

Riyanshi Bohra

352-757-8079 | riyanshibohra@gmail.com | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of Arizona

Tucson, AZ

Master of Science (M.S.) in Data Science; **GPA: 3.8/4.0**

Aug 2023 – May 2025

Coursework: Artificial Intelligence, Machine Learning, Natural Language Processing (NLP), Data Warehousing

Manipal University Jaipur

Jaipur, India

Bachelor of Technology (B.Tech) in Information Technology; **GPA: 3.72/4.0**

Jul 2019 – May 2023

Coursework: Deep Learning, Database Management, Object-Oriented Programming, Statistics, Data Structures

PROFESSIONAL EXPERIENCE

Research Assistant (Data Analyst)

Jan 2024 – Present

Mel and Enid Zuckerman College of Public Health, University of Arizona

Tucson, AZ

- Led a research study on shade accessibility gaps, analyzing **1 billion+** data points across **15** years, generating insights for **3** policy reports that influenced funding decisions impacting **2** million students
- Built dashboards in Tableau to analyze over **30** spatial-temporal trends, visualizing shade distribution patterns across **200+** school districts and enabling data-driven policy decisions
- Examined disparities across **1,200** school playgrounds using ML in Python, identifying a **25%** shade deficit in low-income areas and guiding **\$750K** in funding reallocations for heat resilience

Data Science Intern - Technology Consulting

Jul 2022 – Sep 2022

PwC (PricewaterhouseCoopers)

Mumbai, India

- Engineered ETL pipelines using Python, SQL, and Apache Spark, automating data ingestion and transformation for **15,000+** monthly manufacturing records, reducing cycle time by **45%**
- Trained ML models on **7+** years of manufacturing data to optimize operational workflows, increasing tablet production by **20%** annually (**400K** → **480K units**) and achieving cost savings of **\$50,000/year**
- Designed **5** interactive Power BI dashboards to analyze expenses and track KPIs across **7** production lines, improving decision-making and saving **350+** man-hours/week

TECHNICAL SKILLS

Languages: Python (Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn), R, SQL (MySQL, PostgreSQL), JavaScript

Frameworks: React, Flask, FastAPI, PyTorch, TensorFlow, NLTK, SpaCy, Hugging Face Transformers

BI and Reporting: Power BI, Tableau, Looker, Amazon QuickSight, Microsoft Excel (Pivot Tables, Macros, VBA)

Generative AI: LangChain, LangGraph, VectorDB (Pinecone, Deep Lake), Retrieval-Augmented Generation (RAG)

Developer Tools & Cloud: Git/GitHub, Docker, VS Code, Jupyter Notebook, Cursor, RStudio, AWS, Azure

PROJECT EXPERIENCE

VideoMind AI: Video to Insights | Python, LangChain, Pinecone VectorDB, Streamlit

[GitHub](#)

- Developed a RAG-based video analysis system with multi-video support, capable of processing upto **600** minutes of content with **95%** transcription accuracy and **3x** faster comprehension
- Engineered a hybrid processing pipeline using Pinecone and LangChain, reducing **45-min** videos to **2-min** smart summaries, maintaining search across **4K+** vectors with **sub-200ms** response time

Portfolio+: Virtual Stock Portfolio Management | MySQL, Flask, Python, Streamlit

[GitHub](#)

- Launched a stock portfolio analytics platform, processing **2M+** daily stock updates and managing up to **1,000** user portfolios, ensuring real-time data accuracy with Yahoo Finance API
- Automated transaction logging for **25,000+** buy/sell activities monthly, optimizing query execution with MySQL and improving real-time financial trend analysis with **< 1.5-second** query response time

SafeDrive AI: Real-time Distracted Driving Detection | Python, TensorFlow, Computer Vision

[GitHub](#)

- Deployed a CNN-based classification pipeline using TensorFlow and Keras, analyzing live dash camera feeds to detect and classify **5** types of driver distractions with **92%** accuracy on a **50,000-image** dataset
- Integrated a real-time AI-powered alert system, using edge computing for on-device inference, reducing distraction incidents by **30%** across **200** monitored drivers in a **3-month** pilot study