

# CPSC 330 Lecture 17: Natural Language Processing



# Which metric in what context?

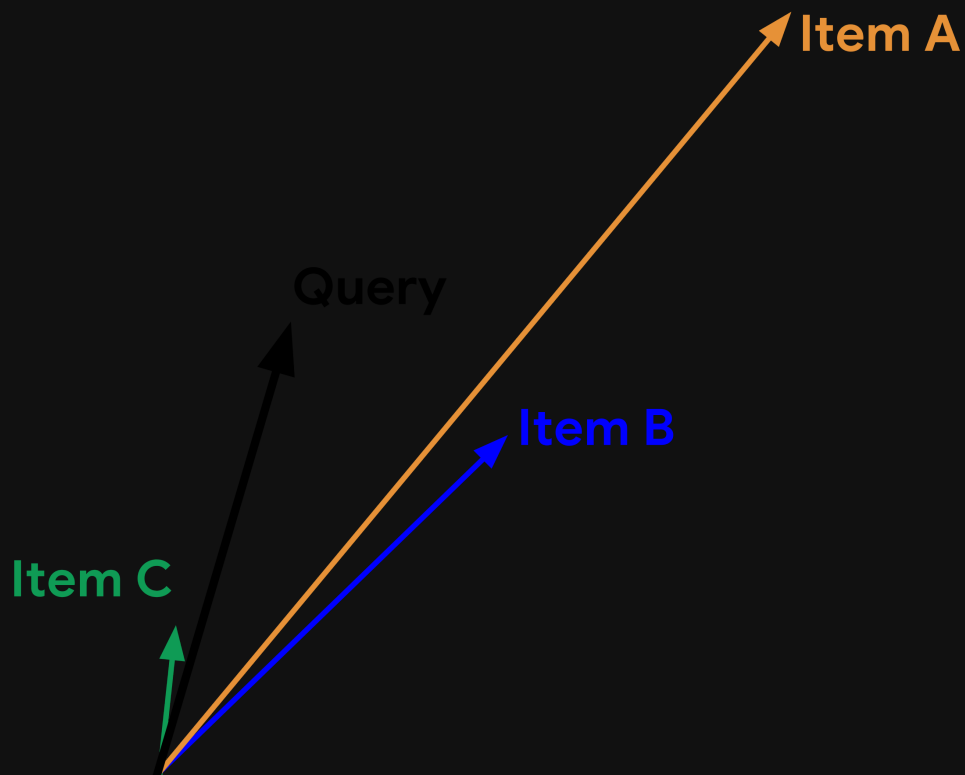
Given a query vector "Query" in the picture below and the three item vectors, determine the ranking of the items for the three similarity measures below:

- Example: Similarity based on Euclidean distance: item B > item C > item A

- **Similarity based on dot product: ?**

- **Cosine similarity: ?**

Adapted from [here](#).





# What is NLP?

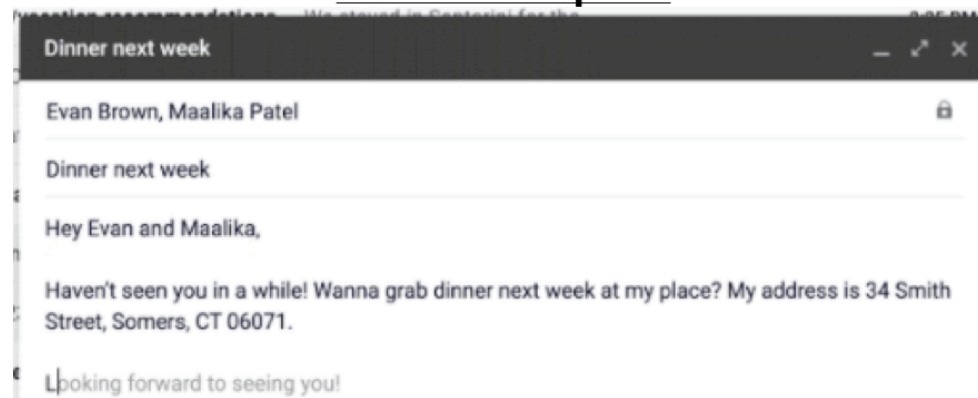
- **Natural Language Processing (NLP)** is a field at the intersection of **computer science, linguistics, and artificial intelligence**.
- It focuses on enabling computers to understand, interpret, and generate human language.

# Examples of NLP applications

## Voice assistants



## Smart compose



## Translation



# Key challenges in NLP

- **Ambiguity:** words can have multiple meanings and meaning depends on previous words/sentences
  - I had toast with *jam* vs We got stuck in a traffic *jam*.
  - If the **baby** does not thrive on raw **milk**, boil **it**.
- **Structure:** syntax and grammar vary widely
  - Time flies like an arrow vs Fruit flies like a banana.
- **World knowledge:** understanding beyond text
  - Olive oil: oil **made from** olives
  - Baby oil: oil **made for** babies

# Goal of this lecture

NLP is a broad field. In this lecture I'll give you a high-level introduction to

- Topic Modeling
- Word embeddings

We will look at Large Language Models (LLM) next class!

# Activity: Context and word meaning

Pair up with the person next to you and try to guess the meanings of two made-up words: **flibbertigibbet** and **groak**.

1. The plot twist was totally unexpected, making it a **flibbertigibbet** experience.
2. Despite its **groak** special effects, the storyline captivated my attention till the end.
3. I found the character development rather **groak**, failing to evoke empathy.
4. The cinematography is **flibbertigibbet**, showcasing breathtaking landscapes.
5. A **groak** narrative that could have been saved with better direction.
6. This movie offers a **flibbertigibbet** blend of humour and action, a must-watch.
7. Sadly, the movie's potential was overshadowed by its **groak** pacing.
8. The soundtrack complemented the film's theme perfectly, adding to its **flibbertigibbet** charm.
9. It's rare to see such a **flibbertigibbet** performance by the lead actor.
10. Despite high expectations, the film turned out to be quite **groak**.
11. **Flibbertigibbet** dialogues and a gripping plot make this movie stand out.
12. The film's **groak** screenplay left much to be desired.