Barun Das

Atlanta, GA | +1 470-838-1164 | bdas31@gatech.edu | linkedin.com/in/barun-das | barunkgp.github.io

Education

GEORGIA INSTITUTE OF TECHNOLOGY *Aug 2021 – May 2023*
MS in Computer Science GPA: 4.0/4.0
Head Teaching Assistant for *CSE 6242: Data and Visual Analytics* consisting of ~1500 students

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR *July 2016 – June 2020*
B.Tech. in Metallurgical & Materials Engineering / Computer Science & Engineering (minor) GPA: 8.97/10 (rank 2nd)

Experience

AUTODESK *Atlanta, GA*
Software Engineering Intern | Neo4J, AWS Neptune, Lambda, Node.js, GraphQL *May 2022 – Aug 2022*

- Created ETL scripts to pull data from three key learning platforms into AWS Neptune, leveraging openCypher for querying – resulting in a highly connected knowledge graph consisting of over 30000 nodes and around 50000 relationships
- Developed a graph-based recommendation system utilizing AWS Neptune which increased monthly active users by 40% (estd)
- Implemented a GraphQL API for the recommendation engine consisting of Lambda functions (built using the Node.js Serverless framework) exposed over AWS AppSync to enable dynamic querying of the knowledge graph
- Provisioned and configured different AWS instances (EC2, Neptune, AppSync, Lambdas) and IAM roles using Cloudformation

GE DIGITAL *Hyderabad, India*
Software Engineering Specialist | Java, SQL, Docker, Jenkins, Kubernetes *August 2020 – August 2021*

- Optimized GE's Kubernetes based solution for creating microservices over AWS – used by over 150 teams globally
- Designed functional platform components as REST-based microservices using Java and Spring Boot (e.g., testing, storing preferences, data layer, etc.) – dockerized and integrated them with a Jenkins CI/CD pipeline to deploy as Kubernetes pods
- Created PDIs and Helm charts for automating deployment of pods on the cluster; migrated build artifacts to JFrog Artifactory
- Visualized pod and system-level statistics for 15 services and components using Grafana and Prometheus
- Conducted performance tests on messaging queues (Artemis, Kafka) using Apache Camel to ensure SLAs were met

Relevant Projects

HUMAN-ASSISTED ACTION RECOGNITION | Ongoing research supervised by Dr. Thomas Ploetz *Jan 2022 - present*

- Aim: to develop a new approach for action recognition that leverages human narration as a weak label to recognize domain-specific actions by utilizing multimodal, transfer learning with limited supervision
- Work so far: recognized labels of interest using BERT and developed a transformer architecture that independently processes speech and video using these labels to beat our baseline *MMAction 2* recognizer on domain-specific actions by 27%
- Current focus and future work: formalizing the architecture, incorporating self-supervision and expanding to more domains while addressing key questions such as improving keyword identification and modality alignment

STORY GENERATION USING KNOWLEDGE GRAPHS AND A GPT-NEO LANGUAGE MODEL| Spacy, VerbNet, NLTK, networkx, GPT-neo *Jan 2022 – Apr 2022*

- Created a knowledge graph from text prompts using Spacy for named entity recognition, Verbnet and NLTK for verb disambiguation and inferring semantic frames and thematic roles, and networkx for visualizing the graph
- Generated new stories from a text prompt using a GPT-neo language model which takes an initial prompt and returns the most relevant candidate according to a scoring function based on the existing knowledge graph

FEDERATED LEARNING BASED RECOMMENDATION SYSTEM | Pytorch, scikit-learn, surprise *Oct 2021 – Dec 2021*

- Worked in a team of 5 to create a novel collaborative and content-based hybrid recommendation model for federated learning without compromising on privacy – also managed to tune performance to improve that of a standard collaborative system
- Configured 2000 edge nodes to send pruned gradients to the central server which consumed them by Federated Averaging
- Designed a demo web-app to visualize the user nodes that received the updated central model with the recommendations

BLOCKBOARD | Python, Javascript (D3.js, Swiper.js) HTML, CSS, Azure *Oct 2021 – Dec 2021*

- Collaborated with 5 teammates to develop a visualization tool that provides simplified, holistic information on Bitcoin (BTC)
- Analyzed correlation among 20k BTC price points, on-chain data, financial metrics, and 3M tweets to develop unique insights
- Combined these data sources to create novel visualizations based on block-time for both laypeople and analysts/traders

Skills

- Languages and frameworks: Python, Java, C/C++, Javascript, Node, D3.js, HTML, CSS, Hadoop, Spark, SQL, GraphQL
- Technologies: Docker, Kubernetes, AWS, GCP, Tableau, Figma, Prometheus, Grafana, Alertmanager, Artemis, Neo4J, Jenkins