**A combination of urinary biomarker panel and PancRISK score for earlier detection of pancreatic cancer: A case–control study**

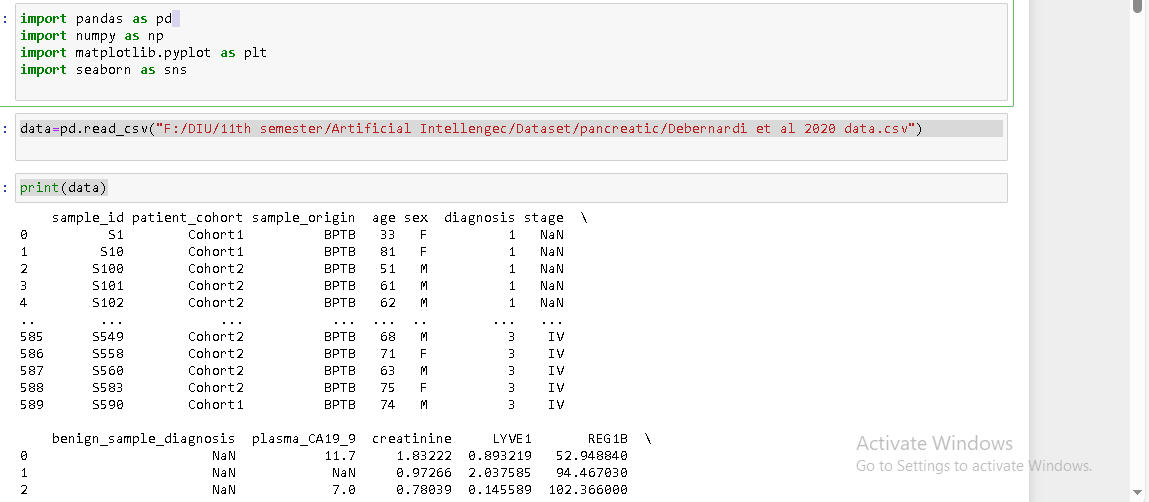
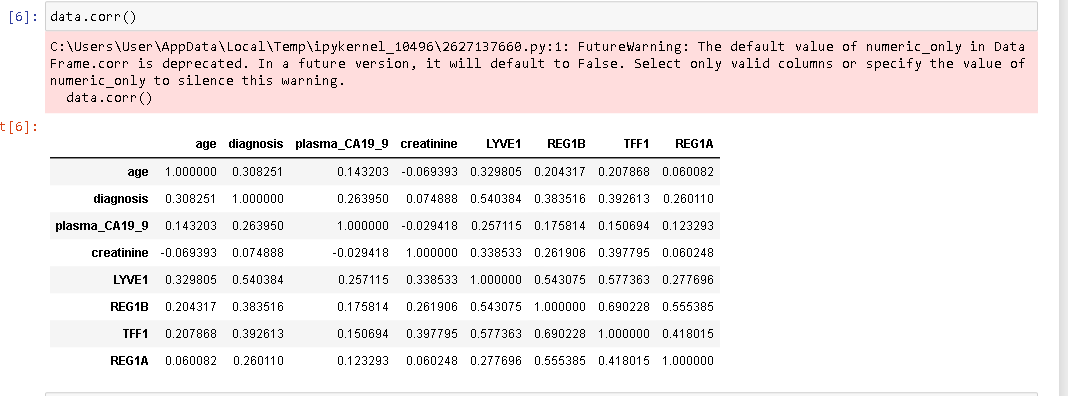
Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal diseases, with an average 5-year survival rate of less than 10%. Unfortunately, the majority of patients have unresectable, locally advanced, or metastatic disease at the time of diagnosis. Moreover, traditional treatments such as chemotherapy, surgery, and radiation have not been shown to significantly improve survival. 9% of patients surviving >5 years. Asymptomatic in its initial stages, PDAC is mostly diagnosed late, when already a locally advanced or metastatic disease, as there are no useful biomarkers for detection in its early stages. The poor outcome is largely due to late diagnosis; however, if the disease is detected at an early stage, when tumours are still small and resectable, 5-year survival can increase to up to 70%.While blood has traditionally been the main source of biomarkers, urine represents a promising alternative biological fluid .It allows a completely non- invasive sampling, high volume collection, and ease of repeated measurements; it has a lesser dynamic range, with a less complex proteome than blood . In this case–control study [32], we utilised retrospectively collected urine and plasma samples for the validation of a urinary biomarker panel previously discovered using mass spectrometry [21]. Of the 590 urine specimens analysed, 183 were from control individuals (control group) who had no known pancreatic conditions or malignancies or history of renal diseases at the time of collection, 208 were from patients with benign hepatobiliary diseases (benign group), and 199 were from PDAC patient

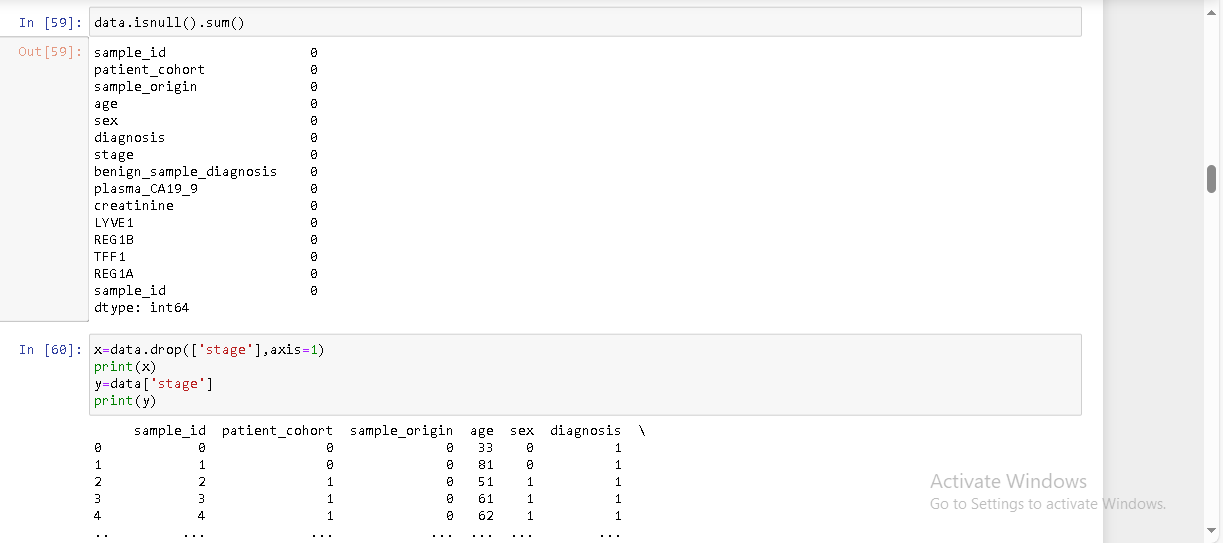
**Methodology:**

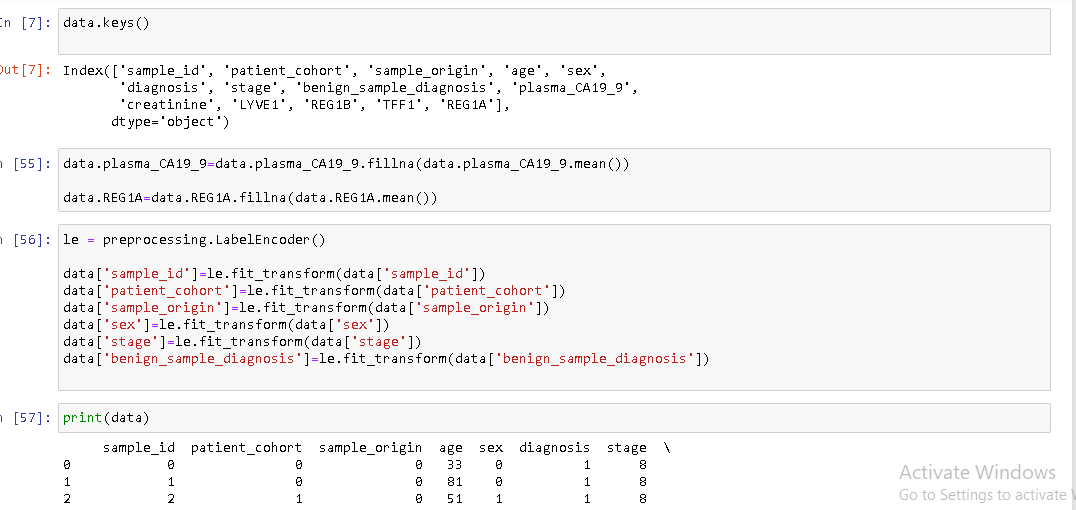
An accurate prediction of pancreatic cancer is necessary for the early stage of treatment and overcoming the mortality rate. This study proposes a machine learning approach to diagnose stroke with imbalanced data more accurately. Five classifiers, including Support Vector Machine, Random Forest, K-nearest Neighbor, Decision Tree, Naïve Bayes, are analyzed in this study. The results show the Random Forest has the highest accuracy of 85%. Decision Tree achieves the second-highest accuracy of 83%, In addition, a user-friendly web app and a user-friendly mobile app are built based on the most accurate model.

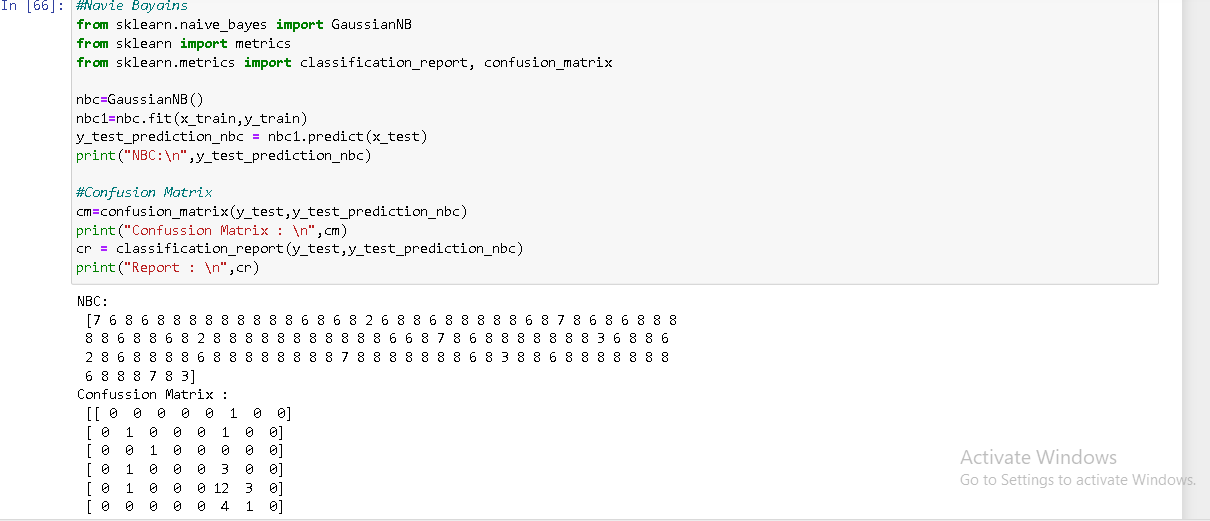
Firstly, we preprocess all the data, we have found some null values that’s why we fill them by using their mean value. Then we convert all the string values into integer values so that we can easily find the accuracy. In the dataset we have 590 rows and 14colums.we divided into two parts. The y variable contains the result so we denoted it as class and the other columns contain features which is stored in variable y. the pre-processed data is divided into 2 parts ,train data and test data. After that, the training data is fed into the different machine learning algorithms to predict the outcome of pancreatic cancer.

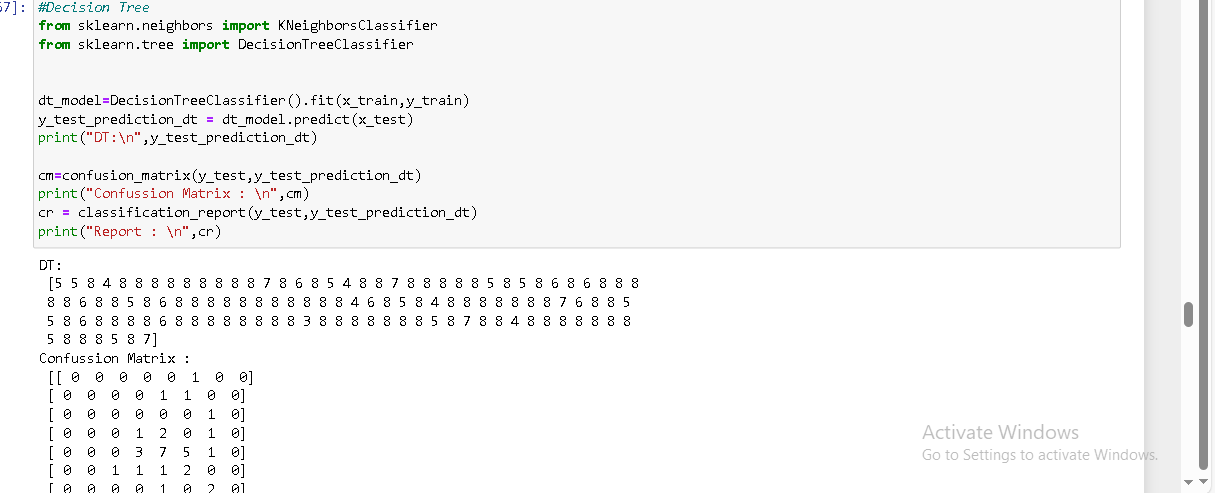
**code :**

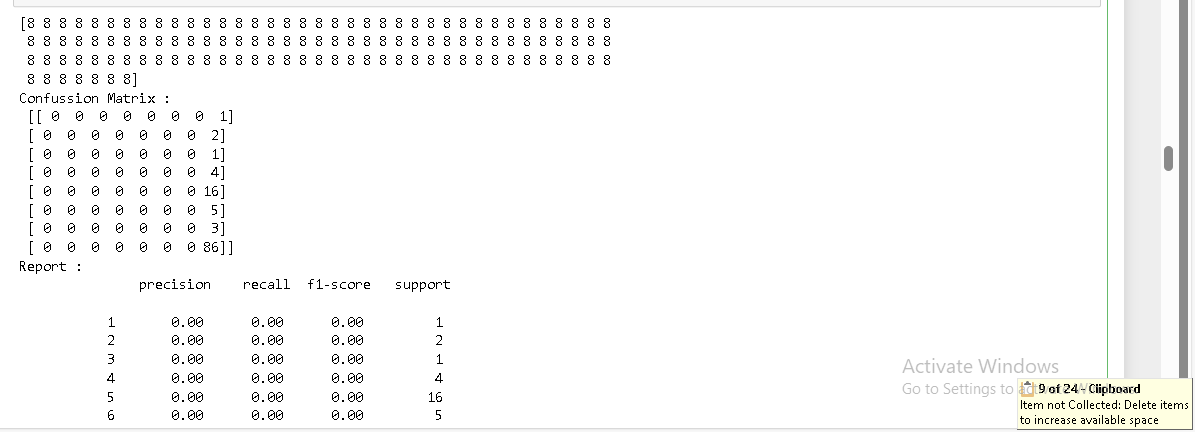
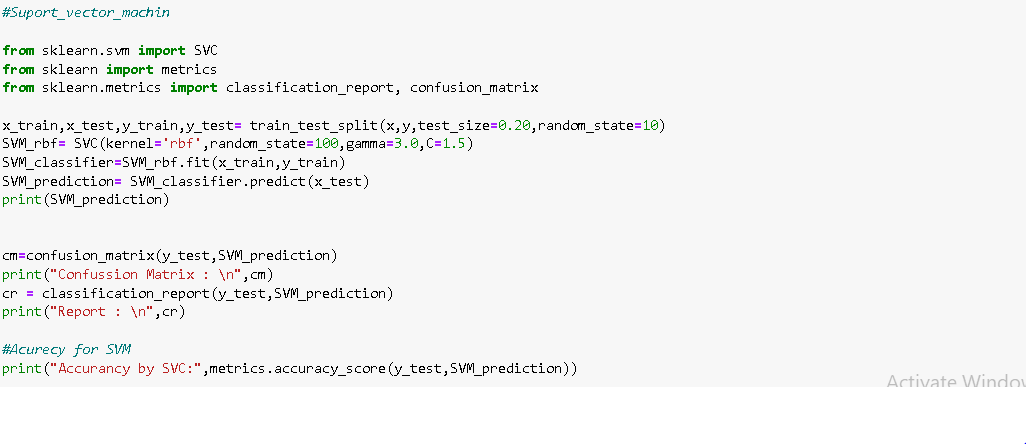
 

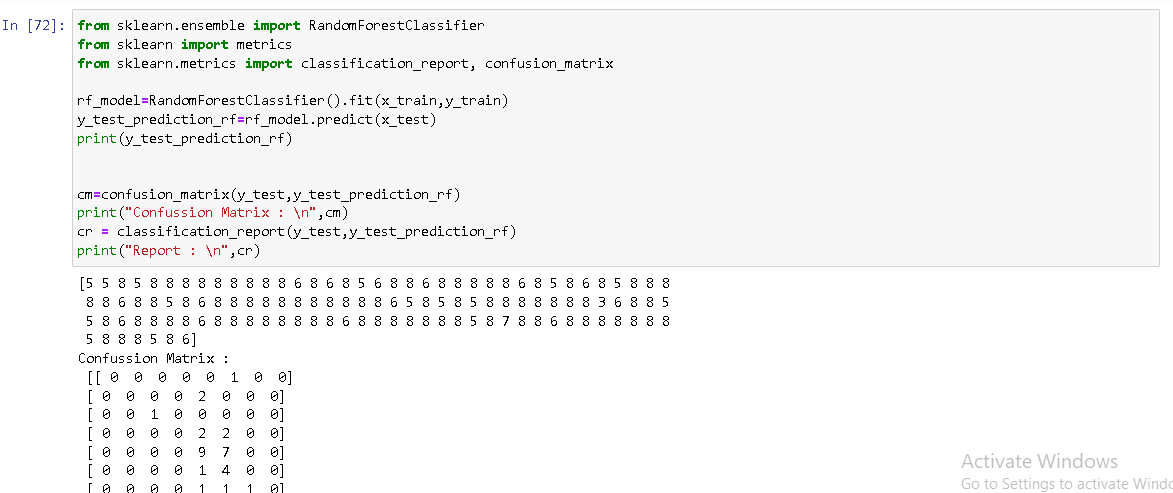


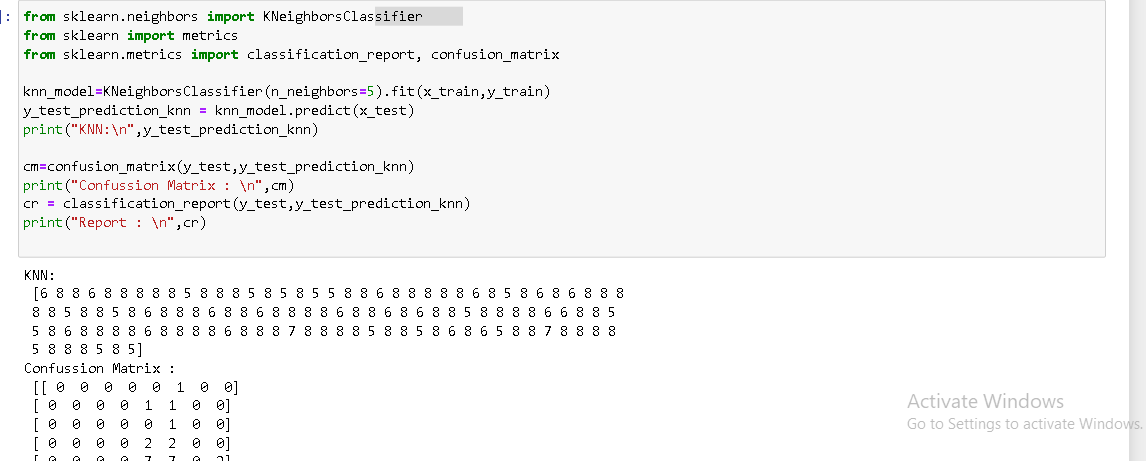


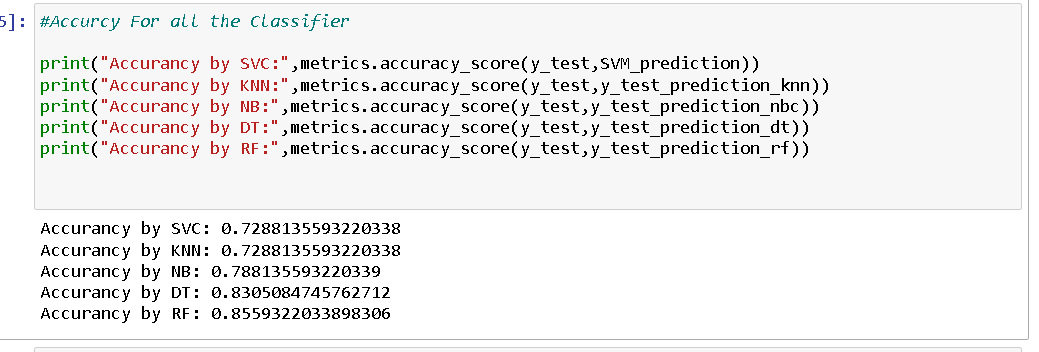




**SVM**







**Result Discussion:**

**We can see the highest accuracy is 85% we get it from Random Forest Classifier. Then Decision Tree give us the 2nd highest accuracy. If we get more data, we could find more accuracy.**

**Conclusions:**

In summary, we successfully validated the performance of our urinary biomarker panel in detecting PDAC earlier. Our PancRISK score is designed to enable further stratification of patients already predisposed to develop PDAC.