

Inverted Index

Sunny Kumar, Kandula.

CSCE 5200 Information Retrieval and Web Search

Dr. Haili, Wang.

9th September 2023

Objective

The current exercise at hand is to draw or represent the given documents using Inverted Index Data Structure.

- Doc 1 new home sales top forecasts
- Doc 2 home sales rise in July
- Doc 3 increase in home sales in July
- Doc 4 July new home sales rise

Solution

We know that an Inverted Index is a data structure that is used to search for text elements belonging to a document,

Where for each term 't', we store a list of all documents that contain 't'.

The unique words in given set of documents are:

new, home, sales, top, forecasts, rise, in, July, increase

Now, let's identify the document IDs to which each of these words belong to, individually:

- New, appears in 1 and 4.
- Home, appears in all the documents.
- Sales, appears in all the documents.
- Top, appears in 1.
- Forecasts, appears in 1.
- Rise, appears in 2 and 4.
- In, appears in 2 and 3
- July, appears in 2, 3 and 4.
- Increase, appears in 3.

Now let's draw the structure:

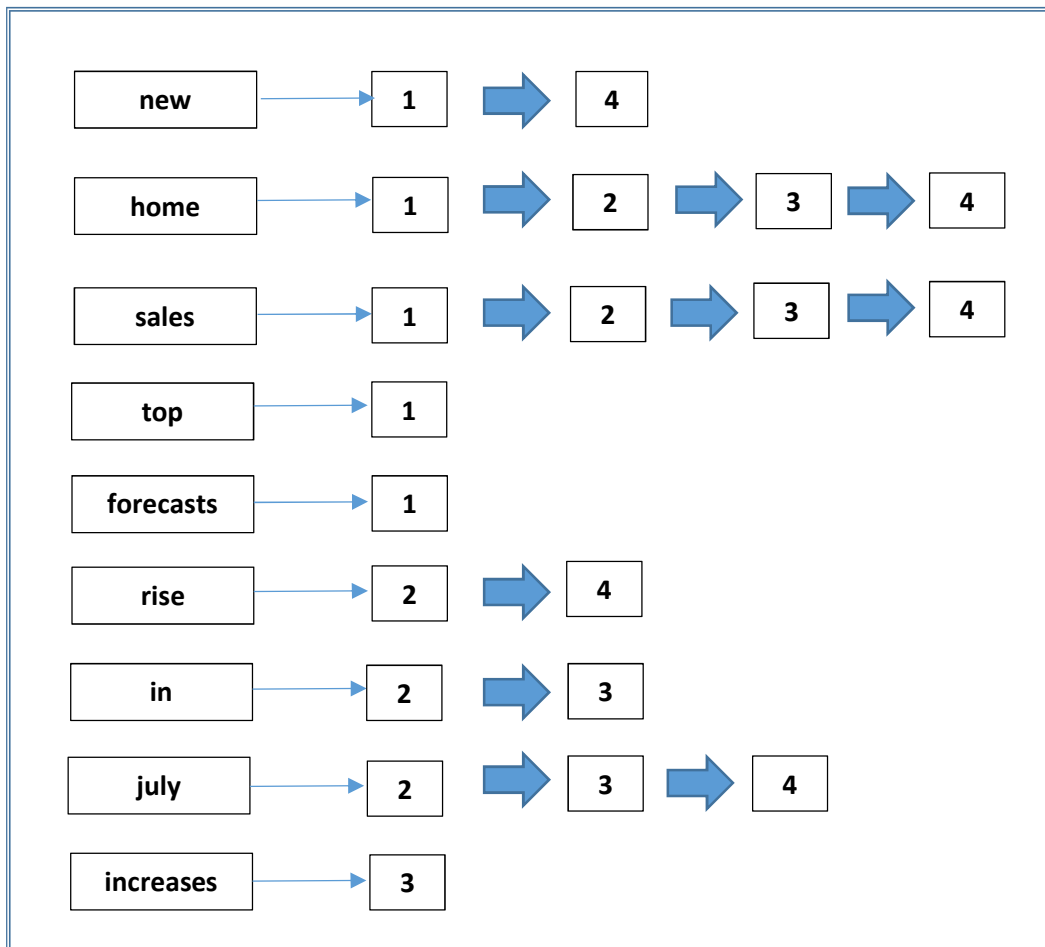


Figure: Inverted Index