



Practical No:-6

Aim:- Classifying Reuters Dataset using DNN,
(Multi-class classification problem).


Theory:-

The Reuters dataset is a collection of short newswire articles, labeled over 46 different topics. It is widely used for text classification and natural language processing tasks. Each article is encoded as a sequence of word indexes & is associated with a single category, making it a multi-class classification problem.

Multi-class classification.

Multi-class Classification is the task of classifying input data into one of three or more classes.

Unlike binary classification which has only two possible outcomes, multi-class classification involves more than 2 classes. In this practical, we use a Deep Neural Network (DNN) to classify news articles into one of the 46 categories.

- 
- Deep Neural Networks (DNN)
- A Deep Neural Network (DNN) is a type of artificial neural network with multiple layers between the input & output layers. DNNs are capable of learning complex patterns in data by adjusting weights through backpropagation. They are highly effective for classification tasks when trained on large, labelled datasets.

Preprocessing steps.

Before feeding the data into the neural network the text data (in the form of sequences of integers) is converted into a binary matrix (one-hot encoded format) using techniques such as:-

- Vectorization of input data.
- one-hot encoding of labels.
- Model Architecture.
The DNN model for this practical consists of
 - Input layer - Accepts the vectorized news article.
 - Hidden layers - Typically 2 layers with ReLU activation functions for non-linearity.
 - Output layer - A dense layer with 46 units, using the softmax activation function



for multi-class classification output.

- Loss Function & Optimizer.

- Loss Function

Categorical cross entropy is used as it is suitable for multi-class classification.

- Optimizer

Adam optimizer is used for efficient and adaptive learning.

- Evaluation Metrics.

The model is evaluated using:

- Accuracy: The proportion of correct predictions.

- Loss - Cross-entropy loss which reflects how far the predicted probabilities are from the true labels.

Conclusion:-

The DNN model successfully classified the Reuters dataset into 46 categories with good accuracy. The practical demonstrated the effectiveness of deep learning in multi-class text classification tasks.