

Prasanna Venkatesh Thambuanantharaman

Raleigh, NC | Github | LinkedIn | pthambu@ncsu.edu | Website | (919)-633-8167

TECHNICAL SKILLS

- **Software Infrastructure:** Docker, Kubernetes, OpenShift, Terraform, Ansible, Cisco ACI, Jenkins, CI/CD Pipelines, Apache Kafka, Snowflake, Apache Airflow, ETL pipelines, Git, Perforce, Software Engineering Best Practices, Microservices Architecture
- **Programming Languages:** Python, R, SQL, Java, C++, JavaScript, HTML, CSS, Shell/Bash scripting
- **Cloud:** S3, Glue, Lambda, Redshift, EC2, GCP BigQuery, Google Compute Engine, Azure Blob Storage, Azure AI Services
- **Databases and Operating Systems:** MySQL, PostgreSQL, MongoDB, DynamoDB, Elasticache, Redshift, Linux, Windows
- **Certifications:** Azure Certified AI Engineer Associate, AWS Certified Cloud Practitioner

EXPERIENCE

Infra Development & Operations Intern, Samsung Advanced Computing Labs, San Jose, California May 2024 - Aug 2024

- Delivered a completely automated Python based solution to manage users in Jira, reducing license requirements by approximately **20%**, through identifying redundancies and removing inactive users. Deployed the solution in a Jenkins pipeline.
- Working to automate the process of tenant creation and configuration in Cisco APIC using the ACI Toolkit, Ansible and Python saving **16 hours** of manual effort per tenant on average.
- Creating a comprehensive data pipeline and service health dashboard to aggregate data from key infrastructure assets and monitor their health. Integrating with PagerDuty API for improved alert mechanisms and resolution times
- Migrating the organization's Jenkins server from dedicated VMs to containers orchestrated by Openshift, thereby reducing costs.

Data Engineer II, Tata Consultancy Services, Chennai, India June 2022 - July 2023

- Designed and architected a data pipeline leveraging Snowflake, SQL, AWS S3, Glue, Lambda, and Apache Airflow, to ingest, load, calculate, and report the carbon emissions of a leading UK grocery chain.
- Spearheaded the implementation of Terraform to automate infrastructure provisioning and configuration on AWS, eliminating human error and reducing time required to provision and configure new resources by **33%**.
- Developed ETL pipelines to source data from financial and ESG reports, creating a unified data model for analytics and reporting.
- Defined the table schemas and used dbt to build a transformation workflow that transformed and tested the data.
- Integrated CircleCI into our development and production environments to speed up our CI/CD and DevOps pipelines by **40%**.
- Collaborated with stakeholders across the supply chain to identify areas of improvement and draft a plan of action that was projected to reduce emissions by over **25%**.

Data Engineer I, Tata Consultancy Services, Chennai, India Sept 2020 - June 2022

- Developed an ETL data pipeline using Python and SQL, that ingested data from over **30,000 energy sensors**, multiple different REST APIs and streamed it into a Kafka queue with **0 data loss**.
- Led a team of 4 Data Engineers to enhance the data acquisition process using threading and asynchronous HTTP requests, achieving a **3x** improvement in processing speeds.
- Implemented data quality assurance best practices and additional redundancy mechanisms to ensure data quality, delivery, and security throughout the pipeline.
- Built Tableau dashboards to analyze and provide recommendations on efficient energy usage, resulting in clients saving over **18%** on energy costs on average.

Software Engineer Intern, Tata Consultancy Services, Chennai, India Feb 2020 - July 2020

- Delivered a network optimization algorithm using Python, PuLP and Pandas to optimize the delivery network and schedules of an intercontinental supply chain, resulting in a reduction of **22%** in shipping and delivery costs.

PROJECTS (Link to all project demos and Github repos can be found on my website)

- **Data Engineering and Analytics:** Constructed an end-to-end data pipeline to analyze NYC Taxi data using Mage AI, Google Compute Engine and BigQuery. Modeled the data into fact and dimension tables, setup the pipeline infrastructure, conducted analysis and created a dashboard using LookerStudio to showcase the analytics.
- **Music Recommendation System:** Implemented a music recommendation system based on the Spotify dataset and outperformed traditional Transformer architectures by utilizing LightGCN, resulting in an AUC score of **97.7%** and Recall of **93.4%**.
- **Neural Networks:** Built a CNN model that recognizes the artist when given a URL of their artwork with a validation accuracy of over **86%**. Trained, containerized, and deployed the model using TensorFlow, Google Cloud Run and Docker.
- **Data Analytics:** Modeled a healthcare dashboard using Tableau to showcase the analysis of flu shots data in Massachusetts. Outlined the table schemas and ingested the data into PostgreSQL database, performed data modeling and conducted analytics.
- **Cloud Architecture:** Designed the cloud architecture of a concert ticket booking website. Employed AWS services like Cognito, EKS, Aurora, DynamoDB, Firewall Manager, achieving a server response time of less than **300ms** during peak load testing.

EDUCATION

- **North Carolina State University, Raleigh, NC** December 2024
Master of Computer Science GPA: 4.0/4.0
- **Amrita Vishwa Vidyapeetham, Coimbatore, India** August 2020
Bachelor of Technology in Computer Science and Engineering GPA: 9.06/10