

# Sreeram Marimuthu

774-535-9143 | smarimuthu@wpi.edu | Worcester, Massachusetts | [LinkedIn](#) | [Portfolio](#)

## SKILLS:

**Programming Languages** | Python (Pandas, NumPy, Scikit-learn, NLTK), Java, SQL

**Machine Learning** | MLflow, Databricks, and H2O AutoML, Supervised & Unsupervised Learning, Natural Language Processing (NLP), Deep Learning (Neural Networks)

**Large Language Models** | BERT, Llama 2, OpenAI (GPT-3.5, GPT-4), Co-Pilot

**Project Management** | Azure DevOps, Jira, Git, Scrum framework, Agile & Waterfall methodologies, Defect Management, API & Headless Testing, Problem Solving, Digital Marketing

**Big Data** | Hadoop (MapReduce), Apache Spark (PySpark, Spark SQL), Apache Pig, ETL

**Data Visualization** | Tableau, Power BI, Plotly

## EXPERIENCE:

### Research Analyst | The City of Worcester, USA | January 2025 – Present (Contract)

- Advancing my in-depth analysis of homeless population & shelter-types data, integrating findings from the Massachusetts Large Cities Health Coalition (MLCHC) for informative resource allocation and policy decisions.
- Collaborate closely with the Massachusetts Public Health Association (MPHA) to align research efforts and share best practices across the 14-largest Massachusetts cities.
- Present insights to stakeholders through data visualizations and evidence-driven recommendations, fostering strategic partnerships and enhancing community well-being.

### Research & Data Intern | The City of Worcester, USA | September 2024 – December 2024

- Partnered with the Commissioner of Health & Human Services on various research initiatives aimed at optimizing public health policies in Worcester.
- Managed comprehensive analyses of police incident reports to map crime patterns over the last 5 months, first responder deployments, and associated costs, guiding the strategic allocation of resources and enhancing the efficiency of homeless support services.
- Supported evidence-based decision-making through reports, presentations, and data visualizations to communicate research findings to stakeholders and the community.

### Machine Learning Engineer | Discern Health, WPI, USA | September 2024 – December 2024

- Graduate Qualifying Project – “Frailty Detection & Management”
- Integrated clinical, socioeconomic, and claims data to identify key frailty indicators—such as comorbidity scores, ADL limitations, ensuring practical interpretability & applicability.
- Executed 10-experimental variations: 3-on Predictive modeling (Frailty Detection via Kim CFI score forecasting) and 7-on Phenotype modeling, spanning entire patient populations, gender-specific & race-specific subgroups, and top-feature subsets to deliver comprehensive, data-driven insights.
- Developed predictive models (using AutoML and Lasso-driven feature selection) to forecast Kim CFI frailty scores, achieving RMSE as low as 0.0052 and  $R^2$  up to 0.9956.
- Performed extensive phenotype modeling on 53-Frailty-related indicators (e.g., Alzheimer’s/Dementia, Rheumatoid Arthritis, General Health), achieving AUCs of up to 0.98 and validating model robustness through calibration curves, feature importance analyses, and multi-strata patient testing.

### Research Study Assistant | Worcester Polytechnic Institute, USA | August 2024 - Present

- Collaborating with Prof. Roe Shraga on a Directed Research project focused on Human-in-the-Loop Data Analysis: HILDA & developing next-generation methods for table union search.
- Investigate metacognitive aspects of human interaction in data discovery—leveraging cognitive psychology and advanced AI techniques—to identify and correct biases that automated systems may overlook.
- Evaluate how emerging large language models (LLMs) necessitate active human oversight to ensure data quality and reliability, focusing on the integration of both synthetic and real-world datasets.

## EDUCATION:

**Worcester Polytechnic Institute** M.S. Data Science | May 2025

GPA: 4.0/4.0

**Anna University**, B.E., Electrical and Electronics Engineering | July 2021

GPA: 3.88/4.0

## RESEARCH AREAS:

Human in the Loop – Data Analysis

Large Language Models | AI-Bias

Healthcare Analytics | Crime Analytics

## STARTUP:

**Co-Founder** | Countdown Events

Established & managed an innovative Event Management organization in Chennai for about 1.5 years.

## CERTIFICATIONS:

**Oracle** – Oracle Cloud Infrastructure Certified Data Science Professional

**Accenture** – Automation Engineer AI & Data Analytics

**IBM** – Data Science Methodologies

## HONORS & AWARDS:

**Best GQP Team - Gold Winner** | GQP-Discern Health | Dec 2024 | Worcester

**KIP Fellowship** | Recognized for dedication to Social Change & Commitment | May 2024 | Worcester

**Gold Medalist** | Achieved Highest Academic Standing in my Department | Dec 2023 | Chennai

**TCS National QT Topper** | Secured a position among Top 10% in the country | Dec 2020 | Chennai

**Dean’s List [2017-18] & [2018-19]** | Aug 2018-19 | Chennai

**Runner Up** | Project Day 2020 | Jan 2020 | Chennai

## COURSEWORK:

ML, NLP, Numerical Methods, Big Data Management & Analytics Statistical Methods, Transforms & Partial Differential Equations

## EXPERIENCE (cont.):

### Data Analyst/Ambassador | Worcester Polytechnic Institute, USA | September 2024 - Present

- Managed administrative tasks, data analysis and initiative development to strengthen ambassador and liaison efforts within the Office of Graduate Studies.
- Lead special projects by leveraging data and research to pinpoint gaps in graduate support services, driving evidence-based improvements for retention and integration.
- Organize interactive events—such as campus tours, webinars, and information sessions—that foster community engagement and ease cultural transitions for new graduate cohorts.

### Research Assistant II | Intrare, WPI, USA | August 2024 – December 2024

- Collaborated with Prof. Andrew Trapp and Intrare (a social startup in Mexico City) to enhance the fairness and efficiency of an AI-driven job matching platform.
- Boosted algorithmic performance by transitioning from synchronous to asynchronous data processing, batching API calls, and refining similarity metrics for more efficient candidate-job matching.
- Provided actionable recommendations to mitigate bias and improve semantic matching from the analysis across 6-experiments focused on demographics (e.g., age, gender, nationality), identifying potential disparities in the hiring process.

### Research Intern | Advocates for Community Transfer, ACT, USA | May 2024 – July 2024

- Conducted geo-spatial and crime data analyses across four major Texas cities, leveraging 911 data, real-time mapping, and U.S. Census insights to inform feasibility reports and policy decisions.
- Uncovered patterns in low-income, high-crime neighborhoods, guiding more targeted interventions and better resource allocation.
- Developed interactive maps and compelling data visualizations to improve stakeholder communication, supporting data-driven decision-making processes.

### Research Assistant | Legendary Legacies, WPI, USA | March 2024 – May 2024 & August 2024

- Performed advanced data analyses on socio-economic and geographic factors influencing crime rates in designated neighborhoods.
- Led a multidisciplinary team—under the Dean's direction—to examine economic data and sociological trends, demonstrating quantifiable potential savings and community benefits for Legendary Legacies.
- Enhanced data visualization through interactive dashboards and created a path-type website detailing the challenges faced by returning citizens, amplifying awareness, and informing program improvements.

### Quality Engineering Analyst | Accenture, India | Dec 2022 – Apr 2023

- Led cross-functional testing initiatives, achieving a 15% reduction in post-deployment issues and a 20% decrease in project-related challenges via pro-active Scrum management.
- Executed API, Automation, and Headless testing on large-scale datasets, ensuring data integrity and functionality alignment with business requirements.
- Collaborated with product owners, designers, and business analysts to identify gaps, refine requirements, and enhance product quality through data-driven insights.
- Produced comprehensive test reports, streamlining defect resolution, and accelerating decision-making for product improvements.

### Application Development Associate | Accenture, India | Aug 2021 – Nov 2022

- Maintained a 0-defect record across 8-consecutive releases through rigorous manual, API, and headless testing procedures.
- Automated critical processes (e.g., contact creation), reducing execution time from several minutes to near instantaneous, resulting in enhanced efficiency.
- Served as the sole Subject Matter Expert for a high-security, customer-facing channel (FIRE), ensuring data compliance and integrity.
- Employed innovative testing strategies, including headless testing and pro-active risk assessment, to maintain a defect-free product environment across ten releases.

## PROJECTS:

### Sentiment Analysis of Amazon Reviews via Text Classification

- Pre-processed the Amazon Fine Food Reviews & implemented ML models using TF, TF-IDF, Bag-of-Words, and word2vec for sentiment classification.
- Employed various BERT pipelines for inference with & without fine-tuning.
- Later applied LoRA techniques to further optimize classification performance.

### Image Generation using GAN

- Implemented GAN architecture using PyTorch, developing Generator & Discriminator classes and optimizing their network structures.
- Conducted extensive experimentation with 16-distinct deep learning models, varying layers, activation functions, optimizers, and loss functions to enhance realism in generated handwritten digit images.
- Fine-tuned the model on the MNIST dataset, adjusting hyperparameters (learning rate, batch size) across multiple epochs to improve image quality.

### Transformative Insights through Spark SQL & MLLib in Retail Operations

- Complex retail transaction data were filtered & transformed using SparkSQL.
- Implemented MLLib workflows to train and evaluate ML models, predicting purchase prices based on customer attributes.
- Derived strategic insights, including customer segmentation and spending pattern analysis.

## LANGUAGES:

English | Tamil | German (A1)

## VOLUNTEERING:

Mentor | WPI

Class Representative | Anna University

Officer | Youth Red Cross Club

Member | Rotaract

School Social Services Leader | SNMSSS