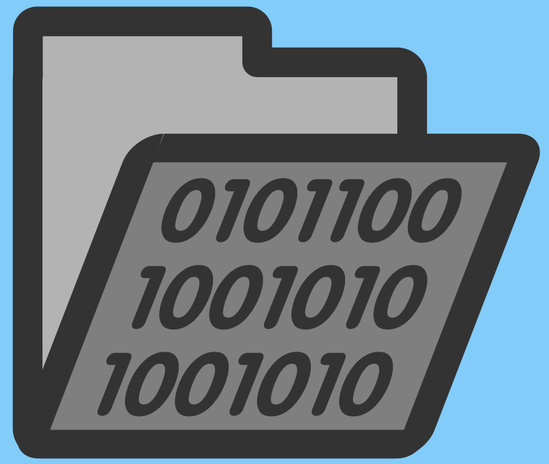
SEARCHDOC

TEAM
BINARYBEATS

THE TEAM - BINARYBEATS

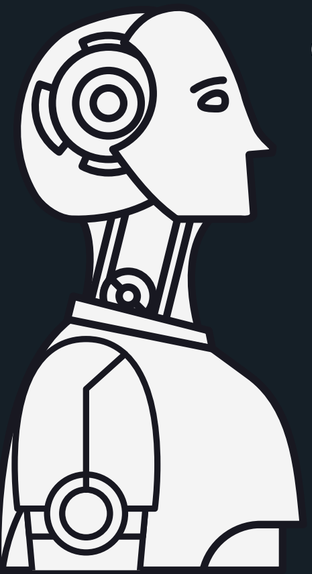
VIDHI BHAMARE
ADITI HINGE
SHREYA WATWE
ANNADA DASH

SY COMP
CUMMINS COLLEGE OF ENGINEERING, PUNE



PROBLEM STATEMENT

There are tonnes of documents out there containing a lot of information. There is a need for a web-based search engine or a bot that can quickly give us the answers by reading the document.

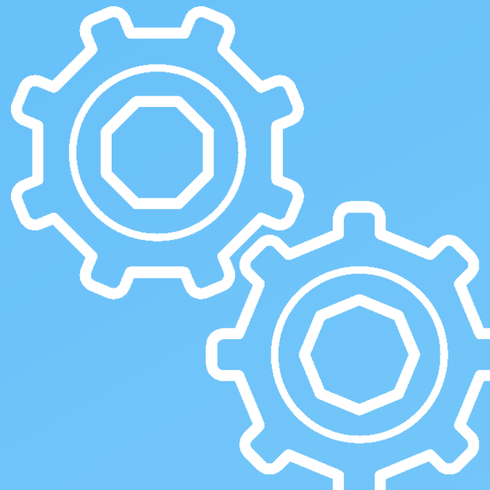


OUR APPROACH

- TAKE A QUERY FROM THE USER
- GIVE AN ANSWER TO THE PARTICULAR QUESTION IN SHORT
- TAKE A QUERY FROM USER TO SUMMARIZE A PARTICULAR TOPIC FROM THE DOCUMENT(USING SUMMARIZER)
- ANSWER THE QUERY OF SUMMARIZATION

HOW DID WE IMPLEMENT?

- We have used the Haystack framework for building search systems. It includes NLP for Search which picks components that perform retrieval, indexing, question answering, reranking, summarization, and much more.
- Implemented this on our own dataset, and searched for queries related to the document. What you get is a concise answer to your question.



TECHNOLOGIES USED

- HAYSTACK FRAMEWORK
- GOOGLE COLAB NOTEBOOK
- PYTHON
- ELASTIC SEARCH

DEMO

FUTURE SCOPE IN TECHNOLOGY

The model can be extended further to many search systems and in various domains. Following are some of the areas we think could use the model:

1. Add more features to the model - input files from user, random QnA generation.
2. Use deep learning models, to improve the quality of output.
3. Integrate the algorithm with chatbot, and provide it in online book reading platforms, like kindle.

FUTURE SCOPE IN REAL LIFE

3. Email services - to access what information was communicated through mails.
4. Generate random questions from text for preparation of various certifications like AZ 900.
5. Effective use of documentation of services Azure, Python, etc.
6.Quick reference to digital study materials.

CHALLENGES THAT WE FACED

REFERENCES

1. https://douran.academy/wp-content/uploads/ebooks/microsoft-certified- asure-fundamentals.pdf - AZ900 online notes
2. https://docs.haystack.deepset.ai/docs - Haystack documentation
3. https://sciresol.s3.us-east-2.amazonaws.com/IJST/Articles/2016/Issue-37/Article43.pdf - Research paper

Thank You!!