VARSHA VISHWAKARMA

Tel: +1 (716)-533-3541 | Email: vvishwak1883@gmail.com
https://www.linkedin.com/in/varsha-vishwakarma-9284958a/ | https://github.com/varsha1883

EXPERIENCE SUMMARY

Machine Learning Engineer with over **6 years** of industry experience and **1 year** as a Research Assistant, specializing in high-impact projects across various sectors, including retail, recommendation systems, consumer banking, revenue generation, and image data analysis. Leading impactful research, focusing on Large Language Models, and bringing a keen understanding of generative techniques for alignment and evaluation methods to enhance customer-facing experiences.

EDUCATION

Master of Professional Studies: Data Science, University at Buffalo, *Buffalo, NY, Dec 2023* **GPA:3.9**
Bachelor of Technology: Electronics & Instrumentations, KIIT University, *India, May 2016* **GPA:3.4**

TECHNICAL SKILLS

Statistical: Data Mining, Data Wrangling, Probability, Hypothesis Testing, Correlations, Association Rule, Decile, Principal Component Analysis, A/B Testing, ETL, Descriptive Statistics
Machine Learning: MLOps, Regression (linear, logistic, regularized, sparse), Classification (SVM, KNN, Decision Trees, Random Forest, XG Boost), Bayesian, GLM, Deep Learning, LSTM, RNN, Clustering, Anomaly Detection, Reinforcement Learning, Computer Vision, Word2Vec, K-Means, Model Monitoring, Generative AI, Encoder and Decoder, Open AI GPT, LLM
Environment & Libraries: Linux, Ubuntu, Windows, GitLab, scikit, Pandas, NumPy, Matplotlib, Dask, NLTK, spaCy, Flask, Gensim, Glove, Tesseract, PySpark, OpenCV, TensorFlow, PyTorch, Keras, PIL, Global Surrogate, LIME, Microsoft Azure, AWS, GCP, Dask, JAX, Docker, Kubernetes, VPC, IAM, RestAPI, YOLO, Rasa-X, SciPy, Databricks, Amazon Sage maker
Programming & Tools: Python, R, SQL, SAS, Big Query, Tableau, Power Bi, PyCharm, Hadoop, HIVE, VS Code, Excel, PowerPoint

PROFESSIONAL EXPERIENCE

Research Intern, University at Buffalo | NY, USA ***Jan 23 – Present***

- Generated realistic morphed facial images using Style**GAN** and MIPGAN, which impact the ability to differentiate them from bogus images.
- Conducted human **annotation** to evaluate the image quality of morphed images, facilitating a comprehensive comparative analysis to understand generative model capabilities, and built a user-friendly interface to provide a seamless platform for annotators to evaluate and assign quality scores to patches of facial images.
- Pioneered the implementation of statistical tests on a facial image quality dataset, annotated by five different annotators, featuring 20% overlapping and 10% duplicate images. Strategically measured annotator consistency and inter-annotator correlation, enhancing the reliability and accuracy of image quality evaluations.
- Engineered a cutting-edge image quality prediction model by integrating **ResNet** for spatial feature extraction and a **Transformer** for capturing long-range dependencies. Enhanced prediction accuracy by 8% compared to conventional models through innovative algorithm optimization.
- Designed an advanced healthcare conversational system using the **Llama2** model using **HuggingFace** transformers and **LangChain**, enhancing patient-provider interactions, automating responses to common inquiries, improving the overall user experience in the healthcare sector, and creating a chatbot to provide real-time responses to user queries, enhancing customer support and user engagement.

Technical Specialist (Data Science / Machine Learning), ***Zensar |*** *Bengaluru, India* ***Mar 21 - Aug 22***

- Led the development of advanced **chatbot** workflows using Rasa Framework for IT support and HR, employing NLP techniques like intent recognition and entity extraction to refine model accuracy and performance.
- Effectively handled a variety of **structured and unstructured data** sources, performing **annotation** tasks to prepare training data for chatbot training and development.
- Integrated **CI/CD** pipelines and **Docker** containers to streamline the deployment process into team workflows to facilitate continuous improvement and rapid iteration of chatbot functionalities.
- Collaborated closely with **cross-functional** teams to ensure chatbot solutions were user-centric, aligning with specific requirements and enhancing overall user interaction and engagement experiences.
- Architected and deployed an end-to-end Machine Learning system on **AWS Cloud** to fully automate the **data visualization** process, achieving a 75% reduction in time and effort for producing interactive visualizations.
- Built the Vinci **Machine Learning framework** for Engineers and Data Scientists to build and deploy models as managed service on **Kubernetes**.

Machine Learning Engineer, ***Client:CIMB Bank |*** *Bengaluru, India* ***Feb 19 - Feb 21***

- Modeled a time series network for **forecasting** ATM cash demand, achieving a 28% reduction in logistic costs for cash replenishment enhancing overall **operational efficiency**, and reducing **logistic expenses.**
- Increased **Mortgage/ CPL** product sales by in the first quarter of model implementation by allowing the sales team to focus on high-value customers with the highest likelihood of product purchase.
- Integrated **fraud risk scoring ML models** mitigated the risk of fraudulent transactions, safeguarded sales integrity, minimized financial losses, targeted high-value customers, and positively impacted the company's bottom line by reducing **fraudulent activities** and increasing sales revenue.

- Applied Market Basket Analysis for detailed customer **purchase behaviour insights**, enabling the creation of tailored bank offers and **product recommendations**, subsequently enhancing **cross-selling** effectiveness and **customer satisfaction**.
- Implemented rigorous **statistical tests** (including t-tests, ANOVA, and regression analysis) and **machine learning ops** on large datasets to identify key **financial trends** and risk factors, leading to more informed strategic decision-making and operational efficiency.
- Extensive hands-on experience and high proficiency with structures, semi-structured and unstructured data, using a broad range of data science programming languages and big data tools including **R**, **Python**, **PySpark**, **SQL**, Scikit Learn, **Hadoop** MapReduce.

Machine Learning Engineer (R&D), ***Accenture*** **|** *Bengaluru, India* ***May 16 - Jan 19***

- Executed web scraping techniques to extract data from various **e-commerce** platforms like Amazon, eBay, and Facebook, ensuring a rich and varied dataset for analysis.
- Integrated Retail Trend model using Google Cloud Platform's natural language API and vision API by analyzing women's garments based on customer interest and forecasts **trend projections** using Recurrent Neural Network (RNN) to provide valuable insights to **retail** clients for **demand planning**, investment, and **warehouse management**.
- Efficiently deployed the model on **Google Cloud Platform**, leveraging **GCP** buckets and **BigQuery** for data storage and management, enhancing operational scalability and data accessibility.
- Developed a deep learning-based **OCR** system for bank document verification with an 82% accuracy rate, implemented in a production environment on **Azure**, ensuring robust and scalable document processing.
- Constructed a Convolution Neural Network (CNN-based) text segmentation and **BERT** classification model, integrating advanced NLP techniques and word embeddings to effectively match resumes with job descriptions for HR recruitment processes. Achieved a 79% accuracy rate by leveraging natural language understanding and semantic analysis, and actively collaborated with HR teams to fine-tune the model for enhanced effectiveness.

PATENTS & PUBLICATIONS

- **Patent Published**: An Imaging System and a Method for Image Quality Enhancement | 202221043339
- **Patent Published**: Method and Device for Performing Data Encryption using Quantum Computing | 202221035778
- **Patent filed**: Data-Driven Method and System for Generating Useful Insights | 202221058762
- **Journal Paper**: ATM Cash Replenishment with Clustering Series (LSTM Network) | May 20 | IJSER | ISSN2229-5518
- **Journal Paper**: Approaches for Offline Cursive Handwritten Character Recognition (OCR) |Jul 19 | IJSR | ART20199819
- **Journal Paper**: Iris Recognition using CNN with Normalization | Nov 19 | IJRAR | IJRAR19K6911