Deep Learning-Based Approach to Automatic Abstractive Summarization of Patent Documents

Authors : Sakshi V. Tantak, Vishap K. Malik, Neelanjney Pilarisetty

Abstract : A patent is an exclusive right granted for an invention. It can be a product or a process that provides an innovative method of doing something, or offers a new technical perspective or solution to a problem. A patent can be obtained by making the technical information and details about the invention publicly available. The patent owner has exclusive rights to prevent or stop anyone from using the patented invention for commercial uses. Any commercial usage, distribution, import or export of a patented invention or product requires the patent owner's consent. It has been observed that the central and important parts of patents are scripted in idiosyncratic and complex linguistic structures that can be difficult to read, comprehend or interpret for the masses. The abstracts of these patents tend to obfuscate the precise nature of the patent instead of clarifying it via direct and simple linguistic constructs. This makes it necessary to have an efficient access to this knowledge via concise and transparent summaries. However, as mentioned above, due to complex and repetitive linguistic constructs and extremely long sentences, common extraction-oriented automatic text summarization methods should not be expected to show a remarkable performance when applied to patent documents. Other, more content-oriented or abstractive summarization techniques are able to perform much better and generate more concise summaries. This paper proposes an efficient summarization system for patents using artificial intelligence, natural language processing and deep learning techniques to condense the knowledge and essential information from a patent document into a single summary that is easier to understand without any redundant formatting and difficult jargon.