

Abstractive Text Summarization

Abstractive summarization is the new state of art method, which generates new sentences that could best represent the whole text. This is better than extractive methods where sentences are just selected from original text for the summary. This demands mostly deep learning techniques.

Extractive summary is choosing specific sentences from the text to compile a summary, while abstractive summary means generating a summary in the computer's own words.

Along with that, there exist numerous subcategories, many unlisted

- Single-document or multi-document means to summarize a single piece of text, or to analyze a collection of texts on different topics, and create a summary that generalizes their opinions.
- Indicative summary captures the general meaning of the text, while informative summary includes all the fine details.

Sequences

Sequenced data is data that takes the form of a list of varying length. Sequences can be difficult for traditional neural networks to process since there is the idea of an order, and the length may vary.

For example, consider the lyrics of a song, a sequence of words. The idea of an order means that certain words naturally come “before” others. It is easy to remember the words in the **normal** order, but much harder to recall the lyrics **backwards**.

In the real world, sequences can be any kind of data of varying length and have a general idea of an order. Some examples are texts, audio recordings, and video recordings. Additionally, we may want to use sequences in the input, output, or even both, in a machine learning application.

Sequential Networks

New network architectures were discovered a few decades ago to deal with sequential data.

RNNs

Recurrent neural networks are a new type of network, in which their layers are used recurrently, or repeatedly. This means the layers are all the same. The network takes in a part of the sequence for each time step and performs some calculation on it. Specifically, for each time step, it uses the previous time step's hidden layer and a new part of the input sequence to make a new output. This is then passed to the next time step, along with the next part of the sequence.

LSTMs

Lstm (Long short term memory) is a type of RNN, which has the ability to choose what's important and what's not, hence the name, as it can be used in both short and long term memory cases.

To summarize a text using an abstractive method, we combine the power of word embeddings and RNNs or LSTMs. This transforms a sequence of text just like neural network transforms a vector

- To build a text summarizer, we first use word embeddings to map our input sequence words to a sequence of vectors.
- Then, we use an autoencoder-like structure to capture the meaning of the passage. Two separate RNNs or LSTMs are trained to encode the sequence into a single matrix or vector, and then to decode the matrix or vector into a transformed sequence of words.
- Lastly, convert the sequence of vectors outputted by the decoder back into words using the word embeddings.