Software requirements specifications documentation for “CDN”

1. Introduction

1.1 Purpose:

The website (CDN) is a tool that caters to the needs of people belonging to various socio-economic groups and divisions in the economy. This website allows the users to visualize and understand the magnitude of the effect the pandemic has brought to the society. Government can use this tool as a virtual advisory body to determine the course of action to be taken, to implement the most optimal solutions and laws. Citizens and netizens can use this tool to gain deeper understanding into the scenario in their locale and understand how cases might be using the prediction models. A news module web-scrapes popular news and provides a summary on the news module page with pictures and links. Businesses and students can use this website to deepen their understanding and even use our website as a secondary resource for developing projects and models of their own. Moreover, the stakeholders can predict the behavior of people from their sentiment towards the lockdowns and take appropriate business decisions to minimise their losses and optimise their returns by making informed and careful decisions. Virtual help desk is an AI powered chat-bot that can answer some questions users may have with respect to the scenario. In summary, our website serves as a platform that presents its users with information and data that could be used in any way the users wish, and the only limit to the applications is their imagination.

1.2 Table of contents:

2. Scope

Data from the Twitter users' tweets would be extracted using web scraping and stored in a database as a structured dataset, and based on their personal details such as gender, age, profession, location of the tweeter, and more, data analysis could be done through various aspects and the results could be shown in different modules and apps described in the Dashboard of the website. The user (government and the stakeholders) will be able to better understand how each demographic population and socio-economic groups feel about the implementations and lock-down.

By combining the data extracted and the analysis with demographic maps, a whole new domain of storytelling and presentation of the results will be opened to us. The government and the interested stakeholders can understand the sentiment of the population on a very specific and local level by observing how various parts of the country react to the lock-down. Moreover, live updates of people's opinions and sentiment will be displayed on the dashboard.

Analyzing the news and current affairs through web scraping would provide additional details and information that will be presented on the dashboard and increase depth to the story-telling aspect. Even summary and relevant information related to the pandemic will be presented in graphical way and also in small textual summaries, compiled from news, expert insights, and tweets.

The tweets module will allow people to view all Tweets from popular politicians, law enforcers, economics, psychologists (behavioral) and more directly in one single place. Relevant tweets related to COVID-19 will be displayed in this module, so that people will have one common place to acquire information on the matter at hand quickly, promptly and without being interrupted by any other unrelated posts and tweets.

The main purpose of the website would be to function as a mediator, combine information from structured and unstructured datasets and analyze them and provide useful, relevant and necessary information to the government and other stakeholders, so that they can better understand the sentiment of the citizens and strategize the laws accordingly, in a way to minimize the spread of the virus and at the same time make people happy and the country functioning normally.

2.1 Initial Functional requirements:

- Provide secure registration and profile management facilities for Customers
- Maintain a secure database for data storage and data analysis and visualization on the website.
- Provide a feedback and support mechanism to understand to needs of the users and assist the users.
- The experts' forum will be used to display the tweets and new articles about the pandemic from reputable and reliable sources.
- A dashboard containing several data visualization modules will be made available on the main page in order to present the users with the information.

- A news module will help the users gain information easily and also access several news articles from the same platform.
- A virtual help desk will help the users interact and gain information easily.
- The live sentiment analysis using VADER will be displayed on the dashboard along with the average sentiment values too in order to understand the way the population feels.
- A simulator will depict the way an epidemic spread among the population to help the users understand, and experiment with the sample population to gain in-depth understanding about the mechanism of the virus spread.

2.2 Initial Nonfunctional requirements:

- Data security: Secure access of confidential data (user's details).
- Accessibility: 24 X 7 availability
- Flexibility: Must be capable of handling peak time orders and demands
- Provision of an advertisement space to catch the customer's attention and behave as a source of revenue.
- In addition to the above-mentioned points, the following are planned to be delivered if deemed necessary.
  - Interactive user interface with multiple and customizable themes will be provided at a later stage based on the demand.
  - The data (raw) will also be made available to the users in CSV format as well as a link for the API.
- This list is subject to change and the final list will be determined by implementation constraints, market forces and most importantly, by end user demands.

2.3 Audience Definitions, Acronyms and Abbreviations

The intended readers of this document are the developers of the site, testers, website owners, managers, and coordinators.

S. no	Abbreviations	Description
1	RAM	Random Access Memory
2	HTTP	Hypertext Transfer Protocol
3	HTTPS	Secure Hypertext Transfer Protocol
4	TCP/IP	Transmission Control Protocol/Internet Protocol
5	NLP	Natural language processing
6	API	Application Programming Interface
7	CSV	Comma Separated Values
8	VADER	Valence Aware Dictionary sEntiment Reasoner

9	CSS	Cascading style sheets
10	JS	JavaScript
11	DB	Database
12	UI	User Interface
13	SQL	Structured Query Language
14	OS	Operating system

2.4 Technologies to be used

Software and Programming languages:

- Python 3: This is one of the most powerful languages around that implements ML algorithms and other tools that would be necessary for the website that is being built. It is important to use such powerful and efficient language with many libraries and modules because data analysis is the most important aspect of our website.
- HTML, XML: Hyper Text Markup Language and Extensible markup Language are the predominant markup languages for web pages. It provides a means to describe the structure of text-based information in a document and to supplement that text with interactive forms, embedded images, and other objects.
- JavaScript: A client-side scripting language used to create dynamic web content and user interface.
- CSS: This is one of the most crucial part in making the website aesthetically more pleasing and attractive.
- Flask: This is a web framework that helps us mount the python scripts onto a webpage and combine it with a database, HTML page and JS.
- Python libraries: All the necessary python libraries for machine learning and data visualization such as scikit learn, matplotlib, NumPy, pandas, seaborn and many more including the tool called VADER which is an NLP tool to analyze the sentiment.

Hardware:

- We are using a local machine that has Python 3 and Flask installed. This development does not require any specific machine with specialized tool set, any regular local machine with adequate amount of processing speed, storage, RAM capacity, and necessary software installed.

2.5 Operating Environment

The website shall operate in all famous browsers, for instance Internet explorer, Google Chrome, Firefox, Safari, and Opera.

2.6 User characteristics

- The user must be capable of using the internet and accessing its resources on a smartphone, a computer or any device that can support browsers to be run.
- The user must also be able to understand the content that would be in English
- The user must be capable of creating an account and be familiar with using websites and browsing on the platform.
- The user must be of age 12 or above and must agree to the terms and conditions to continue to use the website.
- The user can be a government official, businessman, student, common citizen, netizen, developer, blogger, journalist, or any other professionals that wish to access the resources and data, and visualize the same for the purpose of learning more about the situation are allowed to use the website.

2.7 Limitations of the project

- Data visualization and analysis are limited to the tools provided on the website. If the user wishes to process the raw dataset to create more attributes or create new graphs and images, it is on possible on the user interface. The dataset will be made available in both CSV and API forms and the user can use the raw dataset to perform his analysis on his local machine.

---

3. Module description and purpose

3.1 Login

The user enters the website using the login credentials. The user would be required to enter “Username” and “Password” to enter the website. A forgot password option will also be provided to reset the password in case the user has forgotten the password or the username.

3.2 Signup/registration

The user can make his or her account on the website by creating a new profile by signing up via the registration form/webpage. Several fields such as Name, Username, Password, Date of Birth, email address, country and state of residence, and phone number will be asked and will be mandatory.

3.3 Personal data editing

The user can change his or her credentials anytime through this module. The data and profile updating takes place in this module, and an email confirmation will also be sent to the user.

3.4 Data visualization

The data representation will be in terms of graphs, charts and more. The visualization tools will be made available on the dashboard and each of the modules would depict different information. A demographic map will depict the spread of the virus in different regions. A trend of spread will also be displayed in a different module. Also, other forms of data visualization would be depicted on the dashboard along with the live sentiment analysis.

3.5 Experts' opinion forum

This is a forum where Tweets and posts of experts and big stakeholders in the domain of the pandemic are displayed and updated automatically to showcase the perspective of the experts, and hence the name, experts' opinion forum. The purpose of this is to have one go-to place for the opinions of the most important and influential experts for people such as law enforcers and law makers for them to understand and use this page as a virtual advisory tool to implement laws and lockdowns.

3.6 Pandemic simulator

Simulator takes a sample input of connected nodes and r_0 spread of any virus, then pictorially simulates how the virus spreads from a particular infected node in the graph over time.

3.7 Database

The database will store the users' details and also store the data from the APIs to use for data analysis and to build prediction models. The database will be implemented using SQL and the local machines will act as a database; however, a cloud service can also be used in lite version if the cloud storage becomes necessary in the future.

3.8 Support

The support module will provide the users with options to give feedback or ask queries if they are facing some technical issues. The members of the support and customer care would be responsible for replying to the emails, calls, and the feedback forms.

3.9 News module

This interactive tool is meant for simple and basic uses such as requesting details about the epidemic, lockdown details, prevention measures, restrictions, and general pages about the pandemic. The module would also be able to acquire the data from scraping and display it on the UI of the. Only 9 news articles would be displayed at a time to prevent overcrowding of the page and to reduce loading time for the user and improve user experience of the website. All articles are separate and are always and constantly updated to provide the most relevant articles and journals to the user at the time of request.

3.10 Virtual help desk

This interactive tool is meant for simple and basic uses such as requesting details about the epidemic, lockdown details, prevention measures, restrictions, and general questions about the pandemic. In future, the chat bot would be equipped with being able to answer questions directly related to the data analysis and produce data in runtime such as “what is the number of cured cases and total cases in Gujarat”. Moreover, the chatbot would also be able to acquire the data on the sentiment of the user to use for the sentiment analysis part mentioned above. The sentiment of the users is stored along with their profile in the database. This level of intricacy is out of our scope for now, but we could try and implement it in the future.

4. Specific requirements

4.0 Description:

The user interface is common for all user of all backgrounds. Regardless of whether the user is a government official or a student the data visualization would be the same, and the reason for this is that the users may be hesitant to give their occupation or ambiguous about the tools they require from the website, therefore we decided to provide all the tools to all users and the users have the freedom to decide how much of it to make use of depending on their requirements. The datasets are also acquired via API from the link: https://api.covid19india.org/ that provides raw and processed data; however, we would be using the raw data to process our datasets.

4.1 Login

4.1.1 Functional requirements:

- This interface will consist of two mandatory fields namely, “Username” and “Password”. Moreover, there will also be “New User’s Registration” and “Forgot Password” options which will redirect to the “Registration” page in case a user forgets the login credentials.
- If the password entered is correct the user is directed to the main user interface, otherwise an error message is displayed. The user must enter the correct credentials in order to get in.
- If the user enters the wrong login credentials five times continuously, then a popup message will ask if the user would like to change the password and also, in order to prevent bots from getting unwanted access, ask the user to fill a captcha. If the user chooses to change password, then the user is redirected to the “forgot password” page (mentioned below). Else, the user is allowed to try different login credentials for five more times, if the captcha is valid.

4.1.2 Nonfunctional requirements:

- Every unsuccessful attempt must be recorded and stored. If any signs of breach are detected, then the user should receive the information via email.

- The website should be capable enough to handle thousands or even millions of users without affecting its performance
- The software should be portable and moving from one OS to other OS should not create any problem. Moreover, the browser window of any size should be compatible with the website.
- The login page must be simple, elegant, and easier to browse and use for people of various age groups.
- The page should be quick and responsive
- The database must be secure and a back-up database must be available to ensure safety and security of the data.
- The page should be operational every day at all times
- Advertisements on the login and sign-up page should be appropriate, related to the site, and not too large that it occupies the whole page, must be of reasonable size and proportion.

4.2 Registration

4.2.1 Functional requirements:

- The user must enter his or her personal details such as Name, Username, Password, Date of Birth, email address, country and state of residence, and phone number.
- The strength of the password will be shown to make sure the user can build a highly secure and a strong password for their interface. Users will be warned about any mistakes on data format or any other constraints by validation notes and error messages. The password must be of at least 8 characters, at least one numeric and alphabetical value, at least one special character, and at least one capital letter.
- The username and the phone number must be unique. When the button "Save" is clicked, the server will check if the username or the phone number or email is already taken and alert the user.
- The user must also be over the age of 12. An error message will be displayed if the age is under the age of 12.
- A new user would be created if everything is entered correctly and saved.

4.2.2 Nonfunctional requirements:

- Every unsuccessful attempt must be recorded and stored. If any signs of breach are detected, then user should receive the information via email.
- The website should be capable enough to handle thousands or even millions of users without affecting its performance
- The software should be portable and moving from one OS to other OS should not create any problem. Moreover, the browser window of any size should be compatible with the website.
- The sign-up page must be simple, elegant, and easier to browse and use for people of various age groups.

- The page should be quick and responsive
- The database must be secure and back-up database must be available to ensure safety and security of the data.
- The terms and conditions of the site must be made available to the user after the registration
- The page should be operational every day at all times
- Advertisements on the sign-up page should be appropriate, related to content, and not too large that it occupies the whole page, must be of reasonable size and proportion.

4.3 Personal Data Editing

4.3.1 Functional requirements:

- If a user wants to change his or her personal information, he/she can enter his profile by clicking on his name at the top right corner of the main page and he will be directed to the personal details editing page. Once the editing is done, the user must choose to either save or discard the changes. Based on the response the modification would be made.
- This functionality can be used multiple times.
- Once the password is changed, the website will automatically send an email to the user's email ID.
- If any of the attributes are to be modified or changed, the user is required to enter the username and the password to continue.
- The email-ID cannot be changed and if another email-ID is to be used, the user must create a new profile through the registration page.
- The user must manually click on "save" to save the changes made and if the user wishes to discard it, he/she could do that as well. The button will be placed right next to each other at the bottom of the page.
- If multiple changes are done by the same user very often then a captcha is to be introduced after 3 times. This is to prevent bots from changing credentials for users and preventing unwanted and unauthorized access into accounts via hacking or other illicit means.

4.3.2 Nonfunctional requirements:

- The website should handle multiple changes and frequent changes by the users.
- The users must be informed that the credentials are updated every time the credentials are changed.
- The email must be sent immediately and automatically.
- The page should be operational every day at all times
- Advertisements on the credential updating page should be appropriate, related to the content, and not too large that it occupies the whole page, must be of reasonable size and proportion.
- The page should be quick and responsive

- The software should be portable and moving from one OS to other OS should not create any problem. Moreover, the browser window of any size should be compatible with the website.

4.4. data visualization

4.4.1 Functional requirements:

- The data will be available in many formats and types. The demographic map displays the data in an interactive format on a map of the location. The circles on the map will depict the extent to which the CCP virus has spread and the impact and magnitude of the situation is shown by the intensity of the color.
- The data visualization tool also shows the number of current cases, cured, and deceased cases with respect to time in an interactive graphical way. This can be tuned to any state or city in India.
- The live sentiment of the citizens and netizens are also displayed with respect to time using the VADER NLP tool.
- The average responses for each are also displayed separately. Options to allow addition of more data visualization modules would also be given, and as more modules are built, they will be updated on the website.

4.4.2 Nonfunctional requirements:

- The data must be accurate and always up to date.
- The website must be functioning all day long.
- The website must be able to respond immediately and promptly to the users' needs.
- The graphical tools must be able to open on any device and browser and not be affected by the size of the window.
- UI must be interactive and simple to understand for any users.
- Advertisements on the main page (data visualization page) should be appropriate, related to grocery, and not too large that it occupies the whole page, must be of reasonable size and proportion.
- The page should be quick and responsive
- The website should handle multiple changes and frequent changes by the users.

4.5 News module

4.5.1 Functional requirements:

- Client Responsibilities
- Server Responsibilities
- Generic Answer Population
- Information Extraction & key of API
- User Interfaces

4.5.2 Nonfunctional requirements:

- The product should be open-source and use open source API.
- This module must be scraping very relevant information to the pandemic.
- Every article must be individual and seamless and the interface must be small and short
- Limited number of articles should appear and not flood the UI with many and push a lot of burden to the client side.

4.6 Experts' opinion forum

4.6.1 Functional requirements:

- The forum should collect the Tweets of several experts in the domain of Psychology, Law and order, Economics and History via web scraping and display the content on this forum.
- Snippets of relevant news articles and journals are to be displayed as well, on the side next to the tweets.
- The language must remain intact and if available the Tweets can be translated or paraphrased to help the mass understand the expert.
- The snippets must redirect the users to the corresponding tweets or the website on click.
- The time stamp should also be available, and the forum must be updated frequently.
- The users will only be able to view and read the content and they will not be allowed to reply, edit, report, or modify them directly on the website. They must login using their Titter account and be redirected to their profile on twitter and the actions can be performed there.
- The user can choose to view as many snippets as he or she chooses, by clicking the view options button.
- The users will also be able to perform basic filters to streamline their desired content.

4.6.2 Nonfunctional requirements:

- The experts' forum can be used by the law makers to understand the perspective of the experts in one place.
- The content of this webpage must be informative and simple.
- The tweets must be useful and relevant and not abusive or offensive.
- The advertisements must not be displayed on this webpage to ensure that the quality of the webpage is not tainted or damaged.
- The webpage must always display the latest tweets and news articles and must acquire them from credible and reliable sources.
- The news section does not necessarily have to be confined to the domain of India; however, all the content must be applicable to India.
- The data must be accurate and always up to date.
- The website must be functioning all day long.
- The website must be able to respond immediately and promptly to the users' needs.
- The graphical tools must be able to open on any device and browser and not be affected by the size of the window.

- UI must be interactive and simple to understand for any users.
- The source of the news articles and journals, and the tweets must be made available at the bottom of the corresponding items.

4.7 A pandemic simulator

4.7.1 Functional requirements:

- Simulator should call randomized or probabilistic trials and simulate the spread in a connected graph of nodes.
- Should pictorially show every stage/day of spread step by step.
- May also extend to mortality rates within the infected nodes.

4.7.2 Nonfunctional requirements:

- Accuracy: The simulation should depict and map an accurate picture of the overall spread depending on whichever node the user infects.
- Reliability: The plotted picture should be reliable enough and the map should show how the spread progresses.
- Performance: Simulation shouldn't take too long to generate over any given node in the connected graph.

4.8 Database

4.8.1 Functional requirements:

- The database should encrypt the data and store it in an encrypted form.
- The database should take the username, name, contact details, addresses, password, email id and other details that the users use to login in with and store them securely.
- The database should also store the sentiment of the users whenever the data is available.
- The database should also store the API files and CSV files containing the raw data from the Open source API repositories.
- The data visualization tools must be able to access the database and store data whenever needed.
- The ML algorithms must also be able to access the database swiftly and effectively.
- Prediction models are also stored in the database.
- The generated results and outcomes are also stored within the database in order to display them repeatedly without having to process it and perform the same repetitive tasks over and over again with the same dataset.

4.8.2 Nonfunctional requirements:

- The database must be safe and secure, and the data/information displayed must be updated from the database.
- The capacity of the database for each item must be large enough, and save information accumulated over weeks or months. Addition or removal of items must be possible.

- The information about the user is sensitive and vulnerable and therefore it must be secure and private.
- Data integrity, reliability, recoverability, and usability are some of the requirements that must be provided on the site.
- The DB should be operational at all times.
- The data must be highly secure and encoded and backup must be kept elsewhere to restore the data if the database is destroyed or damaged.
- The data must be stored in a hack-resistant and the database must be fortified with highly reliable cyber security systems.

4.9 Support

4.9.1 Functional requirements:

- At the initial stages, the support and maintenance for users are performed via direct contact with the developers as the size of the organization is relatively small and the role of maintenance is also done by the programmers. However, this will eventually come to be a separate entity and specially trained testers and customer care members will be assigned roles to assist the users.
- Phone, email, and feedback forum on the support page are the means of communication between the end users and the developers.

4.9.2 Nonfunctional requirements:

- The email and forms must be reverted back within one working day and assistance via phone call must be done during the working hours on weekdays.
- The assistance can be in terms of using the UI or the website if the user is unable to navigate.
- If the user is facing server crashes or slow updating of data, then the support team can be contacted.
- The customer care and support must be polite, patient and kind to the users.

4.10 Virtual help-desk

4.10.1 Functional requirements:

- Client Responsibilities
- Server Responsibilities
- Response Document Structure
- Generic Question Construction
- Generic Answer Construction
- Generic Answer Population
- Information Extraction & Database
- Supported Question Topics

- User Interfaces

4.10.2 Nonfunctional requirements:

- The product should be open-source and published under MIT license.
- The bot should reply under 30 seconds to users’ queries.
- Talking to the bot should feel like talking to a person (the bot should not reply instantly, 1- 2 second delay between responses).
- The users have to authenticate themselves before being able to query information from the bot.
- The main bot logic should be a separate complete package that can be easily integrated with.