# Triyasha Ghosh Dastidar

New York City, NY
212-920-0905 tg2936@columbia.edu linkedin.com/in/triyasha-gd github.com/triyashaCodes

## Education

**Columbia University** **Sep. 2024 – Dec 2025**
*M.S. in Computer Science; NLP, CV, Practical DL, Applied ML, Cloud Computing; CGPA: 4.22* *NY, USA*

**BITS Pilani** **Sep. 2017 – Jul. 2022**
*B.E. in Computer Science and Engineering and M.Sc.in Chemistry; CGPA: 9.11* *Hyderabad, India*

## Technical Skills

**Programming and Scripting**: Python, Golang, C++, Java, SQL, JavaScript (AngularJS)
**Machine Learning and AI**: PyTorch, TensorFlow, Keras, scikit-learn, Matplotlib
**Cloud and DevOps**: AWS (S3, EC2, RDS, SQS, SNS), Google Cloud Platform (GCP), Docker, Kubernetes, Jenkins
**Development Tools**: Git, VS Code, Jira, Linux

## Experience

**Nutanix** **Aug 2021 – Jul 2024**
*Software Developer* *Bangalore, India*

- **Won 10th Global Nutanix Hackathon** among 250+ teams by developing a **smart personalized security recommendation system** and an **AI model for password vulnerability**, now advancing to a commercial product.
- **Engineered scaled password management and synchronization APIs** in **Golang and Python**, ensuring secure storage and transfer between clusters, **decreasing security breaches by 20%**.
- **Designed and implemented** a **crash-resistant, role-based multi-party authorization workflow**, leveraging **Write-Ahead Logging (WAL)** to increase protection against ransomware attacks by requiring up to **three user approvals** for critical operations.
- **Led migration of 50+ APIs** to a **new gateway**, improving latency and maintaining over **90% regression accuracy** for a unified customer experience.

**AIISC, University of South Carolina** **Oct 2020 – Jan 2022**
*Research Intern* *Columbia, SC*

- Built **BERT-based models** to analyze **10,000+ Reddit posts** for substance abuse analysis, attaining **85% accuracy**, and incorporated **historical sentiment** with time-aware neural models.
- Designed a **gender-sensitive RoBERTa-based system** to detect mental health symptoms in cardiovascular disease, improving symptom and gender-specific recall by **2.5%**.
- Fine-tuned **BERT models** to predict **cognitive decline** using aphasia patient transcripts, achieving **87% accuracy**.

## Projects

**Automatic Prompt Generator and Selection** | *Hugging Face, LangChain, PyTorch* **Sep 2024 - Dec 2024**

- **Developed** an automated prompt generation system integrating advanced reasoning frameworks like **Graph of Thoughts** and **Tree of Thoughts** to improve logical consistency and task-specific performance in LLMs.
- **Engineered** mechanisms for dynamic prompt selection based on **semantic similarity** and **computational efficiency**, reducing latency and increasing response accuracy.

**CarNet: Used Car Price Prediction and Recommendations** | *Python, scikit-learn, Pandas* **Nov 2024 - Dec 2024**

- **Investigated** different machine learning models and their interpretability for **predicting used car prices**.
- **Designed** a recommendation system using **cosine similarity** for car comparisons based on user preferences.

**Co-op Game Match and Recommendation Service** | *AWS, FastAPI, IGDB/Steam APIs* **Sep 2024 - Dec 2024**

- **Developed and deployed Gamezon**, a scalable co-op matchmaking app on AWS EC2 using **FastAPI** with **IGDB and Steam APIs** for game recommendations and player matchmaking.
- **Implemented** a cloud workflow leveraging **Amazon S3** for metadata storage, **Amazon Lambda** for orchestration, **RDS** for database management, **SQS** for queuing, and **SNS** for match notifications.

**KG Embeddings for Pharmaceutical Applications** | *PyKeen, PyTorch, Knowledge Graphs* **Sep 2024 - Dec 2024**

- **Explored** use of **knowledge graph embeddings** for downstream tasks such as **link prediction** and **semantic triple classification** to assess validity of relationships in pharmaceutical datasets.
- **Formulated and evaluated** models with techniques like **TransE**, **DistMult**, and **HolE**, achieving enhanced performance metrics in triple classification tasks.