

Malavika Gowthaman

AI Engineer / Data Scientist / Masters in Data Science

+ 1 647 446 1531 | malavikag115@gmail.com | [LinkedIn.com](#) | [GitHub.com](#)

Scarborough, Ontario, Canada

Data Scientist and AI Engineer with 5 years of experience in data collection, pre-processing, model development and deployment. Expert in Deep Learning, Neural Networks, Machine Learning, and LLMs, with a strong background in IT, Computer Science, mathematics, Software Engineering and data analytics. Skilled in Power BI, Tableau, SQL, ETL Process, ML Flow and Python.

WORK EXPERIENCE

Hope Artificial Intelligence Pvt Ltd - AI Engineer

04/24 – Present

- Developed FitTrack, a Streamlit web app with Python and Google's Gen AI for BMI calculation and personalized recommendation, featuring robust error handling and secure API management. ([GitHub.com](#))
- Created a smoking detection model using PyTorch YOLOv5, achieved 90% accuracy, and documented the process in a Jupyter notebook for future enhancements. ([GitHub.com](#))
- Developed a deep learning model using TensorFlow and VGG19 for cancer classification based on CT scan images. Utilized Docker and GitHub Actions for CI/CD pipelines and deployed the front-end application for this project on the AWS platform. ([GitHub.com](#))
- Built a Random Forest model for personal loan eligibility prediction with 99% accuracy, collaborating with cross-functional teams including finance and product management to deploy the model in a web interface. ([GitHub.com](#))
- Created TensorFlow-based CNN model with 99% accuracy for face mask detection in images and real-time videos, also implemented rigorous data pre-process to improve the data quality and validation techniques. ([GitHub.com](#))
- Developed and deployed an Information Retrieval System Streamlit web application focused on 3 types of documents using Langchain, and successfully launched on AWS EC2 for scalable access. ([GitHub.com](#))

HCL Technologies - Senior Software Developer in Data Science

09/22 – 04/24

- Developed Python code for document classification, enhanced invoice processing, and implemented language detection with 95% accuracy. Utilized MLflow for tracking experiments and managing models.
- Created dashboards for the P2DTM BIM 7D Solutions and led machine learning research with Azure ML, achieving a 20% improvement in forecasting accuracy and gained valuable project management experience.
- Visualized call flow trends and metrics using MS PowerBI and DAX to provide actionable insights from over 50,000 call records. Automated data extraction for over 150 entries with Python and Pypeteer.
- Collaborated with clients to identify business requirements and developed tailored software solutions that met specific business needs.

Xecutesmart Technologies - Data Scientist

06/21 – 09/22

- Developed and deployed disease prediction models using Python, TensorFlow, Keras, and ML algorithms, achieving nearly 90% accuracy. Conducted sentiment analysis on airline data and visualized insights from Twitter comments.
- Analyzed reservation system data with Talend and Power BI, creating dashboards. Designed dimensional data models for banking projects and executed complex SQL queries and ETL processes using Informatica.

Grroom - Software Developer

12/19 – 05/21

- Gained foundational experience as an intern, developed and optimized complex SQL queries, also created interactive dashboards as a full-time Software Developer, which improved data reporting efficiency by 30%.
- Worked on data collection, labelling, and pre-processing in the fashion industry, handled and prepared data from more than 10,000 images to enhance model accuracy and performance.

EDUCATION

Masters of Technology | Anna University,

Tamil Nadu, India

Data Science - 06/2021

Bachelor of Engineering | Anna University,

Tamil Nadu, India

Electronics and Communications Engineering - 05/2019

SKILLS

Technical skills: AI Tools & Frameworks, ML algorithms(Supervised & Unsupervised learning algorithm), Deep Learning Models, Large Language Models, Data Science, Data Analytics, Data collection, Feature extraction, NLP.

Data Visualization & ETL Tools: Tableau, MS Power BI, Informatica Power Center, Talend for Data Integration.

Design and Development tools: AWS EC2, ECR, Azure Machine Learning Studio, AutoML, Visual Studio Code, MS SQL Server, Oracle SQL Developer, CI/CD Pipeline, GitHub Actions, MLFlow.

Database Languages: Oracle SQL, PL/SQL 21c/19c Enterprise edition.

Programming Skills: Python, R Programming (Basics).

Soft Skills: Team Work, Leadership, Good Communication skill, Business Acumen, Complex problem Solver.

CERTIFICATION

- Microsoft Certificated Data Science Associate
- Deep Learning.ai Specialization in Coursera.