Akash Mishra

New York | akash.mishra13@gmail.com | +19294064445 | http://akashmishra.me

Current Graduate Student with over 4 years of work experience

GitHub: - https://github.com/AkashSky1994 LinkedIn: - https://www.linkedin.com/in/akash-mishra1994

Education

New York University, New York — *MS (Computer Engineering)* **September 2021 - May 2023**

Data Science | Machine Learning | Deep Learning | Computer Vision | High Performance Deep Learning | ML in Cybersecurity | Internet Protocol

APJ Abdul Kalam University, India — *BTech (Computer Science and Engineering)* **April 2013 - May 2017**

Experience

Thrasio, NY — ***SDE Intern*** **June 2022 - August 2022**

- Prototyped data ingestion pipelines handling 20 billion rows for Amazon using Snowflake, S3, SQS.
- Implemented the deduping process for such large tables using dbt tool.
- Designed a service to handle incoming data from various sources that helped reduce the developer resources allocated by 50%.

LBB (Little Black Book), India — *Software Engineer* **September 2019 - July 2021**

- Developed search service leveraging Elasticsearch and NLP techniques to provide support for conversational language improving organic user engagement by 60% over the next 6 months.
- Led a dev team to build a smart plug-n-play logistics service using NodeJS, Redis Pub-Sub and OOPs Design Pattern leading to optimized delivery routes and partners consequently reducing the cost by around 25% and average delivery time by at least 1 day.
- Collaborated with finance team to design a distributed finance Ledger microservice which records all microtransactions within the company's ecosystem automating manual finance work, optimizing refund, and payout process reducing the overall cost by 30%.

Bookchor, India — *Software Engineer* **July 2017 - August 2019**

- Built a book recommendation system by implementing Singular value decomposition using prior reading preferences.
- Identified key problems and bottlenecks within the existing architecture and took the initiative to improve the overall performance and scalability of web services by implementing a distributed cache system at the application level resulting in improved API response time by up to 100ms.
- Spearheaded an initiative to propose usage of NoSQL database to the management and later worked on MongoDB for storing unstructured data for analytics and search engine usage.

Projects

Depth Estimation using Self-Supervised Learning from Monocular Images [Link]

- Researched and trained a self-supervised learning model for depth estimation on KITTI dataset by experimenting with various iterations of UNet Architecture involving Encoders (ResNet, ConvNext) and Up-sampling decoders like ESPCN achieving a rms of 3.7
- Enhanced our model using Mask R-CNN pretrained on COCO Dataset to get semantic map which helped us reduce artifact problem in the disparity map and consequently reducing the overall loss for the model to 3.485

Asynchronous Batch Operations NPM Package [Link]

- Authored an Open Sourced NPM Library for asynchronous batch operations which provided a 17% improvement in the execution time and better handling of error catching and response compared to the native Promise.all Method.

Hateful Meme Detection using Multimodal neural network [Link]

- Created various iterations of multimodal neural network to identify hateful intent from the large corpus of memes offered as a part of NeurIPS 2020 Challenge (Proposed by Facebook AI)

Data Analysis for 2018 Stack overflow Developer Survey Data [Link]

- Formulated a compensation model using regularized regression on Salary data from Stack overflow (100k, 300 Column dataset).
- Experimented with various Machine Learning techniques which included PCA, KMeans Clustering and Classification (Random Forest, Support Vector Classifier, Logistic Regression) on a stack overflow survey data

Recommender System Using Spark's ALS Model and LightFM Model [Link]

- Built and evaluated a recommender system on MovieLens large dataset using PySpark's ALS module.
- Proposed a baseline popularity model for cold start strategy.
- Extended the system by implementing and evaluating on LightFM and contrasting the performance with ALS model.

Skills

- ***Language*** **-** Nodejs, Golang, Python, JavaScript **|** ***Database*** **-** MongoDB, SQL (MySQL, PostgreSQL)
- **Other Tools -** Elasticsearch, Solr, AWS (EC2, Lambda, Kinesis, SQS, S3), GCP, Redis, Git, Docker, Snowflake
- **Frameworks:** PyTorch, Scikit learn, SciPy, NumPy, Computer, Pandas, TensorFlow, OpenCV, PySpark, Dask, dbt