

Natural Language Processing (NLP) and Large Language Models (LLMs)

Lecture 8-2: Applications and Finetuning

Chendi Wang (王晨笛)
chendi.wang@xmu.edu.cn

WISE @ XMU

2025 年 5 月 6 日

What can pretrained LLMs do?

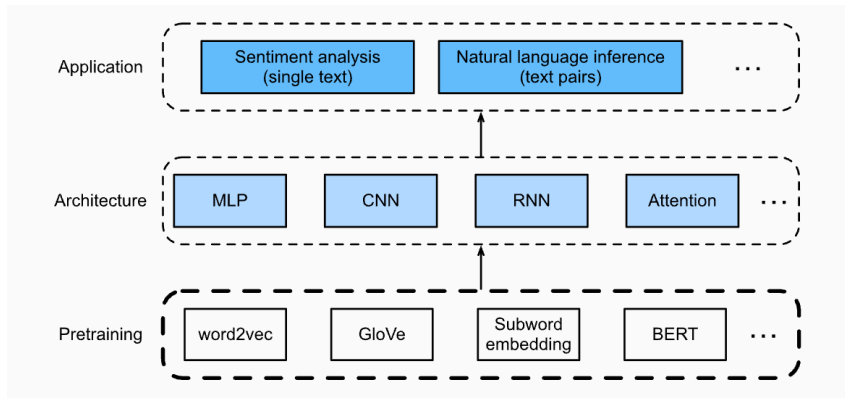


Figure is from d2l

A pipeline of using LLMs

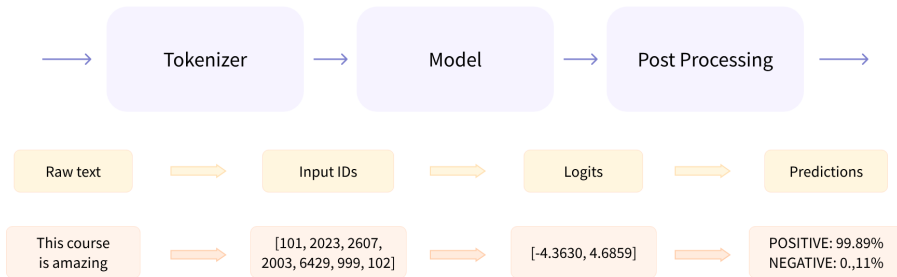


Figure is from a *hugging face* course.

A pipeline of using LLMs

- You may try `pipeline()` in *pipeline.ipynb*
- Still use a mirror

Recap:NER

- **Named-Entity Recognition (NER)**: find and classify names in text.
- **Kobe Bryant** was a legendary basketball player for the **Los Angeles Lakers** from **1996 to 2016**, after completing his high school education in **Philadelphia**.
 - **Kobe Bryant**: Person
 - **Los Angeles Lakers**: Organization
 - **1996 to 2016**: Time
 - **Philadelphia**: Location

NER made easy with Pipelines

- You can now try Named Entity Recognition (NER) using a pipeline.
- Just two lines of code are enough —try it in *NER_pipeline.ipynb*.

① Section 1: Sentiment analysis

② Section 2: natural language inference

Sentiment analysis (a.k.a., opinion mining)

- Determines emotional tone: **positive**, **negative**, or **neutral** in text
- Processes large volumes: emails, chats, social media, reviews
- Automatically detects author's attitude
- Helps improve customer service and brand reputation

Sentiment Analysis (Human Perspective)

- “Great value for money. Highly recommend!”

Sentiment Analysis (Human Perspective)

- “The laptop is slow and overheats frequently. Very disappointed!”

Sentiment Analysis (Human Perspective)

- “This phone has an amazing camera and long battery life, but it’s too expensive.”
- “This phone is too expensive, but it has an amazing camera and long battery life.”

Sentiment Analysis (Human Perspective)

- “The weather is okay today, not too hot.”

Why use a machine for sentiment analysis?

- **Objective insights:** AI-based tools reduce personal bias, offering a consistent and impartial view of customer sentiment.
- **Scalable analysis:** Businesses can process large volumes of unstructured text—such as emails, chat logs, surveys, CRM records, and product reviews—efficiently and cost-effectively using cloud-based solutions.
- **Real-time monitoring:** Sentiment analysis tools enable immediate detection of customer emotions, allowing companies to respond quickly to market trends or potential issues. Alerts can be triggered by specific keywords with negative sentiment.
- **Actionable feedback:** AI systems extract entities (e.g., products or services) associated with negative sentiment, helping businesses make targeted improvements based on authentic customer feedback.

Use cases of sentiment analysis

- **Improve customer service:** Support teams use sentiment analysis to tailor responses based on conversation tone. AI-enabled chatbots detect urgency and escalate critical issues to human agents.
- **Brand monitoring:** Organizations track mentions across social media, forums, blogs, and news sites. Sentiment analysis helps PR teams assess public mood, address complaints, and leverage positive buzz.
- **Market research:** Businesses use sentiment analysis to understand what customers like or dislike. Insights from reviews, surveys, and social media are shared with product teams to guide innovation.
- **Campaign tracking:** Marketers monitor public response to advertising campaigns in real time. If sentiment falls below expectations, they adjust strategies based on sentiment trends and analytics.

Sentiment Analysis via Pipelines

- Try sentiment analysis using *sentiment_pipeline.ipynb*.
- Just two lines of code are enough to use a large language model!

Life's not that simple

- “Oh, I just love when my flight gets delayed for 3 hours with no explanation. It's my favorite part of traveling!”

Irony loves to mess with us

- We have an example of a model that may understand irony in *sentiment_pipeline.ipynb*.

Behind the pipeline

- Reproduce the pipeline yourself in *sentiment_pipeline.ipynb*.

Finetune your model using IMDb

- The Internet Movie Database (IMDb) is an online database of information related to films, television series, podcasts, home videos, video games, ...

IMDb Top 250 movies

As rated by regular IMDb voters.

0 OF 250 WATCHED 0%

250 Titles

Sort by **Ranking** ↑

	1. The Shawshank Redemption 1994 2h 22m R ★ 9.3 (3M) ☆ Rate 🔗 Mark as watched	
	2. The Godfather 1972 2h 55m R ★ 9.2 (2.1M) ☆ Rate 🔗 Mark as watched	
	3. The Dark Knight 2008 2h 32m PG-13 ★ 9.0 (3M) ☆ Rate 🔗 Mark as watched	

Fine-tune Your Model on IMDb

- Try fine-tuning your own model using the IMDb dataset.
- A demo is available in *Sentiment_finetune_bert.ipynb*.

Pretrain a Model on IMDb

- You may try a demo of pretraining an RNN on IMDb in `Sentiment analysis.ipynb`

① Section 1: Sentiment analysis

② Section 2: natural language inference

Natural Language Inference

- Some tasks, such as sentiment analysis, involve classifying a single text sequence into predefined categories.
- However, in many cases, classifying a single sequence is not enough. Instead, we need to **reason** over pairs of text sequences.
- For example, we may want to determine whether one sentence can be inferred from another, or eliminate redundancy by identifying semantically equivalent sentences.
- Natural Language Inference (NLI) studies whether a **hypothesis** can be logically inferred from a **premise**, where both are text sequences.

Types of Relationships in NLI

- NLI determines the logical relationship between a pair of text sequences. These relationships typically fall into three categories:
 - **Entailment:** The hypothesis can be logically inferred from the premise.
 - **Contradiction:** The hypothesis contradicts the premise (i.e., its negation can be inferred).
 - **Neutral:** The hypothesis is neither entailed nor contradicted by the premise.
- NLI is also known as the recognizing textual entailment task.

- Premise: The chef is baking a chocolate cake.
- Hypothesis: The chef is preparing a dessert.

- Premise: The store is closed for renovations.
- Hypothesis: The store is open for business.

- Premise: She is reading a book in the library.
- Hypothesis: She is a college student.