Tushin Kulshreshtha
San Francisco, California 94132
+1 (925) 819-4160 tushink04@gmail.com dextron04.in

Summary

Software Engineer with 4+ years building machine learning-powered internal tools and data processing systems. Expertise in Python, TensorFlow, PyTorch, and C++ for ETL pipelines, analytics dashboards, and ML model deployment. Proven experience developing GenAI applications, real-time data visualization systems, and automated analytics tools that transform massive datasets into actionable engineering insights. Strong background in database optimization (MongoDB, MySQL, PostgreSQL) and production system design.

Experience

SellWizr **June 2025 – Present**
Software Engineer Intern *New York, NY*

- Built comprehensive ETL data processing system using Python and C++ with automated analytics pipelines, processing 100K+ machine-generated records daily to provide engineering teams with real-time insights and performance dashboards.
- Developed internal monitoring tools with Python, Node.js, and MongoDB for engineering productivity analytics, implementing ML-based anomaly detection that reduced system debugging time by 40% through automated pattern recognition.
- Engineered web-based visualization dashboards using JavaScript and REST APIs with comprehensive data analysis capabilities, enabling engineers to make data-driven decisions through interactive charts and automated reporting systems.

Site Service Software **Aug 2024 – Apr 2025**
Full-Stack Software Engineer Intern *San Francisco, USA*

- Architected internal data processing tools using Python, C++, and PostgreSQL with advanced analytics capabilities, building ETL systems that transformed legacy engineering data into modern visualization dashboards used by 50+ engineers daily.
- Developed 20+ REST-based APIs for internal analytics tools using Python and Node.js, implementing database optimization and query performance improvements that reduced data processing time by 30% for engineering teams.
- Built machine learning-powered recommendation system using TensorFlow and Python for internal tool optimization, implementing automated data classification and pattern recognition that improved engineering workflow efficiency by 25%.

MeetX **Nov 2023 – July 2024**
AI/ML Engineer Intern *San Francisco, USA*

- Architected production-ready ML system using PyTorch, TensorFlow, and Python for real-time data analysis, achieving 95% accuracy in pattern recognition for 50K+ concurrent users through advanced deep learning algorithms and model optimization.
- Implemented GenAI and LLM integration using Python with automated prompt engineering and token-based solutions, developing classification models and trend visualization tools that enhanced user engagement by 40% through intelligent content processing.
- Optimized ML model performance using Python frameworks (XGBoost, Scikit-learn) and MongoDB for large-scale data processing, achieving 60% faster inference times through distributed computing and cloud deployment on production systems.

Glitter Fund **Jan 2023 – Dec 2023**
Backend Developer Intern *San Francisco, USA*

- Led development of internal analytics platform using Python, C++, and PostgreSQL for processing financial engineering data, building ETL pipelines and ML-powered risk assessment algorithms that automated manual analysis workflows.
- Engineered real-time data processing systems using Python and MongoDB with advanced analytics capabilities, implementing machine learning models for pattern recognition that improved data quality assessment by 45% for engineering teams.
- Built comprehensive database optimization tools using PostgreSQL and Python with automated migration systems, creating internal documentation and guides that made data processing tools accessible to 20+ engineers across multiple teams.

Technical Projects

AI-Powered Data Processing Platform — Link to GitHub

- Built comprehensive GenAI system using Python, Flask, and OpenAI GPT models with automated content generation and analytics pipelines, implementing prompt engineering and token-based classification for real-world deployment scenarios.
- Developed ETL data processing workflows with Python and machine learning algorithms, creating internal tools for automated content analysis and trend visualization that reduced manual processing time by 80% through intelligent automation.

Advanced Entity Resolution System — Link to GitHub

- Engineered ML-powered data cleaning and entity resolution platform using Python, TensorFlow, and advanced machine learning algorithms for automated duplicate detection and data standardization across large enterprise datasets.
- Implemented real-time entity matching system with configurable ML models and confidence scoring, creating comprehensive audit trails and pattern recognition capabilities that improved data quality assessment by 60% for engineering workflows.

Intelligent Data Scraping and Analytics System — Link to GitHub

- Built sophisticated data extraction and analysis platform using Python, machine learning libraries, and MongoDB with intelligent lead scoring algorithms and automated pattern recognition for financial data processing.
- Developed comprehensive ETL pipelines with Python and advanced analytics capabilities, implementing modular architecture with detailed logging and monitoring systems that process massive amounts of machine-generated data into actionable insights.

Real-Time Communication Platform — Link to GitHub

- Architected scalable messaging system using Node.js, JavaScript, and PostgreSQL with real-time data processing capabilities, implementing WebSocket-based communication and database optimization for high-performance internal tool requirements.
- Built comprehensive web-based dashboard with JavaScript, REST APIs, and responsive design, creating internal documentation and user guides that made the platform accessible to engineering teams across multiple departments.

Technical Skills

Programming & ML: Python (Jupyter), C/C++, TensorFlow, PyTorch, Scikit-learn, Machine Learning, GenAI/LLMs
Data & Analytics: ETL Systems, Data Processing, Analytics Pipelines, Data Visualization, Pattern Recognition
Databases: MongoDB, MySQL, PostgreSQL, Elasticsearch, Database Optimization, Query Performance
Tools & Frameworks: Node.js, JavaScript, Kafka, REST APIs, Docker, Git, Internal Tools Development

Education

San Francisco State University **Expected Fall 2025**
Master of Science in Computer Science — GPA: 3.95/4.0 — Phi Beta Kappa Honor Society *San Francisco, California*
Relevant Coursework: Machine Learning, Data Structures & Algorithms, Database Systems, Software Engineering, Data Science, Computer Vision