**Project Title:** Chronic Kidney Disease: An Analytical Prediction Using Machine Learning

**Objective:**

* To predict Chronic Kidney Disease using each element report of any patient.
* To deal with various null values.
* To complete Data Normalization and segmented for Test and Training cases.
* To find best algorithm application to predict Chronic Kidney Disease.
* To calculate best Accurate among all algorithms using Confusion Matrix.

**User Benefits:**

* The medical sector will be benefited the most from this work.
* On that perspective, maximum users will be the Doctors or Pathologists.
* They can predict patient’s Kidney Disease by using their regular laboratory reports analysis.
* This will help to reduce costs because Kidney Disease report analysis is usually cost higher.
* Can be applicable through IoT in upgraded version of this project.

**Feature in Details:**

We have various features in our project. These are described with details below –

* **Algorithms:** Five algorithms have been used to implement this project. They are:LogisticRegression, K-Nearest Neighbour, Random Forest, Gaussian Naive Bayes, Decision Tree.
* **The versions of different packages:**

numpy: 1.19.2

matplotlib: 3.3.2

pandas: 1.1.3

sklearn: 0.23.2

* **The attenuates of the Data set:** Age, Bp, Sg, Al, Su, Rbc, Pc, Pcc. Ba, Bgr, Bu, Sc, Sod, Pot, Hemo, Pcv, Wc, Rc, Htn, Dm, Cad, Appet, Pe, Ane, Class. The data set has about 25 columns and 400 rows.
* **Train and Test:** 75% of the data set were used to train the model and 25% of the whole dataset was used to test the model. In total 300 rows for training and 100 data set for testing the model.

**Development Tools and Technology:**

The whole model was made using python 3.9.4. The platform that was used for this mode building is Anaconda Jupyter notebook and lab.

**Team Members Contribution Chart:**

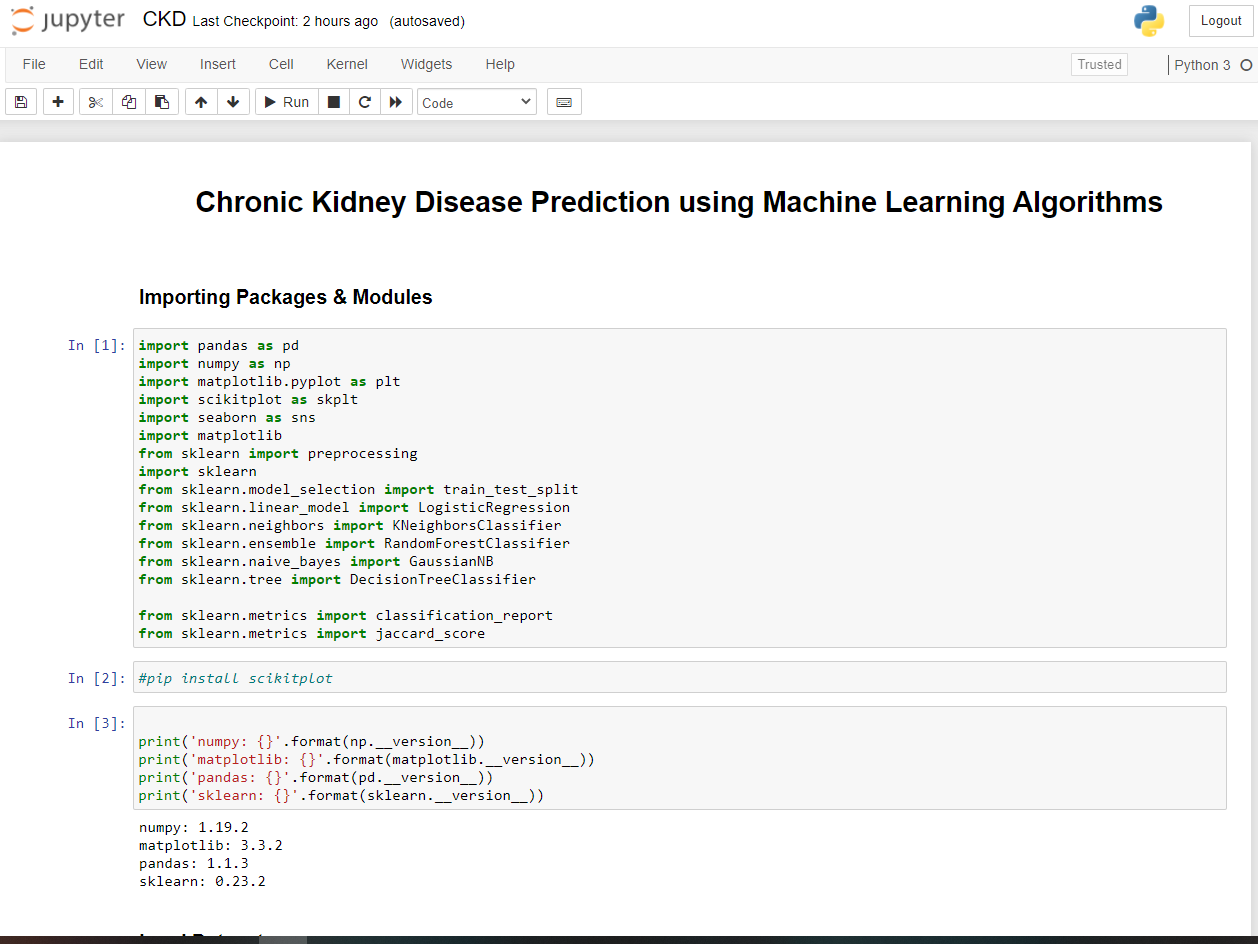
The contribution of the member of this project was same for everyone. This project was built with proper teamwork. Mosfikur Rahaman Leaded the team. Akash Ahmed and Mahafuja Mahin was the co-member of this team. They all had the same contribution for this project.

**Project Timeline:**

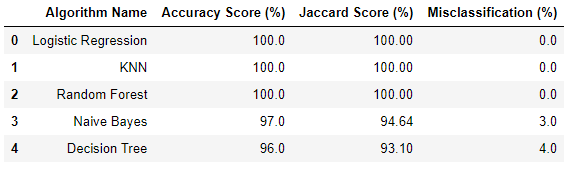
Since we have been working with Machine Learning for quite a while. This project took us only a day to process. The dataset was taken from a valid source and the data was processed using panda’s data frame. The process We have covered general population, high risk persons and all stages of CKD patients with expansive strategies such as early screening, effective CKD registry, prevention and CKD comprehensive care models including cost effectiveness analysis. Five best suited algorithms have been used to train the dataset and predict an accurate value for creating a best model.

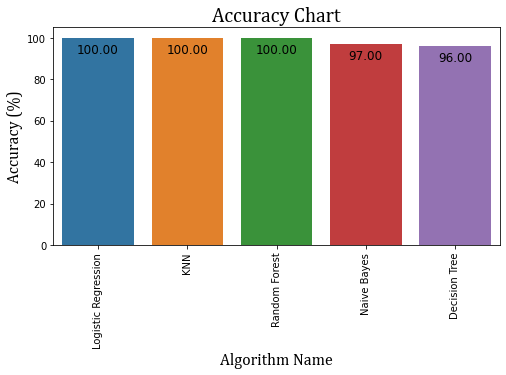
**Screenshot of all Interface (User Interface/Console):**

We worked in Jupyter notebook and jupyter lab. In this environment the whole work can be done easily and short note can be processed very easily. The whole notebook can be converted into a pdf File or any any html file too. This environment is very much useful for developing a Machine Learning Project.

****

**Test Result on Output Screens:**

****



**Source Code (place main/important functions and method first)**

GitHub Link:

**References:**

1. Ahmed, I., Khan, M. S., Paidi, S., Liu, Z., Zhang, C., Liu, Y., ... & Lau, C. (2021). Laser induced breakdown spectroscopy with machine learning reveals lithium-induced electrolyte imbalance in the kidneys. Journal of Pharmaceutical and Biomedical Analysis, 194, 113805.
2. Chaudhuri, A. K., Sinha, D., Banerjee, D. K., & Das, A. (2021). A novel enhanced decision tree model for detecting chronic kidney disease. Network Modeling Analysis in Health Informatics and Bioinformatics, 10(1), 1-22.
3. Balusamy, B. (2021). Utilizing schemes for detecting the threat levels for chronic kidney disease. Materials Today: Proceedings.
4. Dubey, G., Srivastava, Y., Verma, A., & Rai, S. (2021). Chronic Kidney Disease Prediction Using Artificial Neural Network. In Proceedings of International Conference on Big Data, Machine Learning and their Applications (pp. 395-401). Springer, Singapore.
5. Pourhomayoun, M., & Shakibi, M. (2021). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. Smart Health, 20, 100178.

**Conclusion:**

The proposed CKD diagnostic methodology is feasible in terms of data imputation and samples diagnosis. After unsupervised imputation of missing values in the data set by using KNN imputation, the integrated model could achieve a satisfactory accuracy. Hence, we speculate that applying this methodology to the practical diagnosis of CKD would achieve a desirable effect. In addition, this methodology might be applicable to the clinical data of the other diseases in actual medical diagnosis. However, in the process of establishing the model, due to the limitations of the conditions, the available data samples are relatively small, including only 400 samples. Therefore, the generalization performance of the model might be limited. In addition, due to there are only two categories (ckd and notckd) of data samples in the data set, the model cannot diagnose the severity of CKD. In the future, a large number of more complex and representative data will be collected to train the model to improve the generalization performance while enabling it to detect the severity of the disease. We believe that this model will be more and more perfect by the increase of size and quality of the data.