

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

Data Scientist 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 LLM 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: ML Ops, Regression (linear, logistic, regularized, sparse), Classification (SVM, KNN, Decision Trees, Random Forest, XG Boost), Bayesian, Deep Learning, LSTM, RNN, Clustering, Anomaly Detection, Reinforcement Learning, Computer Vision, Word2Vec, K-Means, Model Monitoring, Encoder and Decoder, GAN, BERT, Llama, GPT

Environment & Libraries: Linux, Ubuntu, Windows, GitLab, Pandas, NumPy, Matplotlib, Dask, NLTK, spaCy, Flask, Gensim, Glove, Tesseract, PySpark, OpenCV, TensorFlow, PyTorch, Keras, PIL, Global Surrogate, LIME, Microsoft Azure, AWS, GCP, JAX, Docker, Rasa-X, SciPy, Docker, Databricks, Amazon Sage maker

Programming & Tools: Python, R, SQL, Big Query, Tableau, Power Bi, PyCharm, Hadoop, HIVE, VS Code, Excel, PowerPoint

PROFESSIONAL EXPERIENCE

Data Science Specialist, 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.
- 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 AutoML system on **AWS Cloud** to fully automate the **data visualization** process, achieving a 75% reduction in time and effort for producing interactive visualizations.

Data Scientist, CIMB Bank / Bengaluru, India

Feb 19 - Feb 21

- Handled ML & AI projects focusing on **consumer banking** and was responsible for analytics implementation across marketing and data monetization teams.
- 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 CPL product sales by 12% in the first quarter of model implementation by allowing the sales team to focus on high-value customers more likely to purchase CPL products.
- 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) on large datasets to identify key financial trends and risk factors, leading to more informed strategic decision-making and operational efficiency.
- Provided onsite team training in Machine Learning and Natural Language Processing (NLP), tailoring sessions to address specific needs and fostering increased proficiency among team members.

Machine Learning Engineer, 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 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.

Face Image Quality Assessment and Face Morph Detection

- Generated realistic morphed facial images using StyleGAN and MIPGAN, which impact the ability to differentiate them from bogus images.
- Conducted human annotation to score image quality, including both bogus and morphed images for evaluating the accuracy and effectiveness of FIQA.
- Built a user interface to provide a seamless and user-friendly platform for annotators to evaluate and assign quality scores to patches of facial images.
- Performed comprehensive comparative analysis of bonafide, StyleGAN, and MIPGAN-generated images, employing advanced metrics to assess and contrast their quality and authenticity. This analysis facilitated a deeper understanding of generative model capabilities and the characteristics distinguishing real from synthetic facial imagery.
- 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.

Generative AI

- Designed an advanced healthcare conversational system using the **LLM** model, enhancing patient-provider interactions, automating responses to common inquiries, and improving the overall user experience in the healthcare sector.
- Created a chatbot to provide real-time responses to user queries, enhancing customer support and user engagement.

PATENTS & PUBLICATIONS

- **Patent Published:** An Imaging System and a Method for Image Quality Enhancement | App no. - [202221043339](#)
- **Patent Published:** Method and Device for Performing Data Encryption using Quantum Computing | App no. - [202221035778](#)
- **Patent filed:** Method and System for generating useful insights | App no. - [202221058762](#)
- ATM Cash Replenishment with Clustering Series (LSTM Network) | May 20 | IJSER | [ISSN2229-5518](#)
- Approaches for Offline Cursive Handwritten Character Recognition | Jul 19 | IJSR | [ART20199819](#)
- Iris Recognition using CNN with Normalization | Nov 19 | IJRAR | [IJRAR19K6911](#)