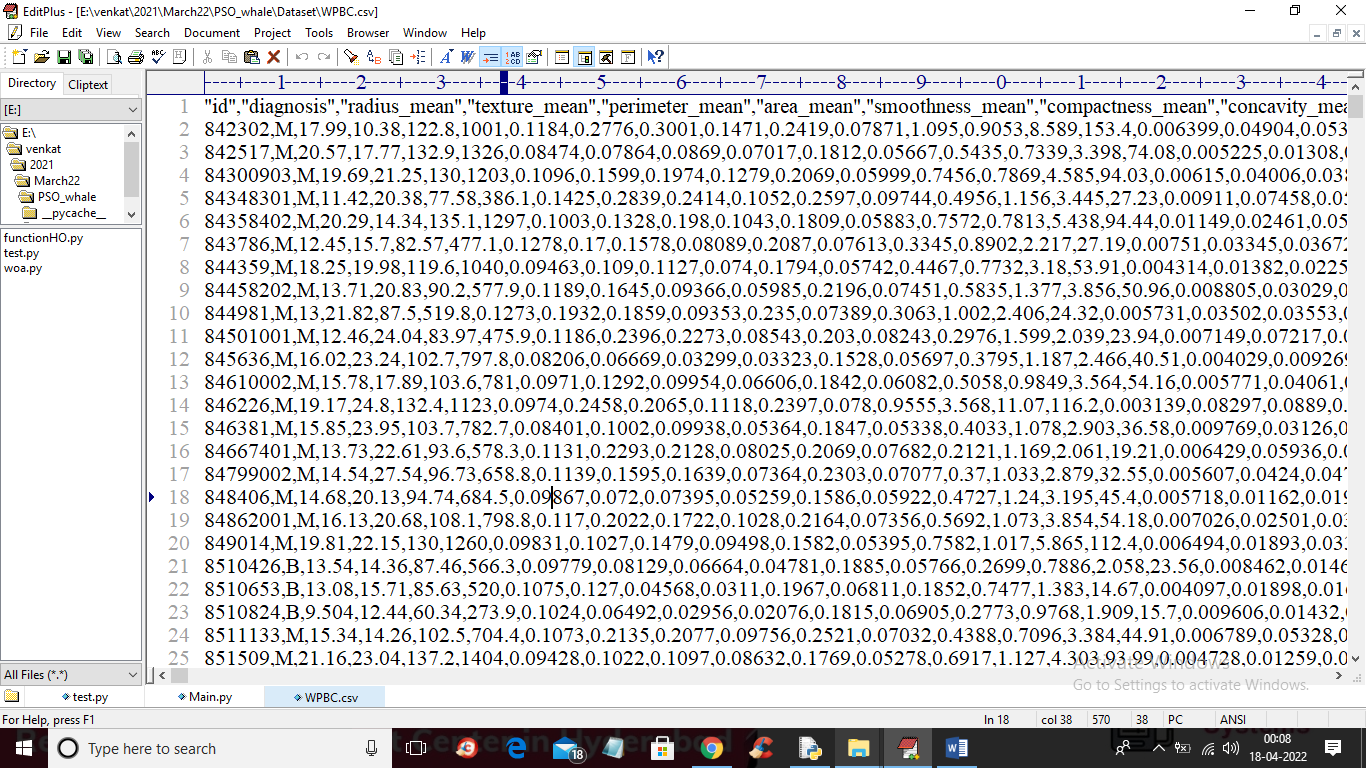
META-HEURISTIC OPTIMIZATION ALGORITHMS BASED FEATURE SELECTION FOR CLINICAL BREAST CANCER DIAGNOSIS

In this project to diagnose breast cancer author is using several machine learning algorithms such as SVM, KNN and decision tree and to optimize features we are using Whale Optimization and PSO optimization algorithms and then compare their performance in terms of execution time, accuracy, precision and recall.

To implement this project we have used WPBC (breast cancer Wisconsin) dataset and it contains 30 features or attributes and below screen showing dataset details



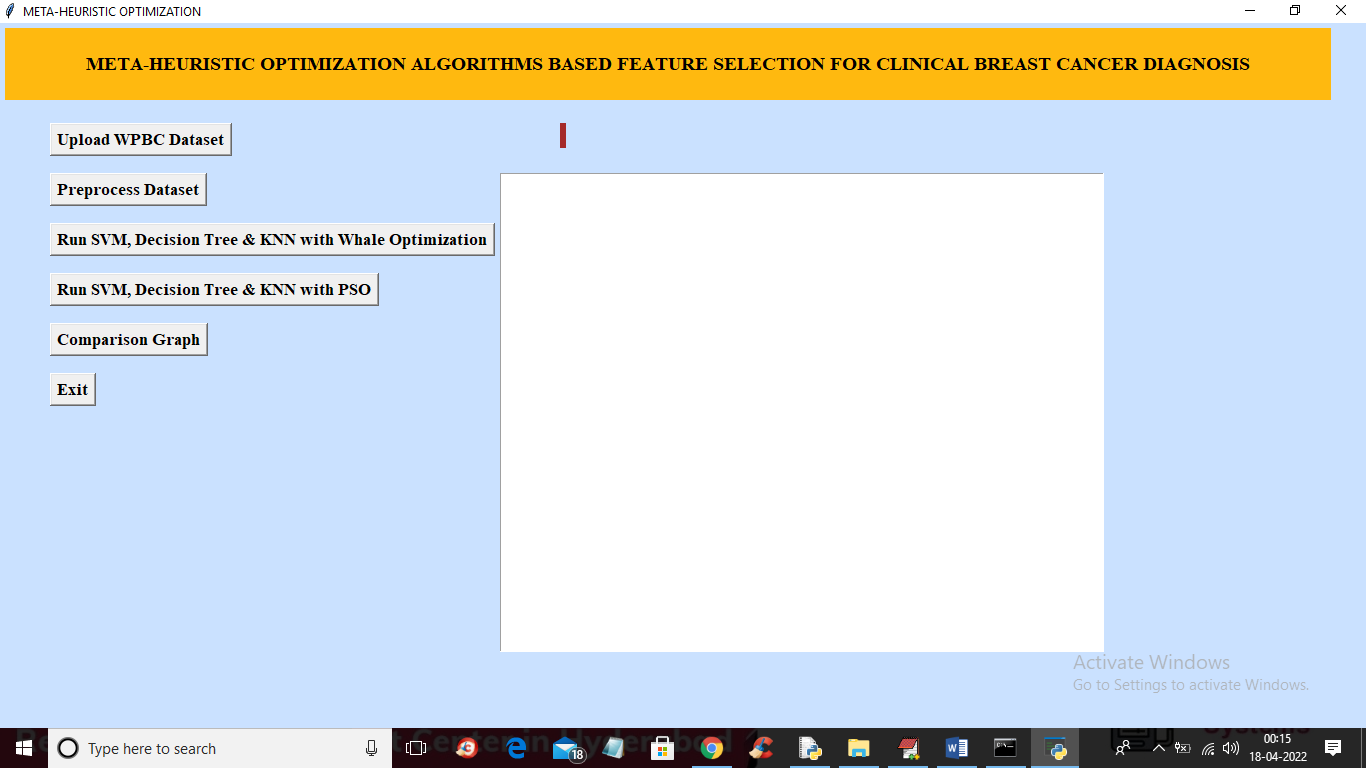
In above screen first row represent dataset column names and remaining rows contains dataset values and in above dataset we have class labels as ‘diagnosis’ which contains values like “B” (benign) and “M” (Malignant). We will use above dataset and then apply whale and PSO algorithm to reduced or select important features and then train with SVM, KNN and decision tree.

To implement this project we have designed following modules

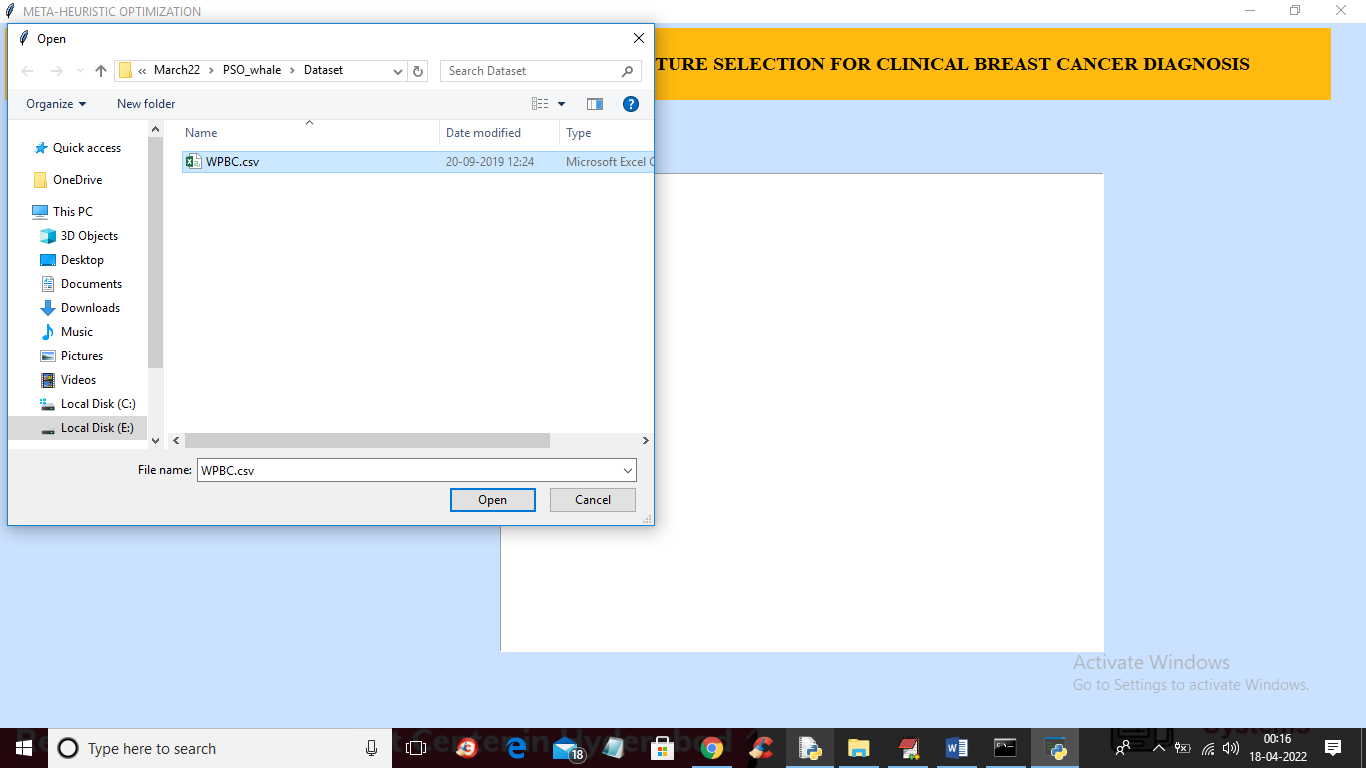
1. Upload WPBC Dataset: using this module we will upload dataset to application
2. Preprocess Dataset: dataset contains missing and non-numeric values and machine learning algorithms will not accept such values so we will apply Preprocessing technique to replace missing values with 0 and then apply label encoding technique to convert non-numeric data to numeric ID’s
3. Run SVM, Decision Tree & KNN with Whale Optimization: above processed features will be input to whale optimization algorithm to select important features and then train with SVM and other algorithms
4. Run SVM, Decision Tree & KNN with PSO: above processed features will be input to PSO optimization algorithm to select important features and then train with SVM and other algorithms
5. Comparison Graph: using this module we will plot accuracy graph between 2 algorithms

SCREEN SHOTS

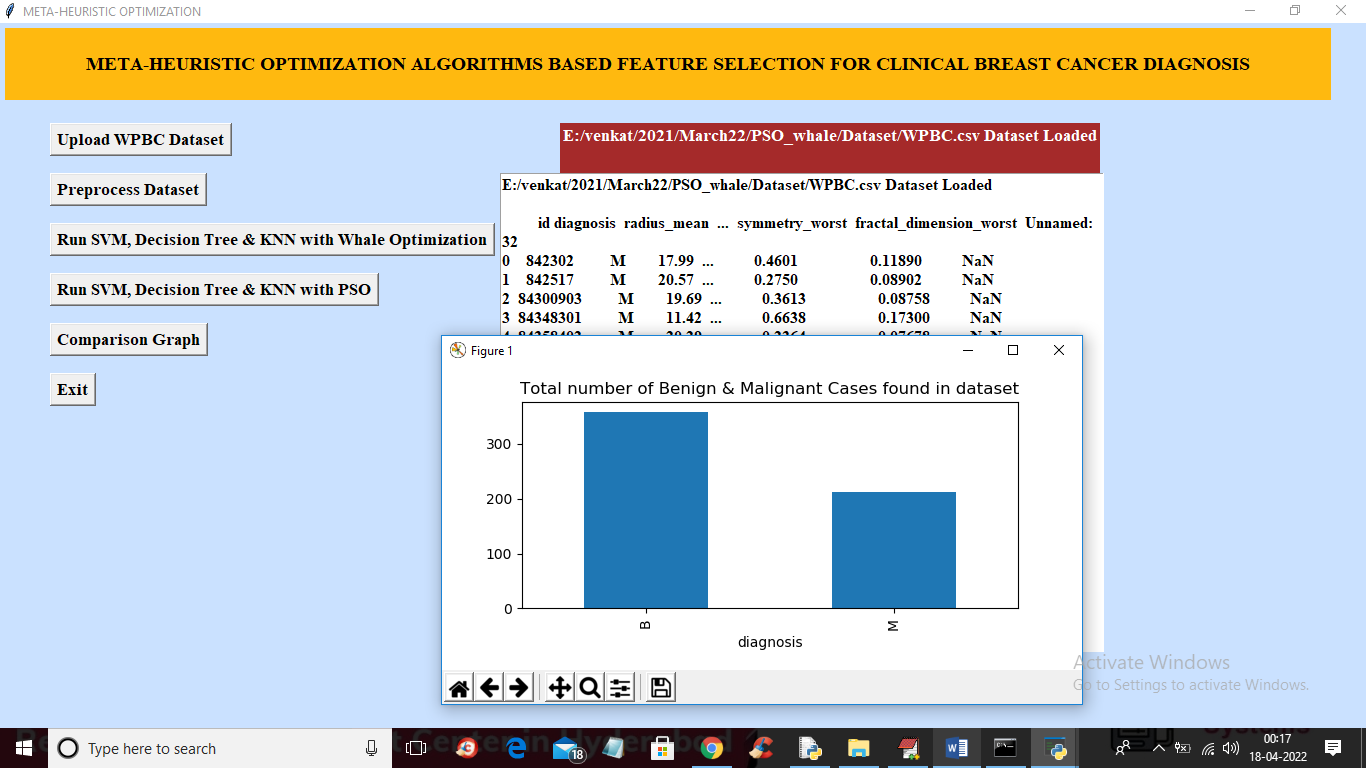
To run project double click on ‘run.bat’ file to get below screen



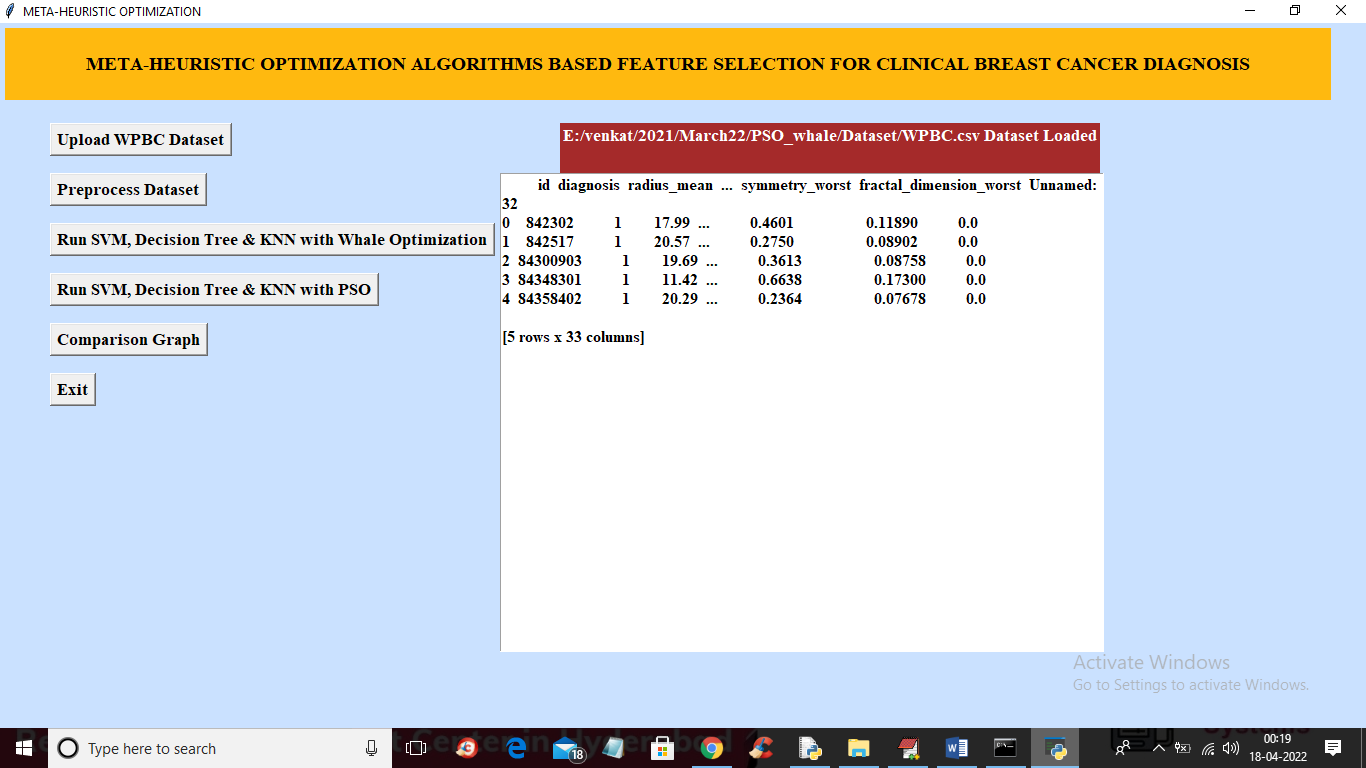
In above screen click on ‘Upload WPBC Dataset’ button to upload dataset and to get below screen



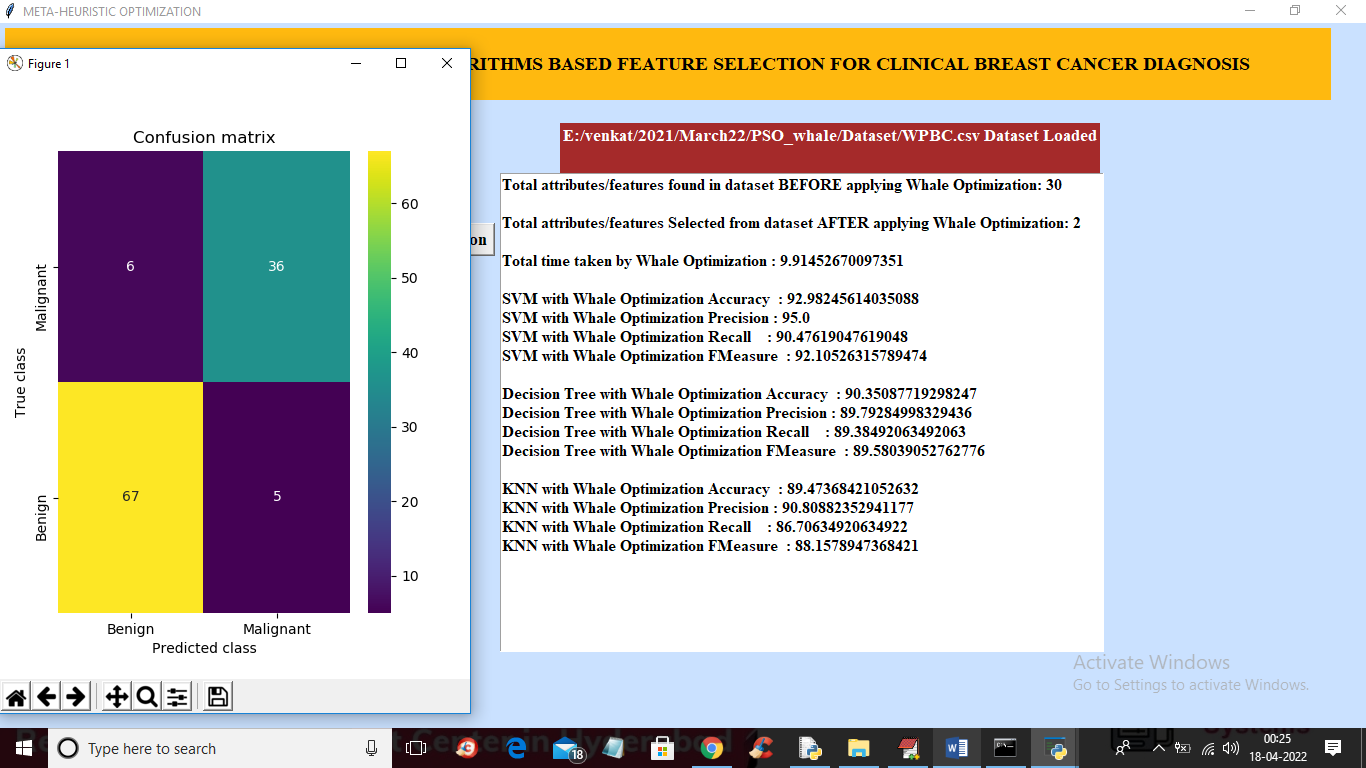
In above screen selecting and uploading dataset folder and then click on ‘Open’ button to load dataset and to get below screen



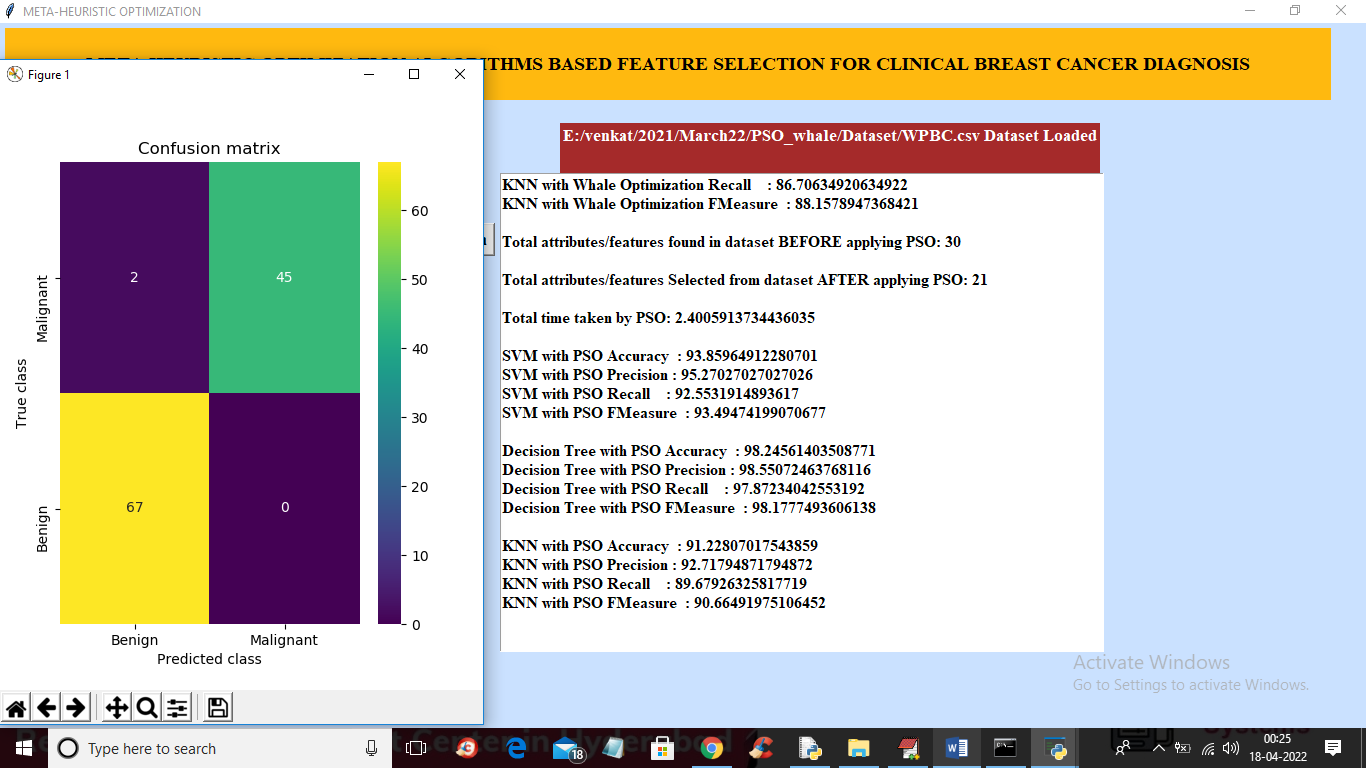
In above screen we can see dataset loaded and in dataset contains non-numeric values and missing NAN values and in above screen we can see graph showing number of benign and malignant cases found in dataset and now close above graph and then click on ‘Preprocess Dataset’ button to replace missing values and convert non-numeric data to numeric data and get below output



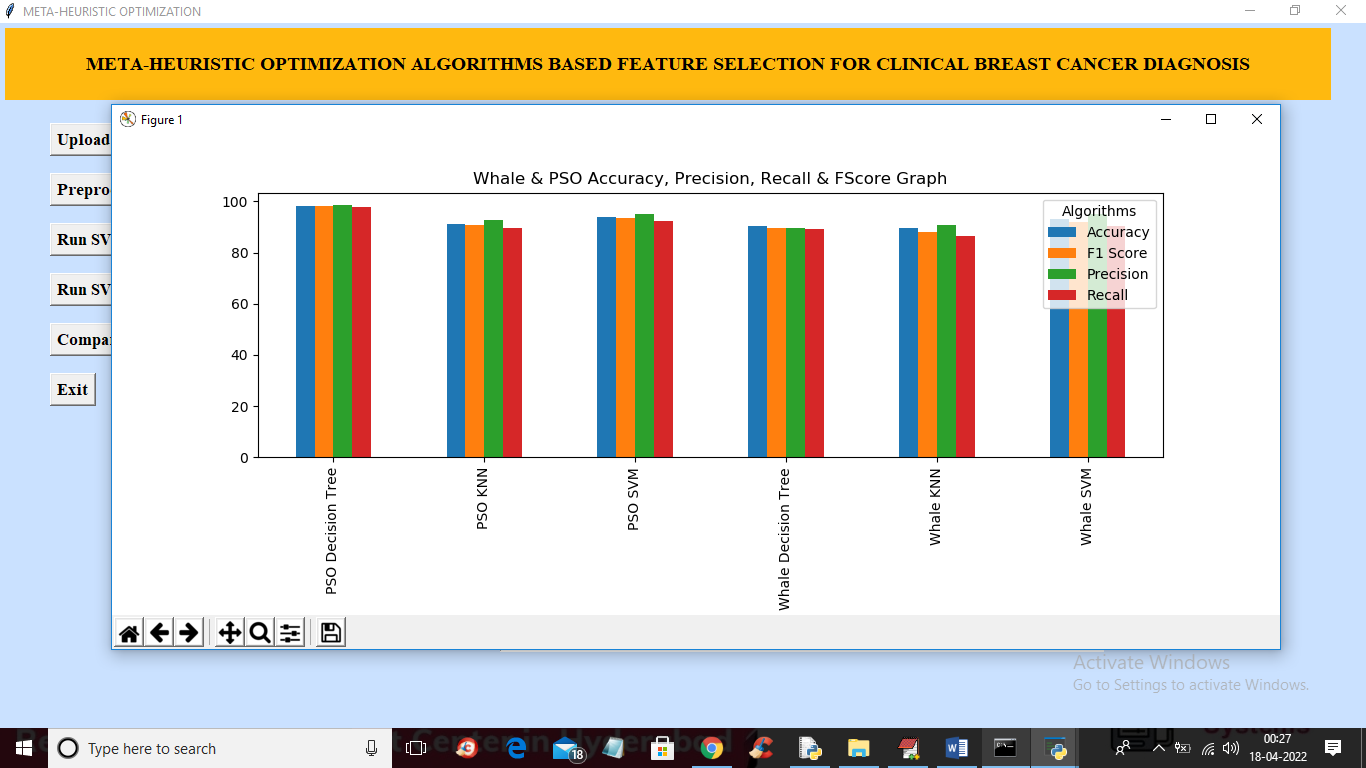
In above screen we can see all values are converted to numeric and now click on ‘Run SVM, Decision Tree & KNN with Whale Optimization’ button to apply whale optimization and train all ML algorithms to get below output



In above screen in first two lines we can see dataset contains 30 attributes and after applying whale we got 2 important attributes and then we can see accuracy of each algorithm on selected features and we can see execution time also and in above screen we can see prediction confusion matrix graph where application predict 67 records as benign correctly and only 6 records are incorrectly predicted and now close above graph and then click on ‘Run SVM, Decision Tree & KNN with PSO’ button to select features with PSO and train all algorithms to get below output

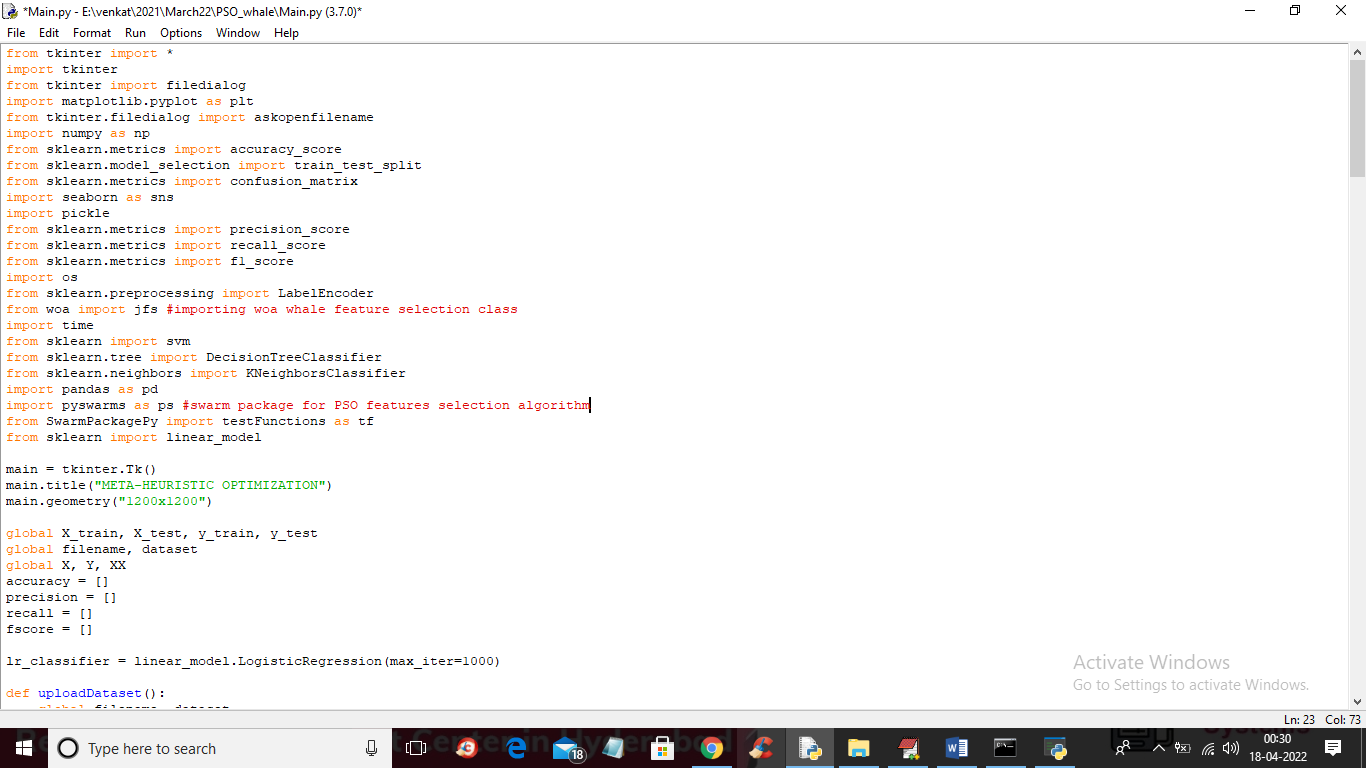


In above screen we can see PSO selected 21 attributes out of 30 and we can see accuracy of each algorithm on selected features and in confusion matrix we can see with PSO 67 records are correctly predicted as Benign and only 2 records are incorrectly predicted and now close above graph and then click on ‘Comparison Graph’ button to get below graph



In above graph x-axis represents algorithm names and y-axis represents accuracy and other values and in above graph each different colour bar represents different metrics like accuracy, precision and etc. In above graph we can see PSO with decision tree give better performance.

In below screen you can see we are using WOA package for whale optimization and PYSWARM package for PSO



In above screen read red colour comments to know about features selection classes