

Introduction

The information content of modern documents is vast and diverse, containing a wealth of information in the form of text. This text content is often structured in a hierarchical manner, with different levels of importance and detail. This hierarchical structure allows for efficient storage and retrieval of information.

The key to effective information retrieval is to be able to accurately and efficiently extract the relevant information from the document. This involves identifying and extracting the main topics and subtopics, as well as the supporting details that are necessary to understand the content.

The goal of information retrieval is to enable users to quickly and accurately find the information they need. This is achieved by using natural language processing techniques to understand the content of the text and to extract the relevant information.

The abstract:

Using language models for generic entity extraction