

# Abin Varghese

[LinkedIn](#) • [Portfolio Website](#) • [GitHub](#)

abinvarghese90@gmail.com • Kingston upon Thames, London, UK, KT2 7BN • +447888704879

## Summary

Experienced ML Engineer with over 2.4 years of commercial experience, specializing in Natural Language Processing (NLP). Proficient in Python and skilled in deploying classical and deep learning applications, with expertise in LSTMs, Transformers, and open-source LLMs like GPT4All. Implemented Federated Machine Learning on clients' IID datasets, driving significant business growth, including raising revenue by a CAGR of 55% in an EdTech startup. Strong track record of leveraging advanced techniques to deliver impactful solutions.

## Education

### Kingston University

MSc, Data Science - First Class (Expected)

London, UK

2023 - 2024

Dissertation - Bridging the Gap: Exploring Privacy-Enhancing Technologies and Facilitating Collaboration between Academia and Industry through Federated Learning, and 'Privacy-as-Code' System Development.

### SRM Institute of Science and Technology

B.Tech, Computer Science and Engineering - First Class with Distinction

Chennai, India

2016 – 2020

## Experience

### INFOSYS LTD.

#### Data Scientist

Mysore, India

12/2021 – 12/2022

- Engineered team leadership: Orchestrated development of an enterprise solution utilising the Attention model to autonomously extract and categorise skills from resumes, leading to substantial cost savings in HR operations at Bangalore and Mysore DC.
- Contributed to ML projects: Enabled BEF's global impact, attracted external clients like GlobalFoundries, boosting profits in 2021-2022.
- Created courses: Trained 5000+ employees in ML, NLP, and stats, indirectly increasing revenue.
- Awarded best author: GPT-3 course in Q2 2022, live to 250,000+ internal employees and external vendors.

### CORIZO

#### Leader for Data Science and Machine Learning

Bangalore, India

02/2022 – Present

- Achieved 55% revenue increase: Led 1000+ students in data science and machine learning, driving growth.
- Headed highest-revenue department: Developed impactful lessons and projects.

### BLOOMTECH

#### Subject Matter Expert for Data Science

California, USA

04/2022 – 06/2022

- Initiated technical projects: Enhanced data-driven capabilities and delivered projects for their clients.

## Projects

- [Applied Privacy Lite](#) → Used Transformers to redact PII from text and implemented Federated Learning on an IID Dataset.
- [Explain Paper](#) → Used GPT4All Open-Source LLM to convert research papers into plain English. Working on deploying it now.
- [US Crime Statistics](#) → Analyzed the US crime statistics in the state of Maryland and deployed it for visualisations.
- [SUPPORT2](#) → Performed extensive EDA and Feature Engineering on the SUPPORT2 health dataset and deployed it.
- [GAN implementation](#) → The goal is to grasp GAN principles for generating new digit images.
- [A-B Testing ML Model Performance](#) → This project compares BERT and DistilBERT models for QA systems via A/B testing.
- [Real Estate Price Prediction with NLP](#) → Created a real estate price prediction model, hosted on Heroku using FastAPI.
- [Credit Card Fraud Detection](#) → Created and deployed using Streamlit 4 robust deep-learning fraud detection models.
- [Flight Ticket Price Predictor](#) → Developed a flight ticket price prediction model using multiple features with over 99% accuracy.
- [Fake News Classifier](#) → Built the LSTM model from scratch to classify fake and legitimate news.
- [Hackathon: Cybertrendz](#) → Special mention for tech-driven privacy solution with policy-to-checklist conversion.
- [Single Pong Player Game](#) → Single-player Pong game with paddle control, score tracking, and specified score limit.
- [Conway's Game of Life](#) → Python implementation of Conway's Game of Life, simulating life-like patterns on a grid.

## Skills

**Expert:** Python 3.X, MySQL, SQL, Classical and Deep Machine Learning, Data Science, NLP, Keras, LLM, Federated Learning.

**Proficient:** Azure, Prompt Engineering, Django, Java, CNN, Linux, TensorFlow, PyTorch, Tableau, AB Testing.

## Leadership, Volunteering and Certifications

[Microsoft Learn Student Ambassador](#) → Community leader representing Microsoft. Conducted workshops on generative AI.

[Applied Privacy Podcast](#) → Hosts Applied Privacy podcast with industry leaders (CEOs, CPOs etc.) on privacy from various perspectives.

[Ukraine Global Taskforce](#) → Data curator aiding Ukraine war logistics, creating live bombing alert heatmap.

[Licenses and Certifications](#) → Please click on this hyperlink to view all the certifications.

*References are available upon request.*