Suryam Gupta

Amherst, MA | (413) 510-7695 | Email | GitHub | LinkedIn

Education

Master of Science, Data Analytics and Computational Social Science Expected May 2026
University of Massachusetts Amherst, Amherst, MA GPA: 4.00/4.00
Relevant Coursework: Machine Learning | Natural Language Processing | Data Science | Quantitative Analysis

Integrated (BS+MS) Master of Science, Mathematics Aug 2019- June 2024
Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat, India CGPA: 8.26/10.00
Relevant Coursework: Probability & Statistics | Algebra | Numerical Analysis | Differential Equations | Optimization

Work Experience

Graduate Tutor & Consultant - Data Analytics and Computational Social Science Sep 2024- Present
University of Massachusetts Amherst, Amherst, MA
- Provide targeted tutoring and consultation to over 60+ graduate students, assisting with Python, R, and statistical methods for Data Science, Quantitative Analysis, and NLP coursework to build their confidence and analytical skills.

Research Experience

A Causal Inference Analysis of Parental Education and Study Time on Academic Success Sep 2024- Dec 2024
University of Massachusetts, Amherst, USA
- Conducted statistical analysis to study the impact on student performance using the Portuguese school dataset, leveraging Directed Acyclic Graphs (DAGs), visualizations, regression modeling, diagnostics, and model evaluation.
- Demonstrated that higher parental education and increased study time had a statistically significant impact on final grades (out of 20), increasing scores by 0.740 and 0.755 points ($p < 0.001$), as validated through hypothesis testing.
- **Tech Stack** - R, RStudio, Quarto, dplyr, ggplot2, dagitty, kableExtra

Evaluating Psychometric Features and Contextual Embeddings of Biomedical Texts May 2023- June 2024
Research Intern, IDSIA, USI/SUPSI, Lugano, Switzerland
- Extracted psychological features (emotions, intensity, LIWC, contextual embeddings) from biomedical subreddits.
- Developed predictive models with XGBoost and multi-layer deep learning architectures to classify texts (anxiety, panic, depression) with an F1-score of 0.894, and employed Large Language Models (LLMs) for causal inferencing.
- **Tech Stack** - transformers, aspect-based-sentiment-analysis, Linguistic Inquiry and Word Count (LIWC), TensorFlow

Analyzing Amazon Reviews, Textual Metrics, and User's Perceived Helpfulness June 2022- Dec 2022
Project Intern, Indian Institute of Management Nagpur (IIMN), Maharashtra, India
- Leveraged Natural Language Processing (NLP) and text pre-processing techniques to analyze 1M+ Amazon reviews for review length, breadth, readability, writing style, sentiment balance, and other engagement metrics.
- Calculated their association with the 'Votes' count to comprehend the features that affect the perceived helpfulness.
- **Tech Stack** - nltk, spaCy, transformers, gensim (Topic Modelling), wordcloud, pyLDAvis, readability

Time Series Forecasting Jan 2022- Mar 2022
Internship Trainee, Adani Industry Cloud, Ahmedabad, Gujarat, India
- Analyzed 4 years of household power consumption data and derived inferences through EDA, graphs, and charts.
- Built a Long Short-Term Memory (LSTM) model to predict future consumption and an interactive Power BI dashboard.
- **Tech Stack** - pandas, NumPy, Matplotlib, Seaborn, scikit-learn, TensorFlow, Microsoft Power BI

Research Publications

Mitrović, S., Frisone, F., **Gupta, S.**, Lucifora, C., Čarapić, D., Schillaci, C., Di Giovanni, S. & Singh, A. (2023b) *Annotating Panic in Social Media using Active Learning, Transformers and Domain Knowledge.* 2023 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE.

Achievements

- Received Swissnex ThinkSwiss Research Scholarships: Asia-Pacific of worth USD 4,500. May 2023

Skills

- **Programming Languages:** Python | R | SQL | Git | C | C++ | Java | MATLAB
- **Softwares:** Jupyter Notebook | Microsoft Power BI | GitHub | Quarto | AMPL | LaTeX | LIWC | Microsoft Excel