MOHAMMED FARHAN BALUCH

Kitchener, ON
+1(226)961-0559 | farhanbaluch.com | baluchm@uwindsor.ca | linkedin.com/in/farhan1301

**I also hold valid reliability security clearance status from Government of Canada*

PROFILE OF SKILLS

- **Proficiency of Machine Learning** demonstrated by a portfolio of comprehensive projects and research endeavors, culminating in presentations at **2** int'l conferences, contributions to **3** peer-reviewed journals, and a dynamic co-op role as an AI Developer with recent applied experience
- **Expertise of Python Programming** with over **4** years of experience, specializing in data analysis, software development, and machine learning, with a proven track record of delivering efficient and scalable solutions
- **Proficiency in SQL** experience denoted through using SQL extensively for data mining, through chatbot project for university which involved heavy usage of SQL in backend for retrieving information from database
- **Data Analysis Skills** exhibited through advanced data analysis Nasscom certification and successful projects like the 'PDF parser tool' involving comprehensive data handling & creating of insightful PowerBI dashboards
- **Web-Development Skills** denoted by working as a Web Development Intern and passing the Microsoft certification 'MTA – HTML & CSS'
- **Collaboration, Presentation and Communication Skills** portrayed through collaborating with several esteemed professors and colleagues worldwide and conducting weekly lab sessions for undergraduate students

TECHNICAL SKILLS

- **Languages/Databases:** Python, PySpark, C, C++, Java, MongoDB, MySQL, NoSQL
- **Machine Learning:** Scikit-learn, TensorFlow, Keras, PyTorch, NLTK, GANs, Pandas, NumPy, Seaborn, Matplotlib, Transformer, BERT, GPT, LSTM, GNN, spaCy, XGBoost, LGBM, SHAP, DQN, VAE, Lime, Optuna
- **Platforms:** Git, Gitlab CI/CD, Docker, Azure DevOps, Azure ML, Databricks, GCP, Apache Hadoop, Spark, OpenAI API, Power Automate, UI Path, RESTful API, Microsoft SQL Server, AWS EC2, SageMaker, Streamlit, A/B testing, Unix, Jira, Power BI, RAG
- **Data Science:** Prediction, Timeseries, Quantile Regression, Outlier Detection, Hypothesis Testing, Statistical modeling, Conversational Agents, Prompt Engineering

EDUCATION

Master of Science Computer Science Artificial Intelligence Stream Co-op Sep 2022 - Present
University of Windsor, Windsor, ON
- Cumulative GPA: **92.2** / 100

Bachelor of Technology in Computer Science Jul 2018 - Jul 2022
Vit Bhopal University, Bhopal, IN
- Cumulative GPA: **8.95** / 10.0

EXPERIENCE

Artificial Intelligence Developer (Co-op), Agri-Foods Canada, Harrow, ON May 2023 - Aug 2023
- Spearheaded the development of a high-precision plant disease identification model, achieving a groundbreaking 98.3% accuracy for powdery mildew detection
- Co-engineered an advanced Agri-Foods chatbot utilizing the Generative AI GPT-4 architecture, integrating Azure ML for enhanced customer interaction and support
- Innovated an internal PDF Parser tool for the HR department, enabling efficient extraction and dashboard visualization of key document data with Spacy and PowerBI
- Optimize model hyper-parameters using Bayesian optimization techniques - random search & gridsearch.
- Fostered collaborative efforts - Standups, Git, and Azure Boards with cross-functional teams within an agile project environment.

Teaching Assistant (Fastlane - DaRMoD), Vector, Remote, ON Jan 2024 – Apr 2024

- Developed and delivered comprehensive tutorials in data readiness, model development, and deployment, effectively simplifying complex AI and machine learning concepts for Startups.
- Provided constructive feedback and evaluations on company projects, leveraging expertise in machine learning to enhance project outcomes.

Graduate Teaching Assistant, University of Windsor, Windsor, ON Sep 2022 - Apr 2023

- Conducted lab sessions and solved students' queries for **50+** students every week for Computer Architecture (COMP-2660) and Systems Programming (COMP-2560) courses
- Grading **50+** students' exam papers and assignments and held 1-on-1 sessions for students in need of extra help

Programming Intern, Coderspacket, Vadodara, India Apr 2021 - Jun 2021

- Developed **3** reusable open-source code packets using JAVA & Python. Packets included facial recognition system, registration system for colleges & chatbot interface
- Created packets were deployed using Docker and downloaded **100+** times by fellow developers for use in bigger open source projects & scaling applications & process improvement
- Performed functional testing and debugging for developed packets
- Worked in the transition of a monolithic app into a microservice architecture using .NET & C# as part of a team

Web Development Intern, Verzeo, Vadodara, India Dec 2019 - Jan 2020

- Responsible for developing and designing login, registration, and other fragmented components using HTML, CSS, JavaScript, Material-UI and React framework.
- Worked closely with team members and actively participated in weekly Agile meetings with Jira
- Led a team of **4** members to execute the back-end functioning of the site using NodeJS, SQL & Git.
- Worked on producing dynamic data graphs for company's website using JavaScript library - D3.js

Paper Reviewer & Program Committee Member, IEEE - AIC, Remote May 2022 - Jun 2022

- Reviewed & critiqued **10** research papers for the IEEE world conference - Applied Intelligence & Computing and gave detailed feedbacks on scope of improvements
- Ensured smooth conduct of conference by holding meetings with other committee members and helping technical team by creating data reports & providing creative solutions for quality improvement

Course Mentor, DEEPLEARNING.AI, Remote Aug 2021 - Dec 2021

- Mentored **50+** students for 'AI for Medicine' course on Coursera & solved programming-related queries & content queries using insightful solutions
- Suggested **4** major improvements in course contents & shared data analytics, and prepared reports for the DLAI team to improve quality, customer satisfaction & service delivery
- Updated technical documentations to enhance course delivery process

PROJECTS

Fine Tuning Llama-2 with Custom Platypus Dataset | *Python, PyTorch, Streamlit* Dec 2023 – Feb 2024

- Employed **Supervised Fine-Tuning (SFT) to adapt Llama 2** for enhanced dialogue generation, leveraging a meticulously curated **custom mini-Platypus dataset** for instruction-based training.
- Applied near-deduplication techniques **using Sentence Transformer embeddings** and FAISS to maintain dataset uniqueness, significantly reducing redundancy without compromising data integrity.
- Pioneered the application of **4-bit precision fine-tuning using QLoRA**, significantly minimizing VRAM usage while maintaining model performance.
- Developed a **text generation pipeline** that incorporates the refined model, demonstrating the practical application of fine-tuning through enhanced text generation capabilities.

Explainable Stock Portfolio Optimization (Thesis project) | *Python, OpenAI Gym* Jan 2023 - Dec 2023

- Pioneered a cutting-edge stock prediction model **integrating Graph Neural Networks (GNN) with Reinforcement Learning (RL)**, enhancing market trend analysis and investment strategy formulation.
- Engineered predictive algorithms by processing vast datasets with advanced techniques like **Node2Vec for embedding generation** and **TD-3** for reinforcement.
- **Enhanced model transparency and interpretability** by incorporating SHAP, LIME, attention networks, and counterfactual explanations, facilitating a deeper understanding of predictive factors and model decisions

Diabetes Readmission Analysis | *Python, R, SQL, PySpark, MLlib* Apr 2023 – July 2023

- Performed **analysis of more than 1,00,000 Clinical Database Patient Records** aiming to understand the factors that are responsible for early readmission of a patient given his/her clinical information
- Implemented Logistic Regression, Lasso Regression, and Random Forest models **using PySpark and MLlib**
- Utilized feature selection methods like **Chi-Square and AIC** (Akaike Information Criterion)
- Engineered a data processing **pipeline with PySpark**, including **VectorAssembler** for feature vectorization
- Achieved a **ROC score of 0.64** with Random Forest indicating a strong predictive capability compared to baseline

Fish Image Classification | *Python, Tensorflow, CUDA* Aug 2023 - Dec 2023

- Executed complex image segmentation techniques on a dataset of **9,000+** images, integrating over **20** distinct attributes to accurately distinguish aquatic species within their ecosystems.
- Benchmarked and analyzed the extraction efficacy of advanced neural network architectures, including **Inception**, **ResNet**, and **EfficientNet**, to ascertain the veracity of the proposed model.
- Employed a rigorous, data-centric methodology to fine-tune a deep learning model, achieving a leading-edge classification accuracy of **99.68%**

PUBLICATIONS

Desai, N. P., **Baluch, M. F.** & Aziz, R. M. (2023). Computer vision model with novel cuckoo search based deep learning approach for classification of fish image. Multimedia Tools and Applications, 1-20.

Baluch, M. F., Patel, S., Aziz, R. M. & Ganie, A. H. (2022). LGBM: a machine learning approach for Ethereum fraud detection. International Journal of Information Technology, 1-11.

Baluch, M. F., Patel, S., Aziz, R. M. & Kumar, P. (2022). A machine learning based approach to detect the Ethereum fraud transactions with limited attributes. Karbala Int. J. Mod. Sci, 8, 139-151.

Desai, N. P., Wadhwani, A., **Baluch, M. F.,** & Mishra, N. (2021, September). A Comparative Assessment Study on Machine Learning Classifiers for Cardiac Arrest Diagnosis and Prediction. In 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-6). IEEE.

AWARDS

Departmental, University of Windsor - $4,000 Jan 2023

Provincial, Ontario Graduate Scholarship - $15,000 Sep 2022

95th /40,000, National Engineering Olympiad 4.0 Jun 2021

Institutional, VIT Bhopal GVSDP Scholarship - $1,000 Jul 2018

CERTIFICATIONS

Machine Learning: Algorithms in the Real-World, Coursera (Specialization) Aug 2021

Data Management & Visualization, Coursera (Wesleyan University) Jul 2021