Sentiment analysis

CS582 - Machine Learning
Professor: Dr. Emdad Khan

Group 2

986191 - Boldkhuu Batbaatar
986264 - Batjargal Bayarsaikhan (Alex)
986159 - Unubold Tumenbayar

Project idea

Sentimental Analysis

- Word2vec model
- Basic classification

Improvement

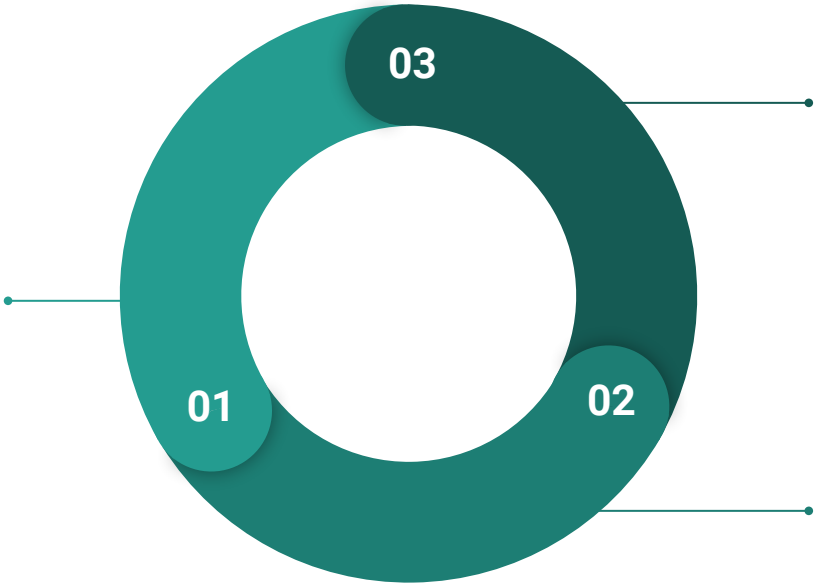
- Recurrent NN
- Long Short Term Memory

Mobile app

Determines if you are optimistic or pessimistic based on your Facebook posts

Project goal

Learn how to use machine learning algorithms in Natural Language Processing



How to apply machine learning model to a mobile application

How to deploy machine learning model as a web service

Word Vectors

Recurrent Neural Networks (RNN)

Long Short-Term Memory (LSTM)

Sentiment analysis

Determining the emotional tone behind a piece of text

Refers to the use of **natural language processing** to identify, extract affective states and subjective information.

This movie is so good.
That book was waste of time.

Natural Language Processing

NLP is all about creating systems that process or *understand* language in order to perform certain tasks.

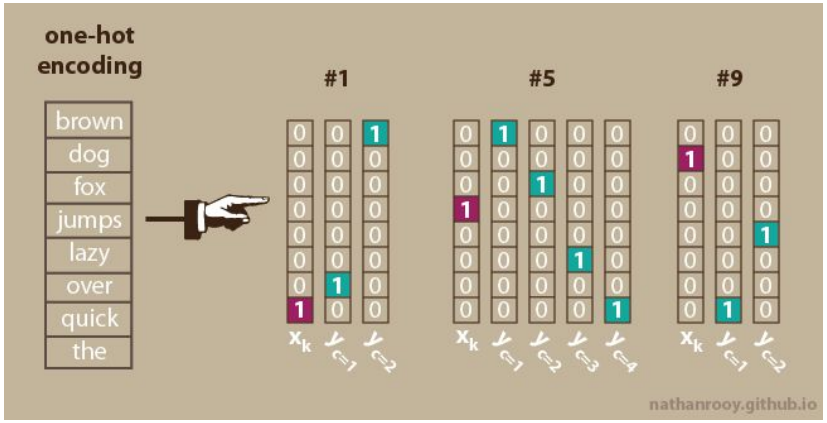
- **Question Answering**
  The main job of technologies like Siri, Alexa, and Cortana
- **Machine Translation**
  Translating a paragraph of text to another language
- **Speech Recognition**
  Having computers recognize spoken words
- **Sentiment Analysis**

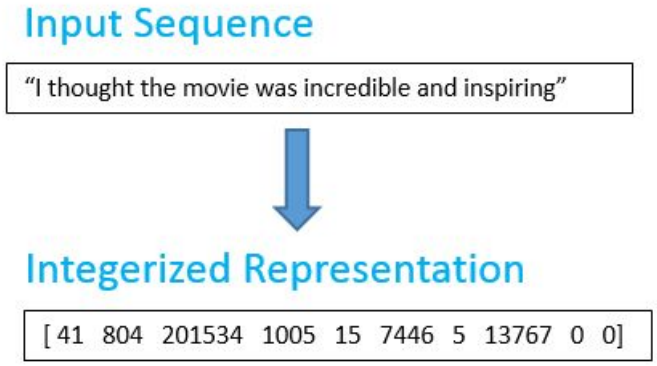
Machine Learning Model

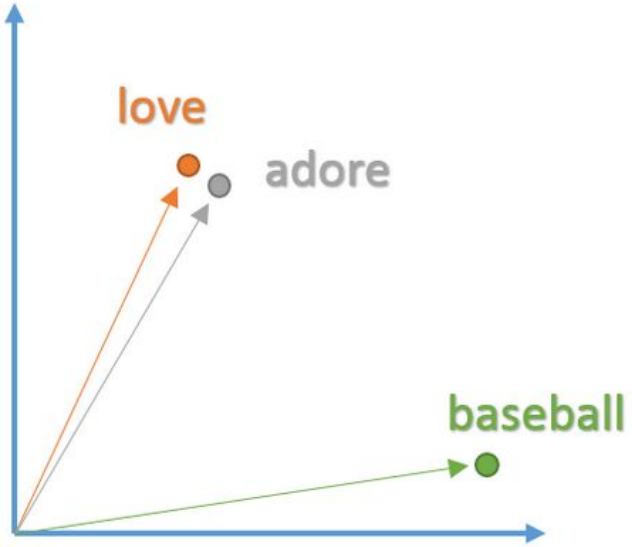
Dataset: IMDB movie reviews

Word Vectors (Word2Vec)

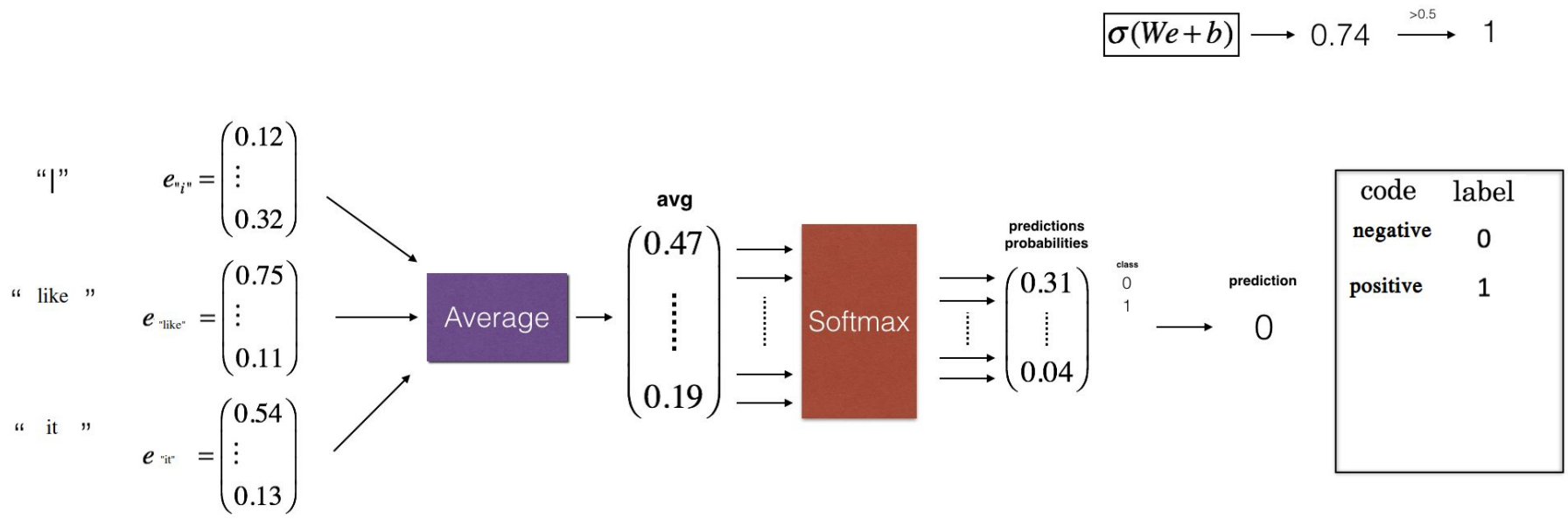
- In order to process text with NLP, it should be represented as a numbers.
- The vector representation of a word is also known as a **word embedding**.
  - Continuous Bag of Words (uses NN)
  - Skip-gram (uses NN)







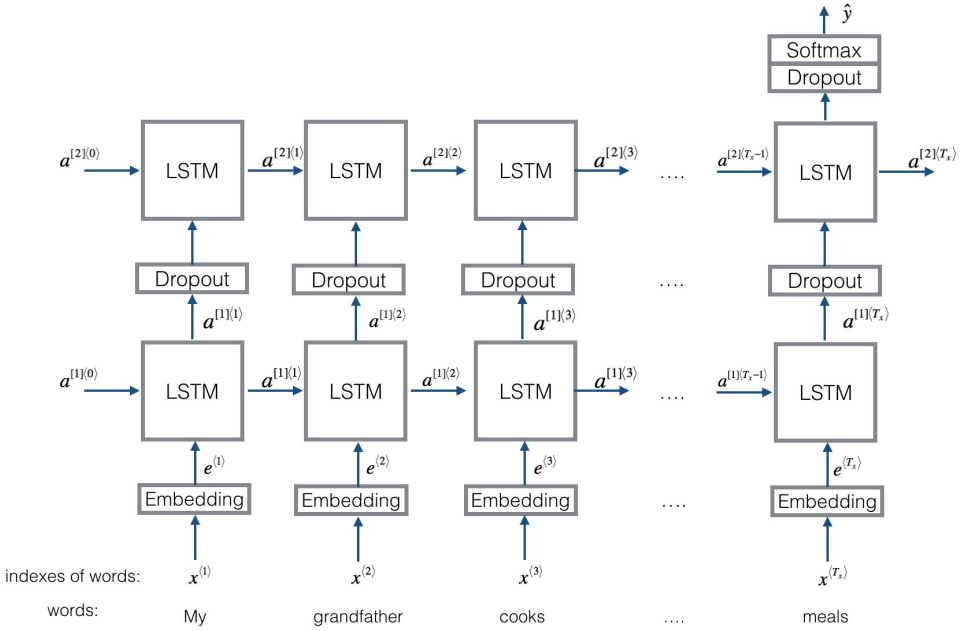
Naive approach with a simple classification



Disadvantage of Naive approach

Recurrent Neural Network (RNN)

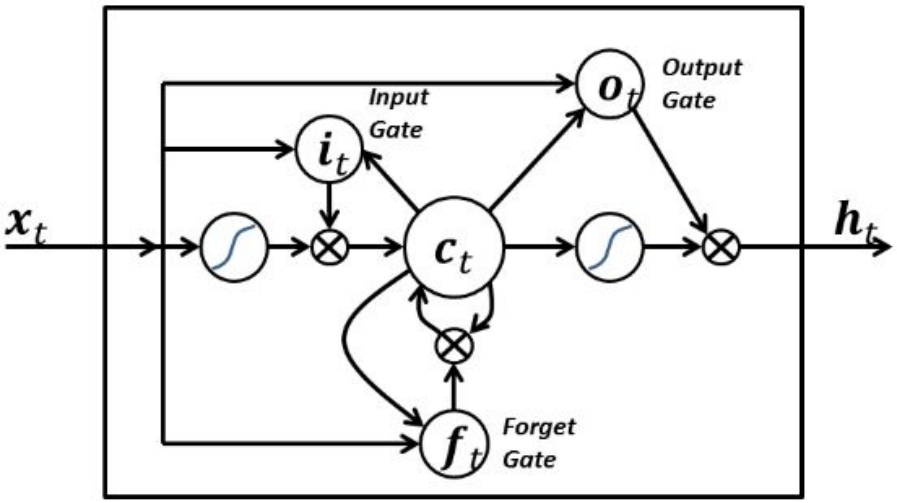
- Each word in a sentence depends greatly on what came before and comes after it.
- In order to account for this dependency, we use a recurrent neural network.



RNN is a class of artificial neural network where connections between nodes form a directed graph along a sequence.

Long Short Term Memory

- LSTMs is addressed for determining which the parts of sentences are important.
- Long Short Term Memory Units are modules that you can place inside of recurrent neural networks.
- RNN composed of LSTM units is

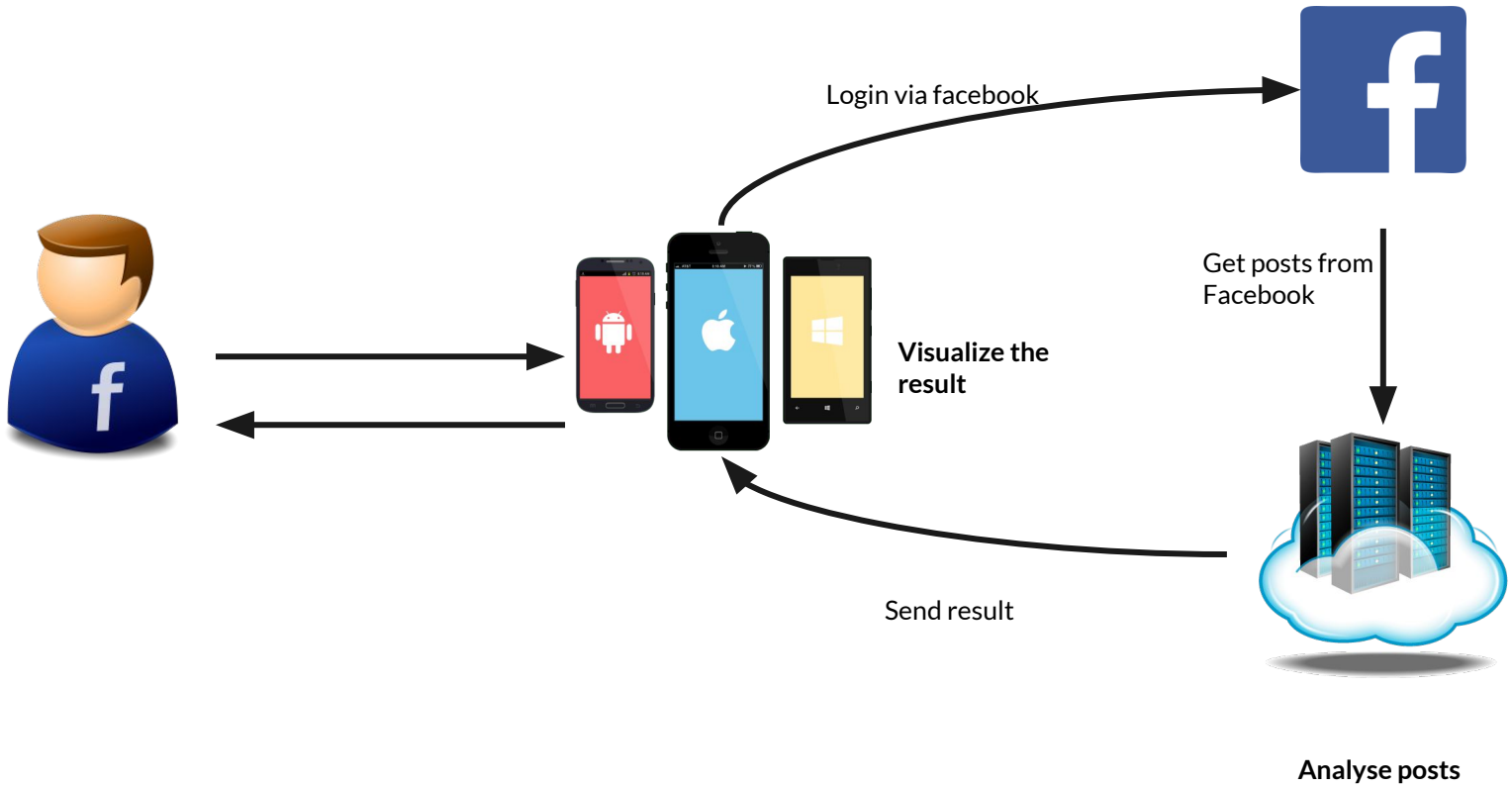


Passage: “The first number is 3. The dog ran in the backyard. The second number is 4.”

Question: “What is the sum of the 2 numbers?”

How to use our model?

Architecture of the app



Problems we faced vs how we solved

Problems	Solutions
First naive approach did not work well.	We used RNN and LSTMs.
It was so challenging to understand the process behind the high level frameworks such as tensorflow.	
It was first experience for us to apply our ML model to real world mobile application.	We deployed our model as a web service and call it using web APIs.

Our novelty (Conclusion)

We can implement sentiment analysis using simple classification with word2vec but it has some drawbacks as we mentioned above.

Therefore, we used RNN and LSTM to make it better,

Demo