Shikhar
https://github.com/Shikhar-S

Email : shikharb@iisc.ac.in
Mobile : 7989962442

Education

- **Indian Institute of Science** — Bengaluru, India
  M.Tech (Research) in Intelligent Systems; GPA: 8.7/10.0 — *Aug. 2019 – Present*
- **Birla Institute of Technology and Science** — Hyderabad, India
  B.E. (Hons.) in Computer Science Engineering; GPA 9.74/10.0 — *Aug. 2014 – July. 2018*
- **City Montessori School** — Lucknow, India
  ISC; Percentage: 98% — *2014*

Experience

- **Myntra Designs Pvt Ltd.** — Remote, India
  Machine Learning Intern — *May 2019 - July 2020*
  - **Size Recommendation**: Implemented an iterative optimisation algorithm to correct product and user size measurements based on user purchase history, improving recommendation coverage by 10-23% and recommendation accuracy by 5-12% across article types while keeping the measurement values interpretable.
    Technologies: Python, Numpy, Pandas, Scikit-learn, Seaborn
- **Media Net** — Mumbai, India
  Platform Engineer — *Aug 2018 - July 2019*
  - **Advertisement Retrieval**: Put a modified Okapi BM25F algorithm into production for retrieving textual ads based on user search keywords, improving coverage over the previous regex-based algorithm, saving company person-hours while keeping the retrieval latency less than 80 ms.
    Technologies: Java, ElasticSearch, Hive, SQL, Word2Vec, TF-IDF
- **Indira Gandhi Centre for Atomic Research** — Kalpakkam, India
  Research Intern — *Summer 2015*
  - **Search Engine**: Implemented Information Retrieval methods like TF-IDF and Fuzzy Retrieval and compared their performance with precision-recall curves. Also developed a GUI for visualizing the results.
    Technologies: Python, SQL, Tkinter

Projects

- **Natural Language to Bash Translation**: Developed an algorithm for translating Natural Language to Bash commands by utilising command Abstract Syntax Tree and Manpage data, resulting in interpretable predictions beating baselines like T5 and Seq2Seq with attention.
  Technologies: Python, Bash, Pandas, Pytorch, Pytorch-Lightning, Seaborn, Transformers
- **Analysis of Competitive Codebases**: Analyzed programs from the competitive programming website Codeforces to predict programmer proficiency, achieving an F-score of 0.62 with logistic regression and Doc2Vec embeddings of linearized Abstract Syntax Tree. Gained insights into embedding relationships across programmers and problems.
  Technologies: Python, Doc2Vec, Scikit-learn
- **Multi-Aspect Sentiment Classification for Online Reviews of Medical Experts**: Reproduced results from Shi et al. on a publically available dataset. Also, visualised the attention distribution over reviews across aspects to obtain important words for the predicted sentiment. Results include an F-score of 0.35-0.41 across aspects.
  Technologies: Python, Matplotlib, Pytorch, Pytorch-Lightning

Publications

- **An extraction based approach to keyword generation and precedence retrieval**: Precedence retrieval involves ranking documents according to their importance for a given query. Developed an algorithm to perform keyword-based summarization of documents and then used these summaries to rank documents. Published in Working Notes of Forum for Information Retrieval Evaluation (FIRE) 2017.

Achievements

- **NTSE Scholarship**: Awarded by Govt of India.
- **INSPIRE Scholarship**: Awarded by Govt of India to top 1% percentile students in ISC.
- **Merit Scholarship**: Top 2% of the batch in BITS Hyderabad.