Approach to Solve the Problem: (Generalize idea of SSC)

1. Prepared problem: First understand the problem and see the data carefully. See all variables and decide independent and dependent variable. After that understand which problem it is classification or predication. Also see what kind of data it is numerical or categorical if numerical or categorical go for ML if data is audio, video then go for DL.
2. Load Libraries: Load important libraries into the python notebook
3. Load Dataset: First read the csv file in the notebook after that decide model data and eliminate

the not important data from the dataset. (see 18 rules of Data Entry is followed or not)

and Rows must be 20 time the column to model any ML Model for prediction or classification.

1. Summarize data: see what data is telling-

Descriptive stat:

Measure of central tendency – Mean, Median, Mode

Frequency

Percentage

Measure of Dispersion- Range, Variance, S.D, Skewness (+/-), kurtosis (Meso, platy, lapto)

Coefficient of Variance.

Inferential Stat:

Correlation

Testing of Hypothesis

Probability

Sampling

Distribution

1. Data Visualization:

Plot scatter plot to see missing data(dropna). Plot histogram for continuous data. Plot correlation to see the correlation between the variable. Box Plot for the outlier detection.

1. Prepare the data:

Handle missing Values and outliers. clean the data. Convert Adherence data yes and no into 1 and 0 for further use. Adherence is the y which is dependent data is the full dataset. Feature selection is the important step in preparing the data. Data transformation is another step to prepared the data generally in the step we standardize or normalize the data according to requirement in this problem I standardize the data for the model development.

1. Evolution of Algorithms:

Spilt out Validation

Spot-checking algorithm(DT and RF, SVM, KNN, Logistic Regression, CNN)

Compare Algorithm

1. Improve Accuracy:

Algorithm tuning and ensemble methods (bagging, boosting)

1. Finalize Model:

Prediction on validation datasets

Create standalone model