tensor with tensorflow
tensor = tf.convert_to_tensor(cases, dtype=tf.float32)

print(tensor)

```

3.3 SciPy

SciPy is a free and open-source Python library used for scientific computing and technical computing. SciPy contains modules for optimization, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, ODE solvers and other tasks common in science and engineering.

```

import scipy as sp
import numpy as np
from scipy import cluster
def f(x):
    return x/y+x**2
x= np.linspace(0,50000,50)
y = 187

import scipy.optimize as opt
[xopt,fopt,gopt,Bopt, func_calls,grad_calls,warnflg] = \
opt.fmin_bfgs(f,x0=100,maxiter=50000,full_output=True)

```

3.4 Matplotlib

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

```

import matplotlib.pyplot as plt
df = covidAfrica[['Country/Region', 'Confirmed', 'Date']]
#filter data to South Africa
df = df[df['Country/Region'] == 'South Africa']
#filter data to June and July 2020
df['Date'] = pd.to_datetime(df['Date'])

```

```
start_date = '2020-06-01'
end_date = '2020-07-31'
mask = (df['Date'] > start_date) & (df['Date'] <= end_date)
df = df.loc[mask]
print(df.head())
```

3.5 Scikit-Learn

Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language.[2] It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy as seen on the notebook attached to this project.

```
import numpy as np
from sklearn.preprocessing import MinMaxScaler
# create a random 10 by 2 array
data = np.random.randint(10, 100, (10, 2))
# transform features by scaling
print(data)
scalar = MinMaxScaler()
new = scalar.fit_transform(data)
print(new)
```

3.6 TensorFlow

TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. Tensorflow is a symbolic math library based on dataflow and differentiable programming.

3.7 Keras

Keras is an API designed for human beings, not machines.

Thus keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library.

4 Conclusion

iiiiiiiiiii

References

Tomas 4, Thorsten Donig 41.2k88 gold badges101101 silver badges123123 bronze badges, and Seyma Kalay 26511 silver badge66 bronze badges, “What’s the correct use of when multiple authors?,” TeX, 01-Nov-1959. [Online]. Available: <https://tex.stackexchange.com/questions/4805/whats-the-correct-use-of-author-when-multiple-authors>. [Accessed: 25-Jan-2021].

J. Brownlee, “Your First Deep Learning Project in Python with Keras Step-By-Step,” Machine Learning Mastery, 15-Sep-2020. [Online]. Available: <https://machinelearningmastery.com/tutorial-first-neural-network-python-keras/>. [Accessed: 25-Jan-2021].

G. Google, “Introduction to TensorFlow —; Machine Learning Crash Course,” Google, 07-Mar-2020. [Online]. Available: <https://developers.google.com/machine-learning/crash-course/first-steps-with-tensorflow/toolkit>. [Accessed: 25-Jan-2021].

J. Brownlee, “A Gentle Introduction to Scikit-Learn,” Machine Learning Mastery, 15-Aug-2020. [Online]. Available: <https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/>. [Accessed: 25-Jan-2021].

Real Python, “Python Plotting With Matplotlib (Guide),” Real Python, 07-Nov-2020. [Online]. Available: <https://realpython.com/python-matplotlib-guide/>. [Accessed: 25-Jan-2021].

J. F. del Río and M. Wiebe, “Array programming with NumPy,” Nature News, 16-Sep-2020. [Online]. Available: <https://www.nature.com/articles/s41586-020-2649-2>. [Accessed: 25-Jan-2021].

A. Bronshtein, “A Quick Introduction to the ‘Pandas’ Python Library,” Medium, 29-Oct-2019. [Online]. Available: <https://towardsdatascience.com/a-quick-introduction-to-the-pandas-python-library-f1b678f34673>. [Accessed: 25-Jan-2021].