

Logistic Regression and KNN model comparison exercise:

For the Pima Indian Diabetes dataset discussed in the class:

- Perform appropriate feature scaling by checking the outliers.
- Don't drop or replace the outliers since it is a healthcare application and we can't lose any patient's information.
- After performing feature scaling, Try a different split ratio (70/30 and 60/40) and check the model performance for both logistic regression and KNN.
- **Note: Please ensure to check the sequence for scaling and splitting. Which should be performed first and then do the process. Please refer to lectures for clarity.**
- Did you notice any chance for model overfitting? (Training accuracy is better but test accuracy is not). Did you notice any significant difference?

Submission:

Please provide neat observations and comments for your code. Submit a .ipynb notebook with your name.

(For example, Ramya_Logistic_KNN_Assignment.ipynb)

Please follow the naming convention.