**Python Project Increment-2**

**Team Members: (Team Number:8)**

1. Sahithya Gadde
2. Aneesh Anne
3. Abdulmuhaymin Ahmad Zakari

**Motivation**:

Under the current circumstance, we the students of UMKC feel the pardon of making a move to participate in solving and studying the well-known virus that has conquered the world entirely. This virus and its dominance are the strongest motivations to point the guns of knowledge and all the resources toward this area and focus currently on this issue that everyone today thinks that it is a major concern.

**Significance:**

According to worldometers.info, there are 142,178 active cases in the united states alone by today, the 29th of March 2020 ("Coronavirus Cases"). The importance of taking an action regarding the pandemic is crystal clear since there is no cure yet for this virus and the spreading rate of it is quietly significant. According to the same resource, the total deaths since the beginning of this virus in china is 33,976 and the number is increasing day by day.

**Objectives:**

The main objectives of this project is to apply what we have studied under this course of machine learning algorithms, applications, and techniques on different datasets available and published online and try to study the essence, the common attributes, and properties of this virus in a way t come up with any patterns that this virus follows.

**Features:**

The project will provide different studies concerning the well-known virus, Covid-19. We will be using AI libraries and machine learning algorithms on datasets of coronavirus as a case study and try to extract features and properties from them. For now, mainly we will use the python programming language to serve this goal with different implemented libraries such as Scikit-learn, TensorFlow, PyTorch, and Keras. The progress of this project will be reported on the milestones set by the instructor with a proper citation to any updates in this regard.

**References:**

After seeing many highly recommended datasets published online that describe and state this virus, the team decided to make the project based on this information and datasets. Many resources provide hourly updated datasets and streams to the public and one of those resources is kaggle.com. The choice of which is the most suitable for this project will be mentioned in the up-comping reports, but the scope of the project is already targeted and the work on these datasets has already been established. 202020

**Increment -1**

**Approach:**

Among many datasets which are available over the internet we have chosen a dataset related to Covid19 and this dataset contains many features related to the patients dragonized with Covid19. Out of all the features available in the dataset we have chosen 64 main features and developed a classification model to predict if the test result of a person tested for coronavirus is positive or negative.

At First, out of the 108 features in the dataset we have dropped 44 irrelevant features and mapped negative as 0 and positive as 1. And then replaced the null values with mean values in each column and performed correlation on the dataset. Then we have assigned the target feature to the variable Y and the remaining features to X. We have divided 0.2% of the data as test data and developed a SVM Model with 90% accuracy.

**Code and Output:**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**Increment- 2**

**Approach:** In this increment, we have further created a Deep Learning Sequential Model with an accuracy 92% and we have also developed some visualizations.

**Screenshots:**

**Output and Code:**

**Visualizations:**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A picture containing screenshot

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A close up of a logo

Description automatically generated**

**Keras Deep Learning Model:**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A close up of a piece of paper

Description automatically generated**

**A screenshot of a social media post

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A close up of text on a white background

Description automatically generated**

**References:**

<https://www.kaggle.com/roche-data-science-coalition/uncover#diagnosis-of-covid-19-and-its-clinical-spectrum.csv>

**Future Work:** Will further work on image classification and the improvements advised in the meeting today.

Contribution:

**Sahithya G:** Worked on Documentation and Machine Learning Models and will work on the improvements advised in the meeting today.

**Aneesh A:** Worked on finding the appropriate dataset, analyzing, preprocessing and cleaning the data and developing the Models and will work on image classification.

**Zakari:** Worked on developing the Deep Learning Model and will continue to work on the model to improve accuracy and will work on image classification.