Sentiment Classification

This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played the game Chrono Cross but out of all of the games I have ever played it has the best music! It backs away from crude keyboarding and takes a fresher step with grate guitars and soulful orchestras. It would impress anyone who cares to listen! ^_^

positive

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

negative

If you have a kid who likes to move the trains by hand, don't spend the extra money on this. I think it's probably a good product, just not for a 3 yr old.

positive

or negative

or neutral

Sentiment Classification

binary positive negative

multiclass positive neutral negative

continuous 1.0 ... 0.0

Binary Classification

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

$$\begin{pmatrix} -0.8364 \\ 0.3455 \\ 1.5644 \\ -0.0012 \\ ... \end{pmatrix}$$

0.847
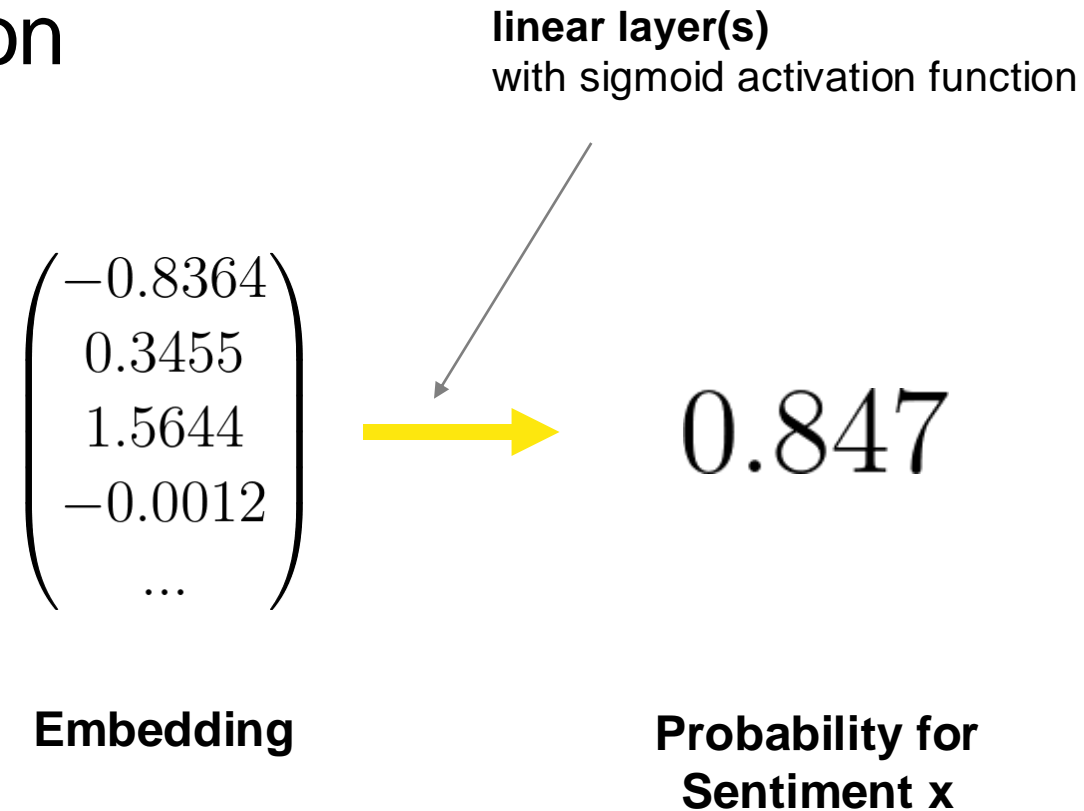
Text

Embedding

Probability for Sentiment x

Binary Classification

linear layer(s)
with sigmoid activation function

$$\begin{pmatrix} -0.8364 \\ 0.3455 \\ 1.5644 \\ -0.0012 \\ ... \end{pmatrix}$$

0.847

Embedding

Probability for Sentiment x

Binary Classification

static representations
(word2vec, GloVe, TF-IDF, ...)

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

$$\begin{pmatrix} -0.8364 \\ 0.3455 \\ 1.5644 \\ -0.0012 \\ ... \end{pmatrix}$$

Text

Embedding

contextual language model
(BERT, Sentence-Transformer, …)

Bidirectional Encoder Representations from Transformers (BERT)

[CLS]	My	name	is	BERT

768

$$\begin{pmatrix} -0.8439 & -1.3409 & 1.9044 & 1.0563 & -0.5644 & \dots \\ 0.3438 & -2.0113 & 3.9534 & 2.0488 & 0.9199 & \dots \\ 1.4931 & -0.8347 & -0.0011 & 1.4566 & 2.0013 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0.4344 & 1.0947 & -0.4267 & -1.5794 & 0.4545 & \dots \end{pmatrix}$$

Embeddings with BERT

- represent word and text semantically (not only syntactically)

- contextualize word embeddings (eg. bank)

- not understandable by humans

- basis for different downstream tasks (eg. sentiment classification)

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.
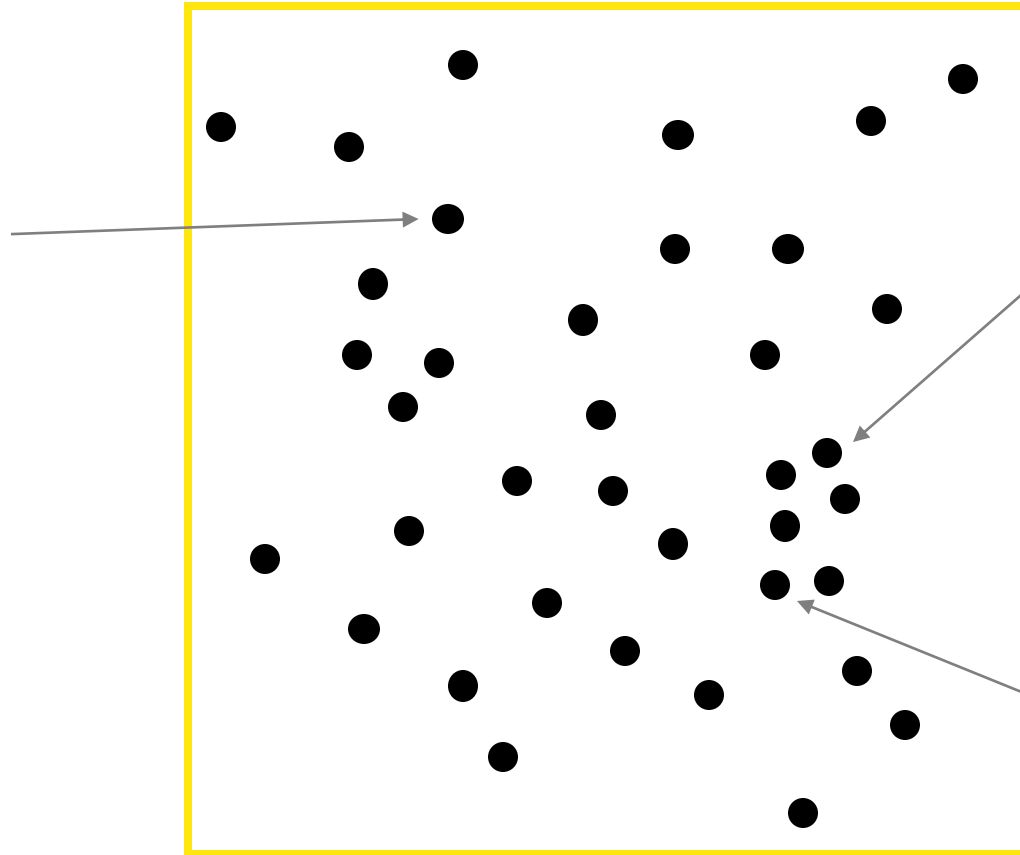
If you have a kid who likes to move the trains by hand, don't spend the extra money on this. I think it's probably a good product, just not for a 3 yr old.

I am quite sure any of you actually taking the time to read this have played the game at least once, and heard at least a few of the tracks here. And whether you were aware of it or not, Mitsuda's music contributed greatly to the mood of every single minute of the whole game. Composed of 3 CDs and quite a few songs …
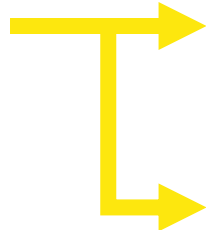
Data Is All You Need

more is generally better
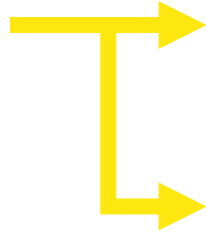
for **Embeddings** ➔ **Unlabelled Text Corpus**
(books, Wikipedia, …)

➔ **Pretrained models from** 🤗 **Hugging Face**

for **Sentiment Classification** ➔ **Supervised Learning**
with sentiment labels

➔ **Semi-Supervised Learning**
with weak sentiment labels

Weak Labelling

I am quite sure any of you actually taking the time to read this have played the game at least once, and heard at least a few of the tracks here. And whether you were aware of it or not, Mitsuda's music contributed greatly to the mood of every single minute of the whole game. Composed of 3 CDs and quite a few songs (I haven't an exact count), all of which are heart-rendering and impressively remarkable, this soundtrack is one I assure you you will not forget. It has everything for every listener -- from fast-paced and energetic (Dancing the Tokage or Termina Home), to slower and more haunting (Dragon God), to purely beautifully composed (Time's Scar), to even some fantastic vocals (Radical Dreamers).This is one of the best videogame soundtracks out there, and surely Mitsuda's best ever. ^_^
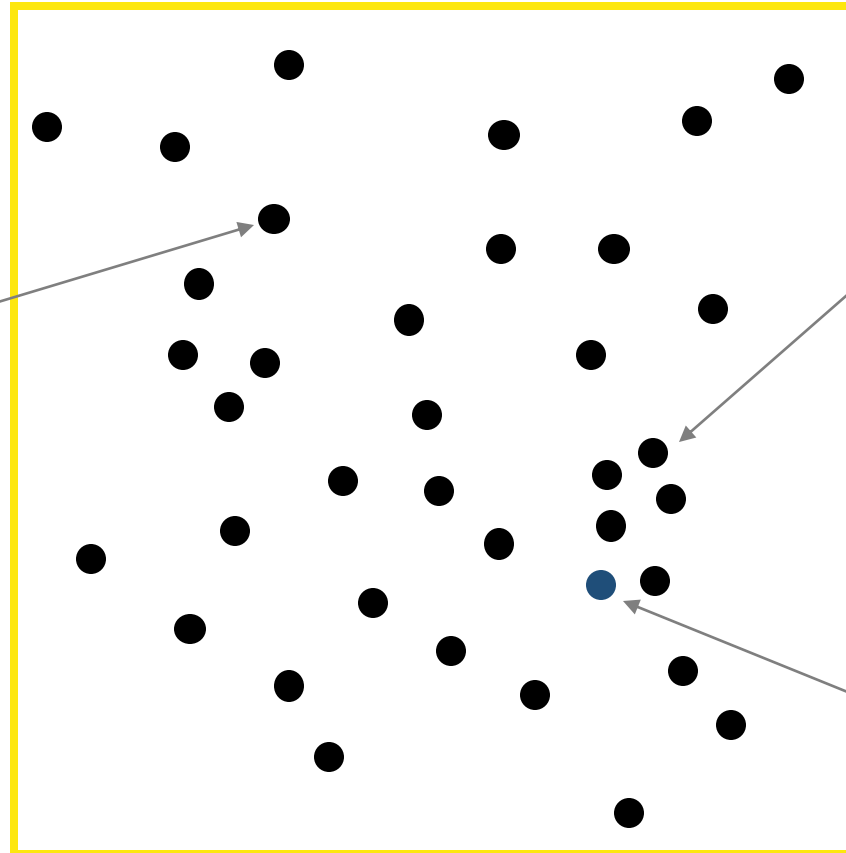
positive or negative

Weak Labelling with KNN

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

negative

This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played …

positive

I am quite sure any of you actually taking the time to read this have played the game at least once, and heard at least a few of the tracks here. And whether you were aware of it or not, Mitsuda's music contributed greatly to the mood of every single minute of the whole game. Composed of 3 CDs and quite a few songs …
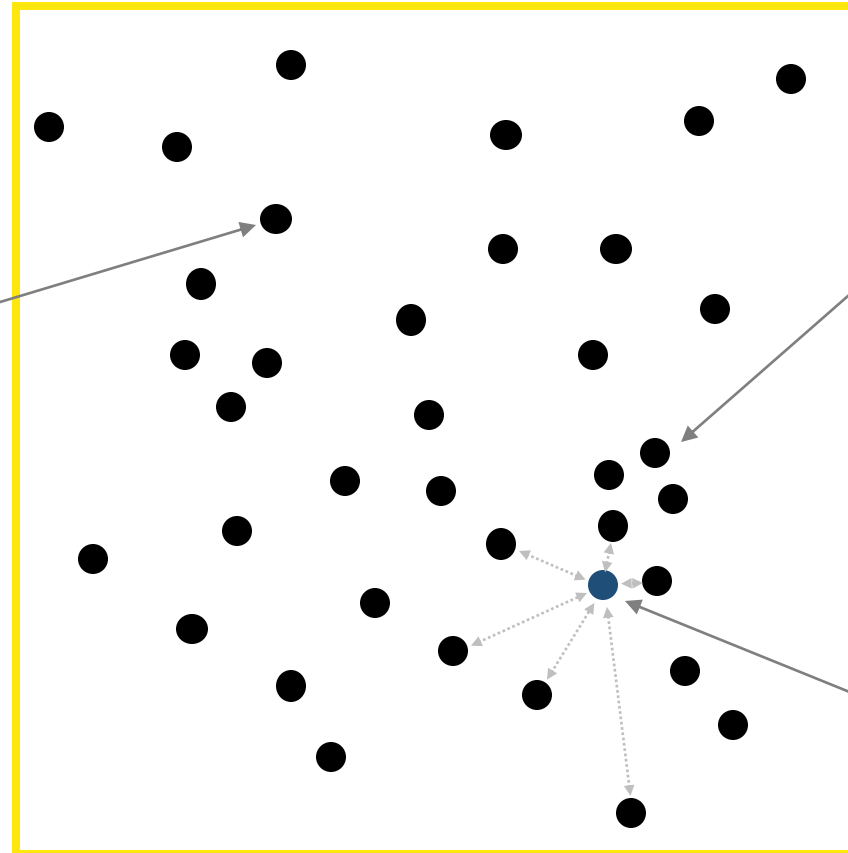
positive   or   negative

Weak Labelling with KNN

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

negative

This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played …

positive

I am quite sure any of you actually taking the time to read this have played the game at least once, and heard at least a few of the tracks here. And whether you were aware of it or not, Mitsuda's music contributed greatly to the mood of every single minute of the whole game. Composed of 3 CDs and quite a few songs …
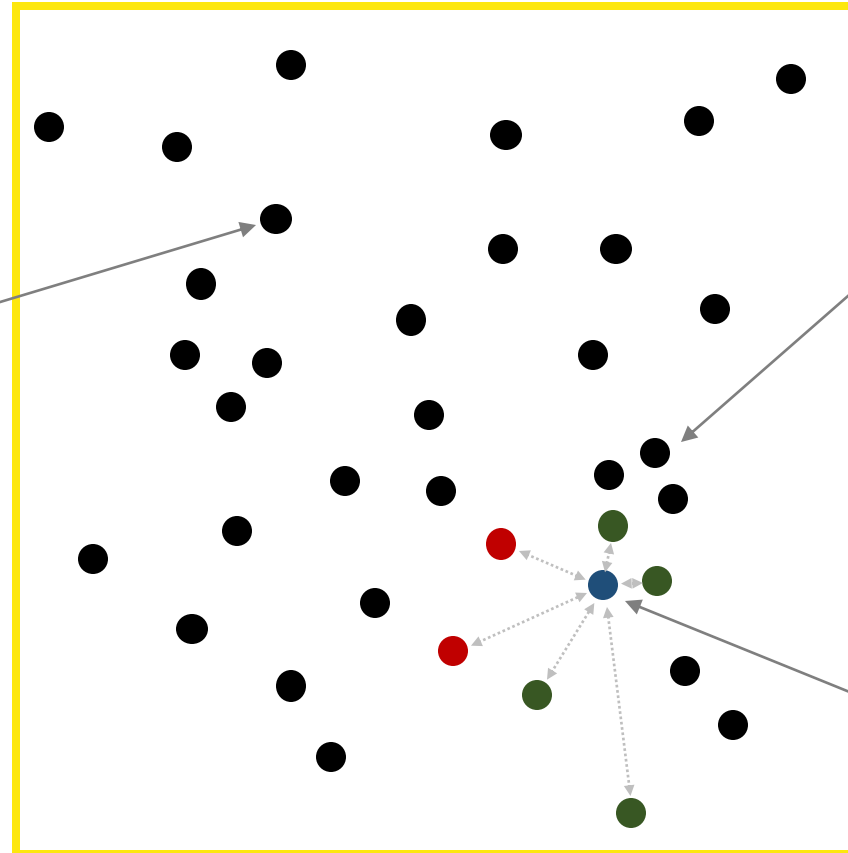
positive   or   negative

Weak Labelling with KNN

A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

negative

This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played …

positive

I am quite sure any of you actually taking the time to read this have played the game at least once, and heard at least a few of the tracks here. And whether you were aware of it or not, Mitsuda's music contributed greatly to the mood of every single minute of the whole game. Composed of 3 CDs and quite a few songs …
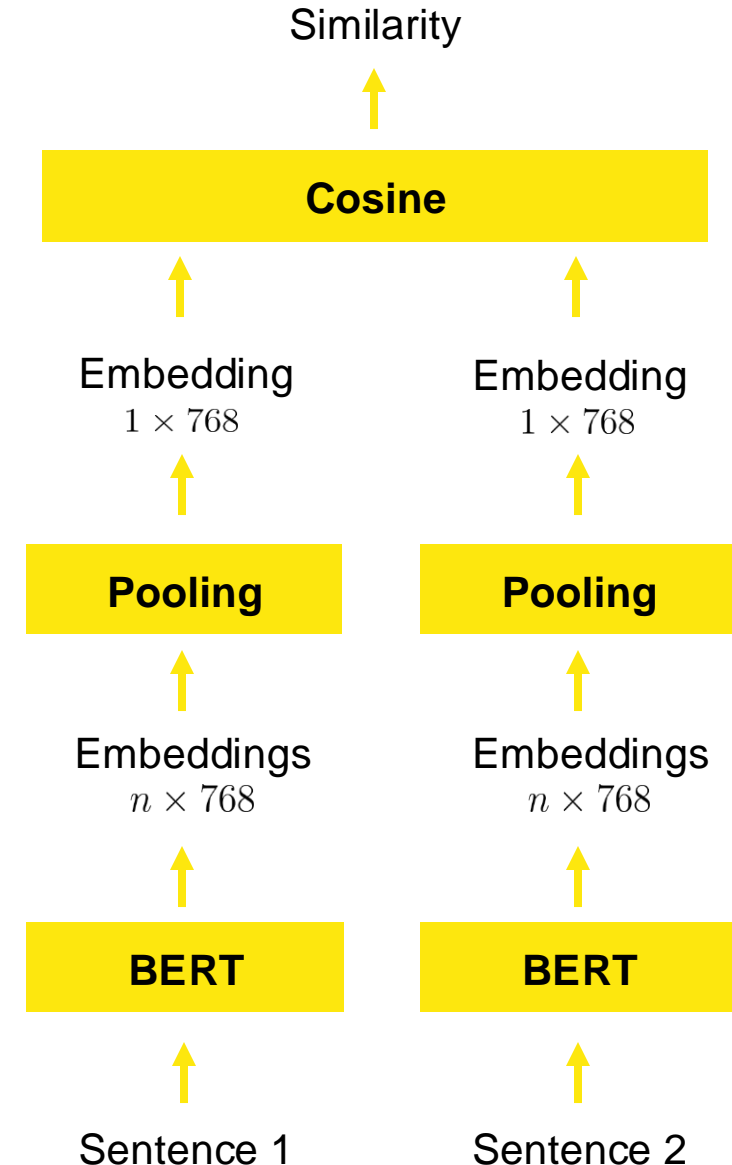
positive   or   negative

Sentence Transformers

- BERT models not suitable for efficient sentence similarity search

- SBERT focuses on comparable sentence embeddings (cosine similarity)

- siamese network based on BERT model

Similarity

Cosine

Embedding
1×768

Embedding
1×768

Pooling

Pooling

Embeddings
$n \times 768$

Embeddings
$n \times 768$

BERT

BERT

Sentence 1

Sentence 2

Mini Challenge Goals

1. understand the concept of embeddings

2. show the effect of more training data

3. show the effect of weak labelling on sentiment classification

4. have fun

Learning Curve
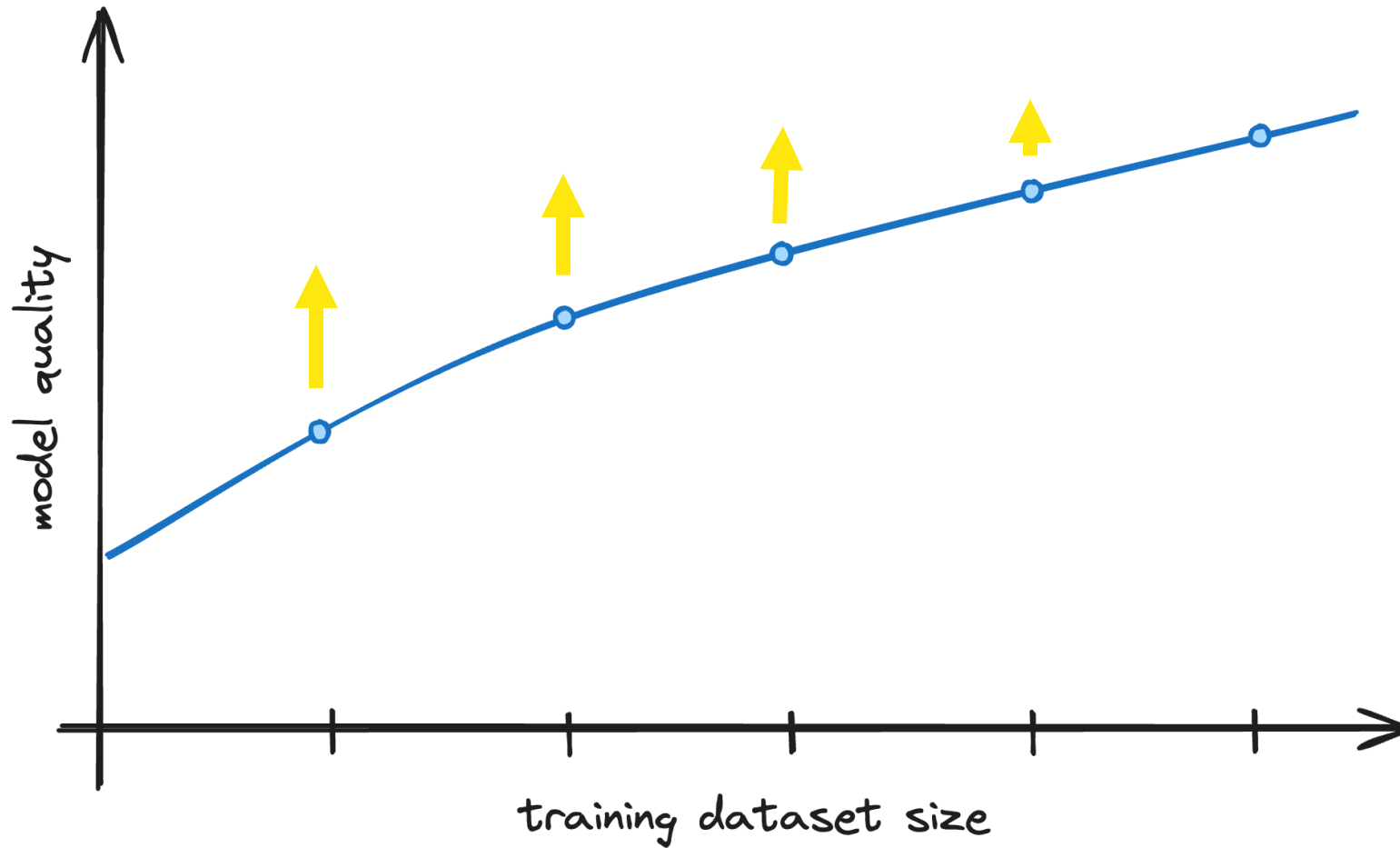
Datasets

- use own or one of the recommended datasets

  (financial_phrasebank, amazon_polarity, sst2)

- minimum samples: 100 labelled, 500 unlabelled

- check for maximum text length

- balance dataset by undersampling

Proposed Approach

1. choose the framework for this work

2. train baseline classification model

3. calculate and analyze embeddings

4. try weak labelling techniques

5. train classification model with additional weak labels

6. compare results and conclude

Infrastructure

- sentiment classification model training is computationally expensive

- possible with Apple Silicon chips or Nvidia GPU with sufficient memory

- alternatively use CSCS