**README: AUTOMATED CLASSIFICATION OF NEWS ARTICLES BY CATEGORY**

**Project Overview:** This project classifies news articles into four categories: Politics, Technology, Entertainment, and Business using machine learning and deep learning models. It also includes sentiment analysis and key phrase extraction, with results visualized through a Streamlit dashboard.

**Dependencies:**

The project requires the following Python libraries:

- `pandas` , `numpy` , `scikit-learn` , `nltk` , `tensorflow` , `transformers` , `rake-nltk` , `streamlit`

**Setup Instructions:**

1. Extract and unzip the folder

2. Install Python and Virtual Environment if required:

- Ensure you have Python 3.7 or higher installed.

- Create a virtual environment and activate it:

```bash

python -m venv env

source env/bin/activate # For Linux/Mac

env\Scripts\activate # For Windows

```

3. Manually Install Dependencies if required:

Run the following commands to install required libraries individually:

```bash

pip install pandas numpy scikit-learn nltk tensorflow transformers rake-nltk streamlit

```

4. Download Dataset if necessary:

- Download the [News Category Dataset](https://www.kaggle.com/datasets/akash14/news-category-dataset/data).

- Save the dataset in the project directory and ensure the file path matches the code.

5. Prepare the Environment:

- Run the following command to download NLTK dependencies (like stopwords):

```python

python -c "import nltk; nltk.download('stopwords'); nltk.download('vader\_lexicon')"

```

**How to Run the Code:**

1. Launch the Streamlit Dashboard: Start the interactive dashboard for predictions and visualizations:

```bash

streamlit run news\_dashboard.py

```

2. Enter the article text in the given box and the prediction will be displayed below