



Report on:

Classification Task

Name: Krishna Kumar Mandal

University ID: 2508742

College ID: np02cs4a240074

Group: L5CG2

Table of Contents

1. Abstract 3

2. Introduction 3

3. Dataset Description 3

4. Methodology 3

5. Results and Discussion..... 3

6. Ethical, Social and Sustainability Impact..... 4

7. Conclusion..... 4

8. References 4

1. Abstract

The report is a supervised machine learning classification example that devoted to determine benign and malignant cases of breast cancer. The example reflects the Sustainable Development Goal 3 (Good Health and Well-being) in that it shows that data-driven models may aid in making early medical diagnoses.

2. Introduction

The report is a supervised machine learning classification example that devoted to determine benign and malignant cases of breast cancer. The example reflects the Sustainable Development Goal 3 (Good Health and Well-being) in that it shows that data-driven models may aid in making early medical diagnoses.

3. Dataset Description

Breast Cancer Wisconsin is a set of 569 cases having 30 quantitative features based on binary images of the cell nucleus of the breast. The target variable will be tumor diagnosis, which will be benign or malignant.

4. Methodology

The dataset was processed prior to the coding level of encoding the class labels, multiples, and standardization of the features. The information was divided into training and testing samples. Multi-layer Perceptron (MLP) classifiers, Logistic Regression and random forests were applied in order to compare the approaches based on linear methods, ensemble methods, and neural network.

5. Results and Discussion

The accuracy, precision, recall, and F1-score were used as the models to assess the performance. The MLP classifier recorded the best performance in the total performance hence showing that it is effective in learning non-linear relationships in the medical data.

6. Ethical, Social and Sustainability Impact

Diagnostic efficiency can be greatly enhanced with the help of machine learning in healthcare, and SDG 3 can be supported by it. Nevertheless, this type of systems has to be utilized ethically, with fairness, transparency, and responsible use in the company of medical workers.

7. Conclusion

The given classification task illustrates that machine learning, especially neural networks, may be successfully used in the medical diagnosis, given the adequate data preprocessing and evaluation.

8. References

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

UCI Machine Learning Repository. Breast Cancer Wisconsin Dataset.