

# Prasanna Venkatesh Thambuanantharaman

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

## EDUCATION

**North Carolina State University**, Master of Computer Science Aug 2023 - May 2025  
Relevant Coursework: Automated Learning & Data Analysis, Neural Networks, Algorithms, Cloud Computing GPA:3.889/4.0  
**Amrita Vishwa Vidyapeetham**, Bachelor of Technology in Computer Science Jun 2016 - Aug 2020  
Relevant Coursework: Data Structures, Database Systems, Machine Learning and Data Mining, Neural Networks GPA:9.06/10

## WORK EXPERIENCE

**Data Engineer II - Tata Consultancy Services**, Chennai, India Jun 2022 - Jul 2023

- Leveraged **Snowflake, CircleCI, SQL, Terraform, AWS S3, Lambda, and AWS Glue** to build a data pipeline that calculated and reported the carbon emissions of Sainsbury's, **saving over 30 man-hours per month of manual efforts**.
- Designed the architecture of the pipeline, performed data modeling, transformation, and configured the cloud infrastructure.
- Headed the Data Analytics and Visualization sub-group and analyzed the emissions data to identify areas of improvement. Built **5+ Tableau dashboards** to visualize and assess the impact of sustainability initiatives.
- Drafted a plan of action based on the insights acquired to **reduce emissions by approximately 35% year over year**.
- Facilitated the onboarding of 8 new data engineers, providing mentorship and conducting knowledge-sharing sessions.

**Data Engineer I - Tata Consultancy Services**, Chennai, India Sep 2020 - Jun 2022

- Developed data acquisition solutions and ETL pipelines by using **Python, SQL, Apache Kafka, and Google Cloud**. **Scaled up the solutions, acquiring and processing data from over 30,000 IoT sensors and multiple APIs** to optimize energy usage.
- Led a team of 4 software engineers to develop a new data acquisition product that employed **threading and asynchronous processing** to improve performance. Accomplished an **improvement of over 3x in acquisition and processing speeds**.
- Included additional redundancy measures to guarantee data delivery and data security in case of failures along the pipeline.
- Performed statistical modeling, hypothesis testing as well as trained and deployed Machine Learning models to recommend energy-saving optimizations which resulted in a **12% reduction in energy cost**.

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

- Created mathematical models and algorithms using **Python, Statistical modeling, Linear Algebra, and Pandas** to solve complex supply chain management problems like full truckload, network optimization, last-mile delivery, etc.
- Produced a proof of concept network optimization algorithm to optimize the delivery network and schedules which **resulted in potential savings of 22% in delivery costs**.

## PROJECTS

### Machine Learning

- Implemented a Music recommendation system using LightGBM, Behavior Sequence Transformer, and LightGCN models on Spotify's Music Streaming Sessions dataset. Achieved an **AUC score of 97.7% and Recall of 93.4%** using LightGCN. [\[Link\]](#)
- 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 **Javascript, Google Cloud Run and Docker**. [\[Github\]](#) and [\[Live Demo\]](#).

### Data Engineering and Cloud Architecture

- 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. [\[Github\]](#) and [\[Dashboard\]](#)
- Designed the cloud architecture of a concert ticket booking website. Employed AWS services like **Cognito, EKS, Aurora, DynamoDB, Firewall Manager** etc. **Achieved a server response time of <300ms** during peak load testing. [\[Github\]](#)

## SKILLS

**Languages:** Python, R, SQL, C, C++, Java, Javascript, HTML, CSS, Shell/Bash.

**Databases and Operating Systems:** MySQL, PostgreSQL, MongoDB, Linux, Windows.

**Tools/Frameworks:** Snowflake, Tableau, PowerBI, Cloud Platforms (AWS, Azure, Google Cloud), Terraform, CircleCI, Kafka, PyTorch, TensorFlow, Mage AI, LookerStudio, Kubernetes, Docker, REST API's, Jenkins, Flask.

**Certifications and MOOCs:** AWS Certified Cloud Practitioner, Applied Data Science with Python, Neural Networks and Deep Learning, Financial Markets (Honors).

## EXTRACURRICULARS

- Served as group leader for over 20 programs in Tata Consultancy Service's #millionhoursofpurpose volunteer campaign.
- Active member of the Data Science Club and Science Policy Pack Club at NC State University.