

# Anjani Ganapathy

tel: +44 79381 80880 • email: anjani.ganapathy@gmail.com

## EDUCATION

UNIVERSITY COLLEGE LONDON, Faculty of Brain Sciences  
Masters in Computational Linguistics

London, United Kingdom  
Expected completion Sept 2025

GEORGIA INSTITUTE OF TECHNOLOGY, College of Engineering  
Bachelor of Science in Industrial and Systems Engineering

Atlanta, Georgia  
May 2019

## EXPERIENCE

Capital One Financial Corporation

Washington, DC

Data Analysis Manager | Card Data Infrastructure

Aug 2023 - Oct 2023

- **Created business value:** Led a team of 5 data analysts to enable adoption of Snowflake depersonalized data ecosystem (DepEco) by >100 business analysts. Enabled external data use in valuation and decisioning models, generating >\$250M in revenue.
- **Directed cross-functional delivery:** Worked across product, tech, and data science to mitigate complex risks in DepEco, leading to a fix in data transfer that resulted in a 2% improvement in DepEco data completeness.
- **Prioritized stakeholder needs:** Translated business strategy and analysis needs into actionable tasks for product and tech teams, leading to key improvements and standardization such as enabling real-time updates to key datasets.

Principal Data Analyst | Card Data Infrastructure

Jan 2022 - Aug 2023

- **Refined process and workload:** Reduced DepEco datasets 20% while maintaining analysis potential, reducing timelines by 1 month.
- **Prioritized data excellence:** Set up validation in Quicksight, enabling escalation and recovery of missing data (5%) in critical datasets.
- **Enabled better credit analysis:** Created dataset using Python to simulate credit policies on card applicants. Onboarded >50 analysts for whom this product improved consistency and speed of analysis.

Senior Data Analyst | Card Data Infrastructure

Jan 2021 - Dec 2021

- **Designed automated monitoring:** Provided insights into 100+ test credit card customer segments. Reports informed decisions such as expanding lending to Puerto Rico and launching products for new-to-credit customers.

Data Analyst | Bank Marketing Experience

Aug 2019 - Dec 2021

## PROJECTS

Dissertation in Computational Linguistics

London, UK

Unsupervised Part-of-Speech tagging for Hindi

Apr 2024 - Present

- **Investigated a neural network's ability to learn syntactic groups:** Implemented a VQ-VAE combined with Gumbel sampling to learn meaningful syntactic groups. Results showed that Hindi BERT embeddings with character embeddings capture Hindi parts of speech more effectively. To be published at UCL as "VQ-VAEs & Gumbel-Softmax for Unsupervised POS Tagging".
- **Used character and word reconstruction to train:** Implemented a character-level LM to predict character sequences and generate an additional loss term for training; this worked in tandem with character embeddings to improve many-to-one accuracy for predictions.

Seminar in Computational Linguistics & NLP

London, UK

Code-Mixed Hinglish to English NMT

Mar 2024 - Present

- **Investigated ability of NMT models to handle code-mixed rather than monolingual text:** Implemented Encoder-Decoder Seq2Seq network in Pytorch and trained on a part human-, part synthetically- generated parallel corpus.
- **Adapted network to use Transformers:** Faced translation challenges with initial architecture despite favorable losses, partly due to limitations in data and computational resources for testing. Evaluation of Transformers in progress; will investigate using existing implementations (e.g. fairseq) if challenges persist.

## SKILLS

- **Programming Languages + Frameworks:** Python, Pandas, PyTorch, TensorFlow, NLTK, SQL, HuggingFace, Sklearn, Ray Tune, LoRA
- **Software & Platforms:** Snowflake, Tableau, AWS Quicksight & S3, Github, Jira
- **Languages:** English – native, Tamil – native, Hindi – native, Spanish – advanced