Software Technology, 2024, pp. 1–2

doi: DOI HERE
Advance Access Publication Date: 17 May 2024
Paper

PAPER

NPHA

LIONTOS GEORGIOS[1,*] and THEODORIDIS TILEMACHOS[2]

[1]Ionian University, Department of Informatics, Corfu, 49100, Ionian Islands, Greece
[*]Corresponding author. email-id.com
FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

In the Software Technology course we should create an online application (with the Python programming language) which will process medical data and more specifically the number of doctors visited by an middle aged patient on average. This research was carried out in order to have the possibility for machine learning algorithms to develop and evolve in terms of their functions.

Key words: sleep, health

Introduction

The rapid development of technology and computer science has brought about significant changes in the way we process and analyse data. In the healthcare context, the need for accurate and efficient data analysis is more urgent than ever. One of the most innovative technologies being harnessed is machine learning, which enables useful conclusions to be drawn from large data sets in ways that were previously unimaginable.

dataset

This is a subset of the NPHA dataset filtered down to develop and validate machine learning algorithms for predicting the number of doctors a survey respondent sees in a year. This dataset's represent seniors who responded to the NPHA survey.

Machine Learning Algorithms

Machine learning is a branch of artificial intelligence that allows computers to learn from data and improve autonomously without the need for explicit programmatic intervention, as for example in the health sector where we report on the rate of interaction between middle aged and doctors.

K-Neigbors

The K-Nearest Neighbors (KNN) algorithm is a simple but powerful machine learning algorithm used for both classification and regression problems. KNN belongs to the class of non-parametric supervised learning algorithms, which means that it makes no assumptions about the underlying distribution of the data.

How KNN works The KNN algorithm works as follows:

Selection of Parameter K: We define a positive integer K, which represents the number of neighbors to consider.

Distance Calculation: for each data point that we want to classify or predict the value of, we calculate its distance from all points in the training set. The most common method of calculating distance is the Euclidean distance, although other distance metrics such as Manhattan or Minkowski can be used.

Finding the K Nearest Neighbors: sort the calculated distances and select the K nearest data points (neighbors).

Feature Selection

A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with an evaluation measure which scores the different feature subsets. The simplest algorithm is to test each possible subset of features finding the one which minimizes the error rate. This is an exhaustive search of the space, and is computationally intractable for all but the smallest of feature sets. The choice of evaluation metric heavily influences the algorithm, and it is these evaluation metrics which distinguish between the three main categories of feature selection algorithms: wrappers, filters and embedded methods.Machine learning algorithms are at the heart of modern artificial intelligence and have revolutionised many areas of science and technology.More specificly,in our subject this algorithm seeks to collect information about the frequency of visits of middle-aged people to a set of doctors.

Results

Applying machine learning to analyse survey data on older people's visits to doctors has demonstrated its usefulness and effectiveness in drawing valuable conclusions. Through the analysis process, critical patterns and trends that can help improve health care for middle-aged were identified.

software release lifecycle

the application developed with the Agile model which include the following stages:

1. Data collection and preparation: Identify our project issue that needs to be addressed and set goals for the project
2. Initial design: Designing a machine learning system is an iterative process. There are generally four main components of the process: project setup, data pipeline, modeling and serving .
3. Interface development: Creating an easy-to-use interface for end users (doctors, nurses, caregivers) to interact with the app.
4. Testing validation: Validation testing is the process of evaluating a new software product to ensure that its performance matches the needs of consumers and in our case both middle-aged and medical professionals.
5. Repeat and improve: Feedback is taken from end users (doctors, nurses) on problems, difficulties or shortcomings in the application.
6. Final delivery: Delivering the final version of the application to customers or users and ensuring its smooth integration into the workflow.

conclusion

The findings of this research suggest that the use of machine learning in health data analysis can offer significant benefits, improving the efficiency and quality of health services provided to older people. Furthermore, the development and implementation of such systems can help in the creation of personalized health programs, contributing to preventive care and overall improvement in the quality of life of middle-aged patients.

Github repository link

https://github.com/inf2021062/University-Project-NPHA

Author contributions statement

Tilemahos Theodoridis: Developed the data upload and overview functionality and developed the machine learning algorithm integration. Giorgos Liontos: Created the 2D/3D visualization features and design the user interface and handled the EDA plots.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions.

References

1. Malani, Preeti N., Kullgren, Jeffrey, and Solway, Erica. National Poll on Healthy Aging. 2017

Diagram UML

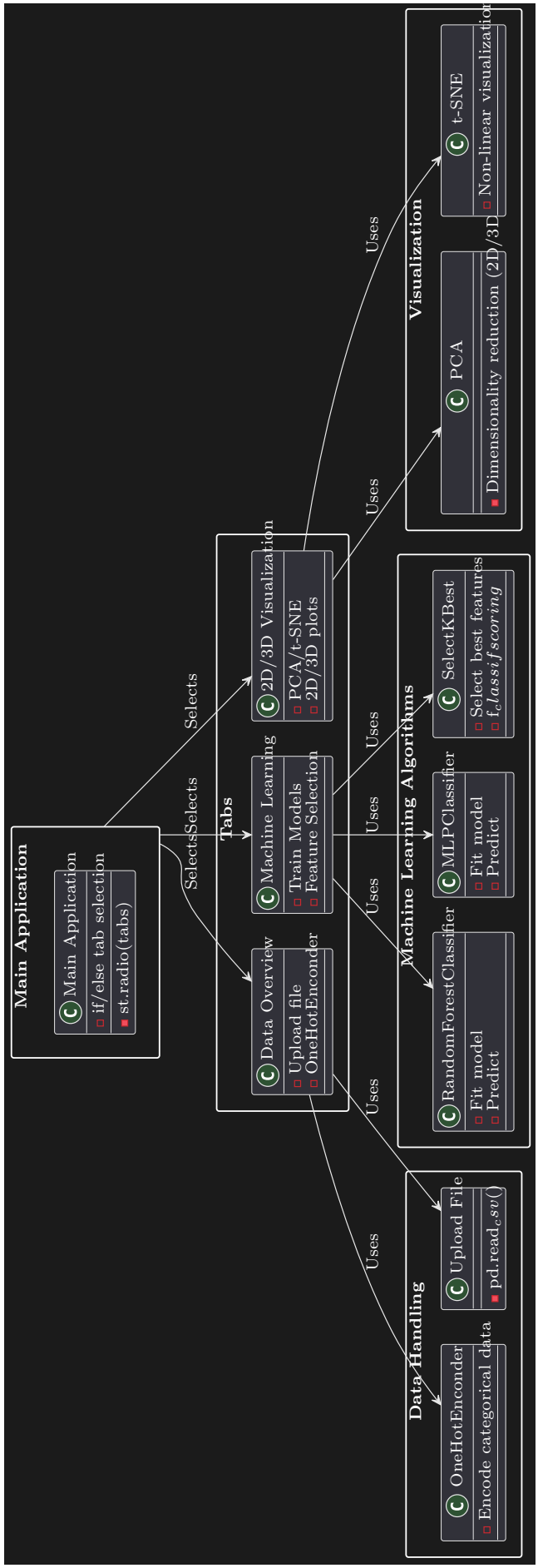

Fig. 1. UML Image