Abstractive Summarization of Social Media Posts
A case study using deep learning

Master Thesis
(Shahbaz Syed)

Bauhaus-Universität Weimar

Examiners:
Prof. Dr. Benno Stein
Prof. Dr. Andreas Jakoby

Supervisors:
Michael Voelske
Dr. Martin Potthast
Motivation
Motivation

Everyday

<table>
<thead>
<tr>
<th>Source</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>4 Billion messages</td>
</tr>
<tr>
<td>Twitter</td>
<td>500 Million tweets</td>
</tr>
<tr>
<td>Email</td>
<td>200 Billion emails</td>
</tr>
<tr>
<td>Blogs</td>
<td>3.4 Million blogs</td>
</tr>
</tbody>
</table>
Motivation

**Extractive vs Abstractive**

The Army Corps of Engineers, rushing to meet President Bush’s promise to protect New Orleans by the start of the 2006 hurricane season, installed defective flood-control pumps last year despite warnings from its own expert that the equipment would fail during a storm, according to documents obtained by The Associated Press.
Motivation

Extractive vs Abstractive

The Army Corps of Engineers, rushing to meet President Bush’s promise to protect New Orleans by the start of the 2006 hurricane season, installed defective flood-control pumps last year despite warnings from its own expert that the equipment would fail during a storm, according to documents obtained by The Associated Press.

Extractive: “Army Corps of Engineers,” “President Bush,” “New Orleans,” and “defective flood-control pumps”
Motivation

Extractive vs Abstractive

The Army Corps of Engineers, rushing to meet President Bush’s promise to protect New Orleans by the start of the 2006 hurricane season, installed defective flood-control pumps last year despite warnings from its own expert that the equipment would fail during a storm, according to documents obtained by The Associated Press.

Extractive: “Army Corps of Engineers,” “President Bush,” “New Orleans,” and “defective flood-control pumps”

Abstractive: “political negligence,” “inadequate protection from floods”
Motivation

Why abstractive using deep learning?
Motivation

Why abstractive using deep learning?

- Automatic feature learning
Motivation

Why abstractive using deep learning?

- Automatic feature learning
- Large amounts of data available compliments deep learning models
Motivation

Why abstractive using deep learning?

● Automatic feature learning
● Large amounts of data available compliments deep learning models
● Low dependency on domain of the texts
Motivation

Why abstractive using deep learning?

- Automatic feature learning
- Large amounts of data available compliments deep learning models
- Low dependency on domain of the texts
- Abstractive summaries are more appealing than text extracts
Motivation

Challenges
Motivation

Challenges

- Deep learning models need clean datasets
Motivation

Challenges

- Deep learning models need clean datasets
- Lack of variety in available datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Size</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gigaword</td>
<td>4 Million</td>
<td>News</td>
</tr>
<tr>
<td>CNN / DailyMail</td>
<td>110,000</td>
<td>News</td>
</tr>
</tbody>
</table>
Motivation

Challenges

- Deep learning models need clean datasets
- Lack of variety in available datasets
- Current evaluation measure (ROUGE) may not be suitable for abstractive summaries

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Size</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gigaword</td>
<td>4 Million</td>
<td>News</td>
</tr>
<tr>
<td>CNN / DailyMail</td>
<td>110,000</td>
<td>News</td>
</tr>
</tbody>
</table>
Contributions
Contributions

- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
Contributions

- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
- Exploring the effectiveness and limitations of a *specific model* for abstractive summarization
Contributions

- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
- Exploring the effectiveness and limitations of a *specific model* for abstractive summarization
- Advocating a need for a better evaluation method than ROUGE for judging abstractive summaries
Reddit
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

Submitted 1 month ago by IFIHADANICKLE4USALL
1000 comments share report

[-] Lostsonofpluto 129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments  share  report

Lostsonofpluto  129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink  embed  report

load more comments (2 replies)
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

Submitted 1 month ago by IfiHadANickel4UsAll
1000 comments share report

Lostsonofpluto 129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink embed report

load more comments (2 replies)
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments share report

[-] Lostsonofpluto 129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink embed report

load more comments (2 replies)
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments share report

[-] Lostsonofpluto 129 points 1 month ago
My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR
My dad believed his family didn’t have a pet monkey

load more comments (2 replies)
Webis-TLDR-17 Corpus
Corpus creation pipeline
Corpus creation pipeline

- Check for presence of tl.{0,3}dr
Corpus creation pipeline

- Check for presence of \texttt{tl\{0,3\}dr}

Example 1

so every student in the world seems to be moving today, i've just been roped into playing van-driver for another couple. but it did get me thinking; if any of you leeds (ish) peeps are desperate to last-minute move some stuff, you should probably send me a pm and we can try and move a van-load this afternoon.

\textbf{tl/dr}: need van this afternoon? pm me a phone number i can sms.

Example 2

kinda drunk!
so i did lie this girl who and i drank a little to much tonight and i kinda told jher i like her and asked her if she felt the same way. she said that is looking for a father figure (she has 2 kids) i said i compeletly understand that and now i move you from the sex list to the awesome fire nd list. and she said that she was happy wityh how i said that. i hope i didnt screw theis freind ship up casue she really is a n awesome person.

\textbf{tl/dr}:uijksd
areyouhumasn
is hard to spoelll wihen you are drubnk
Corpus creation pipeline

- Check for presence of **tl.{0,3}dr**
- **Remove** posts by bots
Corpus creation pipeline

- Check for presence of \texttt{tl.\{0,3\}dr}
- Remove posts by bots
- Remove posts with multiple occurrences of \texttt{tl.\{0,3\}dr}
Corpus creation pipeline

- Check for presence of \texttt{tl.\{0,3\}dr}
- \textbf{Remove} posts by bots
- \textbf{Remove} posts with multiple occurrences of \texttt{tl.\{0,3\}dr}
- \textbf{Preserve} posts containing relevant \texttt{tl.\{0,3\}dr} patterns
Corpus creation pipeline

- Check for presence of `tl.{0,3}dr`
- **Remove** posts by bots
- **Remove** posts with multiple occurrences of `tl.{0,3}dr`
- **Preserve** posts containing relevant `tl.{0,3}dr` patterns
- **Remove** cases where content is shorter than the summary
Corpus creation pipeline

- Check for presence of `tl.{0,3}dr`
- **Remove** posts by bots
- **Remove** posts with multiple occurrences of `tl.{0,3}dr`
- **Preserve** posts containing relevant `tl.{0,3}dr` patterns
- **Remove** cases where content is shorter than the summary
- **Preserve** posts with minimum length of **content 2 words**, **summary 1 word**
Remove posts by bots

your submission has been automatically removed because it does not appear to contain a **tl;dr** or any sort of short summary.

please edit your post to add a (preferably **bolded**) **tl;dr**, ensure that it meets all the other requirements (includes the ages of the people involved and length of relationship) and then [message the moderators](http://www.reddit.com/message/compose?to=%2fr%2frelationships)&amp;subject=please+reapprove) to have it re-approved. if your post already contained a **tl;dr**, please let us know so that we can improve our systems.

*i am a bot, and this action was performed automatically. please [contact the moderators of this subreddit](http://www.reddit.com/message/compose?to=%23relationships) if you have any questions or concerns.*
Remove posts with multiple tl;dr occurrences

I would expect anyone likely to post "tl;dr" wouldn't subscribe to the subreddit, as it would be made clear that long-form content was contained within. It's too bad there isn't some way to make it so that Reddit gives a user an electric shock when they type tl;dr really. Such people should be very ashamed of themselves, bragging about being too lazy (practically and intellectually) to read a detailed comment. I can understand if you don't have the time or motivation to put forth the effort... but to post a comment pointing it out? That's an act of sheer pride in ignorance.
Remove posts with multiple tl;dr occurrences

I would expect anyone likely to post "tl;dr" wouldn't subscribe to the subreddit, as it would be made clear that long-form content was contained within. It's too bad there isn't some way to make it so that Reddit gives a user an electric shock when they type tl;dr really. Such people should be very ashamed of themselves, bragging about being too lazy (practically and intellectually) to read a detailed comment. I can understand if you don't have the time or motivation to put forth the effort... but to post a comment pointing it out? That's an act of sheer pride in ignorance.
Preserve relevant **tl.{0,3}dr** patterns

<table>
<thead>
<tr>
<th>tldr_found</th>
<th>count(tldr_found)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tl;dr</td>
<td>984131</td>
</tr>
<tr>
<td>tldr</td>
<td>285607</td>
</tr>
<tr>
<td>tl:dr</td>
<td>75734</td>
</tr>
<tr>
<td>tly dr</td>
<td>25307</td>
</tr>
<tr>
<td>tl/dr</td>
<td>15893</td>
</tr>
<tr>
<td>tle dr</td>
<td>15432</td>
</tr>
<tr>
<td>tl; dr</td>
<td>10934</td>
</tr>
<tr>
<td>tl, dr</td>
<td>7851</td>
</tr>
<tr>
<td>tl,dr</td>
<td>3943</td>
</tr>
<tr>
<td>tl dr</td>
<td>1339</td>
</tr>
<tr>
<td>tl'dr</td>
<td>1037</td>
</tr>
<tr>
<td>tl-dr</td>
<td>910</td>
</tr>
<tr>
<td>tly, dr</td>
<td>832</td>
</tr>
<tr>
<td>tlsdr</td>
<td>553</td>
</tr>
<tr>
<td>tl-sdr</td>
<td>526</td>
</tr>
<tr>
<td>tle, dr</td>
<td>510</td>
</tr>
</tbody>
</table>
Preserve relevant tl.\{0,3\}dr patterns

<table>
<thead>
<tr>
<th>tldr_found</th>
<th>count(tldr_found)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tl;dr</td>
<td>984131</td>
</tr>
<tr>
<td>tldr</td>
<td>285607</td>
</tr>
<tr>
<td>tl:dr</td>
<td>75734</td>
</tr>
<tr>
<td>tly dr</td>
<td>25307</td>
</tr>
<tr>
<td>tl/dr</td>
<td>15893</td>
</tr>
<tr>
<td>tle dr</td>
<td>15432</td>
</tr>
<tr>
<td>tl; dr</td>
<td>10934</td>
</tr>
<tr>
<td>tl,dr</td>
<td>7851</td>
</tr>
<tr>
<td>tl dr</td>
<td>3943</td>
</tr>
<tr>
<td>tl, dr</td>
<td>1339</td>
</tr>
<tr>
<td>tl'dr</td>
<td>1037</td>
</tr>
<tr>
<td>tl-dr</td>
<td>910</td>
</tr>
<tr>
<td>tly, dr</td>
<td>832</td>
</tr>
<tr>
<td>tlsdr</td>
<td>553</td>
</tr>
<tr>
<td>tl-sdr</td>
<td>526</td>
</tr>
<tr>
<td>tle, dr</td>
<td>510</td>
</tr>
</tbody>
</table>
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments share report

[-] Lostsonofpluto 129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink embed report

load more comments (2 replies)
What are the most ridiculous things you believed to be true as a child?  

submitted 1 month ago by IfiHadANickle4UsAll  
1000 comments  share  report

[-] Lostsonofpluto  129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink  embed  report

load more comments (2 replies)
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments share report

[-] Lostsonofpluto 129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey
What are the most ridiculous things you believed to be true as a child? (self.AskReddit)

submitted 1 month ago by IfIHadANickle4UsAll
1000 comments  share  report

[–] Lostsonofpluto  129 points 1 month ago

My dad spent the better part of his childhood believing his early memories of his family having a pet monkey were just something his kid brain made up. When he asked his parents it turned out they actually did have one and that they gave it to a local zoo before his sister was born.

TL;DR

My dad believed his family didn’t have a pet monkey

permalink  embed  report

load more comments  (2 replies)
Precision of 96% was achieved for 2000 examples manually evaluated.
Filtering process - Base corpus

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>Submissions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw input</td>
<td>286,168,475 (286 Million)</td>
<td>1,659,361,727 (1.6 Billion)</td>
</tr>
<tr>
<td>Contains a tl.{0,3}dr</td>
<td>2,079,035</td>
<td>3,745,320</td>
</tr>
<tr>
<td>Non-bot posts</td>
<td>1,969,413</td>
<td>3,351,597</td>
</tr>
<tr>
<td>Contains a valid tl;dr</td>
<td>1,829,030</td>
<td>2,987,135</td>
</tr>
<tr>
<td>Final pairs (C2, S1)</td>
<td>1,658,567 (~1.6 Million)</td>
<td>2,189,763 (~2.1 Million)</td>
</tr>
</tbody>
</table>
### Filtering process - Base corpus

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>Submissions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw input</td>
<td>286,168,475 (286 Million)</td>
<td>1,659,361,727 (1.6 Billion)</td>
</tr>
<tr>
<td>Contains a tl.{0,3}dr</td>
<td>2,079,035</td>
<td>3,745,320</td>
</tr>
<tr>
<td>Non-bot posts</td>
<td>1,969,413</td>
<td>3,351,597</td>
</tr>
<tr>
<td>Contains a valid tl;dr</td>
<td>1,829,030</td>
<td>2,987,135</td>
</tr>
<tr>
<td>Final pairs (C2, S1)</td>
<td>1,658,567 (~ 1.6 Million)</td>
<td>2,189,763 (~ 2.1 Million)</td>
</tr>
</tbody>
</table>
### Filtering process - Base corpus

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>Submissions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Raw input</strong></td>
<td>286,168,475 (286 Million)</td>
<td>1,659,361,727 (1.6 Billion)</td>
</tr>
<tr>
<td>Contains a tl.{0,3}dr</td>
<td>2,079,035</td>
<td>3,745,320</td>
</tr>
<tr>
<td>Non-bot posts</td>
<td>1,969,413</td>
<td>3,351,597</td>
</tr>
<tr>
<td>Contains a valid tl;dr</td>
<td>1,829,030</td>
<td>2,987,135</td>
</tr>
<tr>
<td><strong>Final pairs (C2, S1)</strong></td>
<td>1,658,567 (~ 1.6 Million)</td>
<td>2,189,763 (~ 2.1 Million)</td>
</tr>
</tbody>
</table>
# Filtering process - Base corpus

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>Submissions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw input</td>
<td>286,168,475 (286 Million)</td>
<td>1,659,361,727 (1.6 Billion)</td>
</tr>
<tr>
<td>Contains a tl.{0,3}dr</td>
<td>2,079,035</td>
<td>3,745,320</td>
</tr>
<tr>
<td>Non-bot posts</td>
<td>1,969,413</td>
<td>3,351,597</td>
</tr>
<tr>
<td>Contains a valid tl;dr</td>
<td>1,829,030</td>
<td>2,987,135</td>
</tr>
<tr>
<td><strong>Final pairs (C2, S1)</strong></td>
<td><strong>1,658,567 (~ 1.6 Million)</strong></td>
<td><strong>2,189,763 (~ 2.1 Million)</strong></td>
</tr>
</tbody>
</table>
Corpus verticals
One purpose - One tl;dr

- Questions
- Vulgar posts
- True summaries
- Topic + tl;dr
- Topic
Asking for help/ opinions

Content

ever since i watched one of the counselors at my camp play tron , i ' ve wanted to build it . the only problem is , i can ' t figure out which colors to build around . the group at my lgs that plays mainly modern is saying i should build r / g because it ' s the best , but i think the money that i would spend on a playset of [[ grove of the burnwillows ]] could be spent elsewhere and have a much bigger impact on the deck and the games i play with it . so my question to you guys is what non-r / g variants of tron could be viable in a competitive modern environment ?

tl;dr

what non-r / g variant should i build ?
my little brother is a sophomore in high-school and he was having a bit of trouble with his crush not liking him so me being the brother i am gave him the advice "dude girls at this age should not be a factor for your happiness just focus on doing you for now " whereas i thought it was good advice he took it the completely other way and has become a huge asshole.

tl;dr gave brother advice about life and now he is an asshole to women
not me but my grandfather was asked if he wanted to come in on an idea with 2 other guys. all he would need was 50 bucks and a little business experience but he said no. today that business is hardee's. he lost out big time.

tl;dr
grandfather had chance to invest in early hardee's, but declined.
doing some traveling this year and i am looking to build the **ultimate travel kit** . . . so far i have a bonavita 0.5l travel kettle and aeropress. looking for a grinder that would maybe fit into the aeropress. this way i can stack them in each other and have a compact travel kit.

**Ultimate travel kit**: what grinder would you recommend that fits in aeropress?
### Sub-datasets across verticals

<table>
<thead>
<tr>
<th>Sub-dataset</th>
<th>Number of (content, summary) pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions</td>
<td>78,710</td>
</tr>
<tr>
<td>Vulgar Posts</td>
<td>299,145</td>
</tr>
<tr>
<td>True Summaries</td>
<td>966,430</td>
</tr>
<tr>
<td>Topic + tl;dr</td>
<td>729,042</td>
</tr>
<tr>
<td>Topic</td>
<td>729,042</td>
</tr>
</tbody>
</table>
Deep Learning for Summarization
Why Deep Learning?
Why Deep Learning?

- Data is not always linearly separable
Why Deep Learning?

- Data is not always linearly separable
Why Deep Learning?

- Deep learning models provide **non-linearity**
Why Deep Learning?

- **Approximate** a function around the datapoints (**Learning**
Why Deep Learning?

- **Approximate** a function around the datapoints (**Learning**)
- Chaining & stacking non-linear units into layers helps the model learn complex relationships (**Deep**)
Why Deep Learning?

- **Approximate** a function around the datapoints (**Learning**)
- Chaining & stacking non-linear units into layers helps the model learn complex relationships (**Deep**)
- Different type of data requires different architecture
Recurrent Neural Network
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What is a RNN?
What happens in a RNN?
What happens in a RNN?
What happens in a RNN?
What happens in a RNN?

\[ s_4 = f(s_3, x_4) \]
What happens in a RNN?

\[ s_4 = f(s_3, x_4) \]

\[ = f(f(s_2, x_3), x_4) \]
What happens in a RNN?

\[
\begin{align*}
s_4 &= f(s_3, x_4) \\
    &= f(f(s_2, x_3), x_4) \\
    &= f(f(f(s_1, x_2), x_3), x_4)
\end{align*}
\]
What happens in a RNN?

\[
s_4 = f(s_3, x_4) \\
= f(f(s_2, x_3), x_4) \\
= f(f(f(s_1, x_2), x_3), x_4) \\
= f(f(f(f(s_0, x_1), x_2), x_3), x_4)
\]
What happens in a state?
What happens in a state?
What happens in a state?

Original embedding space

1 0 0 0 0
0 0 1 0 0
0 0 0 0 1
0 1 0 0 0
0 0 0 1 0

word_1
What happens in a state?

Original embedding space

word_1

embedding_1

1 0 0 0 0
0 0 1 0 0
0 0 0 0 1
0 1 0 0 0
0 0 0 1 0
What happens in a state?

Original embedding space

1 0 0 0 0
0 0 1 0 0
0 0 0 0 1
0 1 0 0 0
0 0 0 1 0

state

embedding_1

word_1
What happens in a state?

Original embedding space

word_1

state

embedding_1

context

Original embedding space

Transformed embedding space

1 0 0 0 0
0 0 1 0 0
0 0 0 0 1
0 1 0 0 0
0 0 0 1 0

0.5 0.6 0 1.8
1.2 0 1.1 1.5 0.9
0 0.5 1.3 0 1.2
1.4 1.6 0 0.3 1.3
0 0.5 0.6 1.2 1.4
1.1 0.4 1.2 0 1.4
Sequence to Sequence model
Sequence to Sequence model

Sind Sie noch wach? → Are you still awake?
Learning to map a sequence to sequence
Learning to map a sequence to sequence

Source sequence (P)

\[ X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \]
Learning to map a sequence to sequence

Source sequence (P)

\[ X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \]

Target sequence (Q)

\[ Y = \{y_1, y_2, y_3, y_4, \ldots, y_m\} \]
Learning to map a sequence to sequence

Source sequence (P)

\[ X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \]

Target sequence (Q)

\[ Y = \{y_1, y_2, y_3, y_4, \ldots, y_m\} \]

Minimize the negative likelihood
Learning to map a sequence to sequence

Source sequence (P)

\[ X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \]

Minimize the negative likelihood

Maximize

\[ p(Y|X, \text{parameters}) \]

Target sequence (Q)

\[ Y = \{y_1, y_2, y_3, y_4, \ldots, y_m\} \]
Sequence to Sequence model
Sequence to
Sequence model

S_word1  S_word2  S_word3  S_word4

Source sequence
Sequence to Sequence model

Source sequence

Emb1  Emb2  Emb3  Emb4
S_word1  S_word2  S_word3  S_word4
Sequence to Sequence model

Source sequence

1 1 1

h 0 1 2 3 h_{final}

state_1 state_2 state_3 state_4

Emb1 Emb2 Emb3 Emb4

S_word1 S_word2 S_word3 S_word4

Sequence to Sequence model
Sequence to Sequence model

Encoder

Source sequence

state_1 → state_2 → state_3 → state_4 → h_{final}

Emb1 → Emb2 → Emb3 → Emb4

S_word1 → S_word2 → S_word3 → S_word4
Sequence to Sequence model

Encoder

Source sequence

C O N T E X T
Sequence to Sequence model

Encoder

Source sequence

S_word1 → h → state_1 → Emb1 → state_2 → h → Emb2 → state_3 → h → Emb3 → state_4 → h → h_final → S_word4

Decoder

Target sequence

T_word1 → h → state_1 → Emb1 → state_2 → h → Emb2 → state_3 → h → Emb3 → state_4 → h → T_word4
Sequence to Sequence model

Encoder

Decoder

Source sequence

Target sequence

Emb1

Emb2

Emb3

Emb4

S_word1

S_word2

S_word3

S_word4

T_word1

T_word2

T_word3

T_word4

1

1

1

1

h

h

h

h

h

h

h

h

state_1

state_2

state_3

state_4

state_1

state_2

state_3

state_4

Emb1

Emb2

Emb3

Emb4

T_word1

T_word2

T_word3

T_word4

SOFTMAX LAYER

CONTEXT
Sequence to Sequence model

Encoder

1 1
h 1
state_1

0
Emb1

S_word1

1
state_2

Emb2

S_word2

2
state_3

Emb3

S_word3

3
state_4

Emb4

S_word4

h
h
h
h

Decoder

SOFTMAX LAYER

1 1 1

h 1
state_1

0
Emb1

T_word1

1
state_2

Emb2

T_word2

2
state_3

Emb3

T_word3

3
state_4

Emb4

T_word4

C O N T E X T
Sequence to Sequence model

Source sequence

Decoder

Target sequence

Softmax Layer

Encoder

this, target is, vocabulary the, containing, some, words, .....
Sequence to Sequence model

Trained end to end

this, target
is, vocabulary
the, containing,
some, words, ….

Source sequence

Encoder

1

1

1

h

0

state_1

Emb1

S_word1

state_2

Emb2

S_word2

state_3

Emb3

S_word3

state_4

Emb4

S_word4

h

h

h

h

state_1

state_2

state_3

state_4

h_final

Decoder

P_word1

P_word2

P_word3

P_word4

1

1

1

T_word1

T_word2

T_word3

T_word4

SOFTMAX LAYER

Trained end to end

Sequence to
Sequence model
Sequence to Sequence model - Generating summary

Source sequence (test)

Encoder

Decoder

Target sequence (ideal/optional)
Sequence to Sequence model - Generating summary

Source sequence (test)

Target sequence (ideal / optional)
Sequence to Sequence model - Generating summary

Source sequence (test)

Encoder

Decoder

Target sequence (ideal / optional)

Summary

this, target is, vocabulary the, containing, some, words, .....
Sequence to Sequence model - Generating summary

Source sequence (test)

Target sequence (ideal / optional)

Softmax Layer

Encoder

Decoder

this, target is, vocabulary the, containing, some, words, .....
Sequence to Sequence model - Generating summary

Source sequence (test):

T_word1, T_word2, T_word3, T_word4

Target sequence (ideal / optional):

P_word1, P_word2, P_word3, P_word4

Encoder:

S_word1, S_word2, S_word3, S_word4

Decoder:

h_0, h_1, h_2, h_3

Summary:

G_word1

SOFTMAX LAYER

Note: The diagram shows the flow of information from the source sequence to the target sequence through the encoder and decoder processes.
Sequence to Sequence model - Generating summary

**Source sequence (test)**

1. S_word1
2. S_word2
3. S_word3
4. S_word4

**Target sequence (ideal / optional)**

1. T_word1
2. T_word2
3. T_word3
4. T_word4

**Summary**

1. G_word1
2. G_word2

**Encoder**

- Emb1
- Emb2
- Emb3
- Emb4

**Decoder**

- Emb1
- Emb2
- Emb3
- Emb4

**Softmax Layer**

**Context**
Sequence to Sequence model - Generating summary

Source sequence *(test)*

```
1 1
h 0
state_1

S_word1
```

```
1 1
h 1
state_2

S_word2
```

```
1 1
h 2
state_3

S_word3
```

```
1 1
h 3
state_4

S_word4
```

Target sequence *(ideal / optional)*

```
T_word1
```

```
T_word2
```

```
T_word3
```

```
T_word4
```

Summary

```
G_word1
```

```
G_word2
```

```
G_word3
```

```
P_word1
```

```
P_word2
```

```
P_word3
```

```
P_word4
```

SOFTMAX LAYER

```
Emb1
```

```
Emb2
```

```
Emb3
```

```
Emb4
```

Encoder

```
1 1
h 0
state_1

Emb1

S_word1
```

```
1 1
h 1
state_2

Emb2

S_word2
```

```
1 1
h 2
state_3

Emb3

S_word3
```

```
1 1
h 3
state_4

Emb4

S_word4
```

Decoder

```
state_1

T_word1
```

```
state_2

T_word2
```

```
state_3

T_word3
```

```
state_4

T_word4
```

**this, target is, vocabulary the, containing, some, words, .....**
Sequence to Sequence model - Generating summary

Source sequence (*test*)

```
Emb1  Emb2  Emb3  Emb4
S_word1  S_word2  S_word3  S_word4
```

```
Encoder
```

```
1 1 1 1
h  h  h  h
0  1  2  3  h_final
```

```
Decoder
```

```
state_1  state_2  state_3  state_4
```

```
P_word1  P_word2  P_word3  P_word4
```

```
SOFTMAX LAYER
```

```
G_word1  G_word2  G_word3  G_word4
```

Target sequence (*ideal / optional*)

```
T_word1  T_word2  T_word3  T_word4
```

Summary

```
this, target is, vocabulary the, containing, some, words, ..... 
```

**Note:** The image includes a flowchart illustrating the process of sequence-to-sequence modeling, with nodes representing states and connections indicating the flow of information from source to target sequences.
Attention!
Sequence to Sequence model

Source sequence

1 1

Encoder

Emb1 Emb2 Emb3 Emb4

S_word1 S_word2 S_word3 S_word4

h 0

state_1

h 1

state_2

h 2

state_3

h 3

state_4

h_fi nal

Decoder

Emb1 Emb2 Emb3 Emb4

T_word1 T_word2 T_word3 T_word4

Summary

this, target is, vocabulary the, containing, some, words, .....
**Sequence to Sequence model**

**Encoder**
- Emb1
- Emb2
- Emb3
- Emb4
- S_word1
- S_word2
- S_word3
- S_word4

**Source sequence**

**Decoder**
- Emb1
- Emb2
- Emb3
- Emb4
- T_word1
- T_word2
- T_word3
- T_word4

**Fixed length**
- h
- state_1
- state_2
- state_3
- state_4

**SOFTMAX LAYER**
- P_word1
- P_word2
- P_word3
- P_word4

Summary
- this, target
- is, vocabulary
- the, containing
- some, words, ....

**Target sequence**
Sequence to Sequence model

**Encoder**
- Input sequence: \( S_{\text{word}1}, S_{\text{word}2}, S_{\text{word}3}, S_{\text{word}4} \)
- Output: \( \text{state}_1, \text{state}_2, \text{state}_3, \text{state}_4, h_{\text{final}} \)

**Decoder**
- Input sequence: \( \text{state}_1, \text{state}_2, \text{state}_3, \text{state}_4 \)
- Output: \( P_{\text{word}1}, P_{\text{word}2}, P_{\text{word}3}, P_{\text{word}4} \)

**Fixed length**

**SOFTMAX LAYER**

**Summary**
- Input: \( \text{target sequence is, vocabulary containing, some, words, .....} \)
- Output: \( \text{Summary} \)
Sequence to Sequence model with attention

this, target is, vocabulary the, containing, some, words, .....
**Sequence to Sequence model with attention**

**Summary**

```
this, target
is, vocabulary
the, containing,
some, words.....
```

**Encoder**

- **Word Annotations**
  - W1
  - W2
  - W3
  - W4

- **State variables**
  - state_1
  - state_2
  - state_3
  - state_4

- **Embeddings**
  - Emb1
  - Emb2
  - Emb3
  - Emb4

- **Source sequence**
  - S_word1
  - S_word2
  - S_word3
  - S_word4

**Decoder**

- **State variables**
  - state_1
  - state_2
  - state_3
  - state_4

- **Embeddings**
  - Emb1
  - Emb2
  - Emb3
  - Emb4

- **Target sequence**
  - T_word1
  - T_word2
  - T_word3
  - T_word4

**SOFTMAX LAYER**
Sequence to Sequence model with attention

Word Annotations

Encoder

Decoder

Source sequence

Target sequence

Summary

1. This, target is, vocabulary the, containing, some, words, .....
Sequence to Sequence model with attention

Word Annotations

Encoder

Source sequence

Decoder

Target sequence

summary

Softmax Layer

Attention Layer

Sequence to Sequence model with attention

Word Annotations

Summary

Source sequence
- S_word1
- S_word2
- S_word3
- S_word4

Target sequence
- T_word1
- T_word2
- T_word3
- T_word4

Encoder
- Emb1
- Emb2
- Emb3
- Emb4

1
1
1

Decoder
- Emb1
- Emb2
- Emb3
- Emb4

1
1
1

Attention Layer

SOFTMAX LAYER

this, target is, vocabulary the, containing, some, words, ....
Sequence to Sequence model with attention

Word Annotations

Encoder

Decoder

SOFTMAX LAYER

Attention Layer

Source sequence

Target sequence

Summary
Sequence to Sequence model with attention

- **Encoder**
  - Source sequence
  - Embeddings (Emb1, Emb2, Emb3, Emb4)
  - States (state_1, state_2, state_3, state_4)

- **Word Annotations**
  - W1, W2, W3, W4

- **Decoder**
  - Target sequence
  - Embeddings (Emb1, Emb2, Emb3, Emb4)
  - States (state_1, state_2, state_3, state_4)

- **Attention Layer**

- **Softmax Layer**

- **Summary**
  - Target sequence
  - Vocabulary containing some words...

- **Diagram notes**
  - this, target is, vocabulary the, containing, some, words, ....
Evaluation
Evaluation

Automatic

Manual
ROUGE for automatic evaluation
ROUGE for automatic evaluation

- ROUGE measures **n-gram overlaps** and **common subsequences** between automatic and ideal summaries

\[
ROUGE - N = \frac{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count(gram_n)}
\]
ROUGE for automatic evaluation

- ROUGE measures **n-gram overlaps** and **common subsequences** between automatic and ideal summaries

\[
ROUGE - N = \frac{\sum_{S \in \{ReferenceSummaries\}} \sum_{\text{gram}_n \in S} \text{Count}_{match}(\text{gram}_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{\text{gram}_n \in S} \text{Count}(\text{gram}_n)}
\]
ROUGE for automatic evaluation

- ROUGE measures **n-gram overlaps** and **common subsequences** between automatic and ideal summaries.
ROUGE for automatic evaluation

- ROUGE measures **n-gram overlaps** and **common subsequences** between automatic and ideal summaries

\[ \text{ROUGE} - N = \frac{\sum_{S \in \{\text{ReferenceSummaries}\}} \sum_{\text{gram}_n \in S} \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S \in \{\text{ReferenceSummaries}\}} \sum_{\text{gram}_n \in S} \text{Count}(\text{gram}_n)} \]

<table>
<thead>
<tr>
<th>Training set</th>
<th>Size</th>
<th>ROUGE-1</th>
<th>ROUGE-2</th>
<th>ROUGE-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webis-TLDR-17</td>
<td>2 Million</td>
<td>17.87</td>
<td>6.11</td>
<td>15.78</td>
</tr>
<tr>
<td>Gigaword (Rush et al. 2015)</td>
<td>4 Million</td>
<td>31.33</td>
<td>11.81</td>
<td>36.38</td>
</tr>
</tbody>
</table>
**ROUGE for automatic evaluation**

- ROUGE measures **n-gram overlaps** and **common subsequences** between automatic and ideal summaries.

\[
ROUGE - N = \frac{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count(gram_n)}
\]

<table>
<thead>
<tr>
<th>Training set</th>
<th>Size</th>
<th>ROUGE-1</th>
<th>ROUGE-2</th>
<th>ROUGE-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webis-TLDR-17</td>
<td>2 Million</td>
<td>17.87</td>
<td>6.11</td>
<td>15.78</td>
</tr>
<tr>
<td>Gigaword (Rush et al. 2015)</td>
<td>4 Million</td>
<td>31.33 ?</td>
<td>11.81 ?</td>
<td>36.38 ?</td>
</tr>
</tbody>
</table>
Manual evaluation
Manual evaluation

- A test set of 200 examples containing 50 examples from each vertical
Manual evaluation

- A test set of 200 examples containing 50 examples from each vertical
- Binary categorization as relevant/irrelevant
Manual evaluation

- A test set of 200 examples containing 50 examples from each vertical
- Binary categorization as relevant/irrelevant
- Manual judgements vs ROUGE scores
Manual judgement vs ROUGE
Manual judgement vs ROUGE
## Manual evaluation
(Cell values are in percentages)

<table>
<thead>
<tr>
<th>Test set /Model</th>
<th>Questions (50)</th>
<th>Vulgar (50)</th>
<th>True(50) summaries</th>
<th>Generic (50)</th>
<th>All (200)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ground truth</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>84</td>
<td>96</td>
</tr>
<tr>
<td>Questions (78,710)</td>
<td>31</td>
<td>50</td>
<td>28</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>Vulgar (299,145)</td>
<td>16</td>
<td>35</td>
<td>14</td>
<td>12</td>
<td>19.5</td>
</tr>
<tr>
<td>True Summaries (966,430)</td>
<td>44</td>
<td>54.9</td>
<td>36</td>
<td>30</td>
<td><strong>41.29</strong></td>
</tr>
<tr>
<td>Topic + tldr (729,042)</td>
<td>44</td>
<td>45</td>
<td>16</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>Topic (729,042)</td>
<td>58</td>
<td>47</td>
<td>50</td>
<td>30</td>
<td><strong>48.5</strong></td>
</tr>
<tr>
<td>Generic (2 Million)</td>
<td><strong>67.3</strong></td>
<td>62</td>
<td>52</td>
<td>34</td>
<td><strong>53.5</strong></td>
</tr>
</tbody>
</table>
### Manual evaluation  (Cell values are in percentages)

<table>
<thead>
<tr>
<th>Test set /Model</th>
<th>Questions (50)</th>
<th>Vulgar (50)</th>
<th>True(50) summaries</th>
<th>Generic (50)</th>
<th>All (200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground truth</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>84</td>
<td>96</td>
</tr>
<tr>
<td>Questions (78,710)</td>
<td>31</td>
<td>50</td>
<td>28</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>Vulgar (299,145)</td>
<td>16</td>
<td>35</td>
<td>14</td>
<td>12</td>
<td>19.5</td>
</tr>
<tr>
<td>True Summaries (966,430)</td>
<td>44</td>
<td>54.9</td>
<td>36</td>
<td>30</td>
<td>41.29</td>
</tr>
<tr>
<td>Topic + tldr (729,042)</td>
<td>44</td>
<td>45</td>
<td>16</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>Topic (729,042)</td>
<td>58</td>
<td>47</td>
<td>50</td>
<td>30</td>
<td>48.5</td>
</tr>
<tr>
<td>Generic (2 Million)</td>
<td>67.3</td>
<td>62</td>
<td>52</td>
<td>34</td>
<td>53.5</td>
</tr>
</tbody>
</table>
Examples
Questions model

ever since i watched one of the counselors at my camp play tron, i’ve wanted to build it. the only problem is, i can’t figure out which colors to build around. the group at my lgs that plays mainly modern is saying i should build r/g because it’s the best, but i think the money that i would spend on a playset of [[grove of the burnwillows]] could be spent elsewhere and have a much bigger impact on the deck and the games i play with it. so my question to you guys is what non-r/g variants of tron could be viable in a competitive modern environment?

what non-r/g variant should i build?

what’s the best build for a new player?
ever since i watched one of the counselors at my camp play tron , i ’ ve wanted to build it . the only problem is , i can ’ t figure out which colors to build around . the group at my lgs that plays mainly modern is saying i should build r / g because it ’ s the best , but i think the money that i would spend on a playset of [[ grove of the burnwillows ]] could be spent elsewhere and have a much bigger impact on the deck and the games i play with it . so my question to you guys is what non-r / g variants of tron could be viable in a competitive modern environment ?

what non-r / g variant should i build ?

what ’ s the best build for a new player ?
just talked to catholic priest on a wedding. guys, have no mercy, please. the shit they preach is a mind killer for our civilization. religious people are fucking dumb, dumb fucks and they spread shit all over the place. i fucking spit on catholics. my whole family is religious and i can’t have intellectual conversation on any subject with them. we can not tolerate any of this shit any longer. i’m picking hitchens and the rest of dawkins for my holidays and i’m going militant then. hoooly shit! sorry for rant but the level of bullshit they can get over with is killing me.

religion = mental retardation.

religion is fucking stupid.
just talked to catholic priest on a wedding. guys, have no mercy, please. the shit they preach is a mind killer for our civilization. religious people are fucking dumb, dumb fucks and they spread shit all over the place. i fucking spit on catholics. my whole family is religious and i can't have intellectual conversation on any subject with them. we can not tolerate any of this shit any longer. i'm picking hitchens and the rest of dawkins for my holidays and i'm going militant then. hoooly shit! sorry for rant but the level of bullshit they can get over with is killing me.

religion = mental retardation.

religion is fucking stupid.
Alright so I just started seeing a girl I’ve known for quite a while and things have been great. We get along really well and just genuinely enjoy each other’s time. Only hiccup is I enjoy smoking copious weed and she well, she hates weed. Her ex-boyfriend apparently let weed take over his life. He was standing her up to go smoke, ditching work to go find weed, and just fucking around basically is what it sounds like. Well she blames the weed for his shortcomings and now judges everyone based on the fact that they smoke weed. She doesn’t care about tobacco, or alcohol either just weed. I know common sense says to just tell her I smoke a ton and tell her it doesn’t impact my life negatively and I still get my shit done. IDK. Basically I’m looking for helpful stories or advice for me if you’ve been in a similar situation. We aren’t super serious at all, and she’s moving states in about two months so it’s not a huge deal but I also want to enjoy our time together while she still lives here.

Been dating a new girl who hates weed because her last boyfriend couldn’t handle smoking and his life. Now I’m torn between telling her how much I smoke and how often or just hiding it until she moves away. Any help is greatly appreciated.

I want to tell a girl I smoke a lot, she hates weed, what do I do?
Did I get underdosed tabs? : tried LSD for the first time yesterday, first took half a tab, then the other half 4 minutes later. Ended up just a bit light headed so 3.5 hours in I took another tab, still got pretty much no visuals.

Question about LSD: I'm a noob and I'd like to know if I did it or not.
Guaranteed Promotion or Possible Future Career Path? : my boss offered me a position she wants to fill this week, while I'm also in consideration for a position with her boss.

need advice on negotiating a promotion.
I finished Path of Daggers earlier this month and took a break to read I Am Pilgrim (I HIGHLY recommend) and I am now ready to start my journey into book 9. One thing, I’ve forgotten a few plot threads. I can’t search for them as possible spoilers, so what do I need to know going in? Oh, and SPOILERS for those not up to here and please, no spoilers for me. I love this series too much for it to be ruined. I know Faile has been taken but I’m not sure about the other main characters whereabouts and smaller character plots. If

where are the characters at and what are they hoping to do? People refreshing my memory would be greatly appreciated.

What do I need to know to start my journey into book 9?
Existing Problems
my wife (teacher) constantly complains about how she can't stand her job, the kids are disrespectful, etc., and how much she hates it. I know my wife, and I know she wants me to say "sweetie why don't you just quit?" but I'm not going to say that under any circumstances. However, I did tell her that I will support her leaving if she has a backup job or if she decides to go back to school (that will help advance her career). I also said that I would pay for her schooling. Her response is that she does not want to go back to school. It's getting to the point where she talks about how much she hates work everyday. It completely ruins the end of the day, as half the time it ends in her crying. I'm not sure what to do.

My (32/m) wife (32/f) complains about her job and wants me to tell her it's okay to quit, but I'm not going to, how do I handle this? wife hates her job and wants me to tell her to quit. I'm not going to tell her to quit, but I also don't know what to do.

I [27/m] don't know how to tell my wife [28/f] that I don't want to quit her job. wife is mad because I don't want to quit.
it was around the holidays a few years ago and i had started smoking for the first time. being somewhat forced to spend "quality time" with my family, i decided to smoke a blunt before going. i get to my parents house, i’m at a solid [9] and my sweet, little, 80 year old grandma was sitting on the couch. i took a seat next to her and she tried to lean her head on my shoulder, being too short she rested it on my bicep. she patted my leg before looking up at me saying “you’re so high”. i nearly jump out of my skin, and look at her a bit side-ways because i had put on new clothes, perfume and even brushed my teeth before coming over. just before i could get the words “how did you know” out of my mouth, she smiled at me, “so high i can’t rest my head on your shoulder.” i laughed nervously and adjusted accordingly so she could lean her head on me.

went to a family gathering super baked, sat next to my grandma who said "you're so high" referring to my height because she couldn't rest her head on my shoulder when i sat up straight. scared the shit outta me.

i got high as fuck and got high as fuck.
it was around the holidays a few years ago and i had started smoking for the first time. being somewhat forced to spend " quality time " with my family, i decided to smoke a blunt before going. i get to my parents house, i'm at a solid 9 and my sweet, little, 80 year old grandma was sitting on the couch. i took a seat next to her and she tried to lean her head on my shoulder, being too short she rested it on my bicep. she patted my leg before looking up at me saying " you're so high " . i nearly jump out of my skin, and look at her a bit side-ways because i had put on new clothes, perfume and even brushed my teeth before coming over. just before i could get the words " how did you know " out of my mouth, she smiled at me, " so high i can't rest my head on your shoulder. " i laughed nervously and adjusted accordingly so she could lean her head on me.

went to a family gathering super baked, sat next to my grandma who said " you're so high " referring to my height because she couldn't rest her head on my shoulder when i sat up straight. scared the shit outta me.

i got high as fuck and got high as fuck.
Conclusion & Future work
• **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
• **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  ○ Amazon product reviews, Clueweb/Wikipedia, Blogs
- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  - Amazon product reviews, Clueweb/Wikipedia, Blogs
- Exploring the effectiveness and limitations of **sequence-to-sequence** models
- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  - Amazon product reviews, Clueweb/Wikipedia, Blogs
- Exploring the effectiveness and limitations of **sequence-to-sequence** models
  - Convolutional networks instead of recurrent
- **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  - Amazon product reviews, Clueweb/Wikipedia, Blogs
- **Exploring the effectiveness and limitations of sequence-to-sequence models**
  - Convolutional networks instead of recurrent
  - Fine tuning the summary generation process
• **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  ○ Amazon product reviews, Clueweb/Wikipedia, Blogs

• **Exploring the effectiveness and limitations of sequence-to-sequence models**
  ○ Convolutional networks instead of recurrent
  ○ Fine tuning the summary generation process
  ○ Bridging the gap between extractive and abstractive
● **Webis-TLDR-17 corpus** for summarization from the social media domain (Reddit)
  ○ Amazon product reviews, Clueweb/Wikipedia, Blogs

● Exploring the effectiveness and limitations of **sequence-to-sequence** models
  ○ Convolutional networks instead of recurrent
  ○ Fine tuning the summary generation process
  ○ Bridging the gap between extractive and abstractive

● **Alternate evaluation method**
Thank You!