PRANAV BANSAL

+1 (425) 279-3200**|**pranavbansal007@gmail.com**|**LinkedIn**|**GitHub**|**Portfolio**|**Seattle, WA

TECHNICAL SKILLS

- **Programming Languages**: C/C++ (8+ years), GO, HTML/CSS, JavaScript, Java, MATLAB, Python, Ruby, SQL/NoSQL
- **Database:** MySQL, PostgreSQL, Google BigQuery, MongoDB, IBM DB2, DynamoDB
- **Technologies:** AWS, CI/CD, Data Analytics, Django, Docker, ExplainableAI, Git, GraphQL Hadoop, Jenkins, Jira, Kafka, Kernel Programming, LLM, Microservices, Natural Language Processing, Node.js, RESTful API and Spring Boot
- **Libraries:** TensorFlow, Keras, Pandas, Scikit-learn, NumPy, PySpark, PyTorch, Matplotlib, OpenCV, SciPy, Steamlit, NLTK

EXPERIENCE

Stony Brook University   *Jan 2024 - Feb 2025*
Research Assistant - GenAI

- Engaged in Explainable AI research on Vision Transformers with **12 attention layers**, reducing model bias and adversarial prompting risks by **30%** through targeted interventions.
- Applied ML techniques (LogitLens, AttentionLens, and TunedLens) to identify the top classes within intermediate layers and Generated heatmaps using Layer-wise Relevance Propagation (LRP), improving model **interpretability by 25%**
- Increased model transparency and reduced adversarial vulnerabilities by **35% through** systematic bias detection, verified using tests like the SHAP score, adversarial robustness tests, and fairness metrics

Brysk   *Nov 2021 - May 2023*
Software Developer – Big Data
Brysk, develops an AI-based grocery retail platform for autonomous checkout stores, similar to Amazon Go Stores.

- Engineered scalable **ETL Data Pipelines** using Google BigQuery, Teradata and Amazon S3 to generate data warehouse, later produced interactive dashboards for retail store inventory and sales reports, reducing manual effort by **4 hours**.
- Optimized already existing ETL jobs for Big Data Processing using **Hadoop** and **Spark** for scalable data pattern and distributed data processing, resulting in **24%** reduction in latency for orders report creation and loading time.
- Designed and deployed Computer Vision AI model for object and action tracking in real time, optimizing **accuracy by 25%** by incorporating features from IoT device inputs including weighing scale, infrared and pressure sensors.

Indian Institute of Technology, Kanpur (IIT-K)   *May 2021 - Aug 2021*
Research Intern – Full Stack

- Developed ML Models: SIR, ARIMA and LSTM based, to **forecast COVID-19 Cases** sourcing data from the India-level time series dataset, consisting of **60M** records.
- Conducted a comparative data analysis of the various ML models, determining that SIR (Susceptible-Infected-Removed) achieved the highest accuracy of **97.6%**.
- Executed frontend and backend for a website, consisting of COVID-19 cases forecast graph for **28** Indian states using **GraphQL** APIs with a **Node.js** backend and a **React** frontend.

PROJECTS AND PUBLICATIONS

Peer-to-Peer File Backup System   *Jan 2024 - Mar 2024*

- Implemented a distributed P2P file storage and backup system (with peers in the same IP network) with maximum backup speed of **410 KB/sec** and retrieval speed of **300 KB/sec**.
- System has features to increase reliability and availability of sensitive data derived from the priority of files being backed up. **Exposure**: Multi-Tiered Systems, Networking, Distributed Storage System, C++, Python.

Crop Disease Detection Using Neural Network and Machine Learning Algorithms   RESEARCH PAPER: IJIRE

- Implemented and compared multiple Neural Networks (including CNN, AlexNet, ResNet, VGG16, VGG19 and InceptionV3) to detect 14 Crop diseases on a PlantVillage Dataset of **30k** images. **AlexNet** excelled with a **99.2%** accuracy.
- Proposed approach employed a combination of NN models to detect **14 diseases** with **98%** accuracy for **24 disease classes.**

Speculative Execution of Distributed System with Commit and Rollback   *Aug 2023 - Dec 2023*

- Implemented **3** custom scalable system calls in the Linux Kernel to enable speculative execution, allowing processes to run without waiting for dependencies, thus increasing scalability and reducing latency by 30%.
- Developed custom application to support these new system calls, resulting in a **2.5x** increase in system speed due to speculative execution. **Exposure**: API, Kernel Programing, Distributed Systems, Operating Systems, gdb, Valgrind

Distributed Storage System with RAFT   *Aug 2023 - Dec 2023*

- Implemented the Raft consensus algorithm to manage replicated state machines, achieving **99.9% data consistency** and fault tolerance in a distributed system with up to 5 nodes.
- Designed and developed a sharded key-value store that handled over **10k requests per second**, ensuring high availability and scalability using the Raft protocol. **Exposure**: API, Kernel Programing, File System, Distributed Systems.

EDUCATION

STONY BROOK UNIVERSITY, New York   *Aug 2023 – May 2025*
Master of Science in Computer Science   ***GPA: 3.85/4***
Coursework: Operating System, Analysis of Algorithms, Data Science Fundamentals, Storage System, Theory of Databases

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, India
Bachelor of Science in Computer Science   ***GPA: 3.7/4***
Coursework: Data Structure and Algorithms, **Distributed System**, Software Engineering, Compiler Design, Computer Architecture, Computer Network, Computer Graphics, Object-Oriented Programming, Database Management System