Akshay Satish

Bangalore, India | akshaysatish1998@gmail.com | 9538005413 | linkedin.com/in/akshay-satish
github.com/akshysatish

Summary

Enthusiastic Data Scientist with experience in working with Multi-Dimensional Data Sets to generate actionable, synthesized data insights. Established ability to solve real-world problems end-to-end, including data modeling, data manipulation, data mining, rigorous explorative data analysis, and machine learning techniques.

Education

Purdue University Aug 2021 – May 2024
Masters in Computer Science West Lafayette, IN

Nitte Meenakshi Institute of Technology Aug 2016 – May 2020
Bachelors in Technology, Computer Science Bangalore, India

Project & Research Experience

Reducing Hallucinations in LLMs, Purdue University Jan 2024 – Apr 2024

- Surveyed 20 medical datasets to identify problems and evaluate hallucination mitigation techniques
- Implemented a combination of prompt engineering and Retrieval Augmented Generation techniques
- Improved quantitative correctness of RAG-LLM responses from 17% to 31% and achieved 95% user-perceived correctness, highlighting the need for consideration in communicating LLM hallucinations

Implementation of a Document Database Jan 2023 - April 2023

- Collaborated with a team of developers over 4 months to design and implement a NoSQL document database from scratch using Python, supporting CRUD operations, B+ Tree indexing, and concurrency
- Improved query speeds by 50% with 200 microseconds to 1 second query time for sizes ranging from 10 documents to 1 million documents, respectively, for complex queries with filtering, projection, and MVCC
- Documented the database architectures and API, including examples of how to use the database for common use cases, to facilitate ease of use and adoption as libraries by other developers on the team

Predicting Earthquake Trends, Purdue University Jan 2022 – Apr 2022

- Extracted and analyzed geospatial data of 18000+ earthquakes to prepare for the prediction model
- Implemented DBSCAN to perform spatial clustering based on the intensity of earthquakes
- Predicted earthquake magnitude trends using LSTM model and visualized the results as a heatmap to communicate high-risk regions based on depth and magnitude

Graph-based keyword extraction Mar 2020 – May 2020

- Built NLP tool to extract relevant keywords and keyphrases from text data via LSTMs in PyTorch
- Applied k-core truss decomposition to build a word graph and extract related keywords to rank the top 10
- Decreased error by 8% as compared to supervised and TF-IDF methods

Technologies

Machine and Deep Learning: Pandas, Scikit-learn, Seaborn, PyTorch, Keras, Tensorflow, Pyspark

Programming: Python, SQL, R, C++, Java, Julia, Latex

Databases: SQL, MySQL, BigQuery, NoSQL

Other: Snowflake, DataBricks, Apache Kafka, Apache Spark

Publications

Graph-Based Keyword Extraction for Twitter Data June 2022

Emerging Research in Computing, Information, Communication and Applications Conference