MODEL BULDING

TRAIN AND TEST THE MODEL USING CLASSIFICATION ALGORITHMS

DATE	09- NOVEMBER- 2022
TEAM MEMBER	B.VENKATA SRAVAN KUMAR B.SUDEEPTH B.AAKASH CH.V.S.YASWANTH z
TEAM ID	PNT2022TMID14539
PROJECT NAME	STATISTICAL MACHINE LEARNING APPROACHES TO LIVER DISEASE PREDICTION

LITERATURE SURVEY

1. Naive Bayes

- Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. They are among the simplest Bayesian network models, but coupled with kernel density estimation, they can achieve higher accuracy levels.
- Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum- likelihood training can be done by evaluating a closed- form expression, which takes linear time, rather

than by expensive iterative approximation as used for many other types of classifiers.

2. ANN

- Artificial Neural networks (ANN) or neural networks are computational algorithms .They intend to simulate the behavior of biological systems composed of neurons. ANNs are computational models inspired by an animals central nervous systems. They are capable of machine learning as well as pattern recognition. These are present as systems of interconnected neurons which can compute values from inputs.
- A neural network is an oriented graph. It consists of nodes which in the biological analogy represent neurons, connected by arcs. It corresponds to dendrites and synapses. Each arc is associated with a weight while at each node. To do the prediction, we need to apply the values received as input by the node and define activation function along the incoming arcs, adjusted by the weights of the arcs.An ANN is trained based on backpropagation algorithm.

Algorithm:

1. XTraining Data set of size m*n

2. yLabels for records in X

3. wThe weights for respective layers

4. lThe number of layers in the neural network,1.L ij ij

5. Dij(l) The error for all l,i,j
    - Reduction in number of deaths due to liver diseases

6. t(l)

   0.For all l,i,j

       ☐ More accurate diagnosis of liver disease by the doctors

  7.For i=1 to m

    al feedforward(x(I),w) al a(L)-y(i) ij+aj

    .t i ij+aj

    .t i

     tij(l) t(l) (l) l+1

   6. if j0 then

    Dij(l) 1 tij(l) +w(l)

    enables Python to be used as an alternative application development language to C++ on all supported platforms

   7. else

ij

Dij(l) 1 tij(l)

including iOS and Android.

F. Spyder Notebook

3.KNN

- Spyder is an open-source cross-platform integrated development environment (IDE) for scientific programming in the Python language. Spyder is extensible with first-party and third-party
- The k-nearest neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems.
- The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other. The KNN algorithm hinges on this assumption being true enough for the algorithm to be useful. KNN captures the idea of similarity (sometimes called distance, proximity, or closeness) with some mathematics we might have learned in our childhood calculating the distance between points on a graph.

4.SVM

Support Vector Machine or SVM algorithm is a simple yet powerful Supervised Machine Learning algorithm that can be used for building both regression and classification models. SVM algorithm can perform really well with both linearly separable and non-linearly separable datasets. Even with a limited amount of data, the support vector machine algorithm does not fail to show its magic.

The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. In short, the hyperplane is (n-1)- D plane for n features.

5. PyQt Library

PyQt is a GUI widgets toolkit. It is a Python interface for Qt, one of the most powerful, and popular cross- platform GUI library. PyQt is a blend of Python programming language and the Qt library. PyQt API is a set of modules containing a large number of classes and functions. While QtCore module contains non-GUI functionality for working with file and directory etc., QtGui module contains all the graphical controls. In addition, there are modules for working with XML (QtXml), SVG (QtSvg), and SQL (QtSql), etc.

For this paper, we have used the PyQt version 5, which is implemented as more than 35 extension modules and

plugins, includes support for interactive tools fr data inspection and embeds Python-specific code quality assurance and introspection instruments, such as Pyflakes, Pylint and Rope.

It is available cross-platform through Anaconda, on Windows, on macOS through MacPorts, and on major Linux distributions. Spyder uses Qt for its GUI and is designed to use either of the PyQt or PySide Python bindings. QtPy, a thin abstraction layer developed by the Spyder project and later adopted by multiple other packages, provides the flexibility to use either backend.