**Scikit-learn workbook**

### References:

1. Scikit-learn.org <https://scikit-learn.org/stable/tutorial/index.html>

2. Wikipedia.com / scikit-learn - <https://en.wikipedia.org/wiki/Scikit-learn>

3. MachineLearningMastery.com / Introduction to scikit-learn - <https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/>

4. UCI Machine Learning Repository - https://archive.ics.uci.edu/ml/datasets/Heart+Disease, accessed 05 November 2021

5. Toward Data Science - https://towardsdatascience.com/heart-disease-prediction-73468d630cfc, accessed 05 November 2021

6. Scikit-learn.org / An Introduction to Machine Learning - <https://scikit-learn.org/stable/tutorial/basic/tutorial.html> , accessed 09 November 2021

7. Scikit-learn.org / Classification - https://scikit-learn.org/stable/modules/sgd.html#classification, accessed 09 November 2021

8. IBM /Supervised vs. Unsupervised Learning: What’s the Difference? <https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning> , access 09 November 2021

9. Pandas.pydata.org, Pandas.dataframe.corr, https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.corr.html, accessed 20 December 2020

10. RealPython.com, NumPy, SciPy, and Pandas: Correlation With Python, https://realpython.com/numpy-scipy-pandas-correlation-python/, accessed 20 December 2020

11. Wikipedia.com, Pearson correlation coefficient , https://en.wikipedia.org/wiki/Pearson\_correlation\_coefficient, accessed 20 December 2020

12. Wikipedia.com, Spearman's rank correlation coefficient, https://en.wikipedia.org/wiki/Spearman%27s\_rank\_correlation\_coefficient,accessed 20 December 2020

13. Wikipedia.com, Kendall rank correlation coefficient, ,accessed 20 December 2020

14. TowardsDataScience.com, Understanding the Confusion Matrix from Scikit-learn, https://towardsdatascience.com/understanding-the-confusion-matrix-from-scikit-learn-c51d88929c79, accessed on 20 November 2021

15. Scikit-learn.org, sklearn.metrics.confusion\_matrix, https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion\_matrix.html, accessed 20 November 2021

16. Scikit-learn.org, Naive Bayes, https://scikit-learn.org/stable/modules/naive\_bayes.html#naive-bayes, accessed 21 November 2021

17. Scikit-learn.org, Gaussian Naive Bayes,https://scikit-learn.org/stable/modules/generated/sklearn.naive\_bayes.GaussianNB.html#sklearn.naive\_bayes.GaussianNB, accessed 21 November 2021

18. MachineLearningMastery.com, Naive Bayes, https://machinelearningmastery.com/naive-bayes-classifier-scratch-python/, accessed 27 November 2021

19. Scikit-learn.org, Decision Trees, https://scikit-learn.org/stable/modules/tree.html, accessed 27 November 2021

20. Kdnuggets.com, Decision Trees Explained, https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html, accessed 27 November 2021

21. Stackabuse.com; Overview of Classification Methods in Python with Scikit-Learn, <https://stackabuse.com/overview-of-classification-methods-in-python-with-scikit-learn/> , accessed 28 December 2021

22. Scikit-learn.org, Logistic regression, <https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression>, accessed 28 December 2021

23. Scikit-learn.org, Confusion Matrix Display, <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.ConfusionMatrixDisplay.html#sklearn.metrics.ConfusionMatrixDisplay.from_predictions>, accessed 20 November

24.

25.

(<i> </i>)

1. creating a new classifier (<i>logreg\_clf</i>) using the LogisticRegression() function

2. fitting (<i>fit()</i>) the logreg\_clf classifier to the test and train data created earlier (known as the estimator for the classification that will ‘fit’ the data to the test dataframe

3. a new variable (<i>y\_pred </i>) is created and applied to the X\_Test variable. Similar to the point above, the estimator <i> ‘predict()</i>’ is used to predict the values of the X-Test data frame.

4. test the accuracy of the Logistic Regression model by using a Confusion Matrix (see below).

5. print statements comparing the predicted output against that of the test data

6. plot the output of the confusion matrix using the ration of predicted / true

**Scipy-Stats ANOVA workbook**

### References:

1. University of Sheffield.ac.uk, Datasets for Teaching, <https://www.sheffield.ac.uk/mash/statistics/datasets>, accessed 01 December 2021

2. Laerd Statistics, One-way ANOVA in SPSS Statistics, <https://statistics.laerd.com/spss-tutorials/one-way-anova-using-spss-statistics.php> , accessed 01 December 2021

3. Scipy.org, Statistical functions (scipy.stats), <https://docs.scipy.org/doc/scipy/reference/stats.html>, accessed 01 December 2021

4. , LaerdStatistics.com, Testing for Normality using SPSS Statistics, <https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php>, accessed 29 December 2021

5. Statistic Solutions.com, The Assumption of Homogeneity of Variance, <https://www.statisticssolutions.com/the-assumption-of-homogeneity-of-variance/>, accessed 29 December 2021

6. TechnologyNetworks.com, One-Way vs Two-Way ANOVA: Differences, Assumptions and Hypotheses, <https://www.technologynetworks.com/informatics/articles/one-way-vs-two-way-anova-definition-differences-assumptions-and-hypotheses-306553>, accessed 29 December 2021

7. LaerdStatistics.com, Independent t-test for two samples, <https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php>, accessed 29 December 2021

8. LaerdStatistics.com, One-way ANOVA (cont...), <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide-4.php>, accessed 29 December 2021

9. LaerdStatistic.com, One-way ANOVA, <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php>, accessed 29 December 2021