

Fnu Tushar

Malvern, PA | +1 (484) 716-7840 | fnu.tushar@outlook.com | linkedin.com/in/fnu-tushar-945648131 | github.com/tus2014ar

PROFESSIONAL SUMMARY

Data scientist building end-to-end ML pipelines and LLM-based systems, with 5+ years of corporate industry experience. MS in Data Analytics, Penn State (Dec 2026); seeking full-time Data Scientist / Applied ML roles.

TECHNICAL SKILLS

Programming: Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, PyTorch, TensorFlow, Keras), SQL, R, Java.

Machine Learning & Statistical Techniques: Regression, Classification, Clustering, NLP, Time Series Forecasting, Feature Engineering, Model Evaluation (AUC-ROC, F1), EDA, Data Analysis, Hypothesis Testing.

LLM & AI Engineering: SBERT, LLM Integration (Gemini, Llama 3), Prompt Engineering, Claude API, AI Governance.

Data Tools & Cloud: Databricks, PySpark, Tableau, Data Warehousing, AWS, GitHub Actions, Docker, MLOps.

AI Productivity Tools: ChatGPT, Julius AI, Cursor IDE, Claude CLI, Claude Code, GitHub MCP, Claude Skills.

WORK EXPERIENCE

Barton Malow

May 2026 – Aug 2026

Intern - Data Scientist

Michigan, United States

- **Trained a linear regression model** to forecast \$53K-\$57K annual Databricks platform spend and flag \$11.6K idle cost.
- **Ran trend analysis on 8,000+ Claude queries** (20+ users) on Databricks: AI's spend share grew 5.5x in 5 months.
- **Built an AI-assisted pipeline:** 211 data-quality checks on 23 tables flagged 11k records; traced a 39% completeness bug.

Pennsylvania State University

Sep 2025 – May 2026

Research Assistant - Data Science

Pennsylvania, United States

- **Built an NLP matching model** (TF-IDF to SBERT) for 120 candidates vs 2,277 projects: 96% accuracy, 122% match gain.
- **Developed a local-first AI compliance monitor** for MyRaina Inc: 5 CMS/NIST rules, Llama 3 explanations, SHA-256 audit log; flags restraint-risk patterns in fall monitoring without sending resident data off-network.

Oracle Corporation

Feb 2022 – Dec 2024

Staff Data Consultant

Bangalore, India

- **Validated billing data** for 600K+ telecom subscribers using SQL checks, anomaly rules and reconciliation: 98%+ accuracy.
- **Created and automated churn and usage KPI dashboards** for weekly leadership reviews, delivered 99% on time.
- **Led onsite testing** for a 15-member team through a 1-month go-live, presenting a declining post-launch bug trend.

Accenture Solutions Pvt Ltd

Jun 2019 – Feb 2022

Data Consultant

Pune, India

- **Improved billing accuracy for 4.5M customers** by resolving 50+ recurring data quality issues in production database.
- **Automated batch bill-run workflows** in SQL: 1 day to 2 hours (12x faster), manual effort down 92%.
- **Designed a SQL data validation framework** (row-level checks, reconciliation, exception flags) that cut discrepancies 60% and shortened month-end close 10%.

EDUCATION

Pennsylvania State University

Pennsylvania, United States

Master of Data Analytics, GPA: 3.96/4

Aug 2025 – Dec 2026

Statistical Analysis, Data-Driven Decision Making, Predictive Analytics, Deep Learning, Natural Language Processing, AI Ethics.

KIIT University

Odisha, India

Bachelor of Technology in Electronics and Instrumentation Engineering, GPA: 3.3/4

May 2015 – May 2019

Control Systems, Digital Electronics, OOP, Data Structures & Algorithms, Artificial Intelligence.

PROJECTS

Osteoarthritis Market Share Forecasting Pipeline

[GitHub](#)

Decision support for an OA injectable brand manager: where Zilretta's visit share is moving, what to forecast and what not to trust, via an ML pipeline, analytics site and a Claude question box on a knowledge graph.

- **Engineered a re-runnable pipeline** that turns IQVIA Excel (72 months, 149K rows) and 171 public datasets into one audited, quality-checked star-schema warehouse.
- **Tested 7 feature groups and 3 models** without leaking future data: logistic regression improved segment forecasts in all 46 test months (+0.44%); nothing beat last month nationally.
- **Implemented a Streamlit site** where an AI assistant answers brand-manager questions from a sourced knowledge graph.

Worldwide Sentiment Analysis of LLM Tweets

[GitHub](#)

Early-warning view of where public trust in LLMs is turning negative, by region and over time, from 305K tweets using NLP scoring, geographic cleanup and time-series forecasting.

- **Cleaned 305K tweets;** normalized locations (aliases, bio inference, geocoding): Other/Unknown 72.8% to 43.9%.
- **Built a sentiment pipeline:** VADER scoring, ridge regression on log-likes and a SARIMAX forecast with weekly seasonality.
- **Delivered an AI-policy brief:** South America fell fastest (0.249 to 0.197); Bard/Gemini scored lowest (0.12-0.14).

PUBLICATIONS

First-authored "Autonomous Multi-Agent Governance: AI Sentinel Framework for Mitigating Psychological Restraints in AI-Driven Fall Management," [Springer CCIS vol. 3126](#), SEET 2026, with industry partner MyRaina Inc.