

Assignment No 4: Vector Space Search Engine

In this assignment, the students will design and develop a search engine based on a normalized vector space model that will provide the ranked retrieval of web pages. The search engine must have back-end and front-end components.

The back-end component contains:

1. The inverted index has the statistics required to compute the normalized weights.
2. Weight matrix of all documents and keywords.
3. Meta files containing information like title, description, URI, and thumbnails of images may be present on the web pages.

The front end provides:

1. Search User Interface (SUI) that enables query formulation and presentation of the search results over a Search Engine Results Page (SERP). The students can use Google SERP as a template.
2. The SERP provides the presentation of search results in a linear list along with pagination. The SERP presents the ranked results along with their relevance to the user query.