

Umang Sharma

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

Github LinkedIn HackerRank

Education

University of Pennsylvania

Masters in Computer and Information Sciences

2024-2026

CGPA: 3.9/4.0

- **Major Coursework** - Machine Learning, Advanced topics in Deep Learning, Computer Vision, Analysis of Algorithms

Thapar Institute of Engineering and Technology

B.Tech. in Computer Engineering

2018-2022

CGPA: 9.46/10

Skills

Programming Skills - Python, R, C, C++, C#, JavaScript, SQL, Perl

Libraries and Frameworks - NextJS, Pandas, PyTorch, Numpy, Scikit-Learn, Keras, TensorFlow, Seaborn, Plotly

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

Work Experience

Decision Analytics Associate - ZS Associates

July 2022 - April 2024

- Employed regression analysis and predictive modelling with Random Forest Classifiers and data visualization, to traverse market landscapes, optimize budgets by 15%, identify key market competitors which resulted in a 20% increase in overall sales performance.
- 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 - ZS Associates

January 2022 - June 2022

- Analyzed market trends for budget allocation using Bayesian Regression, leading to a 27% increase in client budget efficiency.
- Designed and automated various ETL pipelines in SQL and Python, reducing data processing time by 30%, which enhanced overall workflow efficiency.
- Delivered insights via Tableau visualizations and Python Plotly dashboards, boosting decision-making speed by 25% and reducing debugging time by 40%.

Research Experience

Research Assistant – University of Pennsylvania

August 2024 - Present

- Implemented RoBERTa for binary email classification to identify advice-seeking and gossip patterns, achieving 81% F1 score, and developed a multi-class classification pipeline using T5.
- Automated classification for a dataset of over 200,000 emails, enhancing scalability and enabling behavioral pattern analysis.
- Analyzed correlations between gossip frequency and advice-seeking interactions, providing actionable insights into communication behaviors

A Pipeline-based Approach to Generate Meeting Minutes Using TOPSIS

September 2021

- Developed an MCDA-based pipeline solution that integrated seamlessly with existing workflows, resulting in the analysis of 2M+ data points.
- Implemented TOPSIS for sentence ranking, avoiding the training overhead while maintaining high accuracy in extracting key sentences from meeting transcripts, reducing training time by 300%.
- Achieved a higher Rouge-L score than advanced models like BERT and T5 when comparing the transcript summarizations with a 128% better average performance.

Projects

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

Irrigation Water Quality Classifier (Capstone)

January 2021- January 2022

- Lead a team of three to develop a novel dataset and utilized a Random Forest classifier to classify irrigation water quality as suitable or unsuitable with an accuracy of 83%.
- Enhanced the model by employing randomized value correlation techniques and reinforced learning to optimize the system, contributing to sustainable agriculture practices.