

Question Answering For Assessment

Presented By:
Devaiah Ulliyada Arun
Ragith Ayyappan Kutty

INTRODUCTION

What? Answering question based on corpus.

Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Question

What causes precipitation to fall?

Answer Candidate



gravity







INTRODUCTION

What? Answering question based on corpus.

Why? Large amount of data on web. Interested in knowing answer rather than looking into documents.



what is capital of germany

X |  

 All  News  Maps  Images  Videos  More Settings Tools

About 78,50,00,000 results (0.86 seconds)

Germany / Capital



Map data ©2020 GeoBasis-DE/BKG (©2009), Google

Berlin

The capital of Germany is the city state of **Berlin**. It is the seat of the President of Germany, whose official residence is Schloss Bellevue. The Bundesrat ("federal council") is the representation of the Federal States (Bundesländer) of Germany and has its seat at the former Prussian Herrenhaus (House of Lords).

INTRODUCTION

What? Answering question based on corpus.

Why? Large amount of data on web. Interested in knowing answer rather than looking into documents.

How?

- Information Retrieval
- Reading Comprehension



Our Approach

Devaiah

- Query and answer generation: Classical NLP
 - Uses tokenization, summarization, query identification & NER.

Ragith

- Answer generation: Deep Learning
 - Failed Attempt: QA-LSTM
 - Successful attempt: BERT

Classical NLP for QA Generation

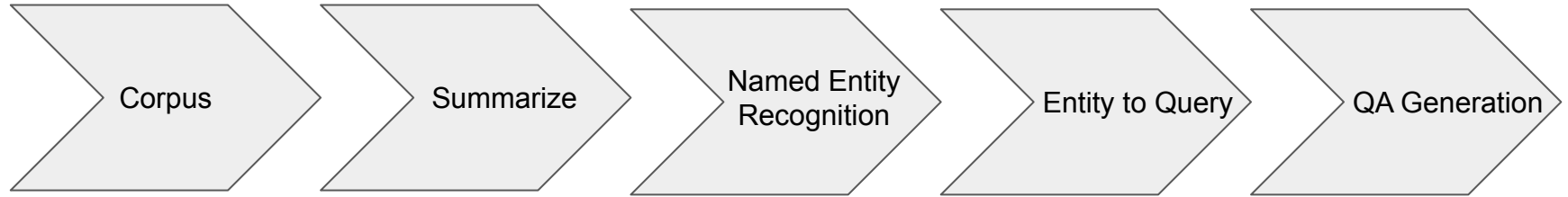


Figure : Pipeline involved for the classical NLP approach for QA Generation. Directed from left to right.

Summarization

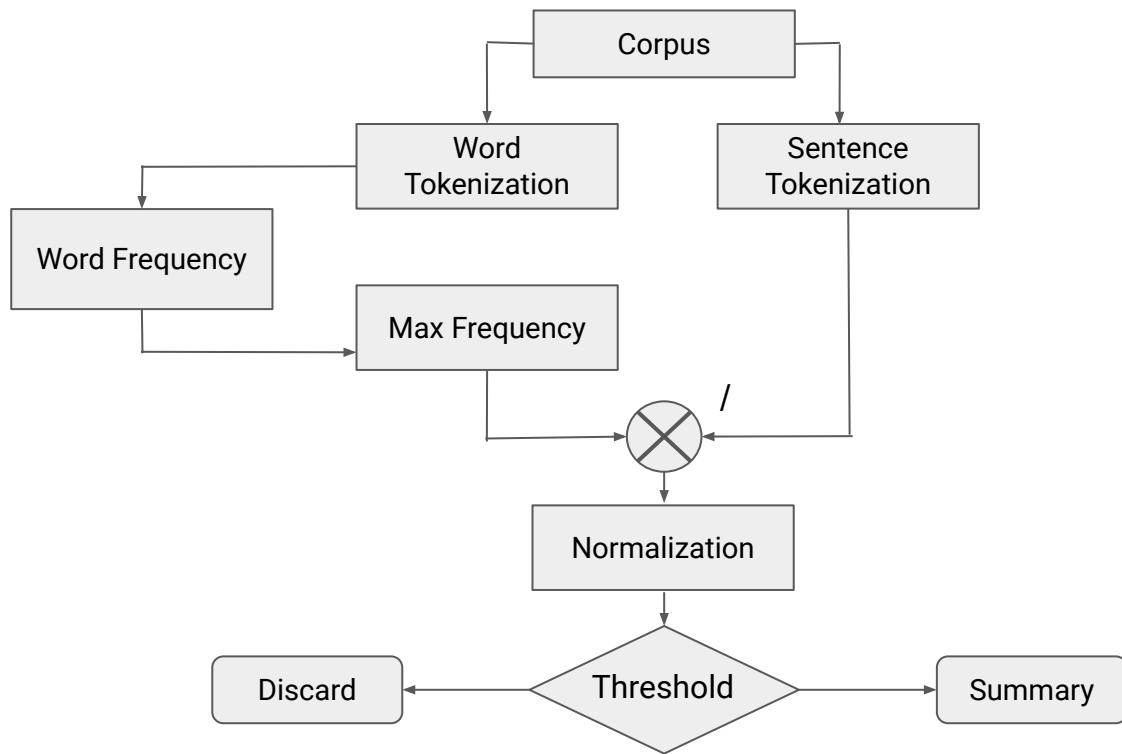


Figure : Flowchart describing the summarization algorithm using spaCy library.

Named Entity Recognition

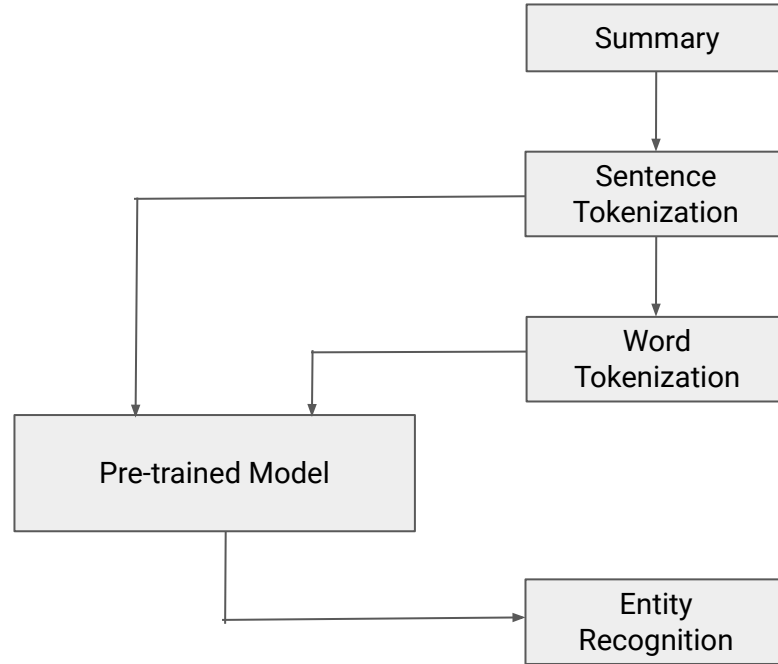


Figure : Flowchart describing the Named Entity Recognition algorithm using spaCy library.

Entity to Query

Entity	Description	Query Tag
PERSON	People, including fictional.	Who
NORP	Nationalities or religious or political groups.	What
FAC	Buildings, airports, highways, bridges, etc.	Where
ORG	Companies, agencies, institutions, etc.	Where
GPE	Countries, cities, states.	Where
LOC	Non-GPE locations, mountain ranges, bodies of water.	Where
PRODUCT	Objects, vehicles, foods, etc. (Not services.)	What
EVENT	Named hurricanes, battles, wars, sports events, etc.	Which
WORK_OF_ART	Titles of books, songs, etc.	What
LAW	Named documents made into laws.	Which
LANGUAGE	Any named language.	Which
DATE	Absolute or relative dates or periods.	When
TIME	Times smaller than a day.	When
PERCENT	Percentage, including "%" .	What percent
MONEY	Monetary values, including unit.	How much
QUANTITY	Measurements, as of weight or distance.	How much
ORDINAL	"first", "second", etc.	Which
CARDINAL	Numerals that do not fall under another type.	How many

Output of the classical approach

[←](#) [→](#) [↺](#) [🏠](#) [localhost:5000](#) [🔍](#) [☆](#) [📌](#) [🔴](#) [🔴](#) [🔵](#) [⚙️](#) [☰](#) [👤](#) [⋮](#)

[About Project](#) [👤 Team](#) [🐙](#)

Question Answer Generation

The Federal city of Bonn is a city on the banks of the Rhine in the German state of North Rhine-Westphalia, with a population of over 300,000. About 24 km south-southeast of Cologne, Bonn is in the southernmost part of the Rhine-Ruhr region, Germany's largest metropolitan area, with over 11 million inhabitants. It is famously known as the birthplace of Ludwig Van Beethoven in 1770. He spent his childhood and teenage years in Bonn. Founded in the 1st century BC as a Roman settlement, Bonn is one of Germany's oldest cities.

Submit

Output of the classical approach

Question : The Federal city of Bonn is a city on the banks of the Rhine in the German state of where, with a population of over 300,000?

Answer : North Rhine-Westphalia

Question : The Federal city of Bonn is a city on the banks of the Rhine in the German state of North Rhine-Westphalia, with a population of how many?

Answer : over 300,000

Question : , , where is in the southernmost part of the Rhine-Ruhr region, Germany's largest metropolitan area, with over 11 million inhabitants?

Answer : Bonn

Question : , , Bonn is in the southernmost part of the where region, Germany's largest metropolitan area, with over 11 million inhabitants?

Answer : Rhine-Ruhr

Question : , , Bonn is in the southernmost part of the Rhine-Ruhr region, where's largest metropolitan area, with over 11 million inhabitants?

Answer : Germany

Question : , , Bonn is in the southernmost part of the Rhine-Ruhr region, Germany's largest metropolitan area, with how many inhabitants?

Answer : over 11 million

Question : , Founded in when BC as a Roman settlement, Bonn is one of Germany's oldest cities?

Answer : the 1st century

Question : , Founded in the 1st century BC as a where settlement, Bonn is one of Germany's oldest cities?

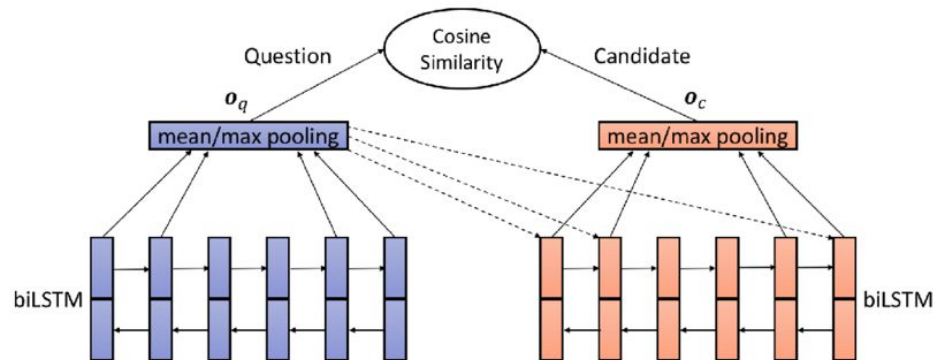
Answer : Roman

Takeaways

- Understanding of NER and Summarization.
- Understanding structure of complex sentences.
- Why semantic recognition is important?
- Understanding Coreference resolution.

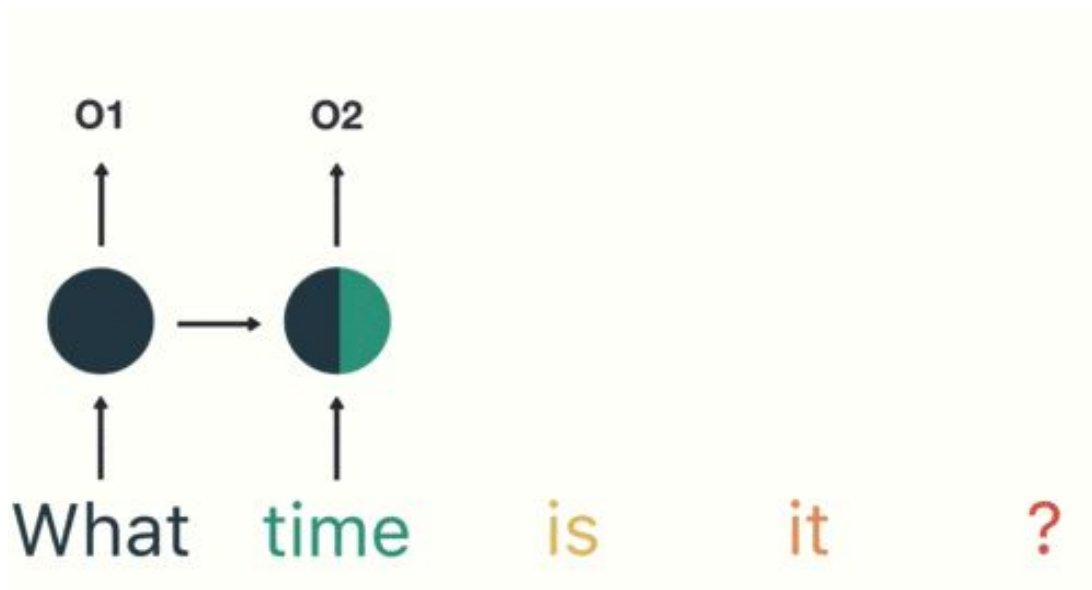
Failed Attempt: QA-LSTM

- BiLSTM: Uses contextual information from future tokens.
- Generates representation for question and answer separately.
- Loss is defined as
 - $L = \max\{0, M - \text{cosine}(q, a+) + \text{cosine}(q, a-)\}$



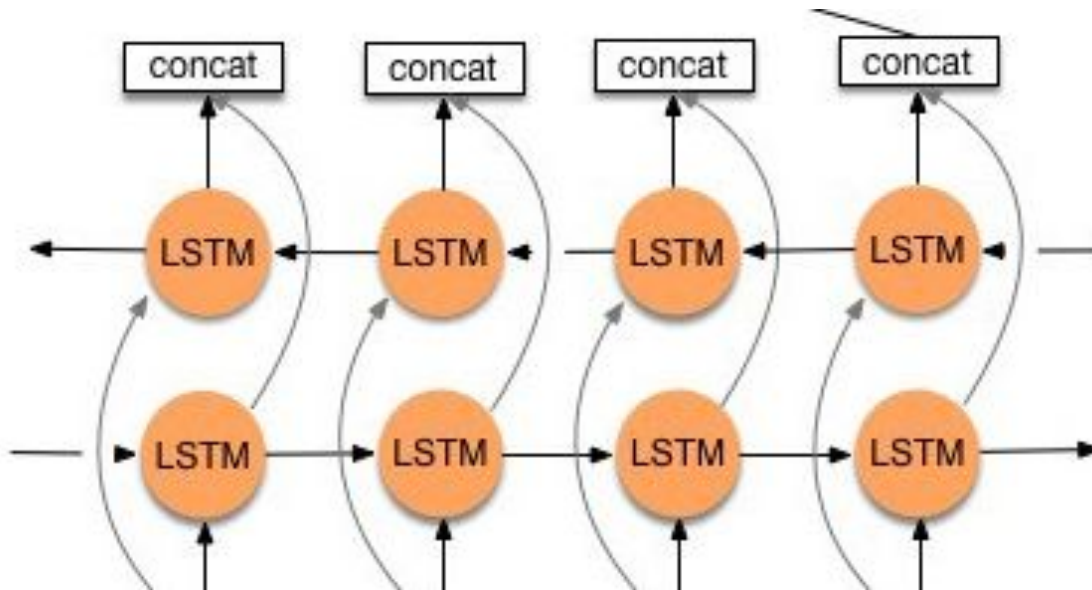
Why QA-LSTM failed?

- Slow to train- requires more training epochs to learn efficiently



Why QA-LSTM failed?

- Not truly Bidirectional



BERT(Bidirectional Encoder Representations from [Transformers](#))

An attention mechanism that learns contextual relations between words (or sub-words) in a text.

BERT(Bidirectional **Encoder** Representations from Transformers)

An attention mechanism that learns contextual relations between words (or sub-words) in a text.

Vanilla Transformer includes two separate mechanisms — an encoder that reads the text input and a decoder that produces a prediction for the task.

BERT(Bidirectional Encoder Representations from Transformers)

An attention mechanism that learns contextual relations between words (or sub-words) in a text.

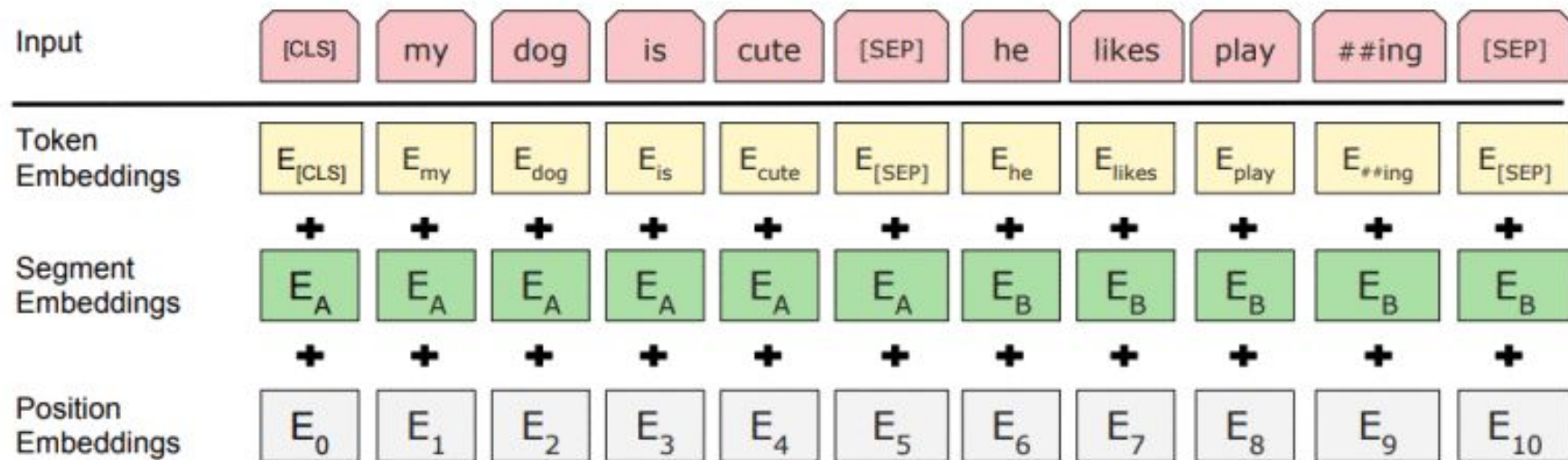
Vanilla Transformer includes two separate mechanisms — an encoder that reads the text input and a decoder that produces a prediction for the task.

Two
Step
Process

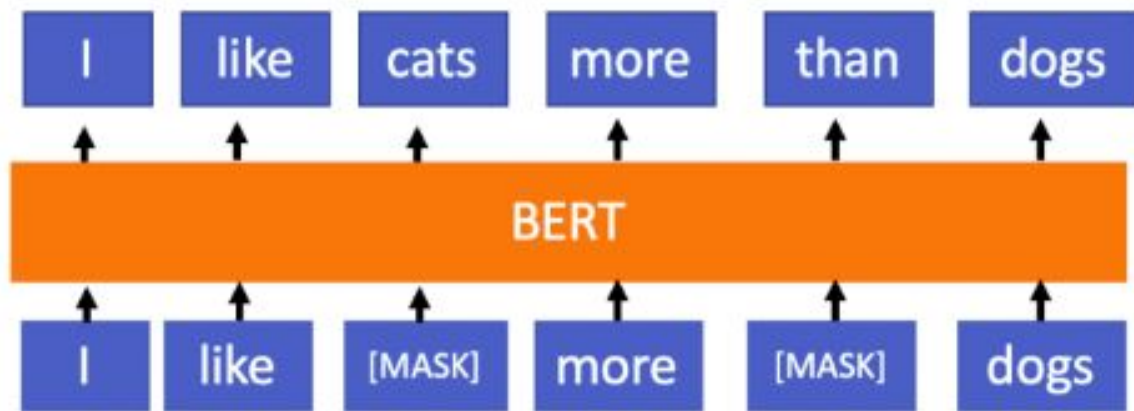
**Encoder:
Learn Language
Model**

Fine tune model

Encoder:Input

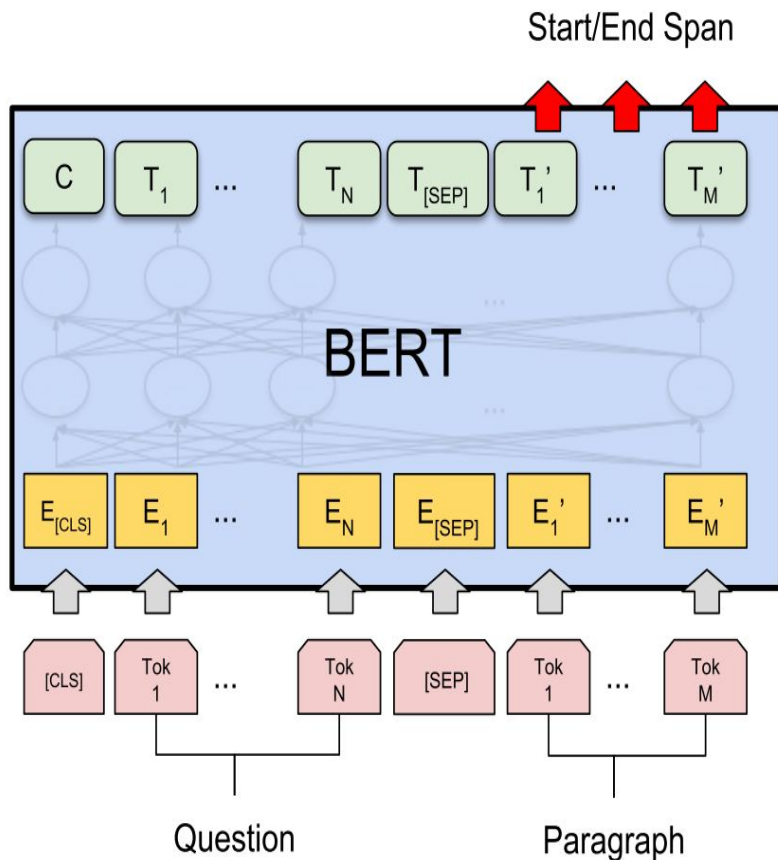


Encoder: Masked Language Model



Randomly mask 15% of token

Fine Tune Model



Article: The Legend of Zelda: Twilight Princess

Paragraph: The Legend of Zelda: Twilight Princess is an action-adventure game developed and published by Nintendo for the **GameCube and Wii** home video game consoles. ... The Wii version was released alongside the console in North America in **November 2006**, and in Japan, Europe, and Australia the following month. The GameCube version was released worldwide in December 2006.

Question 1: What consoles can be used to play Twilight Princess?

Answer: **GameCube and Wii**

Output of Question Generated by Classical Method

Question: The Federal city of Bonn is a city on the banks of the Rhine in the German state of where, with a population of over 300,000?
Answer: North Rhine-Westphalia

Question: The Federal city of Bonn is a city on the banks of the Rhine in the German state of North Rhine-Westphalia, with a population of how many?
Answer: 300,000

Question: where is in the southernmost part of the Rhine-Ruhr region, Germany's largest metropolitan area, with over 11 million inhabitants?
Answer: Bonn

Question: Bonn is in the southernmost part of the where region, Germany's largest metropolitan area, with over 11 million inhabitants?
Answer: Rhine-Ruhr

Question: Bonn is in the southernmost part of the Rhine-Ruhr region, where's largest metropolitan area, with over 11 million inhabitants?
Answer: 300,000

Question: Bonn is in the southernmost part of the Rhine-Ruhr region, Germany's largest metropolitan area, with how many inhabitants?
Answer: 11 million

Question: Founded in when BC as a Roman settlement, Bonn is one of Germany's oldest cities?
Answer: the 1st century BC

Question: Founded in the 1st century BC as a where settlement, Bonn is one of Germany's oldest cities?
Answer: Roman settlement, Bonn is one of Germany's oldest cities..

Question: Founded in the 1st century BC as a Roman settlement, where is one of Germany's oldest cities?
Answer: .

Output : Correct Answers

Question	Answer	Ground Truth
What town was actually granted to the Huguenots on arrival?	James River, at the abandoned Monacan village known as Manakin Town , now in Powhatan County. Some settlers landed in present-day Chesterfield County. On 12 May 1705, the Virginia General Assembly passed an act to naturalise the 148 Huguenots still resident at Manakintown. Of the original 390	Manakin Town
Who was the famous Huguenot theologian and writer in London?	Andrew Lortie (born André Lortie), a leading Huguenot theologian and writer who led the exiled community in London,	Andrew Lortie
When were these settlers naturalized as English colonists?	12 May 1705,	12 May 1705
What leader led the Dutch Revolt and wrote Apologie?	William the Silent , leader of the Dutch (Calvinist) revolt	William the Silent

Output : Incorrect Answers

Question	Answer	Ground Truth
What war in Ireland featured Huguenot regiments?	William of Orange	Williamite war
How many settlers original settled in Manakintown?	hundred	390
Who is the patron saint of the Huguenots?	Netherlands was Pierre Bayle. He started teaching in Rotterdam,	Saint Nicolas
In what field were double and triple expansion engines common?	19th	shipping

Output : Impossible Answers

Question	Answer	Ground Truth
In what year had Manakin Town first been abandoned?	1700	Plausible Answer: 1700
In what year did William the Silent issue his "Apologie"?	(1568–1609).	Plausible Answer: 1568
What nationality was Pierre L'Oyseleur?	William the Silent, leader of the Dutch (Calvinist) revolt	Plausible Answer: Dutch
What percentage of electrical power in the United States is made by generators?	90%	Plausible Answer: 90

Takeaways

- Need of a QA system: google search, text mining of documents.
- How and why transformers perform better than LSTM?
- How to run & finetune Simple Transformer models to do various NLP tasks.

Thank You