A

PROJECT PLAN DOCUMENT

on

Viral Diseases Prediction

Final Year B. Tech. in Computer Science and Engineering

Submitted to

Department of Computer Science and Engineering

Annasaheb Dange College of Engineering & Technology, Ashta, Sangli.

(An Autonomous Institute Affiliated to Shivaji University Kolhapur)

by

Ms. Aditi Mali(20131051)

Roll No 4119

Under the Guidance

ofProf. P. B. More

Academic

Year2023-24

The Project Name: The Viral Diseases Prediction

Brief Description:

The Viral Disease Prediction project is a data-driven initiative aimed at developing and deploying predictive models to forecast the outbreak and spread of viral diseases, such as influenza, COVID-19, Zika, and more. The project leverages advanced machine learning and artificial intelligence techniques to analyze various data sources, including epidemiological data, climate data, travel patterns, and social media trends, to provide timely and accurate predictions regarding the emergence and propagation of viral diseases.

Needs Addressed by the Project:

1. Early Detection: The project addresses the critical need for early detection of viral disease outbreaks. Timely identification of outbreaks is crucial for implementing effective public health measures and mitigating the impact on communities.
2. Public Awareness: Viral Disease Prediction can also contribute to public awareness campaigns by disseminating information on potential disease hotspots and preventive measures, empowering individuals to protect themselves and their communities.
3. Rapid Response: By predicting the spread of viral diseases, public health authorities and healthcare systems can allocate resources more efficiently and respond rapidly to emerging health threats, reducing the overall burden on the healthcare infrastructure.

The name of the project manager and key team member

Project manager	Team member
1. Mr.Ajit Pradnyavant	1. Aditi Mali
	2. Vaishnavi Lokhande
	3. Arushi Billowria

Deliverable of the project

1. Predictive Models: The primary deliverable would be the predictive models themselves. These models would provide forecasts and predictions related to the outbreak, spread, and impact of viral diseases.
2. Data Collection and Integration: A well-organized dataset containing relevant epidemiological, environmental, and social data would be a crucial deliverable. This may involve data cleaning, preprocessing, and integration from various sources.
3. Reports and Documentation: Comprehensive reports outlining the methodology, data sources, model accuracy, and insights derived from the predictions.
4. Dashboard or Interface: A user-friendly dashboard or interface for public health officials, researchers, or the general public to access and visualize the predictions and data.

A list of important reference material

1. Research Papers and Journals:
- "Predictive Modeling of Epidemics: A Comprehensive Review" by Chowell et al.
- "Forecasting Infectious Disease Epidemics: A Review" by Shaman and Karspeck.
2. Public Health Organizations and Datasets and Data Sources:
- World Health Organization (WHO) reports and guidelines on infectious diseases and outbreaks.
- National Institutes of Health (NIH) publications on infectious disease research.
3. Software:
    - Jupyter notebook
    - Python
    - Mysql
    - Flask

A list of definitions and acronyms, if appropriate

1. Epidemiology: The study of how diseases spread and impact populations, including the patterns, causes, and effects of diseases.
2. Predictive Modeling: The use of mathematical and statistical models to forecast future events or trends, such as the spread of viral diseases.
3. Machine Learning (ML): A subset of artificial intelligence (AI) that uses algorithms to enable computers to learn from and make predictions or decisions based on data.

Risk Management

1. Data Quality Assurance :
- Risk: Inaccurate or incomplete data can lead to flawed predictions.
- Mitigation: Implement data validation and cleansing procedures. Ensure data sources are reputable and regularly updated.
2. Model Uncertainty:
- Risk: Predictive models may have inherent uncertainties.
- Mitigation: Communicate model uncertainties transparently. Develop ensemble models for more robust predictions.
3. Overfitting and Model Bias:
- Risk: Models may be overly complex and not generalize well.
- Mitigation: Use appropriate model evaluation techniques. Regularly update models to account for evolving conditions.

Technical process

1. Data Collection: Epidemiological Data: Gather historical and real-time data on disease cases, including location, time, and demographics. Healthcare Data: Access healthcare records and hospital admission data to monitor trends in disease cases.
2. Data Preprocessing: Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies. Standardize data formats and units for consistency.
3. Model Selection: Choose appropriate machine learning or statistical models for disease prediction, such as regression models, time-series models, or deep learning models.

4. Model Training: Split the data into training, validation, and test sets. Train the selected models on the training data while optimizing model parameters.
5. Model Evaluation: Assess the model's performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC for classification tasks.
6. Deployment: Create a website that accepts symptoms as input.

Summary Schedule, Detailed Schedule

Phase 1: Initiation and Planning Project(1 month)

Phase 2: Data Collection and Preparation Project(2 weeks)

Phase 3: Model Development Project(3 weeks)

Phase 4: Model Evaluation and Validation Project(3 weeks)

Phase 5: Deployment and Real-Time Monitoring(2 weeks)

Activity	Timeline
Selection of project	20 July - 15 August
Selection of models	16 August – 3 September
Analysis of model	3 September – 12 September
Technical setup	23 September -30 September
Collaboration with web framework	30 September -13 October
Testing	13 October – 23 October
Deployment	24October – 10 November