Aziz Shameem

azizshameem.2002@gmail.com | Aziz-Shameem | aziz-shameem

Education

Indian Institute of Technology Bombay — Mumbai, India
B.Tech in Electrical Engineering and M.Tech in Artificial Intelligence ***(Cumulative GPA: 9.68/10.0)*** — *Nov'20 – Present*

Publications

- **A. Shameem**, B. Kohli, A. De, "**OPAS: A Neural Approach to Subsequence Matching with Optimized Alignment-Bitstrings**", under review at AISTATS, 2025
- M. Shukla, **A. Shameem**, M. Salzmann, A. Alahi, "**Towards Self-Supervised Covariance Estimation in Deep Heteroscedastic Regression**", under review at ICLR, 2025
- V. Prasad, **A. Shameem**, C. White, S. Nayak, P. Jain, P. Garg, G. Ramakrishnan, "**Speeding up NAS with Adaptive Subset Selection**", AutoML 2024

Research Experience

Variational Auto-Encoders and Diffusions Models — *May'24 – Present*
Summer Internship, Guide: Prof. Alexandre alahi — *EPFL*

- Performed extensive analysis of the **similarities and differences** between two types of Generative Frameworks - **Variational AutoEncoders** and **Diffusion Models**
- Designed experiments to elucidate the **limitations** of VAEs, and suggested **improvements/methods** to overcome them
- Demonstrated methods for better generations by VAEs on **synthetic** as well as **benchmark datasets**
- Working on incorporating the methods for improving **conditional image generation using VAEs**
- Performing analysis of **Latent Diffusion model** used in **Stable Diffusion v2**, to recognize areas of improvements

OPAS : Ordered Permutation-based Alignment for Sequence matching — *Jun'23 – Oct'24*
Guide: Prof. Abir De — *CSE, CMInDS, IIT Bombay*

- Developed a **theoretical framework** based on **lagrangians** for **fast, approximate, constrained sequence matching**
- Tested the developed framework **computationally**, and obtained results on datasets of several **modalities**
- Compared the developed method with several baselines, and demonstrated over **58x speedup** over them
- Compared several **implementation variants, architectures and models** of the proposed system as ablations

Neural Architecture Search and Data Subset Selection — *May'23 – Oct'23*
Guide: Prof. Ganesh Ramakrishnan — *CSE, IIT Bombay*

- Enhanced existing **pruning**-based Neural Architecture Search, experimenting with proven pruning techniques
- Incorporated **adaptive subset selection** to enhance the efficient use of data and obtained competitive results in the same

Adversarial Domain Adaptation — *Jan'23 –May'23*
Guide: Prof. Amit Sethi — *EE, IIT Bombay*

- Developed the complete codebase for fitting a novel model - **WaveMix** for use in a task involving **Domain Adaptation**
- Enhanced the performance of the network to increase its **accuracy** and **generalisability** to unseen datapoints

Professional Experience

Modem Systems Engineering Intern — *May'23 – Jul'23*
Qualcomm Technologies Inc. — *Bengaluru, India*

- Undertook extensive research on enhancing **information-deficit channel estimation** methods for **5G**
- Implemented existing as well as proposed technology in a **python simulator** and verified theoretical results
- Implemented specific **channel models** in the python simulator, to be used for future research and developement

Scholastic Achievements

- Current holder of **Department Rank 1** in **Centre for Machine Intelligence and Data Science** (*Present*)
- Ranked **1st** in Electrical Engineering Dept. (Dual Degree) out of a batch of 100 students for two years (*2022-24*)
- Recipient of the **Summer@EPFL Travel Scholarship** (awarded to 1.6% of the applicants) (*2024*)
- Awarded **Institute Academic Prize** for exceptional performance in academics for **two consecutive years** (*2022, 2023*)
- Secured an **All India Rank of 62** in JEE Mains (Engineering) among 1.3 million candidates (*2020*)
- Secured an **All India Rank of 665** in JEE Advanced among 0.2 million candidates (*2020*)
- Conferred with **AP grade (Advanced Performer)** in **Advanced ML, Linear Algebra, Electronic Devices**
- Ranked in the **national top 1%** in NSEC and selected to appear for Indian National Chemistry Olympiad (*2019*)

Key Technical Projects

Auto-Grad Engine and GPT developement *May'23 – Jul'23*

Self Project *IIT Bombay*

- Built an **Auto-Grad Engine** having **pytorch-like API** from scratch and demonstrated its working on a test bed
- Preformed extensive study on **innate problems** in neural networks pertaining to **gradients and dead neurons**
- Implemented initialization and normalization based methods to bypass problems and demonstrated better results
- Developed a working **GPT-like architecture** from scratch and generated recognizable **Shakespearean text**

Toxicity Detection in Large Language Models *Apr'24 – May'24*

Guide: Prof. Pushpak Bhattacharya(CS772 : DL for NLP) *IIT Bombay*

- Trained a **text toxicity classifier** using **Recurrent Neural Nets**, **LSTMs** and **Transformers**
- **Engineered additional features** from the analysis of textual data in an effort to improve classification of toxicity in texts
- Utilised the trained classifier network as a **filter** to mitigate toxicity in the output of **Large Language Models**

Out of Distribution Detection *Apr'23 – May'23*

Guide: Prof. Sunita Sarawagi(CS726 : Advanced ML) *IIT Bombay*

- Performed a literature survey on **energy-based Out-Of-Distribution Detection** on Computer Vision and Image Datasets.
- Replaced the proposed loss function with a novel **Ranking Loss**, and displayed its superior performance.
- Proposed changes to the documented architecture in an effort to apply the study to **text datasets**

Correlated Multi-Armed Bandits *Mar'23 – Apr'23*

Guide: Prof. Jayakrishnan Nair, Prof. D. Manjunath (EE6106: Online Learning and Optimization) *IIT Bombay*

- Performed an extensive literature review on the formulation of the **correlated multi-armed Bandit** problem
- Applied an existing framework for the **exploitation of correlation** between bandits, on different algorithms and demonstrated **substantial improvements** in the same
- Designed effective visualizations for **comparing** the implemented algorithms and **identification** of the **best arm**

Music Genre Classification using ML and DL *Apr'22 – may'22*

Guide: Prof. Biplab bannerjee(DS303: Intorduction to Machine Learning) *IIT Bombay*

- Implemented **Decision Trees**, **Random Forest**, **Naive Bayes**, **SVM** and **KNN** to classify music based on its genre
- Tried several architectures and hyper-parameters of **Sequential Neural Networks** to classify music based on genre

Machine Learning for COVID-19 Data Analysis *Oct'21 – Nov'21*

Guide: Prof. Amit Sethi(DS203: Programming for Data Science) *IIT Bombay*

- Performed extensive **Exploratory Data Analysis** on **COVID-19** in India and several other countries across two years
- Used **Linear, Polynomial, Lasso, Ridge Regression models** to predict future COVID-19 related casualties
- Implemented **Logistic Regression, SVM, NN, RF** and **GBC** to predict need of ICU admission based on medical factors

Gas Leakage Detection and VSLAM using Nanosaur *Jan'23 – Apr'23*

Guide : Prof. Siddharth Tallur (EE344 : Electronic Design Lab) *IIT Bombay*

- Ideated and designed a mobile robot capable of **toxic gas detection** and **3D reconstruction** of its environment
- Carried out the design and testing of the **Printed Circuit Board** for interfacing the gas sensors with Jetson Nano
- Created **ROS Nodes** for efficient **sensor data accumulation and display**, and got **recognized** for the same

Expense Tracker *May'21 – Jul'21*

Python is Cool, Kids (PyCK) : Summer Course *IIT Bombay*

- Developed an **application** capable of storing and plotting users' expenses in **Python**
- Created an interactive **Front-End** for the project to enable **account creation** and **data entry** into the system using tkinter
- Used Python Libraries **Pandas** and **Matplotlib** to create a **Back-End** to store and revert data as and when required

Other Projects

- Carried out Extensive Data Analysis of **IPL**, and developed systems to predict **scores and outcomes** of matches (*2021*)
- Created basic games like **zig-zag** and **candy-cush** using **Unity and C#**, as part of an online course on GameDev (*2021*)
- Designed and simulated the Schematic and layout of common Amplifiers and Ring oscillator on Cadence (*2022*)
- Designed and implemented the **datapath** and **control logic** for a **16-bit multi-cycle processor** in HDL (*2022*)
- Designed and Impemented a **PID controller** for a **Line Following Robot** as part of Control Systems Laboratory (*2022*)

Technical Skills

Data Analysis/Visualization tools : Pandas, Numpy, Matplotlib, seaborn, plotly, MATLAB, MySQL
Libraries: Pytorch, PyTorch-Geometric, Tensorflow, Numpy, Pandas, SciPy, Seaborn, Scikit-Learn, OpenCV, Matplotlib
Software: Git, LaTeX, MATLAB, OpenCV, GNU Radio, Quartus, NGSpice, Kicad, Arduino, Pycharm, Anaconda, Unity

Positions of Responsibility

Machine Learning Engineer
Data Analysis and Visualisation Team, IITB *Jun'22 – Apr'23*

- Led a team to **successful completion and publication** (On the official public handle) of analysis on **Indian Establishments and Settlements**
- Performed extensive analysis on **Grading Statistics** and **Semester Exchange statistics** of the institute and oversaw its **publication on LinkedIn**
- Implemented the **YOLO version-1** framework from scratch as part of a study project and gave a presentation on the same
- Represented team at a kaggle competition which involved **data pre-processing, model design and training** on NLP tasks
- Delivered a comprehensive presentation on **Loss Functions for Image Restoration** and got recognized for the same

Mentor : Introduction to Machine Intelligence
Seasons of Code, IITB *Apr'22 – Jul'22*

- Prepared and distributed **comprehensive resources** to help beginners understand and appreciate concepts in ML
- Assisted beginners in building trivial **AI systems** using **Reinforcement Learning** in an effort to generate interest in the field

Junior Control Systems Engineer
Hyperloop, IITB *Oct'21 – Apr'22*

- Learned and implemented **sensor data collection** algorithms in MATLAB, based on **Kalman Filter**
- Studied and applied **sensor fusion** to efficiently collect and utilize in-flowing data from designated sensors
- Learned and applied concepts on **PCB design** and **device interfacing** to build compact controllers

Web Developer/Coordinator
Techfest - Asia's largest science and technology festival *Oct'21 – Feb'22*

- Contributed to the creation of **TechFest Website** by developing **hover animations, Bootstrap cards and Carousels**
- Improvised and optimised existing code-base by appropriate fixes and restructuring

Key Coursework

Machine Learning	Advanced ML | Deep Learning for NLP | Learning with Graphs | Mathematical Optimization Techniques | Optimization in ML | Online Learning and Optimization | Introduction to Machine Learning
Commputer Science	Programming for Data Science | Image Processing | Design and Analysis of Algorithms | Data Structures and Algorithms | Digital Systems | Microprocessors | Computer Programming and Utilization
Math and Probability	Markov Chains and Queueing Processes | Applied Linear Algebra | Advanced Probability and Random Processes | Matrix Computations | Estimation and Identification
Online Certifications	Reinforcement Learning Specialization | Python for Computer Vision and Deep Learning | Machine Learning | Unity C# Scripting

Extra Curricular Activities and Other Achievements

TAship	Complex Analysis — Introduction to Machine Learning Linear Algebra — Programming for Data Science
Chess	Secured **First Place** in **9-Up**, a tournament organized by Hostel-9 sports council Represented **Hostel 9** in the **General Championships** at IITB for two consecutive years Received a trophy for **Hounourable Mention** at the **1st Tarabai Shinde Chess Tournament**
Coding	Stood **third** in an **Algorithmic Trading hackathon** organised by **Web and Coding club, IITB**
Tennis	Represented **Hostel 9** in the **General Championships** at IITB for two consecutive years
Miscellaneous	Appointed as the Class Representation of the CMInDS IDDDP batch of 2020 Regular participant of the 5km inter-hostel marathon runs conducted twice a year