

# SHUBHANKAR JOSHI

Corvallis, OR | +1 (541)-250-1125 | joshishu16@gmail.com | linkedin.com/s-joshi16 | github.com/shubhu163

## SUMMARY

Software Engineer with hands-on experience in Python, C++, Java, and RESTful APIs. Skilled in building scalable systems, real-time applications, and automation pipelines. **Delivered a 30% boost in processing efficiency, 20% latency reduction, and 20% model reliability gain across diverse AI-driven projects.**

## SKILLS

- **AI & Machine Learning:** Transformers, CNNs, Optuna, PyTorch Lightning, Wandb, Adversarial Training, Grid Search
- **Programming Languages:** Python, Java, C++, Bash
- **Frameworks & Libraries:** Flask, PyTorch, TensorFlow, Torchvision, Keras, ONNX, OpenCV
- **Cloud, DevOps & Databases:** AWS, Docker, CI/CD, Databricks, MySQL, PostgreSQL, MongoDB
- **Tools & Methodologies:** Git, GitLab, JIRA, Jenkins, RESTful APIs, OOP, Real-time Systems, Software Design, Scalability
- **Soft Skills:** Leadership, Innovation, Strategic Decision-making, Driving Efficiency, Hands-on, Collaborative Work, Reliability

## EXPERIENCE

**Research Assistant**, Oregon State University, Corvallis, OR.

Sept 2023 – Present.

- **Achieved a Pearson correlation of 0.837** in ribosomal sequence prediction by **developing a hybrid LLM architecture** using Transformers and CNNs in **PyTorch**, applied to transcriptome data from 9 human tissues over the first year
- **Boosted data processing efficiency by 20% in under 2 months** by designing and optimizing ETL pipelines using **Python, SAS, and Snowflake**, handling 50M+ ribosomal sequences and enhancing **data warehousing capabilities**.
- Reduced manual validation **efforts by 15%** through automating model deployment using **PyTorch Lightning**
- **Fine-tuned model performance by 10%** through targeted hyperparameter search using **Optuna and grid search**.
- Enabled real-time insights delivery by building **automated Tableau and Looker** pipelines based on user feedback.
- Strengthened system scalability and robustness by developing efficient unstructured data handling modules, including indexing strategies and validation routines for high-volume biomedical records.

**Data Analyst Intern**, Landmark Techedge Pvt Ltd, Pune, India.

Aug 2022 – May 2023

- **Improved data efficiency by 25%** over 3 months by **building automated ETL pipelines** using SQL Server, SSIS, and ADF, deployed on the internal production analytics platform.
- **Drove a 19% increase** in sales of Q4 2022 by creating **Tableau** dashboards of 5GB of daily data, to support strategy decisions.
- Enhanced forecasting accuracy and speed by automating data integration with **Python, Alteryx, and BigQuery**.
- **Increased scalability by 25%** by optimizing inventory management scripts, applied in **supply-chain tracking systems**.
- Delivered reliable in-house tools by **collaborating cross-functionally** on performance-boosting solutions.

**Software Engineer & Operations Lead**, Samarth Enterprises, Pune, India.

Nov 2020 – July 2022

- Reduced inventory tracking time **by 30% within 5 months** by building automation tools with **Python and Excel VBA**.
- **Elevated defect detection by 10%** by creating **Power BI dashboards** for **real-time QC on wire harnesses and regulators**.
- **Led 10+ technicians** over a year during ramp-ups, managing tasks, escalations, and **Piaggio delivery timelines**.

## EDUCATION

**Master of Science in Computer Science - GPA – 4.0**

Sept 2023 – Dec 2025

Oregon State University, Corvallis, OR.

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

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

Aug 2019 – May 2023

Savitribai Phule Pune University, Pune, India.

**Relevant Coursework:** Data Structures, Web Technology, Project Management, Data Science, Networks & Statistics.

## PROJECTS

**Bird Classification Robustness** | Python, PyTorch, Torchattacks.

April 2024 – June 2024.

- **Achieved 98.5% precision and 98.14% F1** by developing a robust bird classifier using PyTorch and Torchattacks.
- Ensured model reliability by performing chunk-based training, adversarial testing, and rigorous evaluation.

**DDOS Attack Detection System** | Python, Tensorflow, Scikit-Learn.

Jan 2024 – March 2024

- **Achieved 93.57% F1 score** by implementing a GAN-based defense on CIC-IDS datasets for DDoS detection in NIDS.
- **Raised adversarial detection accuracy from 5% to 40%** using WGAN and feature extraction on 20 key attributes.

**Real-Time Chat Application** | React, Node.js, WebSocket, MongoDB

Sept 2023 – Dec 2023

- Built a **real-time chat application** with group chats, message history, and user presence.
- Cut message latency by 20% by integrating **Redis caching into the backend** a **WebSocket-based** chat app.

## RECOGNITION

- Co-authored a paper on ensemble-based ocular disease recognition, published in **IEEE Xplore (ICCUBEA 2023)**.
- Led a 20-member team as **Aaroh Music Club Head**, organized events, and performed as a **singer**. ([sing.shubhankar.com](https://sing.shubhankar.com))