Saksham Bassi

About / LinkedIn / Github / 737-288-2873 / sb7787@nyu.edu 40 Newport Parkway Jersey City, NJ

Education

- **New York University - Courant Institute of Mathematical Sciences** GPA - 3.625/4.0
  Master of Science in Computer Science - focus on Machine Learning *May 2023*

Skills

- **Languages**: Python, C++, Node.js, Java, Angular, HTML, CSS, R, Git
- **Databases**: SQL, MySQL, MongoDB, CosmosDB, DynamoDB, PostgreSQL
- **Tools**: HuggingFace, PyTorch, Keras, Tensorflow, Apache Spark, Airflow, Flask, Google Cloud Platform, Matplotlib
- Recommendation Systems, Natural Language Processing, Machine Learning, Computer Vision, Time-series Modeling

Experience

- **CILVR Lab, NYU** New York, New York
  Research Assistant *Sept 2022 - Present*
  - NLP research on improving generalization in cross-lingual models with Duygu Ataman and Kyunghyun Cho.
- **Amazon** Seattle, WA
  Software Development Engineer Intern *May 2022 - Aug 2022*
  - Implemented a new notification service that sends push notification when a payment is failed in Amazon Care.
  - Designed a fast and simplified system design for the service, and wrote unit tests to make it fault-tolerant.
  - Aim to fix 95% failed payments using this service. [*Java, AWS Lambda, DynamoDB*]
- **Glance, InMobi Group** Remote
  Data Scientist II *Apr 2021 - Aug 2021*
  - Built exploration-feed ML infrastructure for millions of daily users to improve recommendations.
  - Integrated Node2Vec model to learn user embeddings and automated execution through Airflow.
  - Improved user engagement by 30% using recommendations from top-k creators. [*Python, PySpark, Airflow*]
- **HSBC** Pune, India
  Software Engineer *July 2019 - Apr 2021*
  - Devised Logging architecture to monitor application and on-premise logs for critical alerting.
  - Saved $3 Million annually which were spent on proprietary logging software. [*Fluentd, Grafana, GCP, Python*]
  - Created alert policies & automated VM processes using Python APIs for GCP.
  - Spearheaded development of a system to map and send Financial documents to clients plus mailing service.
- **Tata Institute of Fundamental Research** Bangalore, India
  Machine Learning Research *Jan 2019 - June 2019*
  - Performed modeling and factorization of time-series data using statistical features in domains like astronomy, finance, mathematics and commodity prices. [published paper]
- **Inter-University Center for Astronomy and Astrophysics** Pune, India
  Deep Learning Research *Feb 2018 - Dec 2018*
  - Formalized a method to classify variable star classes in sky surveys using CNN-LSTM models and eliminated the need of pre-processing large time-series datasets. [published article]
  - Accelerated computational time by a factor of 10 and produced instant results from raw light curves.

Publications

- **A learning algorithm for time-series based on statistical features**. *International Journal of Advances in Engineering Sciences and Applied Mathematics* [Link]
- **Deep Learning Diagnosis of Skin Lesions**. In *International Conference on Computing, Communication and Networking Technologies* [Link]

Projects

- **Causal-inference**: Deep models (Residual & Transformers) for high-dimensional simulated causal effect [code]
- **HuBERT-disfluency**: Working on End-to-end disfluency removal using audio + fine-tuned BERT model
- **Audio Denoiser**: Developed Autoencoder to denoise audio signals and implemented CNN on its spectrograms to qualitatively assess industrial processes. [*Python, Flask, Keras, HTML, CSS*]