

Get To The Point Summarization with Pointer-Generator Networks

MICANS INFOTECH

ABSTRACT

- ▶ Neural sequence-to-sequence models have provided a viable new approach for abstractive text summarization (meaning they are not restricted to simply selecting and rearranging passages from the original text).
- ▶ However, these models have two shortcomings: they are liable to reproduce factual details inaccurately, and they tend to repeat themselves.
- ▶ In this work we propose a novel architecture that augments the standard sequence-to-sequence attentional model in two orthogonal ways.
- ▶ First, we use a hybrid pointer-generator network that can copy words from the source text via pointing, which aids accurate reproduction of information, while retaining the ability to produce novel words through the generator.
- ▶ Second, we use coverage to keep track of what has been summarized, which discourages repetition

EXISTING

- ▶ Summarization is the task of condensing a piece of text to a shorter version that contains the main information from the original.
- ▶ There are two broad approaches to summarization: extractive and abstractive.
- ▶ Extractive methods assemble summaries exclusively from passages (usually whole sentences) taken directly from the source text, while abstractive methods may generate novel words and phrases not featured in the source text – as a human-written abstract usually does.

CONTD..

- ▶ Due to the difficulty of abstractive summarization, the great majority of past work has been extractive. However, the recent success of sequence-to-sequence models, in which recurrent neural networks
- ▶ (RNNs) both read and freely generate text, has made abstractive summarization viable.
- ▶ Though these systems are promising, they exhibit undesirable behavior such as inaccurately reproducing factual details, an inability to deal with out-of-vocabulary (OOV) words, and repeating themselves

DISADVANTAGE

- ▶ Neural sequence-to-sequence models have provided a viable new approach for abstractive text summarization (meaning they are not restricted to simply selecting and rearranging passages from the original text).
- ▶ However, these models have two shortcomings: they are liable to reproduce factual details inaccurately, and they tend to repeat themselves.

PROPOSED

- ▶ we present an architecture that addresses these three issues in the context of multi sentence summaries. While most recent abstractive work has focused on headline generation tasks (reducing one or two sentences to a single headline),
- ▶ we believe that longer-text summarization is both more challenging (requiring higher levels of abstraction while avoiding repetition) and ultimately more useful.
- ▶ Therefore we apply our model to the recently-introduced CNN/Daily Mail dataset which contains news articles (39sentences on average) paired with multi-sentence summaries, and show that we outperform the state-of-the-art abstractive system by at least 2 ROUGE points.
- ▶ Our hybrid pointer-generator network facilitates copying words from the source text via pointing, which improves accuracy and handling of OOV words, while retaining the ability to generate new words.
- ▶ The network, which can be viewed as a balance between extractive and abstractive approaches, is similar Forced-Attention Sentence Compression, that were applied to short-text summarization.
- ▶ We propose a novel variant of the coverage vector from Neural Machine Translation, which we use to track and control coverage of the source document. We show that coverage is remarkably effective for eliminating repetition.

ADVANTAGE

- ▶ We apply our model to the CNN / Daily Mail summarization task, outperforming the current abstractive state-of-the-art by at least 2 ROUGE points
- ▶ hybrid pointer-generator architecture with coverage, and showed that it reduces inaccuracies and repetition. We applied our model to a new and challenging long-text dataset, and significantly outperformed the abstractive state-of-the-art result.

HARDWARE REQUIREMENTS

- ▶ Processor – Pentium -III
- ▶ Speed – 1.1 Ghz
- ▶ RAM – 256 MB(min)
- ▶ Hard Disk – 20 GB
- ▶ Floppy Drive – 1.44 MB
- ▶ Key Board – Standard Windows Keyboard
- ▶ Mouse – Two or Three Button Mouse
- ▶ Monitor – SVGA

MICANS INFOTECH

SOFTWARE REQUIREMENTS

- ▶ Operating System : Windows 8
- ▶ Front End : Java /DOTNET
- ▶ Database : Mysql/HEIDISQL

MICANS INFOTECH

CONCLUSION

- ▶ In this work we presented a hybrid pointer-generator architecture with coverage, and showed that it reduces inaccuracies and repetition.
- ▶ We applied our model to a new and challenging long-text dataset, and significantly outperformed the abstractive state-of-the-art result.
- ▶ Our model exhibits many abstractive abilities, but attaining higher levels of abstraction remains an open re-search question.

REFERENCE

- ▶ Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Ben-gio. 2015. Neural machine translation by jointly learning to align and translate. In International Conference on Learning Representations.
- ▶ Qian Chen, Xiaodan Zhu, Zhenhua Ling, Si Wei, and Hui Jiang. 2016. Distraction-based neural networks for modeling documents. In International Joint Conference on Artificial Intelligence.
- ▶ Jackie Chi Kit Cheung and Gerald Penn. 2014. Unsupervised sentence enhancement for automatic summarization. In Empirical Methods in Natural Language Processing.
- ▶ Sumit Chopra, Michael Auli, and Alexander M. Rush. 2016. Abstractive sentence summarization with attentive recurrent neural networks. In North American Chapter of the Association for Computational Linguistics.
- ▶ Michael Denkowski and Alon Lavie. 2014. Meteor universal: Language specific translation evaluation for any target language. In EACL 2014 Workshop on Statistical Machine Translation.

MICANS INFO TECH