PIYUSH PANWAR

Software Engineer

📞 +91-8306968232   @ piyush2002panwar@gmail.com   📍 Gurugram, India   in Linkedin   ⌥ Github   Leetcode

EXPERIENCE

AI/ML Intern
Detoxio AI ↗

📅 June 2025 – Sept 2025   📍 Remote

- Developed **AI agent configurations** for model testing, enhancing the efficiency of **red teaming** processes
- Conducted extensive testing on models from **Ollama** and **Hugging Face**, focusing on advanced learning techniques such as **Jailbreaking** and adversarial attacks
- Automated testing procedures using **NOX**, significantly reducing manual effort and increasing testing throughput by 40%

Research Intern
Statistical Modeling (Academic) ↗

📅 Oct 2025 – Present   📍 Remote/Academic

- **Pioneered** a Monte Carlo simulation study to estimate **Stress-Strength Reliability (SSR)** for the **Xgamma-Exponential (Xg-E) distribution**.
- **Implemented** advanced **MLE** techniques under **GPHC** and **Debugged** a fundamental error in the core mathematical expression for SSR.
- **Optimized** the simulation framework using **Python** and **Numba (JIT)**, achieving high precision across **10,000 replications** and reducing computation time.
- **Analyzed** simulation outputs using Matplotlib and Pandas to visualize reliability trends, validating theoretical models against empirical data.

Open Source Contributor (ArviZ)
ArviZ – Bayesian Analysis Library ↗

📅 Jan 2024 – Present   📍 Remote

- **Spearheaded** core computational features (e.g., implemented `bayes_factor()` and developed `plot_ppc_intervals()`) for advanced **Bayesian Model Comparison** and validation.
- **Led** architectural refactoring initiatives, migrating plotting modules for improved **modularity, scalability, and clean separation of concerns**.
- **Reinforced** test reliability and code stability by significantly **extending Pytest coverage** and implementing **Test-Driven Development (TDD)** across statistical features.
- **Enhanced** user accessibility across visualization backends and ensured code quality through refactoring of core statistical logic (`KDE`) and adherence to open-source standards.

Detailed Contributions: View Portfolio of 8+ Merged Pull Requests ↗

EDUCATION

B.Tech. (CSE) - 8.62 CGPA
Polaris School of Technology (Starex University)

📅 2023 – 2027   📍 Gurugram, Haryana

TECHNICAL SKILLS

- **Languages:** Python (Numba/JIT), SQL, Java, JavaScript
- **AI & GenAI:** Hugging Face, Ollama, XGBoost, Scikit-learn, Optuna, Model Validation
- **Statistics & Research:** Bayesian Inference (ArviZ), MLE, Survival Analysis, Pytest, Hypothesis Testing
- **Data Science:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Bokeh, Streamlit
- **Web & Tools:** Git, Docker, Sphinx, REST APIs, Postman, React.js, HTML/CSS
- **Core Concepts:** Data Structures, Algorithms, Code Refactoring, Statistical Analysis

ACHIEVEMENTS

- **"First Runner-Up in Hack With Uttarakhand (Team Code Heist)** - led a 36-hour offline hackathon with innovative problem solving and teamwork.'
- Contributed to **[Open Source]** initiatives including **HacktoberFest, GirlScript**, enhancing real-world software development skills.

PROJECTS

Credit Risk Prediction System ↗

- **Python | Scikit-learn | XGBoost | Optuna**
- Developed an end-to-end credit risk scoring engine that classifies borrowers into four risk tiers, predicting default probabilities and assigning credit scores (300–900).
- Engineered critical financial features (e.g., Loan-to-Income, Utilization Ratios) and leveraged **Optuna** for Bayesian hyperparameter optimization to maximize model AUC-ROC.
- Designed and deployed an dashboard using **Streamlit**, providing real-time risk analytics and AI insights for loan approval decisions.

Health Insurance Premium Prediction ↗

- **Python | XGBoost | Linear Reg. | Statistics**
- Built a high-precision prediction model achieving **98% accuracy** addressing high-error through strategic data ("Young" vs "Rest").
- Significantly reduced extreme prediction errors (from 27% to 2%) by conceptualizing and integrating a domain-informed **"Genetical Risk"** feature for the under-25 demographic.
- Utilized **Variance Inflation Factor (VIF)** analysis to eliminate multicollinearity and implemented a hybrid approach to ensure model stability and interpretability.