**NEHRU INSTITUTE OF ENGINEERING AND TECHNOLOGY COIMBATORE-641105**

Pravin Raj B ,Pulla Reddy Y, Vamshi krishna K ,Thirumaran G.K, Adithyan E.

### Title :- Public Health Awareness

### Algorithm:-

# Machine Learniing & Linear Regression:-

The device analysis of the illness based on the prediction made through the algorithm used by the device. The Naive Bayes Classifier is used to make intelligent health predictions, and it is implemented in this way.

The increase of the health problems and the need for effective medical health care have led to an investigation of machine learning that can be applied in health problems. This paper presents a recent systematic review of machine learning approaches in predicting health problems. Furthermore, we will discuss the challenges, limitations, and future directions for the application of machine learning in the health field.

We collect research articles and studies that are related to the machine learning approaches in predicting mental health problems by searching reliable databases. Moreover, we adhere to the PRISMA methodology in conducting this systematic review. We include a total of 30 research articles in this review after the screening and identification processes.

Then, we categorize the collected research articles based on the mental health problems such as schizophrenia, bipolar disorder, anxiety and depression, posttraumatic stress disorder, and mental health problems among children. Discussing the findings, we reflect on the challenges and limitations faced by the researchers on machine learning in health problems. Additionally, we provide concrete recommendations on the potential future research and development of applying machine learning in the health field.

# Data Set :-

https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey

# PROCEDURE:

FOR PREDICT THE FUTURE CHANGES IN PUBLIC HEALTH AWARENESS USING RANDOM FOREST.

STEPS INVOLVED IN RANDOM FOREST:

# 1. DATA COLLECTION AND PREPARATION:

\* Data collection is the primary step in preparation for analysis. Gather historical data that includes both predictor variable (features) and the target variables (what you want to predict). In this we predict the future awareness that a company is about to make.

Clean and preprocess the data by handling the process

2. DATA SPLITTING:

\* Split the dataset into two parts (training and testing) or (validation set). The training set used to train the Random Forest model and testing set will be used to evaluate its performance

3. FEATURE SELECTION:

\* In large number of Datas. It includes such as Timestamp, Age, Reviews, Country, etc.

# 4. RANDOM FOREST MODEL TRANING:

\* Train the random forest model using the Training data. This involves building an ensemble of decision trees.

Random Forest is an ensemble of multiple decision trees. Each tree is trained on random subset of the features.

This randomness reduces overfitting.

# 5. MODEL EVALUATIONN:

\* Use the testing set to evaluate the model’s performance. Common evaluation metrics include MAE, MSE, or R-squared for regression tasks and accuracy, precision, recall.

# 6. PREDICT FUTURE VALUES:

\* Once the model is trained and evaluated, you can use it to make predictions on new, unseen data for future time periods.

# 7. MONITERING AND DELPOYMENT:

\* Moniter the real-time changes to continuous predictions. And deploy the model in real-time if your satisfied with the performance and requirements.