TREC Deep Learning Track

Nick Craswell (Microsoft), Bhaskar Mitra (Microsoft and UCL), Emine Yilmaz (UCL), Daniel Campos (Microsoft and UW)
Deep Learning Track

• Deep models learn an intricate representation of a large-scale dataset
• Here: Adhoc ranking dataset with full Q+D text and Q-D judgments
  • Task 1: Given Q and top1k D, rerank D for relevance
  • Task 2: Given Q and corpus of D, retrieve/rank D for relevance
• What’s new: Large set of judged training queries Q and Q-D
• Research questions:
  • Does the large training set help? Beat strong baselines?
  • Do tasks 1 and 2 require different technology?
  • Will we find elegant/fast/effective solutions to task 2 (end-to-end retrieval)?
  • Task 1 also has a non-TREC leaderboard, will results agree with TREC (judged) eval?
  • Do metrics on sparsely judged many queries approximate deeply judged few queries? Which is better?
Dataset: MS-MARCO

http://www.msmarco.org/

<table>
<thead>
<tr>
<th>Ranking dataset</th>
<th>Count</th>
<th>File size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus</td>
<td>8,841,823 passages</td>
<td>2.9G</td>
</tr>
<tr>
<td>Training queries</td>
<td>808,731 queries</td>
<td>33.5M</td>
</tr>
<tr>
<td>Training qrels</td>
<td>532,761 judgments on 502,939 of training q</td>
<td>10M</td>
</tr>
</tbody>
</table>

**MS-MARCO Dataset Timeline**

- **December 2016**: MS-MARCO QnA dataset release
- **April 2018**: Reaches 1 million queries
- **July 2018**: Define ranking dataset. Internal challenge in MS. Top three systems: Deep, LTR, handcrafted
- **October 2018**: Ranking dataset released at [http://www.msmarco.org](http://www.msmarco.org)
- **November 2018**: Establish leaderboard for (re)ranking challenge
- **Next**: Run TREC 2019 ranking challenge, with NIST evaluation
- **Next**: Keyword Extraction Dataset/Competition
Creating the MS-MARCO data

All Bing queries en-US 100%

Anonymized/rewritten 100%

Filtered: QnA queries in English (non-Adult) ~10%

Scrape top-10 passages for a sample of queries

Judge: Answer available? Write answer. is_selected
503k training queries

Further filter/rewrite answer [182k human answers]
MS-MARCO

- Starting Sept 2016
- Queries Overall streaming
  - Privacy and Legal filter and/or anonymization and/or rewriting
- Filter
  - Isadult
  - Foreign
  - Classifier ml (based on judges MS-MARCO if the query was answerable)
  - Other classifiers Xia/Tang
- Scrape Bing for top-10 passages
- Judge
  - Answer available? No.
  - Answer available can’t summarize
  - Answer available can summarize
  - Can’t judge / junk
  - Write answer
  - Select passage (is_selected)
  - “16”
  - 1,030,000
- If too much like copy+paste, rewrite for smart speaker
  - “there are 16 teaspoons in a cup”
  - 182,000 answers

- V2.0 two datasets
  - 900k Answer(aka QA)
  - 130k Wellformed answer (Aka NLGEN)
- V2.1 merged (April 23)
  - Wellformed split evenly and not always available
  - 800k Train
  - 100k dev
  - 100k eval
- Passage Reranking Task(October 26th)
  - 8.8 m passages
  - Given Top 1000 from BM25, rerank by relevance.
  - 651,221 relevance labels
- KeyWord Extraction Dataset(ETA March 2019)
MS-MARCO ranking dataset

- Pool the passages 8.8 million
- 1 million Queries
  - 550k having an answer and is_selected
  - 1.4 million positive in total
  - BM25 top1000
- Training file with triples
  - train.triples (272gb(full) and 27gb(small) QID, Positive PID, Negative PID.
- Top1000 files
  - Top1000 passages from BM25
  - Test set top-1000 return top-10 6.8k queries

- 3.2 million docs clean body
  - Based on the 3.5 million urls the passages come from.
  - Use as metadata, train embeddings, etc.
Results So Far

• MSFT Hackathon
  • 21 teams
  • Top3 got close scores
    • Deep (.27)
    • Handcrafted (.24)
    • Traditional ML (.22)

• External ->
  • 7 results
grep -w queries.train.tsv fall | egrep -w rome

115577  date of the fall of rome
214518  how did rome fall apart
394523  in what year did the fall of ancient rome occur
584175  what caused the fall of roman empire
920193  what was the fall of the roman empire
924037  what were the factors that led to the fall of the roman empire
941441  when did the eastern roman empire fall
1009300  which event usually marks fall of western roman empire
1063096  why did roman empire fall
1140059  which diseases affected the fall of the roman empire

155577  fall of rome
214518  fall of the roman empire
394523  the fall of rome
584175  the fall of the roman empire
920193  fall of roman empire
924037  when did the roman empire fall
941441  why did the roman empire fall
1009300  rise and fall of the roman empire
1063096  when did rome fall

why did the roman empire in the east survive the fall of rome
www.ushistory.org/civ/6f.asp.
1952 the decline and fall of the roman empire book
what lead to the fall of rome
roman fall
fall airline fares 2018 rome
books about the fall of the roman empire
what led rome to fall quizlet diocletian reforms
the roman republic to the fall of the roman empire timeline
what were the effects of the fall of rome
gibbons rise and fall of the roman empire

ancient rome: the rise and fall of an empire
fall of the roman empire
fall of the roman republic
how did the western roman empire fall
the decline and fall of the roman empire pdf
the fall of rome
the fall of the roman empire
the fall of the roman empire film
what were the effects of the fall of rome
which of the following correctly describes prince odoacer's role in the fall of the western roman empire? prince odoacer forced augustulus to hand over his crown, and germanic kingdoms claimed the land. prince odoacer is credited with overtaking western rome and creating a thriving economy and culture. prince odoacer served under emperor augustulus during the collapse of the western roman empire. prince odoacer was e
That being said, the Fall of Rome is normally either dated to the Sack of Rome by Alaric the Goth on August 24th A.D. 410, or the resignation of the last Roman Emperor of the Western Empire, Romulus Augustus, on September 4th, A.D. 476.

For the film, see The Fall of the Roman Empire (film). The Fall of the Western Roman Empire (also called Fall of the Roman Empire or Fall of Rome) was the process of decline in the Western Roman Empire in which it failed to enforce its rule, and its vast territory was divided into several successor polities.

History >> Ancient Rome. Rome ruled much of Europe around the Mediterranean for over 1000 years. However, the inner workings of the Roman Empire began to decline starting around 200 AD. By 400 AD Rome was struggling under the weight of its giant empire. The city of Rome finally fell in 476 AD. Rome reached its peak of power in the 2nd century around the year 117 AD under the rule of the great Roman emperor Trajan.

The Fall of the Western Roman Empire in which it failed to enforce its rule, and its vast territory was divided into several successor polities.

394523 in what year did the fall of ancient rome occur

193179 History >> Ancient Rome. Rome ruled much of Europe around the Mediterranean for over 1000 years. However, the inner workings of the Roman Empire began to decline starting around 200 AD. By 400 AD Rome was struggling under the weight of its giant empire. The city of Rome finally fell in 476 AD. Rome reached its peak of power in the 2nd century around the year 117 AD under the rule of the great Roman emperor Trajan.

920193 what was the fall of the roman empire

3424901 Causes for the Fall of the Roman Empire-Antagonism between the Senate and the Emperor One of the main causes for the Fall of the Roman Empire was the Antagonism between the Senate and the Emperor. One of the main causes for the Fall of the Roman Empire was the Slave Labor. The number of slaves increased dramatically during the first 2 centuries of the Roman Empire.

941441 when did the eastern roman empire fall

3137531 Its capital city was Constantinople (or New Rome). It represented an administrative division of the Roman Empire, but after the fall of the western part it survived until the fall of Constantinople in 1453. The Eastern Empire was also called the Byzantium Empire. The term Byzantine Empire dates from the 17th century.

1009300 which event usually marks fall of western roman empire

4145016 The event that most commonly marks the end of the Western Roman Empire is the overthrow of the Roman Emperor Romulus Augustus by the German Foederati.

1063096 why did roman empire fall

2563452 The Western Empire was ruled by Rome, the Eastern Empire was ruled by Constantinople. The fall of Rome discussed here is referring to the Western Roman Empire which was ruled by Rome. The Eastern Roman Empire became known as the Byzantium Empire and remained in power for another 1000 years. The City of Rome is Sacked. The city of Rome was thought by many to be unconquerable. However, in 410 AD, a Germanic barbarian tribe called the Visigoths invaded the city. The Eastern Roman Empire, or Byzantium, fell in 1453 to the Ottoman Empire. 2 Many poor people were glad to see Rome fall. 3 They were starving to death while being taxed heavily by Rome. 4 Near the end of the Roman Empire, the city of Rome was no longer the capital.

No Judgment:

584175 what caused the fall of roman empire

924037 what were the factors that led to the fall of the roman empire

1140059 which diseases affected the fall of the roman empire
How many languages does the president speak?

How many kids does Trump have?

How many educated men voted for Trump?

How many days has Trump taken off?

How many counties won by Trump?

How many bankruptcy cases has Trump filed?

How long has Trump been in office?

How did Trump win the election?

Has Trump ever had surgery?

Does Trump own all the Trump buildings?

Does Stan Smith support Trump?

Does Mitch McConnell support Trump?

Definition of the word Trump

How much does it cost to stay at Trump Hotel?

How tall is Trump Tower?

How popular is Trump today?

President Trump's health

How much does it cost for security for Trump?

How many times has Trump gone bankrupt?

How many people has Trump fired in his presidency?

How many of Trump's staffers have left since he took office?

Is Trump leaving the White House?

Is Trump getting impeached?

Is the First Lady on vacation with Trump?

Is Trump doing a good job?

How to show support for Trump?

How tall is Trump?

How popular is Trump today?

Is Trump's latest tweets?

Where was Trump born?

What type of plastic surgery has Trump had?

How much does Trump weigh?

What do the psychics say about Trump?

What is the narrative of the story of Trump?

What was the percentage of young white voter for Trump?

What is Trump on social media today?

What is the Trump Card?

What is Trump's psychology?

What is Trump's latest tweets?

What is Trump's religion?

When did Trump marry?

What is President Trump's age?

Is Trump like?

What is the nationality of Trump?

What is the象征 of Trump?

What is Trump's latest tweets?

What is the象征 of Trump?

What is Trump's religion?

When and where was Donald Trump born?

What’s the feud between Trump and Rosie?

What’s President Trump’s approval rating?

What are his next steps?

What type of helicopter does Trump own?

What type is Trump’s private jet?

What does President Trump say about climate warming?

What do women think of Trump?

What is a Trump-like?

What is Trump’s big announcement?

What is the Trump Card?

What is the象征 of Trump?

What is the likelihood Trump gets impeached?

What is the first name of Trump?

What is President Trump’s salary as president?

What is the likelihood Trump gets impeached?

What is the likelihood Trump gets impeached?

What is the象征 of Trump?

What is the象征 of Trump?

What is the likelihood Trump gets impeached?

What is the象征 of Trump?

What is the象征 of Trump?

What is the象征 of Trump?

What is the象征 of Trump?

What is the likelihood Trump gets impeached?

What is the likelihood Trump gets impeached?

What is the likelihood Trump gets impeached?

What is the likelihood Trump gets impeached?
Evaluation

• 6800 test queries (held out)
  • Reranking task, you are given top1k to rerank
  • Submit top 1000 ranking for each query

• Evaluation 1: Use MS-MARCO’s is_selected label
  • Metric: Mean reciprocal rank of is_selected passage

• Evaluation 2: Use TREC pooled evaluation on ~50 queries
  • 3-level judging: 0) Not an answer, 1) Has an answer, 2) Nice readable answer
  • Metrics: AP, NDCG, RBP, ...

• Evaluation question: Do they agree? Which is better?
Possible directions for future years

- Find more uses of the 1 million MS-MARCO queries
- Document ranking rather than passage ranking
- More ways of getting training labels
- Tasks beyond adhoc ranking
- New Queries/New Community Proposed Questions.