

SHUBHANKAR JOSHI

☎ +1 (541)-250-1125 ✉ 16.shubhankar@gmail.com 💻 [s-joshi16](#) 📁 [portfolio](#) 🌐 [shubhu163](#)

SUMMARY

Driven AI Engineer skilled in machine learning, data analytics, and bioinformatics. A problem solver, and team player proficient in Python, PyTorch, SQL, and Tableau, seeking roles in the field of Data Science.

EDUCATION

Oregon State University, Corvallis, OR Sept 2023 – May 2025

Master of Science in Computer Science - **GPA - 3.81**

Relevant Coursework: Machine Learning, Deep Learning, DBMS, Optimization for AI, Algorithms, Cyber Security, & NLP

Savitribai Phule Pune University, Pune, India Aug 2019 – May 2023

Bachelor of Technology in Instrumentation & Control - **CGPA - 9.26 / 10**

Relevant Coursework: Data Structures, Web Technology, Project Management, Data Science, Computer Networks & Maths.

EXPERIENCE

Research Assistant, Oregon State University, Corvallis, OR Sept 2023 – Present

- Analyzed **50M+ ribosomal sequences** from 9 diverse human tissues, performing transcriptome alignment and extracting key insights on protein translation using **Bowtie and Pysam**, enhancing **Data Warehousing** capabilities.
- Streamlined **ETL** pipeline, reducing **data processing** time by **30%** with **SAS** and **Python**.
- Developed a hybrid **LLM** with transformer and CNN layers for ribosomal count predictions, applying **Data Modeling**.
- Automated project pipeline with **PyTorch Lightning**, enabling efficient model training and testing.
- Achieved a Pearson correlation coefficient of **0.837** between predicted and true ribosomal coverage in the **test set**, demonstrating strong **Predictive Analytics**.
- Integrated loss functions like **Poisson and multinomial NLL**, leading to a **5%** increase in model convergence speed.
- Optimized model parameters using **grid search** and **Optuna**, leading to a **10%** increase in model performance metrics.
- Automated **loss metric logging** and visualization using **Wandb**, resulting in a **15%** reduction in manual validation efforts.
- Generated **correlation heatmaps** for shape and count levels across tissues, identifying **key patterns with Power BI**.

Data Analyst Intern, Landmark Techedge Pvt Ltd, Pune, India Aug 2022 – May 2023

- Extracted & managed **5GB** daily production data from **MySQL, HDFS, and Excel** to create **Dashboards on Tableau**.
- Pre-processed data with **Python** and **SQL queries** & analyzed it through **Pandas** to calculate **10+** key performance indicators (**KPIs**) for decision making.
- Collaborated in **agile teams** to update real-time **data visualization dashboards** which increased **sales by 19%**.
- Co-developed a data pipeline architecture in a **team of five**, optimizing data flow for a **power management dashboard**.

PROJECTS

Coffee Shop Sales Dashboard 📈 | SQL, Excel, PowerBi. June 2024 – July 2024

- Created **Power BI dashboard** analyzing **10,000+** sales transactions, revealing trends and store performance.
- Used **SQL** to process data, identifying a **15% weekend sales increase** and **20% top product growth**.

Bird Classification Robustness 📈 | Python, PyTorch, Torchattacks. Jan 2024 – March 2024

- Developed a robust **bird species classifier using a ResNet** with a model **precision** of **98.53%** and **F1 score** of **98.14%**.
- Data Cleaning, Feature Engineering, Feature Selection, Model Training, Model Evaluation & Deployment

Car Insurance Sales Dashboard 📈 | SQL, Excel, Tableau June 2023 – Aug 2023

- Designed a Tableau dashboard handling **30,000+** insurance records, and established **10+ pivotal KPIs**.
- Integrated **Slicers, filters, and parameters** for intuitive data exploration in **Tableau**.

SKILLS

Skills: AI & ML, Data Analytics, Frontend, Bioinformatics, Project Management, Communication Skills.

Languages: Linux, Python, Java, C, C++, JavaScript, SQL, HTML, CSS.

Developer Tools: VS Code, Jupyter Notebook, MATLAB, Figma, PowerBI, Tableau, Excel, API.

Libraries/Frameworks: Pytorch, Numpy, Pandas, GitHub, Docker, ReactJS, NodeJS, Git, Pysam, Bowtie.

EXTRACURRICULAR & PUBLICATIONS

Led a **20-member team** as Aaroh Music Club Head, organized events, and performed as a **singer**. ([sing.shubhankar](#))

Automated Street Lights using Solar Energy : Published in National Level E-Conference on Science & Technology 2019.

Ocular Disease Recognition using Ensemble Techniques : Published on **IEEE Xplore** in **ICCUBEA - 2023**.

User Survey of Music Streaming Apps : Selected in **HWWE 2023** and currently in process of publication.