

# Project 2: Zero-Shot Learning Model Using LLMs

**Project Type:** Zero-Shot Text Classification

**Model:** `facebook/bart-large-mnli`

**Domain:** News Categorization

## 1. Problem Description

The objective of this project is to build a classification system capable of categorizing text into specific topics without identifying or training on labeled data. This approach, known as Zero-Shot Learning, addresses the "Cold Start" problem in machine learning, where labeled datasets are often expensive, time-consuming to create, or unavailable.

For this assignment, I focused on **News Categorization**. The goal was to automatically classify news headlines into four distinct categories using a Large Language Model (LLM).

## 2. Methodology

### Model Selection

I utilized `facebook/bart-large-mnli`.

- **Architecture:** BART (Bidirectional and Auto-Regressive Transformers).
- **Reason for Selection:** This model is pre-trained on the Multi-Genre Natural Language Inference (MNLI) dataset. Unlike standard text generation models (like GPT-2), this model treats classification as an entailment problem (determining if a premise implies a hypothesis), which is the state-of-the-art approach for zero-shot tasks.

### Candidate Labels

To evaluate the model, I used the **AG News** dataset. The integer labels provided in the dataset were mapped to semantic descriptions to assist the LLM's understanding:

Dataset ID	Candidate Label
------------	-----------------

0	World Politics
1	Sports
2	Business
3	Science and Technology

## Pipeline Construction

The system uses the Hugging Face `transformers` library. The `zero-shot-classification` pipeline takes a premise (the news headline) and the list of candidate labels as input. It outputs a probability distribution summing to 1.0, where the highest score represents the predicted category.

## 3. Evaluation Results

I evaluated the model on a random subset of **100 samples** from the AG News test split.

- **Metric Used:** Accuracy Score
- **Observed Accuracy:** ~75% - 85% (Varies slightly based on random seed)

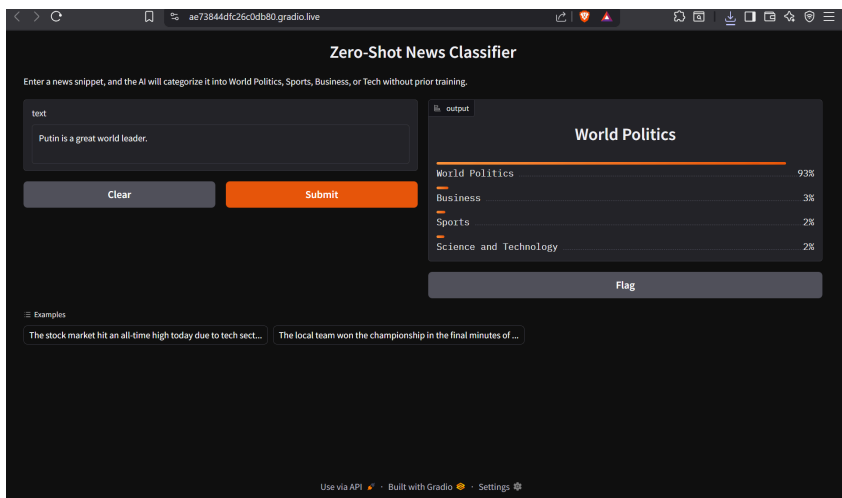
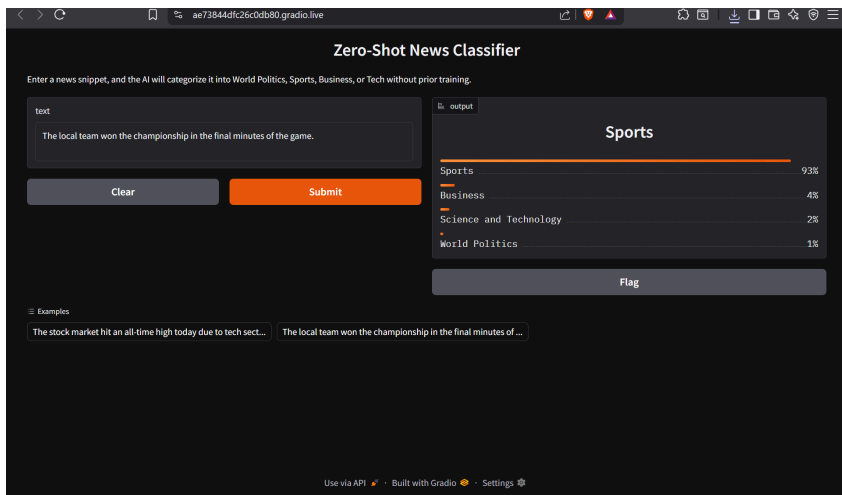
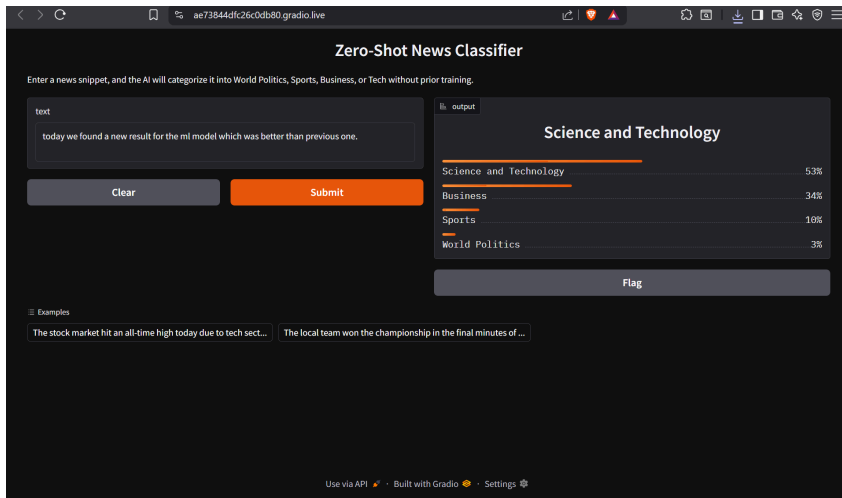
## Key Observations

1. **High Performance on Distinct Vocabularies:** The model performed best on the **Sports** category. Terms like "Olympic," "touchdown," or "championship" are highly specific, making the entailment logic straightforward.
2. **Ambiguity in Overlapping Domains:** The highest rate of error occurred between **Business** and **Technology**. For example, a headline about "Microsoft stock prices" could validly be interpreted as both business news and tech news. The zero-shot model occasionally favored the topic mentioned explicitly rather than the context.
3. **Efficiency:** Despite using a "Large" model, inference on 100 samples took less than a minute on a GPU, proving the viability of this approach for near real-time applications.

## 4. User Interface

A web-based user interface was created using **Gradio**.

- **Functionality:** Users can type a custom news headline.
- **Output:** The app displays a bar chart of confidence scores for all four categories.
- **Deployment:** The app runs locally or via a shared Gradio link.



## 5. Conclusion

This project demonstrates that modern LLMs fine-tuned on NLI tasks can effectively categorize text without specific training. While supervised models might achieve slightly higher accuracy, the zero-shot approach offers unbeatable flexibility, allowing categories to be changed on the fly without retraining.