Shreyansh Singh

+91-9654070844 shreyansh.pettswood@gmail.com shreyansh26 shreyansh26 Homepage

Education

IIT (BHU) Varanasi, India **Jul 2016 – May 2020**

Bachelor of Technology (B.Tech.) in Computer Science and Engineering *CGPA - 9.57/10.0*

Member of the Club of Programmers and founder of the InfoSec Club

Experience

Mastercard (AI Garage) **Aug 2020 – Present**

Analyst - Artificial Intelligence *Gurugram, Haryana*

- Formulated and developed a graph-based representation learning algorithm for cards and merchants to detect transaction frauds which showed considerable performance boost ($\approx$ 6% AUCPR) over existing algorithms and demonstrated good trade-off in terms of training time vs. performance. The work resulted in a successful research publication.
- Designed and implemented a novel memory-efficient ($\approx$ 80% reduction in memory usage than the state-of-the-art model) tabular GAN architecture for the enhanced Threatscan product by Mastercard. The model is capable of handling datasets with large cardinality. This work resulted in a successful research publication.
- Currently working on an AI-based approach to improve the security of Mastercard APIs from a behavioral usage standpoint along with anomaly detection.

Samsung Research Institute - Bangalore **May 2019 – Jul 2019**

Summer Intern *Bengaluru, Karnataka*

- Implemented and simulated the MAS5G architecture, a new 5G mobility scheme, published in IEEE FiCloud, 2019.
- Locally deployed and tested a proof-of-concept version of the architecture using Kubernetes, Cassandra and Node.js.

C3i Center, IIT Kanpur **Dec 2018 – Jan 2019**

Research Intern *Kanpur, Uttar Pradesh*

- Developed a system to classify Linux executables as malware or benign using static and dynamic analysis techniques.
- Achieved $\approx$ 96% accuracy for the task and deployed the entire pipeline on their internal Malware Analysis system.

Innoplexus AG **May 2018 – Jul 2018**

Data Science Intern *Pune, Maharashtra*

- Worked on an OCR + NLP-based model to extract and label segments of text from PDFs. Completely revamped the existing pipeline to make an 80% faster and more accurate ($\approx$ 92%) system.
- Experimented with image processing methods and Faster-RCNN model for detection and extraction of tables from PDFs.

Publications

- **MeTGAN: Memory efficient Tabular GAN for high cardinality categorical datasets** - ***Shreyansh Singh****, Kanishka Kayathwal, Hardik Wadhwa and Gaurav Dhama* at the 28$^{\text{th}}$ International Conference on Neural Information Processing (ICONIP), 2021
- **CuRL: Coupled Representation Learning of Cards and Merchants to Detect Transaction Frauds** - *Maitrey Gramopadhye***,* ***Shreyansh Singh*****, Kushagra Agarwal, Nitish Srivasatava, Alok Singh, Siddhartha Asthana and Ankur Arora* at the 30$^{\text{th}}$ International Conference on Artificial Neural Networks (ICANN), 2021 (* $\equiv$ Equal contribution)
- **IIT (BHU) Varanasi at MSR-SRST 2018: A Language Model Based Approach for Natural Language Generation** - ***Shreyansh Singh****, Avi Chawla, Ayush Sharma and A.K. Singh* in Proceedings of the 1st Workshop on Multilingual Surface Realisation at the 56$^{\text{th}}$ Association for Computational Linguistics (ACL), 2018

Projects

Annotated ML Papers | Blog **Apr 2021 - Present**

- Annotated research papers from the field of natural language processing (NLP) and uploaded to Github to facilitate reading research papers less daunting for newcomers in the field.
- Wrote blog posts to accompany each annotated paper to present the ideas of the paper concisely. The initiative received good and encouraging feedback from readers. The GitHub repository currently has 75 stars.

ML Optimizers in JAX | *Python, JAX* **Jun 2021**

- Implemented some popular ML optimizers from scratch and used them with the MSE loss to train a simple linear regression model on a dataset of 5 features.
- Utilized JAX to perform the differentiation of the loss function w.r.t to the weights and the bias without explicitly writing their derivatives as a separate function to improve generalizability for other loss functions.

Privacy-preserving Deep Learning for Medical Image Classification | *Python, Tensorflow* **Nov 2019**

- Implemented a system to get deep learning model predictions on chest X-ray images for pneumonia classification in a privacy-preserving manner.
- Secure predictions involved the use of secure multiparty computation. Differential privacy was used to preserve privacy.
- The trained and encrypted model was deployed locally to serve predictions which could be queried by the client.

Technical Skills

Languages: Python, C/C++, Javascript, SQL, HTML/CSS, Bash
Technologies/Frameworks: Pytorch, Tensorflow, JAX, PySpark, Git, Django, Docker, Kubernetes

Achievements/Extracurriculars

- US patent on a reinforcement learning and natural language processing based approach to recommend relevant charities to people based on the news article they are reading.
- Earned silver medal for ranking in the top 5% (115^{th} among 2426 teams) while participating solo in the Kaggle Shopee Price Match Guarantee competition.
- Ranked 55^{th} (top 10%) in the Multi-dataset Time Series Anomaly Detection challenge, KDD Cup - 2021.
- Ranked 8^{th} in AI Blitz#6 and 10^{th} in AI Blitz#7 competitions organized by AIcrowd.
- Ranked 15^{th} in CryptoHack CTF, a modern-day cryptography focused Capture the Flag event.
- Recipient of the student scholarship to attend Black Hat Asia 2019 in Singapore in which 100 students were selected from 82 countries.
- Event coordinator and problem setter for the Capture the Flag event of Technex'19, the technical fest of IIT (BHU) Varanasi and Codefest'19, the departmental fest of the CSE department.

Scholastic Achievements

- Secured all India rank 576 in JEE Advanced 2016 among 0.2 million candidates and all India rank 125 (99.99 percentile) in JEE Mains 2016.
- Secured all India rank 116 in the Kishore Vaigyanik Protsahan Yojana (KVPY) examination 2015.
- Awarded NTSE scholarship through National Talent Search Examination (NTSE) 2014 wherein 1000 meritorious students are selected at the class 10^{th} (all India) level.
- Top 1% ($\approx$ top 300) in India in each of the National Standard Examinations in Physics, Chemistry and Astronomy (NSEP, NSEC, NSEA) in 2015 and 2016.