Umang **Sharma**

(267) 249 9980 umangsharma.cs@gmail.com

Website Github LinkedIn HackerRank

Education

University of Pennsylvania 2024-2026
Masters in Computer and Information Sciences CGPA: 3.9/4.0

- **Major Coursework -** Software Systems, Advanced topics in Deep Learning, Computer Vision, Network Systems, Analysis of Algorithms

Thapar Institute of Engineering and Technology 2018-2022
B.Tech. in Computer Engineering CGPA: 9.46/10

Skills

Programming Skills - Python, R, C, C++, C#, JavaScript, PHP, SQL, Perl
Libraries and Frameworks - ReactJS, Pandas, PyTorch, Numpy, Scikit-Learn, Keras, TensorFlow, Seaborn, Plotly, Django, Flask, Node
Databases and SQL - Snowflake, Airflow, BigQuery, MySQL, PostgreSQL, FireStore, MongoDB
Methodologies and Concepts - Machine Learning, Data Analytics, Business Intelligence, Recommendation Systems, Reinforcement Learning, Learning Optimization, Data Science, NLP, Deep Learning, Time Series Forecasting, Excel, PowerBI, Statistics, REST APIs

Work Experience

AI and Data Systems Intern – NeedleSpotter Inc. June 2025 – August 2025

- Built an end-to-end dashboard to fetch, display, and edit AI-generated tweets from medical literature with real-time modification capabilities.
- Engineered a multi-stage filtering system leveraging scoring rubrics (clinical relevance, impact, journal quality) to pre-select PubMed articles for downstream LLM generation.
- Engineered a feedback collection system within the dashboard interface, capturing human edits and preferences to create training datasets for **reinforcement learning from human feedback (RLHF)** model improvement.
- Implemented an iterative model training loop using edited tweet data to fine-tune **upstream Generative Models**, achieving improved content quality through continuous learning from user modifications and editorial feedback patterns.
- Implemented a plugin-based architecture for **automated article rewriting** using GPT APIs, designed for tunable readability and tone control based on user persona (e.g., oncologists vs patients).

Decision Analytics Associate (Software & AI Solutions) - ZS Associates July 2022 - April 2024

- Developed and deployed scalable ML models (DeepAR, TFTs, LSTMs) using Python, integrating them into production APIs with CI/CD pipelines, reducing manual forecasting by 50%.
- Built and optimized ETL pipelines in Python and SQL, improving data processing efficiency by 40% with parallelized workflows and Airflow automation.
- Designed and maintained full-stack applications using Flask, Django, and React and Tableau, enabling interactive data visualization and real-time insights.
- Engineered innovative software applications utilizing Deep learning models including Temporal Fusion Transformers, DeepAR and LSTMs getting correct predictions 80% of the time.
- Spearheaded a project to create interactive Tableau dashboards that visualized data analysis results from advanced machine learning models, enhancing decision-making processes for clients and increasing engagement by 30%.

Decision Analytics Associate Intern (Data & Software Engineering) - ZS Associates January 2022 - June 2022

- Automated ETL workflows using Python and SQL, reducing data processing time by 30% and improving data reliability.
- Integrated CI/CD workflows for ML models using Docker and GitHub Actions, streamlining deployment and cutting debugging time by 40%.
- Enhanced API performance for real-time data access, reducing response time and improving system scalability.
- Leveraged Bayesian regression models to analyze market trends, leading to a 27% increase in budget allocation efficiency.

Research Experience

Research Assistant – University of Pennsylvania August 2024 - Present

- Engineered a RoBERTa-based binary classifier for advice-seeking and gossip detection in emails, achieving an 81% F1 score, automating manual classification and increasing processing speed by 5x.
- Automated classification for 200,000+ emails, reducing manual workload by 80% and enabling large-scale behavioral pattern analysis.
- Discovered correlations between gossip frequency and advice-seeking interactions, leading to data-driven policy recommendations on workplace communication strategies.

Projects

Distributed CHORD System for P2P Networks – University of Pennsylvania January 2025 – April 2025

- Developed a distributed hash table (DHT) based on the CHORD protocol to enable decentralized storage and lookup with logarithmic hop guarantees.
- Simulated node join/leave and consistent hashing, maintaining fault tolerance and key redistribution correctness.

Autonomous Zooming Cameraman - University of Pennsylvania August 2024 - December 2024

- Developed and fine-tuned YOLOv11 model for basketball detection with improved accuracy (1280 image size, 100 epochs), integrated with CSRT tracking algorithm for continuous 30-frame ball tracking
- Designed and implemented a multi-stage video processing pipeline incorporating object detection, Gaussian blur-based heatmap generation, and temporal smoothing with 30-frame rolling buffer for professional-grade autonomous sports recording
- Created comprehensive post-game analysis system featuring team movement visualization through heatmaps, ball possession statistics, and automated commentary generation using OpenAI GPT-Vision API integration