

# Barun Das

Atlanta, GA | 470-838-1164 | [barun.das.23@gmail.com](mailto:barun.das.23@gmail.com) | [LinkedIn](#) | [GitHub](#)

## EDUCATION

---

<b>Georgia Institute of Technology</b> <i>Masters of Science in Computer Science</i>	Aug 2021 – May 2023 GPA: 4.0/4.0
<b>Indian Institute of Technology Kharagpur</b> <i>Bachelors of Technology in Materials Engineering, Minor in Computer Science</i>	July 2016 – June 2020 GPA: 9.34/10.00

## EXPERIENCE

---

<b>Software Engineering Intern</b> <i>Autodesk</i> <i>Built a content recommendation system on an AWS Neptune knowledge graph consisting of over 30000 nodes and 50000 relationships, which increased monthly active users by 40%.</i>	May 2022 – Aug 2022 <i>Atlanta, GA</i>
<ul style="list-style-type: none"><li>Created ETL scripts to ingest data from three data warehouses into AWS Neptune, using gremlin for querying</li><li>Overcame performance bottlenecks on AWS Neptune by rewriting advanced openCypher queries in gremlin</li><li>Developed a GraphQL API for the recommendation engine, consisting of Node.js Lambda functions exposed over AWS AppSync, enabling other developers to dynamically query the knowledge graph</li><li>Provisioned and configured different AWS instances and IAM roles using Cloudformation</li></ul>	
<b>Software Engineer</b> <i>General Electric</i> <i>Developed a Kubernetes platform used by over 150 orgs at GE. I designed and integrated new platform services with our Jenkins CI/CD pipeline, migrated binaries and build artifacts to JFrog Artifactory for better dependency management, and designed operational dashboards for core platform services using Grafana and Prometheus</i>	Aug 2020 – Aug 2021 <i>Hyderabad, India</i>
<ul style="list-style-type: none"><li>Designed platform components (data layers, notification service) as REST microservices using Java + Spring Boot</li><li>Reduced deployment times by 25% by automating concurrent deployments of multiple platform layers</li><li>Wrote custom controllers for Kubernetes resources and implemented pod auto scaling to maintain performance</li><li>Conducted performance tests on messaging queues (Artemis, Kafka) using Apache Camel to ensure SLAs were met</li></ul>	

## PROJECTS

---

<b>HAAR</b>   <i>Python, Pytorch, openCV, BERT</i> <i>Designed a self-supervised approach for action recognition in egocentric view. This is a multimodal model that uses RGB, flow and narration embeddings to recognize action classes which helps reduce dependence on expensive frame-level annotations. It is also easier to finetune to newer datasets because the action map can be learned. (<a href="#">GitHub</a>)</i>	Aug 2022 – Present
<ul style="list-style-type: none"><li>Designed scalable data pipelines and infrastructure to aid in the extraction and training of 20 million RGB frames and 100+ hrs of 1080p video, with associated narrations</li><li>Improved training times by optimizing infrastructure utilization through multiprocessing and CUDA operations</li><li>Paper in submission at <i>AAAI 2024</i></li></ul>	
<b>DoGe</b>   <i>PyTorch, openCV, GradCAM</i> <i>Improved domain generalization (DG) robustness to distribution shifts on image datasets, as measured by OOD-Bench. We also analyzed the effect of different scheduling algorithms on classification accuracy (<a href="#">GitHub</a>)</i>	Sept 2022 – Dec 2022
<ul style="list-style-type: none"><li>Designed a new learning objective based on gradient muting and regularization that improved DG accuracy</li><li>Outperformed state-of-the-art baselines on multiple image datasets like CMNIST, TerraIncognita and Camelyon</li></ul>	
<b>Blockboard</b>   <i>Flask, Python, pandas, scikit, Javascript, Azure, MySQL</i> <i>Developed an online tool for Bitcoin visualizations based on block time. We scraped trading APIs to gather 20000 rows of BTC price points, on-chain data and financial metrics. We also analyzed the tweet sentiment for 3 million tweets on BTC during the same period (<a href="#">Demo video</a>)</i>	Aug 2021 – Dec 2021
<ul style="list-style-type: none"><li>Designed interactive and responsive price visualizations using Flask and D3.js, backed by MySQL (Azure)</li><li>Analyzed correlation among different data points and time windows to uncover insights on BTC price movements</li></ul>	

## TECHNICAL SKILLS

---

**Languages:** Java, Python, JavaScript, Bash, SQL, HTML/CSS, openCypher, gremlin  
**Frameworks/Libraries:** PyTorch, pandas, scikit, openCV, CUDA, NLTK, Spark, D3.js, Spring, React.js  
**Technologies:** Git, Docker, Kubernetes, Jenkins, AWS, GCP, Databricks, Maven, Jira