

TA-Bot

An application of NLP

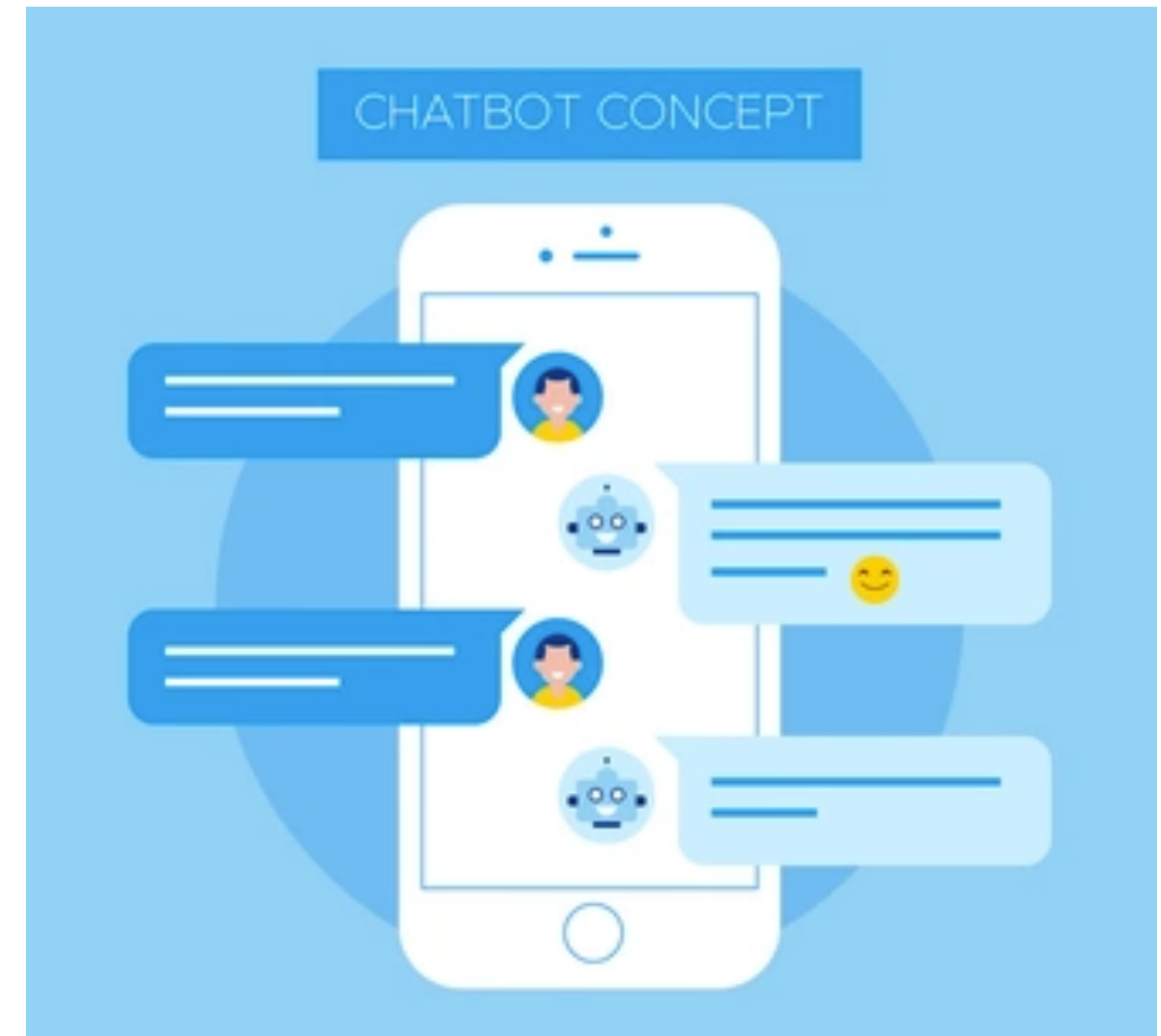
Olga, Surojit, Arjun and Harsh

Introduction

Chatbot as application of NLP

Goal: create a tool that automatically answers frequently asked questions posed by Univ.AI students, e.g.,

- What is the deadline for exercise submission?
- Will the sessions be recorded?
- Where to submit the homework?



Datasets

SQuAD2.0
The Stanford Question Answering Dataset

Squad version 2.0 contains ~100,000 questions

Train set

Dev Set

We are not using the 'unanswered' questions.

Univ.AI

Univ.ai Frequently Asked Questions

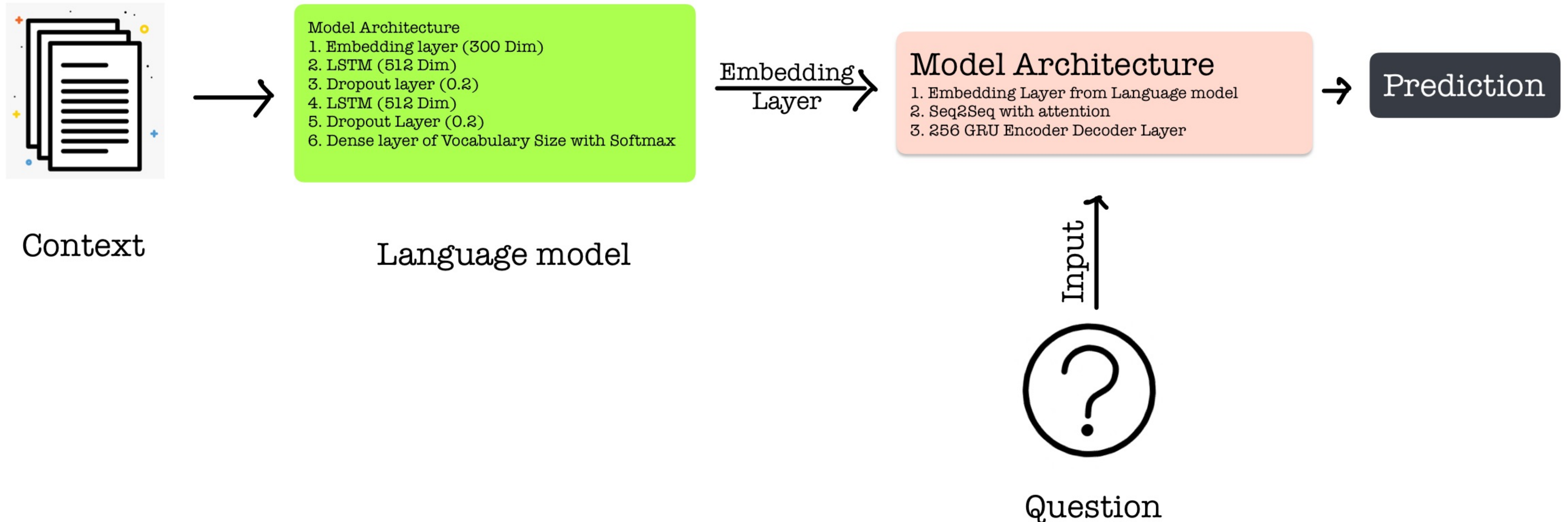
Faq.csv

Test.csv

Context (added manually)

Baseline Approach

RNN model with sentence piece word embedding



Results for baseline model

Squad Dataset results

5650
Which contestant reportedly received a death threat on the show?
The predicted answer is: automatic mps
The actual answer is: John Stevens
17587
What nation does the Marshall Islands vote similarly to in the General Assembly?
The predicted answer is: the north african states
The actual answer is: the United States
4
In which decade did Beyonce become famous?
The predicted answer is: the early 1970s
The actual answer is: late 1990s
4199
The placement of deciduous trees on the Equator facing side of a building can have a negative effect on solar availability in which season?
The predicted answer is: spring
The actual answer is: winter
10616
In what year did Bell get some of Edison's patents?
The predicted answer is: 2013
The actual answer is: 1879

Test Dataset results

Will the pre-class session be recorded?
The predicted answer is: mechanical
What is the deadline for quiz submission?
The predicted answer is: digital
What is the deadline for exercise submission?
The predicted answer is: shiva
How many hours do I need to complete this course?
The predicted answer is: time forda
Who will grade the exercise?
The predicted answer is: perfective heart failure
Why is the auto-grader failing me?
The predicted answer is: ?? frow triangle
Do I do the exercises individually?
The predicted answer is: the meaning of the expression of the entire world
Is the lab compulsory?
The predicted answer is: the ?? orfel ?? and the orogenna
Will the sessions be recorded?
The predicted answer is: computing
Can I have access to the recorded videos?
The predicted answer is: the crusader
Where are the recordings?
The predicted answer is: salvia ?? d ?? otia ?? and rabilia
Where can I ask questions regarding reading material
The predicted answer is: the lower level of the process
Where to find course material?
The predicted answer is: the teko and pious
Do we have homework?
The predicted answer is: dhot
Where can I find the homework?
The predicted answer is: standing lower elevations
Where to submit the homework?
The predicted answer is: covers ?? distance ?? orchow protection ?? or other countries

Transformer based Model

ALBERT Model

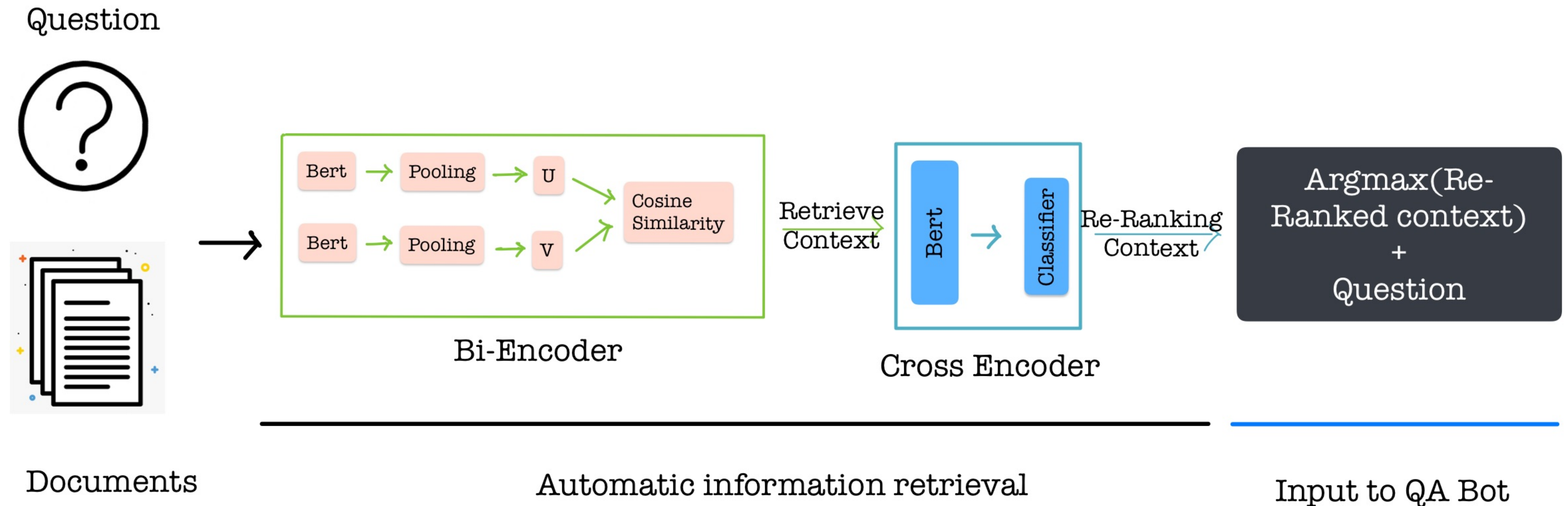
1. Using a pertained ALBERT Model from Hugging Face
2. Training the Model On Squad Dataset for 3 epochs.
3. Fine Tuning the model on Faq dataset with context handpicked for each question.

1	Question	Answer
2	will preclass session be recorded	preclass quiz is to just check your understanding of reading
3	what is deadline for quiz submission	5 pm on day of following lecture
4	what is deadline for exercise submission	5 pm on day of following lecture
5	how many hours do i need to complete this course	15 hours per week
6	who will grade exercise	exercises are autograded once you click submit button
7	why is autograder failing me	autograder does accept alternative solutions

Problem - Context is added manually, we want to solve this.

Application of Retrieve and ReRank

Automating context retrieval process



We are only using the cross encoder part because of low data

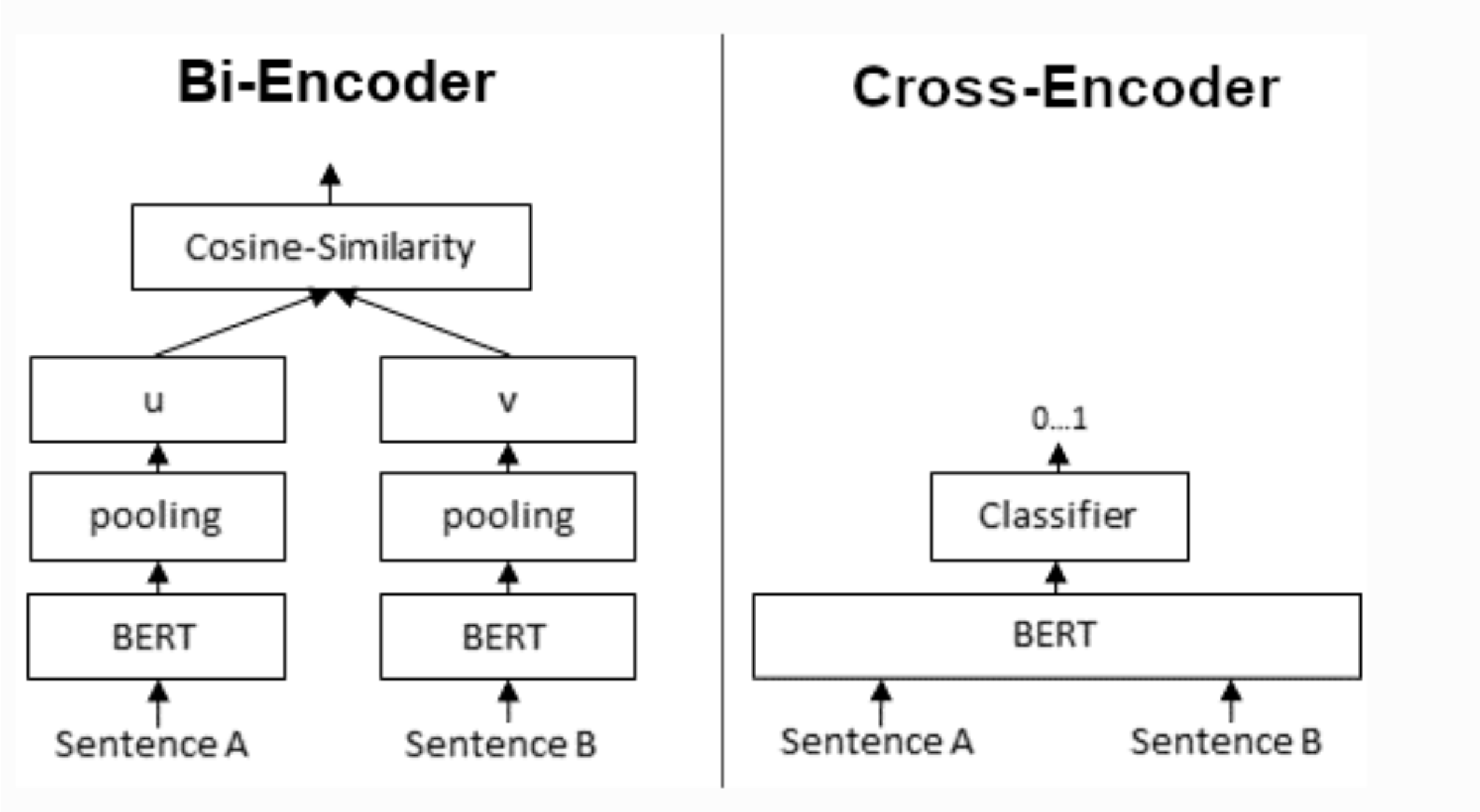
Cross Encoders

Bi-Encoder & Cross Encoder

Sentence Embeddings can be compared using Cosine similarity.

Fast for large number of documents

Used for ranking sentences



Sentences are passed simultaneously
No sentence embedding is produced.

Produces output value between 0-1
Indicating similarity between sentences

Slow for large number of documents

Used to Reordering Ranked Documents

Results for ALBERT

With Cross encoder and fine tuning on faq dataset

1	Question	Answer
2	will preclass session be recorded	course schedule may include readings in course textbooks
3	what is deadline for quiz submission	5 pm on day of following lecture
4	what is deadline for exercise submission	5 pm on day of following lecture
5	how many hours do i need to complete this course	our courses need time commitment of about 15 hours per week from enrolled students
6	who will grade exercise	exercises are autograded once you click submit button
7	why is autograder failing me	autograder does not accept alternative solutions

Short Comings of Albert model

1. Does not perform text generation
2. Points out the part of the context that the model thinks has the answer
3. Might not give good results in no answer situations

T-5

A text generative approach to chat-bot

1	Questions	Answers
2	answer_me: will the pre-class session be recorded?	<pad> no. all a a a aa a aa a a a aa a a a a aa a a a aa a
3	answer_me: what is the deadline for quiz submission?	<pad> 5 pm on thea 5 pm on the a a a homework.
4	answer_me: what is the deadline for exercise submission?	<pad> 5 pm on thea homework.
5	answer_me: how many hours do i need to complete this course?	<pad> about 15 hours per week.
6	answer_me: who will grade the exercise?	<pad> unfortunately, you and your a your a exercises will be announced during the course.
7	answer_me: why is the auto-grader failing me?	<pad> unfortunately, you not begrader is not begrading is not begrader is not begrading is a a a a a a a a a a a

1. The T-5 model is not performing well with the given dataset.
2. We think it requires more data for fine-tuning.

Conclusion

Our take away from the project

1. Our best model is Albert with Cross Encoder for information Retrieval.
2. If we have more data for fine-tuning than T-5 model may perform better

Future Work

List of improvements to be done

1. Scraping data from all Univ.ai pages.
2. Using NLP Data augmentation techniques.
3. Adversarial Training for a more robust model.

Index

1. Introduction
2. Dataset
3. Baseline Approach
4. Transformer Based Model - ALBERT
5. Cross Encoder
6. Application of Cross Encoder
7. Shortcomings of ALBERT
8. T5 - text generative model
9. Conclusion
10. Future work