

Suryam Gupta

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

EDUCATION

| | |
|---|---------------------|
| Master of Science, Data Analytics and Computational Social Science | Expected May 2026 |
| <i>University of Massachusetts Amherst, Amherst, MA</i> | GPA: 4.00/4.00 |
| <i>Relevant Coursework: Machine Learning Natural Language Processing Data Science Quantitative Analysis</i> | |
| Integrated (BS+MS) Master of Science, Mathematics | Aug 2019- June 2024 |
| <i>Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat, India</i> | CGPA: 8.26/10.00 |
| <i>Relevant Coursework: Probability & Statistics Algebra Numerical Analysis Differential Equations Optimization</i> | |

SKILLS

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

WORK EXPERIENCE

| | |
|--|-------------------|
| Graduate Research Assistant | Feb 2025- Present |
| <i>University of Massachusetts Amherst, Amherst, MA</i> | |
| <ul style="list-style-type: none">• Developing a machine learning model to automate the detection of politically motivated edits on Wikipedia, identifying whether modifications were made to enhance or undermine a politician's public image. | |
| Graduate Tutor & Consultant - Data Analytics and Computational Social Science | Sep 2024- Present |
| <i>University of Massachusetts Amherst, Amherst, MA</i> | |
| <ul style="list-style-type: none">• 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 |
| <i>University of Massachusetts, Amherst, USA</i> | |
| <ul style="list-style-type: none">• 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 |
| <i>Research Intern, IDSIA, USI/SUPSI, Lugano, Switzerland</i> | |
| <ul style="list-style-type: none">• 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 |
| <i>Project Intern, Indian Institute of Management Nagpur (IIMN), Maharashtra, India</i> | |
| <ul style="list-style-type: none">• 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 | |
| Linear & Integer Optimization using Python PuLP and AMPL | May 2022- July 2022 |
| <i>Summer Intern, Indian Institute of Management Ahmedabad (IIMA), Gujarat, India</i> | |
| <ul style="list-style-type: none">• Worked with 16 real-life optimization problems and constructed their mathematical models.• Found their optimal solution using CBC & CPLEX solver by employing AMPL software and PuLP library of Python.• Tech Stack - PuLP, NetworkX, AMPL, Microsoft Excel, pandas | |

Credit Card Fraud Detection

Jan 2022- May 2022

Project Intern, Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India

- Applied data mining, cleaning, and wrangling techniques to clean and pre-process the data.
- Performed Exploratory Data Analysis (EDA) to discover meaningful patterns and insights related to fraud detection.
- Built a loan approval prediction model using Logistic Regression with 91.92% accuracy.
- **Tech Stack** - pandas, NumPy, Matplotlib, Seaborn, scikit-learn

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

CERTIFICATIONS & ONLINE COURSES

- | | |
|---|----------|
| • R Programming (Johns Hopkins University) - Coursera | Nov 2022 |
| • Data Analysis with R Programming (Google) - Coursera | Nov 2022 |
| • Machine Learning (Andrew Ng's, Stanford) - Coursera | Oct 2021 |
| • SQL for Data Science (University of California, Davis) - Coursera | Aug 2021 |
| • Excel Basics for Data Analysis (IBM) - Coursera | May 2021 |
| • Introduction to Data Analysis Using Excel (Rice University) - Coursera | May 2021 |
| • Mastering Data Structures & Algorithms using C and C++ - Udemy | Apr 2021 |
| • Machine Learning - Unscool | Mar 2021 |
| • Basic Java and Python - Unscool | Feb 2021 |

CLUBS

Society for Cultivation of Science and Humanities (SCOSH), SVNIT | Senior Core Member Sept 2020- Mar 2022

Student chapter that aims to develop and nurture scientific temperament among students

- Organized and delivered a lecture on Python libraries; NumPy and pandas, to 100+ students.
- Organized and managed various events such as InQuest, Prakriti, BEYOND, Science Hunt, and academic workshops (Python, MATLAB) & webinars.
- Held the position of Magazine and Newsletter Editor of the club.

Kirdaar, Centre for Human Resource Development (CHRD), SVNIT

Nov 2019- May 2020

Drama club under the student chapter that inculcates aesthetic values in the students and creates human resources.

- Participated in Nukkad Natak (Street Play) to raise awareness on the topic 'Nepotism.'

Youth for Planet, SVNIT | Senior Member

Sept 2019- Mar 2020

Student-driven initiative addressing environmental causes, fostering sustainable development, and addressing climate issues through various impactful projects.

- Collected funds for the Bihar Flood Relief in 2019.
- Conducted campus-wide waste cleanup drives and tree-planting campaigns.
- Part of educational visits to city greenhouses and Morning Sports Culture for better student health.

TALKS & LEADERSHIP ROLES

- Delivered a talk on “Getting started with Natural Language Processing (NLP) and its applications” in the 59th InterAct (Internship Experience and Research Activities) seminar organized by the Department of Mathematics, SVNIT. 2023
- Delivered a talk on “Real-World Application of Linear and Integer Optimization using Python” in the 53rd InterAct seminar, SVNIT. 2022
- Class Representative (CR) of the SVNIT batch for 2 years. 2019—2021

LANGUAGES

| | |
|----------|-------------------------------|
| English | Full Professional Proficiency |
| Hindi | Full Professional Proficiency |
| Sanskrit | Academic |

HOBBIES

Reading, Chess, Guitar, Traveling