# Neural Summarization by Extracting Sentences and Words

**Jianpeng Cheng** **Mirella Lapata**
ILCC, School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB
jianpeng.cheng@ed.ac.uk mlap@inf.ed.ac.uk



## Abstract

Traditional approaches to extractive summarization rely heavily on human-engineered features. In this work we propose a data-driven approach based on neural networks and continuous sentence features. We develop a general framework for single-document summarization composed of a hierarchical document encoder and an attention-based extractor. This architecture allows us to develop different classes of summarization models which can extract sentences or words. We train our models on large scale corpora containing hundreds of thousands of document-summary pairs[1]. Experimental results on two summarization datasets demonstrate that our models obtain results comparable to the state of the art without any access to linguistic annotation.

## 1 Introduction

The need to access and digest large amounts of textual data has provided strong impetus to develop automatic summarization systems aiming to create shorter versions of one or more documents, whilst preserving their information content. Much effort in automatic summarization has been devoted to sentence extraction, where a summary is created by identifying and subsequently concatenating the most salient text units in a document.

Most extractive methods to date identify sentences based on human-engineered features. These include surface features such as sentence position and length (Radev et al., 2004), the words in the title, the presence of proper nouns, content features such as word frequency (Nenkova et al., 2006), and event features such as action nouns (Filatova and Hatzivassiloglou, 2004). Sentences are typically assigned a score indicating the strength of presence of these features. Several methods have been used in order to select the summary sentences ranging from binary classifiers (Kupiec et al., 1995), to hidden Markov models (Conroy and O'Leary, 2001), graph-based algorithms (Erkan and Radev, 2004; Mihalcea, 2005), and integer linear programming (Woodsend and Lapata, 2010).

In this work we propose a data-driven approach to summarization based on neural networks and continuous sentence features. There has been a surge of interest recently in repurposing sequence transduction neural network architectures for NLP tasks such as machine translation (Sutskever et al., 2014), question answering (Hermann et al., 2015), and sentence compression (Rush et al., 2015). Central to these approaches is an encoder-decoder architecture modeled by recurrent neural networks. The encoder reads the source sequence into a list of continuous-space representations from which the decoder generates the target sequence. An attention mechanism (Bahdanau et al., 2015) is often used to locate the region of focus during decoding.

We develop a general framework for single-document summarization which can be used to extract sentences or words. Our model includes a neural network-based hierarchical document reader or encoder and an attention-based content extractor. The role of the reader is to derive the meaning representation of a document based on its sentences and their constituent words. Our models adopt a variant of neural attention to extract sentences or words. Contrary to previous work where attention is an *intermediate* step used to blend hidden units of an encoder to a vector propagating additional information to the decoder, our model applies attention *directly* to select sentences or words of the input document as the output summary. Similar neural attention architectures have been previously used for geometry reasoning (Vinyals et al., 2015), under the name *Pointer Networks*.

[1] Resources are available for download at http://homepages.inf.ed.ac.uk/s1537177/resources.html