Covid Companion

Rishika Ahuja   
*Department of Computer Engineering*   
*St. Francis Institute of Technology*Mumbai, India  
rishikaahuja5@student.sfit.ac.in

Anashwara Kurien   
*Department of Computer Engineering*   
*St. Francis Institute of Technology*Mumbai, India  
anukurien7@student.sfit.ac.in

Ann Zachariah   
*Department of Computer Engineering*   
*St. Francis Institute of Technology*Mumbai, India  
annzachariah1020@student.sfit.ac.in

Jayashri Mittal

Assistant Professor

*Department of Computer Engineering*   
*St. Francis Institute of Technology*Mumbai, India  
jayashrimittal@gmail.com

Aishwarya John   
*Department of Computer Engineering*   
*St. Francis Institute of Technology*Mumbai, India  
ashjohn127@student.sfit.ac.in

***Abstract*—The COVID-19 outbreak, declared as a pandemic by the World Health Organization (WHO), rapidly spread across the globe spreading havoc in its wake. Having suddenly distorted each and every person's routine lifestyle, this pandemic has induced a considerable degree of fear, worry and concern in the population at large. COVID-19 is putting our mental health at risk since it has been proven stressful for plenty of people. Mental health includes our emotional, psychological, and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make healthy choices. Being social beings, humans were not meant to live in isolation. Community is critical for us to thrive, especially for someone with mental illness who is already experiencing the common symptoms of loneliness and isolation. As the death toll and hardships due to COVID-19 continue to rise, the number of people who are experiencing elevated and prolonged fear and anxiety appears to be growing as well[4]. Thus there has been an exponential increase in the number of people suffering from mental disorders with the pandemic taking over the globe.**

**In this project we have tried to quantify the effect of COVID-19 on the mental health of the users of our system and then suggest to them some basic ways to help them cope with it.**

***Keywords— Covid-19, Mental Health, Anxiety, Depression, Mental Health screener, Covid Anxiety Scale(CAS), Long Short-Term Memory (LSTM), Principal Component Analysis (PCA)***

# Introduction

The COVID-19 outbreak, which first emerged in China, has been declared as a pandemic by the World Health Organization (WHO). As the coronavirus pandemic rapidly sweeps across the world, it is inducing a considerable degree of fear, worry and concern in the population at large. COVID-19 is putting our mental health at risk since it has been proven stressful for plenty of people[1].A person's mental health affects how they handle stress, relate to one another and make decisions.It also influences the way individuals look at themselves, their lives and others in their lives.Studies show that at least 1 in 5 children and adolescents have a mental health disorder at any given time. Yet, fewer than one in five of these children receive the mental health services they need. Among young people, at least 1 in every 10 has a serious emotional disturbance at any given time. We’re social beings, and we are not meant to live in isolation. Community is critical for us to thrive, especially for someone with mental illness who is already experiencing the common symptoms of loneliness and isolation.Thus there has been an exponential increase in the number of people suffering from mental disorders with the COVID-19 pandemic wreaking havoc all over the globe.

Therefore, our project aims to judge the effect of COVID-19 on the mental health of the users and then suggest to them some basic ways to help them cope with it.

## Description

As COVID-19 has occurred suddenly and is highly contagious, this will inevitably cause people anxiety, depression, etc. In such times the study on the public psychological states and its related factors during the COVID-19 outbreak is of practical significance[1]. Mental health includes our emotional, psychological, and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make healthy choices. Mental health is important at every stage of life, from childhood and adolescence through adulthood. Mental health is important because it can help you to cope with the stresses of life,be physically healthy,have good relationships,make meaningful contributions to your community,work productively and realize your full potential.Mental health is also important because it can affect physical health. For example, mental disorders can raise your risk for physical health problems such as [stroke](https://medlineplus.gov/stroke.html), [type 2 diabetes](https://medlineplus.gov/diabetestype2.html), and [heart disease](https://medlineplus.gov/heartdiseases.html). Over time,mental health can change when one may be dealing with a difficult situation, such as the current COVID-19 scenario. The situation may wear a person out and overwhelm their ability to cope with it. This can worsen their mental health.This project is aimed at tackling the sudden downgrading of mental health of the general population that is brought about by the onset of the global pandemic. It has a huge societal impact and is the need of the hour.

*B. Problem Formulation*

To create a companion system for an individual which helps to analyze the mental health of the being during Covid-19 pandemic and provides precise suggestions based on their mental condition.

*C. Motivation*

With the sudden outbreak of COVID-19 which caused complete closure of schools, colleges, work and social life in general and the infectious power of the virus, it is inevitable to see a drastic rise in anxiety, depression and other stress reactions.[1] With the new limitations on daily life and social activities for an unknown period of time, the population will inevitably suffer from stress and anxiety and eventually may lose confidence in life, ultimately taking a toll on the mental health of society. Thus to help every individual through these socially distanced times we decided to take up the project to make a Covid Companion app. An App that will help individuals assess themselves as per their habit inputs if they are likely to have any sort of mental health problems and thus suggest a solution to help cope through the problems in ways that may bring significant changes to one’s daily lives.

*D. Proposed Solution*

In this project, a model is prepared based on the dataset which contains the mental health information of people during the Covid-19 pandemic[2]. The dataset is cleaned and processed using the appropriate algorithms and the relevant features from the dataset are selected which helps in training the model. The model then forms the basis for the COVID Companion app. The user's features are collected in the app and given to the previously trained model. The mental health of the user is analyzed and suggestions are given which help to improve the mental health of the user.

*E. Scope of the Project*

The global Coronavirus Disease 2019 (COVID-19) pandemic affected millions of people and forced the mobilization of governments worldwide.Such sudden changes have a direct impact on a person's mental health.This project aims to increase the mental wellbeing of the users and reduce their depressive symptoms. We aim to increase mental health awareness among the general public and help them cope with the stress and hardships caused by the sudden onset of this pandemic. Mental health awareness increases the chances for early intervention, which can result in a fast recovery. Awareness reduces negative adjectives that have been set to describe our people with a mental illness. By raising awareness, mental health can now be seen as an illness. These illnesses can be managed by treatment. [Mental disorders](https://medlineplus.gov/mentaldisorders.html) are serious conditions which can affect your thinking, mood, and behavior. They may be occasional or long-lasting. They can affect your ability to relate to others and function each day. But there are treatments. People with mental disorders can get better, and many of them recover completely. Our project hopes to build an app that proves to be a worthy companion to the mental health of those suffering in these hard times by helping them access the state of their mental well being and providing simple measures to help them cope with it and begin their journey on the path of recovery.

# Review of Literature

**[1]** **Title:** Study on the public psychological states and its related factors during the outbreak of coronavirus disease 2019 (COVID-19) in some regions of China

**Data:**

600 valid questionnaires were received. The Self-Rating Anxiety Scale (SAS) and the Self-Rating Depression Scale (SDS) were used.

A total of 605 psychological state questionnaires were distributed to the general population through online questionnaires from February 6 to 9, 2020. 600 valid questionnaires were received, and the response rate was 99.17%.600 valid answers, resulting in a 100% effective rate. Inclusion criteria include the following: (1) 18 years old and above and (2) completed questionnaire. Exclusion criteria include the following: (1) 17 years old and below and (2) questionnaire responses are not logical.

**Method:**

In this study, the Self-Rating Anxiety Scale (SAS) and the Self-Rating Depression Scale (SDS)were used.

The Self-Rating Scale questionnaire was completed by the following survey items according to the unified guidance methods. The contents include the following: (1) General information includes name, gender, age, education level, occupation and residence; (2) SAS is used to evaluate the subjective feelings of anxiety; and (3) SDS is used to measure the degree of depression.

The two independent self-assessment scales mainly assess the frequency of symptoms of the respondents in the past week, each with 20 items, using a 4-level score (1 for a little of the time, 2 for some of the time, 3 for a good part of the time and 4 for most of the time).

Standard score = 1.25 × total score.

Anxiety levels were graded as the following: standard score below 50 = non-anxiety; 50–59 = mild anxiety; 60–69 = moderate anxiety; and above 70 = severe anxiety.

Depression levels were graded as the following: standard score below 53 = non-depression; 53–62 = mild depression; 63–72 = moderate depression; and above 73 = severe depression.

The data were organized and analyzed using SPSS 23.0 software. The surveyed population was divided into anxiety groups and non-anxiety groups according to the SAS scoring criteria. According to the SDS scoring standard, the surveyed population was divided into depression groups and non-depression groups.

The analysis of the relationship between gender, age, education level, occupation, region and anxiety or depression initially used the chi-square test. The variables with p < 0.1 were entered in the multiple logistic regression analysis model. The correlation between SAS and SDS standard scores was analyzed by Spearman correlation analysis, and p < 0.05 on double sides was statistically significant.

**Algorithm:**Logistic Regression

**[2] Title:** COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak

**Authors:** Yuki Yamada, Dominik-Borna Ćepulić, Tao Coll-Martín, Stéphane Debove, Guillaume Gautreau, Hyemin Han,Jesper Rasmussen, Thao P. Tran, Giovanni A. Travaglino, COVIDiSTRESS Global Survey Consortium & Andreas Lieberoth

**Data:**

This N = 173,426 social science dataset was collected through the collaborative COVIDiSTRESS Global Survey – an open science effort to improve understanding of the human experiences of the 2020 COVID-19 pandemic between 30th March and 30th May, 2020.

The dataset contains demographic background variables as well as measures of Asian Disease Problem, perceived stress (PSS-10), availability of social provisions (SPS-10), trust in various authorities, trust in governmental measures to contain the virus (OECD trust), personality traits (BFF-15), information behaviours, agreement with the level of government intervention, and compliance with preventive measures, along with a rich pool of exploratory variables and written experiences

**Method:**

3,426 people accessed an online survey link to provide their experiences over a period of 62 days (30th March to 30th May. The stored dataset represents 125,306 people who met inclusion criteria (18 years of age and older and gave informed consent)

**[3] Title:** Coronavirus Anxiety Scale: A brief mental health screener for COVID-19 related anxiety

**Authors:** Sherman A. Lee

**Abstract:**

As the coronavirus pandemic rapidly sweeps across the world, it is inducing a considerable degree of fear, worry and concern in the population at large and among certain groups in particular, such as older adults, care providers and people with underlying health conditions. In public mental health terms, the main psychological impact to date is elevated rates of stress or anxiety which has not been adequately addressed. But as new measures and impacts are introduced – especially quarantine and its effects on many people’s usual activities, routines or livelihoods – levels of loneliness, depression, harmful alcohol and drug use, and self-harm or suicidal behavior are also expected to rise.

**Data:**

The dataset was created through an online survey and a total of 775 adults took that survey from 11-30 March 2020 out of which 446 men and 329 women with a combined mean age of 32.72 years. Participants were asked to report their age, gender, ethnicity, education, current residency, coronavirus diagnosis, and history of anxiety. Because the study focused on anxiety about the coronavirus, participants also had to have spent at least one hour during the past two weeks thinking about and/or watching media about the coronavirus, as well as have experienced significant anxiety, fear, or worry about the disease outbreak.

For the survey, participants were asked to choose, using a 5-point time anchored scale (0=not at all to 4=nearly every day over 2 weeks), the rating of “2” as the item response. This item was embedded into the questionnaire to eliminate participants who may threaten the integrity of the study’s results by not appropriately attending to the questionnaire’s content. The objective of this study was to develop and evaluate the properties of the Coronavirus Anxiety Scale (CAS), which is a brief mental health screener to identify probable cases of dysfunctional anxiety associated with the COVID-19 crisis.

Methods Used:

In this study Coronavirus Anxiety Scale was used to identify cases of dysfunctional anxiety. This 5-item scale, which was based on 775 adults with anxiety over the coronavirus, demonstrated solid reliability and validity. A pool of 20 candidate items was created based on the psychology of fear and anxiety literature where each item captures a unique manifestation of this particular form of anxiety that includes cognitive (i.e., repetitive thinking; worry; processing biases; dreaming; planning), behavioral (i.e., dysfunctional activities; avoidance; compulsive behaviors), emotional (i.e., fear; anxiety; anger), and physiological (i.e., sleep disturbances; somatic distress; tonic immobility;) dimensions of coronavirus anxiety.

To address the influences of sampling error, an internal replicability approach was employed by subjecting one half of the study's data to a principal component analysis and the other half to a series of confirmatory factor analyses using bias-corrected bootstrap Maximum Likelihood estimations. The PCA was used to identify the five most robust and representative symptoms of coronavirus anxiety, while a confirmatory factor analysis CFA was used to test replicability of the PCA results. The entire data set was then used to examine the construct validity and diagnostic viability of the coronavirus anxiety symptoms using a series of correlations and receiver operating characteristic (ROC) analyses, respectively.

Properties of the CAS items from the principal components analysis are dizziness, sleep disturbances, tonic immobility, appetite loss and abdominal distress. These symptoms had to be extracted from the first component of the PCA because they account for the highest possible squared correlations among the item pool.

Twenty coronavirus anxiety symptoms were subjected to a PCA with Varimax rotation with the first component accounting for 59.85% of total variance. Specifically, pattern/structure coefficients ranged from 0. 81 to 0. 84, communality coefficients ranged from 0. 74 to 0. 79, and cross-loadings ranged from 0. 23 to 0.29. These symptoms assess distinct, physiological reactions of fear and anxiety related to the coronavirus and are highly reliable as a cluster.

CFA was run to test whether or not the five symptoms identified in the previous PCA cohered together into a single, coronavirus anxiety construct where multigroup CFAs were run to examine if the coronavirus anxiety construct was being measured the same way across the demographic variables of age (18–29 vs 30 and older), gender (women vs men), and race (Whites vs non-Whites). Analysis of the parameter estimates revealed that sleep disturbances and appetite changes were much stronger indicators of the coronavirus anxiety construct for Whites than for the non-Whites. Although the strength of these two indicators were different between the races, the measurement of coronavirus anxiety was still valid for both groups.

CAS total scores were correlated with measures of disability, distress, and coping, to examine the validity of the construct and explore its relationship with relevant attitudes and demographic factors. CAS scores were strongly, positively correlated with functional impairment, alcohol or drug coping, negative religious coping, extreme hopelessness, and passive suicidal ideation, in support of this instrument’s construct validity as a measure of dysfunctional anxiety.

Receiver operating characteristic analyses were used to evaluate the diagnostic viability of the CAS as a mental health screening tool, as well as determine a cut score that best distinguishes individuals who experience clinically significant impairment because of coronavirus anxiety from those who were also anxious but not disabled by the pandemic. A CAS score 9 for ROC optimally classified adults as having or not having dysfunctional levels of anxiety with a false positive rate of 15%. Thus, these results support the CAS as a diagnostically accurate mental health screening tool with strong classification features.

CAS scores were strongly, positively associated with functional impairment, alcohol or drug coping, negative religious coping, extreme hopelessness, and passive suicidal ideation. The results of this study also support the CAS as a useful mental health screener, as its diagnostic qualities (90% sensitivity and 85% specificity) are comparable to other psychiatric screening tests.

**[4] Title:** Measuring coronaphobia: the psychological basis of the Coronavirus Anxiety Scale

**Authors:** Sherman A. Lee

As the death toll and hardships due to COVID-19 continue to rise, the number of people who are experiencing elevated and prolonged fear and anxiety appears to be growing as well. Although research has shown that many people eventually adapt to the threats of a pandemic and become less anxious and afraid, a significant number of people will not adjust and will instead develop long-lasting emotional problems.

To help address this growing mental health concern, I created the Coronavirus Anxiety Scale (CAS). The CAS was designed to identify individuals suffering from dysfunctional anxiety about the coronavirus, otherwise known as “corona phobia”.

The CAS is composed of five somatic-based symptoms of fear and anxiety that are triggered by thoughts or information about the coronavirus. Two of the CAS symptoms, dizziness and tonic immobility, seem to capture the physiological reactions of elevated fear to coronavirus related stimuli. Not surprisingly, these symptoms are also commonly reported during extremely frightening situations such as in panic attacks and physical assaults.

Available research suggests that the CAS exhibits solid psychometric and diagnostic qualities. Statistical analyses of the CAS using two independent samples have supported factorial validity and diagnostic accuracy of this mental health screener. A body of empirical research demonstrating the reliability and validity of the CAS is absolutely essential for this instrument to be considered scientifically legitimate and appropriate for clinical use.

Finally, studies examining the appropriateness of the CAS for use in specific groups, such as children, and its application using different methods of assessment, such as clinical interviewing, are also important for the advancement of this measure and the study of the corona phobia construct.

**[5] Title:**  JMIR Mental Health - Measuring COVID-19 Related Anxiety in Parents: Psychometric Comparison of Four Different Inventories

**Dataset:**

This dataset was prepared during the outbreak of the covid-19 pandemic. The purpose of this study is to compare the distributions, validities, and reliabilities of four different COVID-19 anxiety scales: Fear of COVID-19 Scale, Coronavirus Anxiety Scale, Pandemic Anxiety Scale, and one subscale of the COVID Stress Scales. A total of 1526 individuals started the web-based survey, resulting in a final sample of 515 parents after data cleaning (see below). Participants were predominantly mothers (465/515, 90.3%) with a university degree (307/515, 59.6%). A share of 19.8% participants (102/515) came from Carinthia. The parents were aged 18-58 years (mean 34.95 years, SD 5.39). The majority were employed (285/515, 55.3%) or worked in the household (180/515, 34.9%). In terms of family status, 27.4% participants (141/515) were unmarried and 68.5% (353/515) were married. At the time of the survey, 94.4% participants (486/515) were in a partnership. Four participants (0.8%) stated that they had confirmed COVID-19 infection.

**Method:** Participants were recruited on the web during a 6-week period (June 29 to August 9, 2020), mainly via social media in parenting or child-related Facebook groups and message boards. The evaluation of the HTTP referrers showed that half of the final sample (50.48%) found the survey through Facebook. Participants were required to be ≥18 years of age. The survey took an average of 22 minutes to complete. They were provided with a separate link to enter the raffle that could not be connected with the data from the study, which was anonymous.

All descriptive and correlational analyses were performed using SPSS version 25 (IBM Corporation). A two-tailed P value <.05 was considered statistically significant. Interpretation Guidelines for Pearson correlations were, r=0.10 considered to be a small correlation, r=0.30 a medium correlation, and r=0.50 a large correlation. For the EFA model fit, the root mean square error of approximation (RMSEA) was calculated. An item response theory (IRT) approach to provide measures for item discriminability and difficulty. A graded response model was used. This model is an extension of the two-parameter logistic model that is applicable for ordered polytomous variable data (eg, Likert scales). A sample size of N=500 is recommended for accurate parameter estimation. Marginal maximum likelihood estimation was used for estimation of the parameters. Item discrimination (alpha) and item difficulty (beta) for each scale separately based on the initial proposed unidimensional factor structures of COVID-19–related anxiety scales were calculated. Alpha values ≤0.64 were considered as low item discrimination, values between 0.65 and 1.34 as moderate, and values ≥1.35 as high.

**[6]Title :** Predicting fear and perceived health during the COVID-19 pandemic using machine learning: A cross-national longitudinal study

**Algorithm:**

For each of the predicted variables (i.e., fear and health), we fit two types of machine learning models, one being a linear model (LASSO [least absolute shrinkage and selection operator] and one being a non-linear model (ERT [Extremely Randomized Trees]).

**Accuracy:**

Fear of the virus:

LASSO: R2avg = .35, R2median = .35, p< .001;

ERT: R2avg = .32, R2median = .36, p< .001; Ntrials = 896

Perceived health:

LASSO: R2avg = .05, R2median = .08, p = .002, Ntrials = 896;

ERT: R2avg = .05, R2median = .09, p = .003, Ntrials = 896

**Limitations:**

It is unclear to which extent people with psychiatric disorders and medical conditions participated in the current study. Indeed, this may have distorting effects on perceived threat and health.

**[7] Title :** Artificial intelligence in prediction of mental health disorders induced by the COVID-19 pandemic among health care workers

**Proposed methodology:**

Canonical correlation analysis(CCA).

PCA algorithm

Clustering

ANN

SVM

Decision Trees

K nearest neighbours

# Challenges Identified

The challenges identified are as follows:

* The mental health of a user should be quantified objectively and accurately for proper care to be provided.
* In cases where the mental health is found to be critical care should be taken that the user is referred to a psychiatrist on an urgent basis.
* The suggestions need to be carefully curated to help the users.
* If the users find the suggestions to be triggering in any manner, the suggestions need to be changed immediately.

# Problem Definition

To create a companion system for an individual which helps to analyze the mental health of the being during Covid-19 pandemic and provides precise suggestions based on their mental condition

# Proposed System Methodology

A. *Architectural Design*

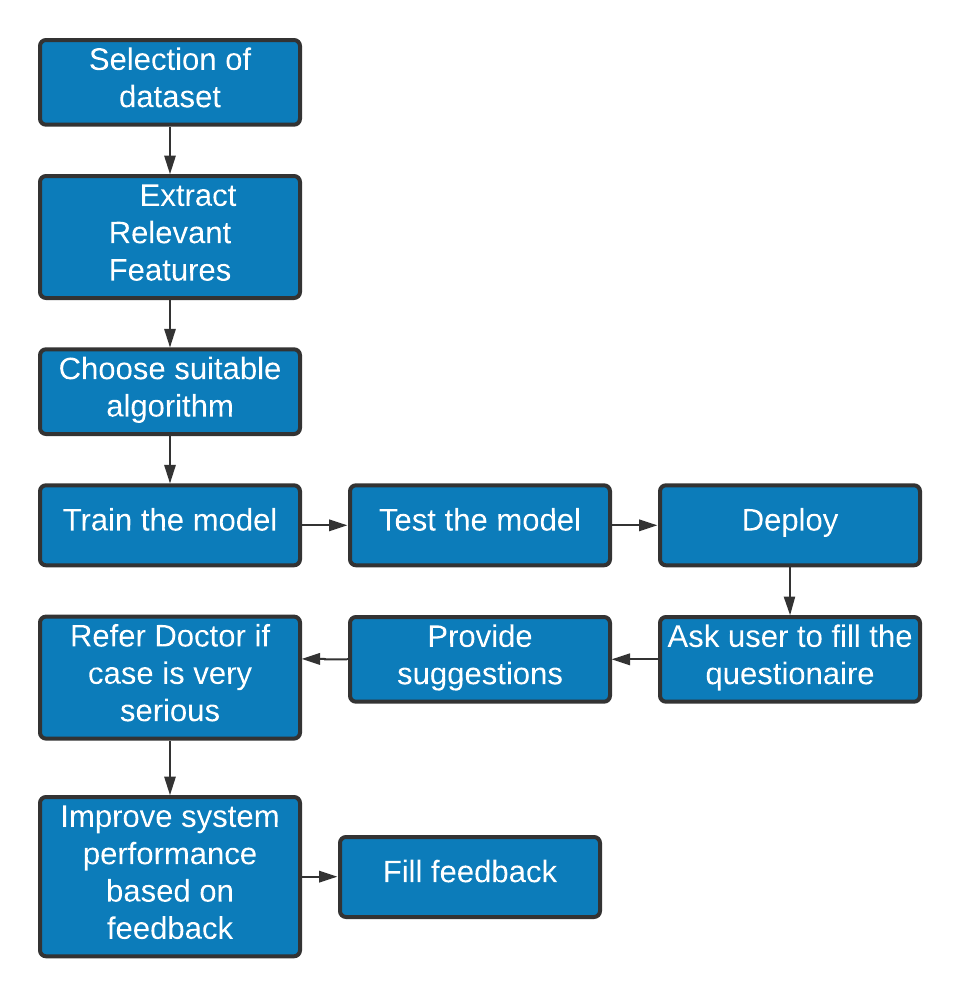
**

Fig 5.1 Architecture diagram

Fig 5.1 depicts the architecture diagram for our system. The first step in this is to select a suitable dataset from the many available options online. Upon selecting a suitable dataset, the suitable features from it will be extracted and an appropriate algorithm for the training of the model will be selected. Once the algorithm is selected the model will be trained, tested and deployed.

The trained model is then deployed in our system where it will take as input from the users via a questionnaire and analyse the mental health of the user on the basis of which suggestions will be provided. In a scenario where the user is found to be having critical mental health the user will be referred to a licensed psychiatrist for proper care. The users will have the capabilities to provide feedback based on the performance of the system which can be utilized later for making the system better and improving its performance.

## B. Proposed Algorithm

**Data Cleaning and Preprocessing:**

* **Null Values were padded to zero**
* **For Feature Selection PCA (Principal Component Analysis) was used:**

The Dataset used for this project has 125306 rows. Large datasets such as this one are often difficult to interpret. Between Principal component analysis (PCA) and Independent Component Analysis(ICA) we chose PCA as ICA gives rise to underfitting in the model. PCA is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance. Finding such new variables, the principal components, reduces to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, not a priori, hence making PCA an adaptive data analysis technique.

* **For Model Training Long Short Term Memory (LSTM)**

LSTM networks are an extension of recurrent neural networks (RNNs) mainly introduced to handle situations where RNNs fail. RNN fails to store information for a longer period of time. RNNs are absolutely incapable of handling such “long-term dependencies”. Another issue with RNN is the vanishing gradient problem.

LSTM has been so designed that the vanishing gradient problem is almost completely removed, while the training model is left unaltered. LSTMs provide us with a large range of parameters such as learning rates, and input and output biases. Hence, no need for fine adjustments. The complexity to update each weight is reduced to O(1). With LSTMs, there is no need to keep a finite number of states from beforehand as required in the hidden Markov model (HMM).

A Long Short Term Memory Network consists of four different gates for different purposes as described below:-

**Forget Gate(f):** It determines to what extent to forget the previous data.

**Input Gate(i):** It determines the extent of information to be written onto the Internal Cell State.

**Input Modulation Gate(g):** It is often considered as a sub-part of the input gate and much literature on LSTM does not even mention it and assume it is inside the Input gate. It is used to modulate the information that the Input gate will write onto the Internal State Cell by adding non-linearity to the information and making the information **Zero-mean**. This is done to reduce the learning time as Zero-mean input has faster convergence. Although this gate’s actions are less important than the others and are often treated as a finesse-providing concept, it is good practice to include this gate in the structure of the LSTM unit.

**Output Gate(o):** It determines what output(next Hidden State) to generate from the current Internal Cell State.

*B. Working of the Project*

* *PCA*

STEP 1: STANDARDIZATION

The aim of this step is to standardize the range of the continuous initial variables so that each one of them contributes equally to the analysis. More specifically, the reason why it is critical to perform standardization prior to PCA, is that the latter is quite sensitive regarding the variances of the initial variables. That is, if there are large differences between the ranges of initial variables, those variables with larger ranges will dominate over those with small ranges (For example, a variable that ranges between 0 and 100 will dominate over a variable that ranges between 0 and 1), which will lead to biased results. So, transforming the data to comparable scales can prevent this problem.

STEP 2: COVARIANCE MATRIX COMPUTATION

The aim of this step is to understand how the variables of the input data set are varying from the mean with respect to each other, or in other words, to see if there is any relationship between them. Because sometimes, variables are highly correlated in such a way that they contain redundant information. So, in order to identify these correlations, we compute the covariance matrix.

### STEP 3: COMPUTE THE EIGENVECTORS AND EIGENVALUES OF THE COVARIANCE MATRIX TO IDENTIFY THE PRINCIPAL COMPONENTS

Eigenvectors and eigenvalues are the linear algebra concepts that we need to compute from the covariance matrix in order to determine the *principal components* of the data.

### STEP 4: FEATURE VECTOR

### Computing the eigenvectors and ordering them by their eigenvalues in descending order, allow us to find the principal components in order of significance. In this step, what we do is, to choose whether to keep all these components or discard those of lesser significance (of low eigenvalues), and form with the remaining ones a matrix of vectors that we call Feature vector.

### 

### STEP 5: RECAST THE DATA ALONG THE PRINCIPAL COMPONENTS AXES

### In the previous steps, apart from standardization, you do not make any changes on the data, you just select the principal components and form the feature vector, but the input data set remains always in terms of the original axes (i.e, in terms of the initial variables).In this step, which is the last one, the aim is to use the feature vector formed using the eigenvectors of the covariance matrix, to reorient the data from the original axes to the ones represented by the principal components (hence the name Principal Components Analysis). This can be done by multiplying the transpose of the original data set by the transpose of the feature vector.

* *LSTM*

Take input the current input, the previous hidden state, and the previous internal cell state.Calculate the values of the four different gates by following the below steps:-

1. For each gate, calculate the parameterized vectors for the current input and the previous hidden state by element-wise multiplication with the concerned vector with the respective weights for each gate.
2. Apply the respective activation function for each gate element-wise on the parameterized vectors. Below given is the list of the gates with the activation function to be applied for the gate.
3. Calculate the current internal cell state by first calculating the element-wise multiplication vector of the input gate and the input modulation gate, then calculate the element-wise multiplication vector of the forget gate and the previous internal cell state and then adding the two vectors.
4. Calculate the current hidden state by first taking the element-wise hyperbolic tangent of the current internal cell state vector and then performing element-wise multiplication with the output gate.

# Performance Evaluation Parameter

For the performance evaluation of the project we aim to use Confusion matrix and classification report.

A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The matrix compares the actual target values with those predicted by the machine learning model. This gives us a holistic view of how well our classification model is performing and what kinds of errors it is making.

The classification report visualizer displays the precision, recall, F1, and support scores for the model.There are four ways to check if the predictions are right or wrong:

1. **TN / True Negative**: the case was negative and predicted negative
2. **TP / True Positive**: the case was positive and predicted positive
3. **FN / False Negative**: the case was positive but predicted negative
4. **FP / False Positive**: the case was negative but predicted positive

**Precision — *What percent of your predictions were correct?*** Precision is the ability of a classifier not to label an instance positive that is actually negative. For each class, it is defined as the ratio of true positives to the sum of a true positive and false positive.Precision:- Accuracy of positive predictions.Precision = TP/(TP + FP)

**Recall — *What percent of the positive cases did you catch?*** Recall is the ability of a classifier to find all positive instances. For each class it is defined as the ratio of true positives to the sum of true positives and false negatives. Recall:- Fraction of positives that were correctly identified. Recall = TP/(TP+FN)

**F1 score — *What percent of positive predictions were correct?*** The F1 score is a weighted harmonic mean of precision and recall such that the best score is 1.0 and the worst is 0.0. F1 scores are lower than accuracy measures as they embed precision and recall into their computation. As a rule of thumb, the weighted average of F1 should be used to compare classifier models, not global accuracy. F1 Score = 2\*(Recall \* Precision) / (Recall + Precision)

**Support**

Support is the number of actual occurrences of the class in the specified dataset. Imbalanced support in the training data may indicate structural weaknesses in the reported scores of the classifier and could indicate the need for stratified sampling or rebalancing. Support doesn’t change between models but instead diagnoses the evaluation process.

# Experimental Setup

*Software Requirements:*

* 1. Operating System: Android
  2. Application : Android Studio, Colab Notebook
  3. Programming Language : Python & Kotlin

*Hardware Requirements:*

A.CPU: Intel i5/Ryzen 5 or above

B. Clock speed: 2.5 GHz or above

C. GPU: NVIDIA Geforce GTX 960 or above

D. RAM size: 8GB or above

E. Hard Disk capacity: 100GB or above

A Database is to be maintained in order to store the login and registration details of the users and also the feedback provided by them.

The Test set and train set will be decided during the training phase of the project according to the performance of the model evaluated by the decided parameters.

##### References

1. Yenan Wang, Yu Di, Junjie Ye & Wenbin Wei, “Study on the public psychological states and its related factors during the outbreak of coronavirus disease 2019 (COVID-19) in some regions of China, Psychology, Health & Medicine”, 2020.
2. Yamada, Y., Ćepulić, DB., Coll-Martín, T. *et al.,* “COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak.”  *Sci Data* 8**,** 2021.
3. Sherman A. Lee, “Coronavirus Anxiety Scale: A brief mental health screener for COVID-19 related anxiety, Death Studies”, 2020.
4. Sherman A. Lee, “Measuring coronaphobia: The psychological basis of the Coronavirus Anxiety Scale”, Dusunen Adam: The Journal of Psychiatry and Neurological Sciences”,2020.
5. Kubb C, Foran HM, “Measuring COVID-19 Related Anxiety in Parents: Psychometric Comparison of Four Different Inventories”, *JMIR Mental Health*, 2020.
6. Eder SJ, Steyrl D, Stefanczyk MM, Pieniak M, Martínez Molina J, Pešout O, et al., “Predicting fear and perceived health during the COVID-19 pandemic using machine learning: A cross-national longitudinal stud”,2021.
7. Ćosić, Krešimir et al. “Artificial intelligence in prediction of mental health disorders induced by the COVID-19 pandemic among health care workers.” *Croatian medical journal vol. 61,3*, 2020.