Lecture 12: Text classification
Sentiment analysis

LING-351 Language Technology and LLMs

Instructor: Hakyung Sung

October 2, 2025

Table of contents

Review

Last class

- Text classification
- Conceptual understanding of how neural network works

Text classification

What do we need?

- Corpus to train

Text classification

What do we need?

- Corpus to train
- Label(s) to predict

Text classification

What do we need?

- Corpus to train
- Label(s) to predict
- Neural network to train the texts (from the corpus) and labels

Python Tutorial

Python Tutorial

We will explore how to implement this in Python.

- Pre-built code is provided
- Opportunity to learn how to train a neural network model
- Optional make-up chance for those who missed a previous assignment (+5 points)

Python Tutorial

Task: sentiment analysis on short movie review texts

- *"This film is terrible!"* → Negative? Positive?

Python Tutorial

Task: sentiment analysis on short movie review texts

- *"This film is terrible!"* → Negative? Positive?
- *"This film is not terrible, it's great!"* → Negative? Positive?

Python Tutorial

Task: sentiment analysis on short movie review texts

- *"This film is terrible!"* → Negative? Positive?
- *"This film is not terrible, it's great!"* → Negative? Positive?
- Corpus: the IMDB dataset, a collection of movie reviews labeled as positive or negative

Python Tutorial

Task: sentiment analysis on short movie review texts

- *"This film is terrible!"* → Negative? Positive?
- *"This film is not terrible, it's great!"* → Negative? Positive?
- Corpus: the IMDB dataset, a collection of movie reviews labeled as positive or negative
- Model: a pre-trained Transformer (BERT) fine-tuned as a sentiment classifier on this corpus

Reminder

Reminder

By October 2nd (**TODAY**)

1. Review the sample papers on the course website (*https://hksung.github.io/Fall25_LING351/materials/*)
2. Add your names to the shared sheet (*https://docs.google.com/spreadsheets/d/1on8icHoXUsj74m1UNEHk8CycHEAmVH1nRsUatpn9xYc/edit?usp=sharing*) - *First come first served*
3. **Choose one paper you like best**
4. **Groups will be announced next week**