**Task 1- News Article Categorization**

**Introduction: -**

This program is designed to automatically categorize news articles into predefined topics using text feature extraction (TF-IDF) and a multi-class classification model (Neural Network). The program also provides a basic command-line interface for processing articles.**Key Features: -**Text Feature Extraction: Using TF-IDF to transform text data into numerical features.Multi-Class Classification Model: Utilizes a neural network for categorizing news articles into multiple predefined topics.Command-Line Interface: Allowing users to type the input articles for classification and receive the predicted category.

**We have used following Python libraries in the program: -**

* Numpy
* pandas
* scikit-learn
* tensorflow
* re
* string
* json

**Dataset: -**

The dataset is a CSV file named news-article-categories.csv with the following structure:Title: Column containing the titles of the news articles.Body: Column containing the bodies of the news articles.Category: Column containing the category labels for each article.

**Program Structure: -**

The program consists of the following components:

Data Preprocessing

1. Loading and Cleaning Data:

* Remove punctuation and digits from the text.
* Convert text to lowercase.
* Combine title and body columns into a single text column.

1. Text Feature Extraction:

* Use TF-IDF to convert text data into numerical features.

Model Training

* Model: Train a neural network model for multi-class classification.
* Evaluation: Evaluate the model's performance using accuracy on the test set.

Command-Line Interface

* Input: Users can input text for category prediction.
* Output: The program predicts the category of the input text.

**Usage: -**

* Run the Program: Execute the script to start the command-line interface.
* Enter Text: Input text into the command line when prompted.
* Predict Category: The program predicts the category of the input text and displays it.
* Exit: Type 'exit' to quit the program.
* Accuracy: The model's accuracy on the test set is displayed after training.
* CLI: The command-line interface allows users to input text and receive predicted categories interactively.