

Mihir Kumar

+91-7428872048 | mihir.kumar@alumni.iitd.ac.in | [Mihir Kumar](#)

EDUCATION

Indian Institute of Technology Delhi

July'16 - July'20

B.Tech in Electrical(Power & Automation)

Coursework: Intro. to Computer Science, Data Structures and Algorithms, Signals and Systems, Computer Architecture, Digital Image Processing, Probability and Stochastic Probability, Embedded Systems

WORK EXPERIENCE

Data & Backend Developer, Soul AI

June'24 - Current

Software Engineer

Hyderabad, India

- Designed and developed scalable **multi-database SQL editor and console** in **Python**, streamlining AI research data collection. Used daily by 5,000 freelancers and internal team, handling over 100,000 queries efficiently.
- Implemented dynamic data lake connections and SQL language support, improving query platform functionality.

Data Science & Analytics, BlinkIt

Sep'20 - Sep'21

Data Engineer

Gurugram, India

- Developed and managed a robust data pipeline for the marketing team, leveraging **Python** and **MySQL** within **JupyterHub**. This pipeline efficiently updated the **Redash** database, ensuring seamless data access across teams.
- Created multiple data dashboards on **Redash** querying and structuring techniques to present valuable insights.
- Implemented parallel computations and optimized thread assignment, achieving a **10% reduction** in pipeline processing time. Further enhanced dashboard load times to reach **near-instantaneous** ($\sim 100\text{ms}$) performance.

Credit Risk Analyst, CitiCorp

May'19 - July'19

Summer Intern, Machine Learning

Mumbai, India

- Increased the risk discrimination of the variable to **4-fold** for corporate and **2-fold** for other commercial clients
- Leveraged Python and sophisticated machine learning algorithms using Random Forest classifier to accurately identify high-risk customers by analyzing transaction history. *Implemented in production and received a PPO*

PROJECTS

Data Deduplication On Cloud Storage || Java, Cloud Computing, Chunking, Hashing

- Developed a comprehensive solution for data deduplication in cloud storage, utilizing Java programming language and advanced cloud computing techniques.
- Implemented a range of chunking and storage algorithms including probabilistic extreme binning algorithms, fixed length chunking, variable length chunking. Seamlessly integrated algorithms like Rabin-Karp rolling hash and bipartite graph to optimize the deduplication process.

Vessel Segmentation Using Morphology || Python, Digital Image Processing, Morphological Operations

- Implemented a vessel segmentation pipeline in Python for analyzing retina maps. The project aimed to identify and map the vessels of the eye while removing optic disks. To achieve this, Applied morphological operations to smooth the veins, examined the contours, and employed relevant conditions for accurate selection.
- Implemented operations that filtered out any patches or non-vein patterns, ensuring the reliability of the segmentation process.

Solving 8-Puzzle Game || Java, Dijkstra's Algorithm, Graph Traversal

- Conceptualized a graph-based approach, where the possible states of the puzzle were treated as nodes, and the valid moves as edges, effectively representing the puzzle as a graph. I then employed Dijkstra's Algorithm for efficient graph traversal to find the optimal path from a given random start state to an assigned random end state.

Nursing Home Web Portal || Javascript, MySQL database

- Designed the front-end & database for a comprehensive web application to search, view & compare nursing homes
- Implemented 4 types of users with different access privileges.Used views,indexes & triggers for efficient execution

SKILLS & INTERESTS

Technical Skills	C/C++, Python, Bash, MATLAB, SQL, NodeJs, Java, SAS
Softwares	AWS, Eclipse, Arduino, Autodesk, Tina TI, SAS EG, Tectia, Jupyter
All India Rank	Obtained a rank 776 out of 1.5 million students across India in JEE (advanced)