|  |
| --- |
| **Project Title: ML model to analyse and predict human illness** |
| **Members:**  **1.Vishwanath K R**  **2.Amrutha A** |
| **Abstract** |
| **Introduction:**  In this project using python modules, dataset the machine learning model can analyse and predict the human disease/illness.  **Process/Approach:**   * **Gathering data**: Source is from Kaggle, dataset comprises of two CSV files one for training and the other for testing the models. * **Cleaning the data**: In our dataset all the columns are numerical and the target column i.e., prognosis is a string type and is encoded to numerical form using a label encoder. * **Model building**: The cleaned data is used to train all the four models and we will be using confusion matrix to determine the quality of the model. * **K-Fold Cross Validation technique** for training of the model on k-1 subsets and the remaining one subset is used to evaluate the model performance. * **Support Vector Classifier** this algorithm tries to find an optimal hyperplane that accurately separates the samples into different categories in hyperspace. * **Gaussian Naive Bayes Classifier** is a probabilistic machine learning algorithm that internally uses Bayes Theorem to classify the data points.   Equation: P(A|B) = P(B|A) \* P(A) / P(B)   * **Random Forest Classifier i**n a random forest classifier, all the internal decision trees are weak learners, the outputs of these weak decision trees are combined i.e., mode of all the predictions is as the final prediction. * **Bernoulli Naïve Bayes Classifier** is another useful naïve Bayes model. The assumption in this model is that the features are binary (0s and 1s) in nature. * **Inference**: After training the four models we will be predicting the disease for the input symptoms by combining the predictions of all the models (by taking mode). This makes our end prediction much more accurate.   **Performance Measures:**  The quality of all 4 models is same as to the one observed on the confusion matrix below:    **Final Result:** |