

(ML)

Assignment-2

Part-1

Ans-1) In classification, the model is fully trained using the training data and then it is evaluated on test data before being used to perform prediction on new unseen data. For instance, an algorithm can learn to predict whether a given email is spam or ham (not spam). Classifier is a predictive modelling problem where the class label is anticipated for a specific example of input data.

Eg - food is organised by type in grocery aisles, and our school day is organised by subjects, class periods and students.

Ans-2) Difference b/w binary classification & multiclass classification:-

Binary classification.Multiclass classification.

No. of classes

It is a classification of two groups, i.e. classifies objects in at most two classes.

There may be any no. of classes in it, i.e. classifies the object into more than two classes.

Algorithms used

- Logistic Regression.
- k-nearest neighbours
- Decision trees.
- Support vector
- ~~naive~~ naive bayes

- k-nearest neighbors
- Decision trees
- Naive bayes
- Random forest
- Gradient boosting

Ex:-

- Email spam
- Churn prediction
- Conversion prediction.

- Face classification
- Plant species
- Optical character recognition

Ans-3) A decision boundary is defined as a hypersurface used by pattern classifiers to classify input patterns into different regions based on certain criteria. In the context of complex-valued neurons, the decision boundaries for the real and imaginary parts intersect orthogonally.

If the boundary is well-defined and clearly separates the classes, the model's predictions, accuracy increases. However, if the decision boundary is too precise or overfits the training data, it might not generalize well to new data.

- Ans-4)
- Step 1:- Define the problem
 - Step 2:- Gather and preprocess data
 - Step 3:- Feature selection and engineering
 - Step 4:- Choose an algorithm.
 - Step 5:- Model training and Evaluation.
 - Step 6:- Hyperparameter tuning
 - Step 7:- Interpretability and Explicability
 - Step 8:- Deployment.
 - Step 9:- Continuous Monitoring and Maintenance
 - Step 10:- Conclusion.

Define problem - Every machine learning project begins with a well-defined problem. Start by understanding what you want to accomplish with your classification model.

Gather data and preprocess - Collect and clean your data addressing issues like missing values, outliers and noisy data.

Feature selection and engineering - Creating new features or transforming existing ones to improve your model's predictive power.

Choosing of algorithm - Select the right classification algo depend on dataset and problem type.

Model Training and Evaluation - Train your chosen model using training dataset.

Hyperparameter tuning - Most algorithms need fine tuning to optimize model performance. Tools like grid search or random search can help find the best combination of hyperparameters.

Interpretability and Explicability - Understanding how model arrives at its prediction is crucial, especially in critical applications.

Deployment - Once the model performs satisfactory, it has to be deployed in real world environment.

5) Evaluating a classifier's performance is important bcoz it helps to understand how well the model is performing and how it compares to other models.

The most common method to measure a classifier's per

performance is using the Accuracy metric. Here, we compare the actual and predicted class of each input data point, and each match counts for one correct prediction. Accuracy is then given as the no. of correct predictions divided by total no. of predictions.

Rajeshwari
Choudhury
22053180