Project Objectives:

With the spread of the COVID-19 virus worldwide, the pressure it has placed on the public in various fields cannot be ignored. For individuals, it is necessary to solve the problems caused by the COVID-19 virus in health, psychology, life and work. For social groups, the spread of the COVID-19 virus between countries and regions, its spread trend since the outbreak, and the impact it has and will have on society are all worthy of attention. In this challenge, our team want to do some meaningful data visualizations and deeper analytics about the COVID-19 pandemic.

Based on the above facts, this project team hopes to develop such an information platform: It provides real-time, accurate, usable, and meaningful data, which cover various aspects of the COVID-19 epidemic that the society and the public are concerned about. The representation of these data should be easy to read, understand and use. It is deployed on the AWS cloud computing service platform to achieve real-time supervision and flexible allocation of resources. It provides an online browsing interface for users in the form of a website. This website should provide support for different platforms, such as PC and mobile. Considering these needs of users, we decided to create a website framework to host our application. Therefore this application is not only the website for visualizing the epidemic data, but also provides more useful information behind the huge data. We believe that this application can help users have more good understanding from the data.

Project Main Functions:

According to the brief requirements of this project, the content finally presented to users by this project is as follows:

1.A real-time map representing the distribution of COVID-19 epidemic situation. It should look like a map of the world, with regions and countries distinguished by colors according to the current epidemic situation. Each map block that can be selected should display the following information: The current number of diagnoses, the cumulative number of diagnoses, the number of cured, the number of deaths.

2.A series of charts on the development trend of the epidemic. According to the data provided by the above map, a corresponding chart is generated according to time and area. Including:

historical diffusion trend chart, and provide machine learning support to generate future trend chart; daily data update chart; regional comparison function, etc.

3.Hot news or topics related to COVID-19. According to the user's regional choice, provide corresponding epidemic information, including but not limited to news, time comment, community discussion topics, etc.

4.Area search function. It is convenient for users to quickly find the areas they want to view, or popular areas.

Project Plan:

This is the schedule of our group(based on table 2):
Complete the drafts of front-end pages. **Due**: 11/06/2020
Complete data collection and data pre-processing , **Due**: 12/06/2020
Completed construction model framework. **Due**: 15/06/2020
Complete the analysis of the COVID-19 data. **Due**: 05/07/2020
Complete the analysis of news, public opinion, video and other content related to COVID-19
Due: 05/08/2020
Complete project deployment and go-live **Due**: 08/08/2020

Roles:

Xuesong Zhang: Xuesong is good at web development and java programming, so Xuesong 's work is mainly focused on back-end development, development of interface, calling databases and the development of AWS. Participate in part of machine learning, data scraping, data analyzing (with other team members).

Nan Xiao: Xiao is the main person in charge of front-end design. At the same time, he is also responsible for data collection, data pre-processing, and data visualization tasks. Our team will use the flask framework to build our application, so we need a person who can balance the front-end design and data visualisation by python programing. In website design part, Xiao will use geographic GeoJSON data and presents it in a Flask application that uses Mapbox.
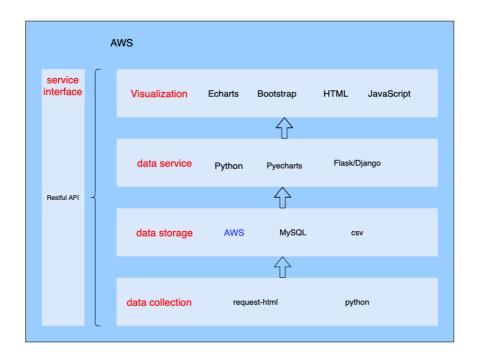
Qiangang Zhu: Zhu has extensive experience in machine translation related projects, and he is good at machine learning and deep learning. He will also be responsible for data processing, data scraping, and data analysis in the project. Considered that Zhu is good at Python and java programming, when other team members encounter programming problems, he also acts as a technical guide.

Yanlin Ren: Yanlin is mainly responsible for the deployment of Cloud and data visualization (with Nan Xiao). Similar to other team members, Yanlin will also use Python for data-related tasks: Data processing, data analysis, model evaluation. She is also the spokesperson of the group, responsible for the management of this group.

PS: This is the current rough division of roles, and we will also adjust it to specific work needs in the project.

Architectures:

Our application is deployed on the AWS cloud computing service platform to achieve real-time supervision and flexible allocation of resources. It provides an online browsing interface for users in the form of a website. This website should provide support for different platforms, such as PC and mobile. In order to have responsive layout for this website, we would use bootstrap framework to design. Considering this project includes a big amount of data analysis with python programming, we picked Flask framework to implement. Although Flask is actually not an MVC framework. It is a minimalistic framework which gives developers a lot of freedom in how structure applications, but MVC pattern is a very good fit for what Flask provides. Therefore we are intended to apply MVC pattern in our application. When users requests to view a page by entering some keyword that they like, the application matches the keyword to a predefined page, The controller action uses the models to retrieve all of the necessary data from a database, places the data in an array, and loads a view, passing along the data structure. Finally, the view accesses the structure of data and uses it to render the requested page, which is then presented to the user in their browser. The whole architectures are showed in below.

Data Plan:

Based on the description of the project in the Objectives section, the project team believes that the following data is necessary:

1. Since the beginning of the epidemic, information on the number of diagnoses, the number of deaths, and the number of people cured in various countries/regions has been counted. Such data can be obtained from the official website of the health committees of various countries, or real-time data provided by some portals, or from the open source database on Github. This information is the basic information required by the project, used to generate epidemic maps and charts, and learning algorithms used for time series forecasting to generate forecast trend charts.

2. News, articles, and community forum posts. These data can be obtained from the News API website or the tracking API of social media. The data should be subject to text analysis to distinguish ranking criteria such as relevance and real-time.

We need to choose some appropriate dataset, which need to meet the following characteristics: data reliability, data diversity, notions of "agreement". These are the data we currently choose. We will evaluate whether they are suitable in the later stage, and it is possible to add more datasets to meet the requirements of our project.

2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository by Johns Hopkins CSSE

New Time-series dataset:
- time_series_covid19_confirmed_global.csv ([Link Raw File](#))
- time_series_covid19_deaths_global ([Link Raw File](#))

Github:

Here are Github account of team member:

Name	User name
Yanlin Ren	YanlinRen
Nan Xiao	eVyOptJacob
Qiangang Zhu	qiangangZ
Xuesong Zhang	ZXuesong

In order to grasp the real-time progress, the team members will upload their own documents or codes to github every week.

Team Management:

The team members take turns to record the weekly meeting, checking the progress and evaluating together. The assignment of tasks will be based on the skills of different team members, but we should try our best to avoid the delay of the whole plan because a certain part of the task is not completed and ensure that each team member has participation and makes contribution in every part of this project. The team meeting is held in Zoom every Wednesday and Sunday. The meeting time can also be flexibly changed according to the weekly tasks. For the rest of the time, if we need to communicate and discuss our project with each other, we will use "wechat" or "whatsup".