ANKITA SHELKE

AI/ML DEVELOPER - Data Analysis, Machine Learning, Generative AI

✉ ankitashelke969@gmail.com ☎ +1 (250)-879-4098 ⚲ Canada in LinkedIn ⚙ GitHub ⚬ Portfolio

SKILLS

- **Programming & Data Tools**: Proficient in Python, R, SQL, Jupyter, VS Code, Google Colab, Git, GitHub.
- **Frameworks & Libraries**: PyTorch, TensorFlow, LangChain, LangGraph, Hugging Face, Pandas, NumPy.
- **GenAI & LLM Systems**: Large language models (LLMs), prompt engineering, RAG, retrieval techniques.
- **Cloud & MLOps**: AWS, Docker, Streamlit, ECS, ECR, AWS Fargate, CI/CD pipelines, GitHub Actions.
- **Visualization & Reporting**: Power BI dashboards, data aggregation, KPI reporting, analytics derivation.

WORK EXPERIENCE

AI/Data Science Intern   January 2025 – Present
INeuron   *Remote*

- Implemented LLM-based RAG document intelligence platform, integrating document chat, semantic search, and automated comparison analytics workflows, enabling faster document retrieval and interactive analysis.
- Constructed modular production pipelines for ingestion, embedding generation, and inference using Python scripting, enabling multi-document chat capabilities & scalable workflow processing across document collections.
- Deployed containerized AI services on AWS Fargate (ECS/ECR) with automated CI/CD via GitHub Actions, ensuring consistent deployment processes and reproducible LLM operations with integrated monitoring.
- Collaborated with cross-functional teams to translate business requirements into technical specifications, optimizing solution architecture design and improving AI feature adoption within deployment cycles.

Research Assistant I - Science   October 2024 – December 2024
Thompson Rivers University   *British Columbia*

- Implemented LLM and RAG-based NLP workflows to summarize academic research papers, reducing manual literature review time by 60% while generating structured, publication-ready insights using Python pipelines.
- Co-authored inflation research publication, performing statistical data analysis, data wrangling, synthesis from 200+ structured, unstructured sources, producing reproducible datasets, publication ready drafts using Python.
- Orchestrated RAG retrieval system for academic PDFs using vector embeddings, Chroma database, and NLP indexing, enabling rapid query responses and efficient large-scale literature information extraction.
- Formulated AI-assisted analytical for extraction of research methodologies and key findings, utilizing NLP pipelines, entity recognition, and improving faculty productivity and reducing manual review time by 55%.

Graduate Research Assistant   January 2023 – September 2024
Thompson Rivers University   *British Columbia*

- Collected, cleaned, and preprocessed over 50 years of macroeconomic datasets using R, ensuring data integrity, normalization, and structured inputs for predictive models analyzing inflation trends.
- Implemented ARIMA, and ensemble ML algorithms, performing systematic cross-validation and improving forecast model accuracy by 29% and producing reproducible outputs for macroeconomic trend predictions.
- Managed comprehensive statistical time series analysis, calculating correlations, lag variables between inflation, macroeconomic indicators, directing optimal model selection.
- Devised hybrid forecasting methodology combining machine learning models, statistical time-series techniques, increasing predictive accuracy by 29%, generating actionable, data-driven insights to support policy decisions.

Teaching Assistant   January 2023 – April 2024
Thompson Rivers University   *British Columbia*

- Mentored 30+ undergraduate students in ML and deep learning, simplifying neural networks, ensemble methods, improving assignment submission accuracy to 92% and enhancing coding implementation proficiency metrics.
- Performed lab sessions with Python, scikit learn, TensorFlow, transforming theoretical ML concepts into hands on coding exercises, improving student model implementation proficiency scores by 38% across evaluated tasks.
- Evaluated 150+ assignments and ML projects, providing detailed algorithmic feedback to improve model accuracy and reinforce students' understanding of supervised, unsupervised, and ensemble learning techniques.
- Guided student capstone projects from dataset selection to model evaluation, reinforcing applied statistical analysis, Machine Learning model optimization, and predictive performance assessment.

Intern Developer   July 2021 – December 2021
Roughnext REI (OPC) Private Limited   *India*

- Constructed sentiment analysis models using Amazon Kindle Reviews dataset (10,000+ samples), implementing Python-based text preprocessing workflows including tokenization, stop-word removal, and NLP pipelines.

- Fabricated and compared feature extraction techniques including Bag of Words, TF-IDF vectorization, Word2Vec embeddings, analyzing algorithmic effectiveness across machine learning classifiers, improving accuracy by 25%.
- Benchmarked six algorithm-feature combinations, demonstrating that ensemble methods outperformed probabilistic approaches, while incorporating Word2Vec embeddings improved Naive Bayes accuracy by 25%.
- Applied evaluation methodologies for sentiment classification, analyzing precision, recall, and F1 score metrics, documenting experimental design, and providing insights into optimal feature-algorithm pairings.

Junior Python Developer   **October 2019 – May 2020**
Maxgen Technology Private Limited   *India*

- Architected student management system using Python, SQL databases, and REST APIs to automate 1,000+ administrative tasks like student record keeping and reducing manual paperwork for educational institutions.
- Programmed 5+ interactive features for tracking student information and generating reports, enabling school administrators to streamline daily operations, improve data visibility, and enhance overall workflow efficiency.
- Gained experience in Python programming fundamentals, writing modular code for handling 500+ student records, interactive user interfaces, applying best practices in software development for educational applications.
- Collaborated closely with cross-functional 5+ team members to analyze business requirements, implement functional software solutions, ensuring alignment with educational sector needs and optimizing operational efficiency.

Systems Engineer   **March 2015 – August 2019**
Tata Consultancy Services   *India*

- Analyzed SBI National Pension System transaction datasets, writing optimized SQL queries to extract financial data from bank servers, leveraging indexing and query optimization, reducing issue resolution time by 35%.
- Created 10+ comprehensive transaction reports and interactive financial dashboards using SQL data extraction, Power BI, enabling clients to make informed investment decision, supporting management with business insights.
- Maintained 99.9% uptime for 10+ critical pension management applications on virtual servers and cloud infrastructure through daily health monitoring, system diagnostics, digitized alerts, and maintenance protocols.
- Modified Java-based API code to enhance security protocols, integrating OAuth authentication, encryption across 5+ modules to revoke unauthorized permissions, ensure data protection compliance for sensitive information.

PROJECT EXPERIENCE

Optimising Canada's Inflation: Time-Series and ML Models   **January 2023 – September 2024**
Individual Research Project, British Columbia   Link

- Implemented 8 forecasting models including ARIMA, neural networks, ensemble algorithms, performing systematic cross-validation, statistical evaluation, enhancing prediction accuracy by 29% across scenarios.

KidSciBot – Intelligent Science Tutor for Kids   **October 2024 – December 2024**
Individual Project, British Columbia   Link

- Engineered an intelligent question classification system leveraging LangGraph workflows, routing science queries through RAG retrieval, web APIs, and LLM embeddings, improving chat-bot response accuracy.

Sales Performance Analysis of Amazon   **March 2024 – May 2024**
Individual Project, British Columbia   Link

- Designed sales analytics dashboard in MS Excel analyzing 2024 Amazon revenue data of 9000+ transactions, providing insights into regional sales metrics, product categories, interactive trends for business decisions.

EDUCATION

Master of Science in Data Science   September 2022 – September 2024
Thompson Rivers University, Kamloops
Post Graduation Diploma in Applied Statistics   June 2020 – June 2021
Indira Gandhi National Open University, Distance Learning, India
Bachelor of Technology in Information Technology   June 2010 – June 2014
Techno India College of Technology, India

CERTIFICATIONS

- **Python 101 for Data Science** - Cognitive Class
- **Intermediate SQL for Data Science** - LinkedIn Learning
- **Fundamentals of Deep Learning** - NVIDIA