

DIVYANSHU RAJ

(+1) 623-206-5233 | divyanshuraj.6815@gmail.com | Tempe, Arizona, USA
[linkedin.com/in/divyanshu-raj](https://www.linkedin.com/in/divyanshu-raj) | divraz.github.io | medium.com/@divyanshu-raj

EDUCATION

Master of Science (Thesis), Computer Science

Arizona State University, Tempe (USA)

Aug'22 - May'24

Bachelor of Technology, Information Technology

Indian Institute of Information Technology, Allahabad (India)

Jul'13 - Jun'17

WORK EXPERIENCE

Software Development Engineer Intern, Amazon, Tempe, USA

May'23 - Aug'23

- Designed the architecture for Brand Customer Reviews (BCR) Auto Reply that establishes contact between a seller and the buyer of the product with critical reviews, automatically in an event-driven fashion with an AWS CICD pipeline.
- Reduced average latency from 200 ms to 2 ms on Opensearch queries used by BCR Auto Reply architecture.
- Handled the Migration of existing BCR architecture with detailed documentation and roll-back plan for all regions.
- Implementation of BCR Auto Reply architecture with S3 Bucket, AWS SQS Queue, AWS Serverless Lambda, AWS Dynamo database, and Elastic search with complete code coverage through unit tests and Integration tests in JAVA.
- Projected to benefit 1000+ customers over a sum of \$100k on a monthly basis as they used a paid third-party tool to automate this before. The total cost for this feature on Amazon's end is \$200 monthly with the optimized architecture.

Software Development Engineer, Streamoid Technologies, Bangalore, India

Jul'17 - Aug'22

- Designed, and developed a Data Ingestion pipeline using API first approach. FastAPI docker images deployed on Google Cloud Run for autoscaling, Google Datastore as a decoupling layer, and time series support. 90% reduction in operations.
- Enabled 20% faster data ingestion by clients (~20M entries daily) with an 80% increase in frequency, and 90% reduction in MySQL database CRUD functionality using delta updates based on time series data.
- Prepared system architecture and implemented a real-time scalable Data Processing pipeline, an event-driven, push-based architecture with a priority queue. [\[My Blog\]](#). Used RabbitMQ clusters, Redis, Google Kubernetes Engine, KEDA, AWS Lambda, and Google Cloud Run for a scalable processing pipeline. The pipeline is 20% faster than previous and has achieved a 40% cost reduction to \$0.0001 per product.
- Semi-Supervised Text Classification, NER, and Text Generation with Transformers in Fashion space. Fine-tuned BERT with MLM and GPT2 with CLM with 15 Million fashion sentences and released the language models for downstream tasks.
- Improved and implemented a Transformers-based Text Classification pipeline that used a fine-tuned BERT model as a base layer with 72 classes and achieved 98% model accuracy, and 95% real-world accuracy. [\[My Blog\]](#)

RESEARCH EXPERIENCE

Research Assistant, Logos Robotics Lab, Arizona State University, Tempe, USA

Aug'22 - Present

- [“Language-Conditioned Change-Point Detection to Identify Sub-Tasks in Robotics Domain”](#) published at [Articulate Robots workshop at RSS 2023](#). Continuing Thesis on this work. (<https://arxiv.org/abs/2309.00743>)
- Strategy to prune Text Datasets to achieve SOTA using Transformer embeddings (Identifying Quality data points)
 - Used SNLI and GLUE datasets. Pruned them using 4 strategies (random, easy, hard, and equal ratio) and evaluate the model's accuracy. Uses, context vector, Kmeans clustering, RoBERTa model. [\[My Blog\]](#)
- Training an Inverse semantic model for plan execution with Robotic-Human Interaction using T5. Designed an experiment with mini-grid, reinforcement learning, and BART to achieve this.
- Design and Train a Quantum Machine Learning Model to train on Iris Dataset. [\[My Blog\]](#)

ACTIVITIES & ACHIEVEMENTS

- Writer for “Towards Data Science” and “Analytics Vidhya” on medium. Several articles have been published with them.
- All India ranked **193 in ACM ICPC 2014** competitive coding competition, Asia, Amritapuri site online round.
- Actively trained dozens of Company employees on Applied Deep Learning and Machine Learning Operations.
- Teaching Assistant in “Perception in Robotics” Grad course, and “Data Structures and Algorithms” Undergraduate course.

SKILLS

System Design █ Software Development █ Backend Development █ AWS █ Microservices Architecture █ Load Testing █ Data Ingestion █ Database Management █ Event-Driven Architecture █ Machine Learning █ Deep Learning █ Quantum Machine Learning █ Cloud Platforms █ Data Analysis █ Version Control █ Cloud Computing █ Teaching and Training █ Research █ Technical Writing █ Competitive Coding █ Git █ AWS CICD Pipeline █ AWS Opensearch █ AWS S3 Bucket █ AWS SQS Queue █ AWS Serverless Lambda █ AWS Dynamo Database █ AWS SNS Topic █ FastAPI █ Docker █ Google Cloud Run █ Google Datastore █ MySQL █ RabbitMQ █ Redis █ Google Kubernetes Engine █ Google Logging █ MongoDB █ Solr █ Robotics █ NLP █ Transformers █ BERT █ GPT █ T5 █ JAVA █ Python █ C++ █ LaTeX █ Communication █ Problem-Solving █ Teamwork █ Leadership █ Adaptability █ Time Management █ Critical Thinking █ Creativity █ Research Skills █ Writing Skills