Deep Learning for Text Summarization: A Brief Survey

B. Tauro, M. Serry, T. Itani and T. Salah, Heriot – Watt University

**Abstract** - Deep learning has been widely adopted in recent years to provide superior alternatives to traditional text summarization techniques, in this paper we evaluate sample papers that utilize deep learning for text summarization weither extractive or abstractive In order to evaluate overall challenges and trends in the field.

**Index Terms** - Text Summarization, Abstractive Text Summarization, Extractive Text Summarization, Deep Learning.

—————————— ◆ ——————————

# Introduction

*Automatic text summarization* is the task of producing a concise summary while preserving key information and overall meaning [1]. Previously, text summarization was dominated by traditional machine learning techniques [2]. In recent years, Deep learning has demonstrated more promising results to pre-existing approaches [3], Namely in both the extractive and abstractive paradigms. In this paper we briefly review the usage of deep learning models from sample papers to discuss motivations, challenges, and overall trends in the field.

# Models Used

A close up of a logo

Description automatically generatedIn this section, we review the deep learning models that were used in our sample papers.

## Restricted Boltzmann machines

Restricted Boltzmann machines (RBMs) are probabilistic graphical models that can be interpreted as stochastic neural networks. They are a variation of Boltzmann machines that learn more efficiently [4]. They have various applications outside of text summarization, including dimensionality reduction, topic modeling and feature learning [5].

A close up of a map

Description automatically generated

Figure 2 - The basic structure of a deep auto encoder.

Figure 1 - The basic structure of restricted Boltzmann machine (RBM).

In an RBM, we have layers where no two units within the same group are connected, in contrast to a BM where each node is connected to all other nodes symmetrically [5].

## Deep Auto-Encoders

A Deep Auto-Encoder is feed forward neural network with the fuction of reconstructing an input x [7]. They are a deep generative model that learns a binary encoding capable of fast and efficient retrieval times in comparison to latent analysis and and when used as a filtering mechanism for more traditional approaches such as TF-IDF [7].

They are a feed forward network with the exact number of nodes in both the input and output layers, a compressed bottle neck layer is used as a gating mechanism between the encoding and decoding process of the model. This allows the network to recreate the input from sparse features while also recognizing the most vital features in the input space [6].

# Experiments

# Evaluation

# Conclusions

# End Sections

## Appendices

## References

1. Andrew Turpin, Yohannes Tsegay, David Hawking, and Hugh E Williams. 2007. Fast generation of result snippets in web search. In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 127–134.
2. Dong, Y., 2018. A survey on neural network-based summarization methods. arXiv preprint arXiv:1804.04589.
3. K˚ageb¨ack, M., Mogren, O., Tahmasebi, N., and Dubhashi, D. (2014). Extractive summarization using continuous vector space models. In Proceedings of the 2nd Workshop on Continuous vector space space models and their Compositionality (CVSC)@ EACL, pages 31–39.
4. Fischer, Asja & Igel, Christian. (2012). An Introduction to Restricted Boltzmann Machines. 14-36. 10.1007/978-3-642-33275-3\_2.
5. A. Sharma, " Restricted Boltzmann Machines — Simplified", October 2nd, 2018. [Online]. Available: https://towardsdatascience.com/restricted-boltzmann-machines-simplified-eab1e5878976. [Accessed May 4, 2020].
6. Hinton, G. and Salakhutdinov, R. (2011), Discovering Binary Codes for Documents by Learning Deep Generative Models. Topics in Cognitive Science, 3: 74-91. doi:10.1111/j.1756-8765.2010. 01109.x
7. Mahmood Yousefi-Azar, Len Hamey, Text summarization using unsupervised deep learning, Expert Systems with Applications, Volume 68, 2017, Pages 93-105, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2016.10.017. (http://www.sciencedirect.com/science/article/pii/S0957417416305486)
8. J Dragomir R Radev, Eduard Hovy, and Kathleen McKeown. 2002. Introduction to the special issue on summarization. Computational linguistics 28, 4 (2002), 399–408.