

Research Area: Data Analytics, AI to model epidemic patterns and disease dynamics

Project Title: Forecasting the probability of spread of COVID-19 considering Community Hygiene parameters

Faculty: Dr. Malaya Dutta Borah,
Assistant Professor, Dept. of Computer Science and Engineering,
NIT Silchar, Assam.

Project Summary:

COVID-19 pandemic created a huge threat to global public health. The main challenge is to prevent and control of spreading COVID-19 from local to global sphere since as on date there is no vaccine for it. One of the key factor of prevention and controlling of such diseases is to maintain community hygiene, a foundation of good health, clean environment and to main social distancing.

Community hygiene plays a big role in spread of COVID-19. The COVID-19 basically spreads from human to human i.e. when a healthy person comes in touch to an infected person; there is a high probability of getting infected.

The personal hygiene parameters such as covering the mouth with handkerchief/ tissue paper while coughing/ sneezing. This personal hygiene is the atomic unit of community hygiene. It this work, grass root level study of Community hygiene will be conducted. The data collected from the field work will be feed into Machine learning techniques to forecast the probability of spread/infection of COVID-19 in a particular locality.

For the ground work, three categories of COVID-19 infection zone i.e. Green, Orange and Red will be taken into consideration. This work will also focus on ecological variable and its impact to ensure effective information flows.

Objectives:

- To identify the community hygiene parameters related COVID-19 infection in Red zone and Orange zone.
- To perform comparative study on community hygiene parameters in Green, Orange and Red zone.
- To develop mathematical model for COVID-19 spread
- To forecasting the probability of spread of COVID-19 using Machine learning techniques from the community hygiene parameters.

Methodology:

The main focus of this research is to Forecasting the probability of spread of COVID-19 considering Community Hygiene parameters. In this respect, the work is divided into **three main modules** as follows:

- (1) Data Preparation:
 - Collection of data from three categories of COVID-19 infection zone i.e. Green, Orange and Red
 - Data gathering, cleansing, integration, transformation and visualization of data by focusing trends, Seasonality, Irregularity and Cyclicity of data.

- (2) Forecast Modelling Methods:
 - Classical Methods: Exponential Smoothing Model, Autoregressive and Moving Average (ARIMA) method.
 - Machine learning Methods: Long Short-Term Memory (LSTM), Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) models

- (3) Evaluation / Validation of Model
 - Forecasting Performance Measures
 - Cross-validation methods
 - Evaluating Model Accuracy: Mean Absolute Percent Error method

Expected Outcome: This research work will be beneficial not only to prevent the spread / infection of COVID 19 but also other pandemic diseases in future.