

# Anusha Lihala

✉ anusha.lihala@gmail.com

🌐 anushalihala

🌐 anusha-lihala

🌐 @anusha.lihala

## Professional Summary

---

Experienced Machine Learning Engineer and Data Architect with expertise in machine learning, NLP, semantic graph solutions, and advanced data architecture. Skilled in integrating AI solutions, optimizing knowledge systems, and delivering data-driven insights through innovation and collaboration.

## Skills

---

**Graph Technologies:** Neo4j, RDF, OWL, Cypher

**Machine Learning Libraries:** Autogen, LangChain, LlamaIndex, PyTorch, TensorFlow, scikit-learn, pandas, matplotlib

**NLP:** LLMs, Agents, RAG, finetuning (including PEFT), RLHF

**Programming:** Python, JavaScript (React, Angular), NodeJS

**Deployment:** AWS, Azure, GCP, Databricks

## Experience

---

09/2023 – 12/2024

**Software Engineer, LivePerson**

- Developed a support conversation to knowledge article generation system leveraging LLMs, employing multi-agent workflows to optimize information accuracy and reduce hallucinations.
- Conducted advanced prompt engineering to refine NLP outputs, ensuring relevance and quality by excluding context-specific or sensitive details.
- Designed and implemented methodologies for clustering and aggregating knowledge articles using LLMs and similarity metrics (e.g., cosine similarity with sentence embeddings).
- Devised research methodologies using LDA, cosine similarity, and KL divergence to evaluate article grouping and improve knowledge retrieval.
- Built and deployed supervised classifiers for intent detection and conversation categorization.
- Utilized Databricks for data management, creating views, and automating environment replication for seamless deployment using Databricks Asset Bundles.

06/2022 – 12/2022

**Software Engineer, Microsoft Ireland Research**

Implemented features in C# and ReactJS, maintained and improved efficiency of pipelines in ADF and Scope, monitored system performance and provided live site support.

08/2021 – 04/2022

**Data Scientist and Data Engineer, Koneksys**

- Implemented a Neo4j data model to integrate and merge product attributes from disparate sources.
- Created advanced Cypher algorithms for leveraging hierarchical relationships and enriching knowledge graphs.
- Created robust web scraper using NodeJS and Puppeteer with a 90% reduction in scraping time.
- Developed NLP pipelines to resolve entity matching issues across e-commerce platforms.
- Led a team of four to implement scalable scraping and end to end data pipeline solutions.

01/2020 – 04/2020

**AI Officer, Trinity College Dublin GSU - Demystifying AI conference**

Shortlisted research papers based on quality and relevance to conference

04/2019 – 08/2019

**Research and Development Intern, Nihilent Analytics Ltd**

Increased sentiment analysis accuracy by 8% in social media analytics using NLP and feature engineering.

## Education

---

2019 – 2020

**MSc Computer Science - Intelligent Systems, Trinity College Dublin**

**Graduated with Distinction; Top 5% in Machine Learning.**

*Relevant Modules: Knowledge & Data Engineering, Machine Learning, Text Analytics, Artificial Intelligence, Information Retrieval, Data Visualization*

2015 – 2018

**Bachelor of Computer Applications (BCA), Birla Institute of Technology**

**First Class with Distinction (89.9%)**

## Professional Development

---

- 2024 **Graph Neural Networks, Zak Jost**  
Message passing paradigm, label propagation, Graph Convolutional Network and Graph Attention Network.
- 2024 **Generative AI with Large Language Models, deeplearning.ai and AWS**  
Generative AI lifecycle with a focus on Parameter Efficient Fine-tuning (PEFT), Reinforcement learning from human feedback (RLHF), and ReAct prompting.
- 2024 **LangChain for LLM Application Development, deeplearning.ai and LangChain**  
LangChain for prompts, parsing, memory, chains, question answering, and agents.
- 2022 **Neo4J Certified Professional, Neo4J GraphAcademy**  
Neo4j property graph model, Cypher queries, graph data modelling and importing data.
- 2019 **Functional Programming in Scala Specialisation, EPFL (4 courses)**  
Effectively using functional programming principles and parallel collections in Scala, and manipulating data with Spark.
- 2018 **Data Scientist Nanodegree, Udacity (Term 1)**  
Cleaning and pre-processing data, engineering relevant features, reducing data dimensionality and using machine learning models to make data-informed decisions.
- 2018 **Deep Learning Specialisation, deeplearning.ai (5 courses)**  
Included CNNs and RNNs with variants such as LSTMs, optimization algorithms, and strategies for handling bias and variance including regularization. Projects using TensorFlow.

## Projects

---

- 2023 **Dish Ratings Extraction, (Personal Project: NLP and Prompt Engineering)**  
Developed an NLP pipeline to extract sentiment and aggregate and rank dish ratings from restaurant reviews using advanced prompt engineering with GPT4.
- 2022 **Narrative Recommender System, (Personal Project)**  
Built a Neo4j-based recommender system using Sorensen-Dice similarity and attributes for user activity and narratives.
- 2020 **LOCATIO: Ontology for Information Overload, (MSc Dissertation)**  
Developed an ontology to mitigate information overload by knowledge design and personalisation. Conducted simulations using adaptive user models to demonstrate enhanced knowledge assimilation through use of ontology.
- 2019 **Propaganda Detector, (Personal Project: NLP)**  
Built a deep learning classifier for propaganda detection using TensorFlow and BERT.
- 2018 **Image Classifier, (MOOC Project: Computer Vision)**  
Built an image classifier using PyTorch which uses pre-trained models with transfer learning to reduce training time and deliver accurate predictions.
- 2018 **Customer Segments Identification, (MOOC Project: Machine Learning)**  
Applied preprocessing, dimensionality reduction (PCA), and clustering techniques (KMeans) on population and customer demographic and spending data to identify most promising population segments for customer outreach.

## Publications

---

- "A Step Towards General NLP with Dynamic Memory Networks," *Towards Data Science*, April 2019.
- "Explaining Document Classifications with Counterfactuals," *Towards Data Science*, April 2019.
- "Attention and its Different Forms," *Towards Data Science*, March 2019.