

VINMATHI IYAPPAN

California | +1 510-589-9951 | vinmathi.iyappan@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Enthusiastic Data Scientist with 7 years of professional background, specializing in Healthcare, Manufacturing and Logistics industries. Proven track record of improving operational efficiency, forecasting accuracy, and decision-making.

EDUCATION

Master of Science in Business Analytics

California State University, East Bay

May 2025

GPA: 3.95/4

Bachelor's, Electronics and Instrumentation Engineering

Anna University

May 2015

GPA: 9.04/10

PROFESSIONAL EXPERIENCE

California Department of Public Health

Sept 2024 – Dec 2024

Data Scientist Intern

- Replaced manual data extraction with an AI-powered automation solution using **Power Apps**, achieving 95% model accuracy, eliminating 10 hours per week and scaling the solution across multiple teams.
- Initiated and developed the team's first **SQL** database with 30 tables, migrating Excel-based data into structured **ETL** pipelines to support large-scale public health data.

Essilor Luxottica – Oakley

May 2024 - Aug 2024

Data Scientist Intern

- Single-handedly implemented a **Power BI** dashboard to replace Excel-based reports, improving warehouse monitoring and enhancing decision-making for senior leadership, resulting in an estimated \$150K in annual cost savings.
- Enhanced Backorder AI **machine learning** model by applying **feature engineering** to 1TB of 'Available to Promise' data, improving prediction accuracy by 25% and reducing stockout frequency by 6% in the prototype phase.

Thryve Digital Health LLP

Mar 2021 - Aug 2022

Data Analyst

- Resolved claim pricing discrepancies by addressing critical data handling issues using **COBOL**, resolving 50 high-priority tickets and reducing revenue leakage by 8%.
- Identified and resolved an initialization issue in legacy COBOL-based claim eligibility code across 30 programs, preventing incorrect claim denials and manual rework. This saved 100+ hours of resource time and estimated ~\$80K annually.

Cognizant Technology Solutions

Nov 2015 - Feb 2021

Mainframe Developer

- Tuned a high CPU-consuming **DB2** batch job, reducing execution time from 58 minutes to 0.5 minutes and CPU usage by 98%, cutting mainframe processing costs by \$150K+ per year.
- Designed and implemented an ETL workflow to extract and process data from the **IMS** database using **Informatica Power Exchange** Change data capture reducing data latency by 60%.
- Contributed to a 20% reduction in post-release defects by identifying potential issues during the **peer review** process.

RESEARCH EXPERIENCE

Research Assistant

Title: [Using Utility driven clustering Analyzing United States prescription drug expenditure](#)

- With prescription drug expenditure rising annually, this research aims to optimize healthcare resource allocation and guide policy decisions for cost reduction.
- To understand spending patterns, wrangled 500k+ records, applied **dimensionality reduction** and utility-based **clustering**
- Discovered potential cost-saving opportunities that could reduce drug expenditure by ~5%, informing targeted policy recommendations

Title: [Using LLM to understand the Impact of Epilepsy disclosure or concealment on Workplace Outcomes](#)

- Analyzing 540K+ reddit forum discussions to understand workplace-related struggles of individuals with epilepsy using Natural Language Processing (**NLP**), **sentiment analysis** and Large Language Models (**LLMs**).
- Performed **hypothesis** testing to explore the consequences of disclosure vs. concealment of Epilepsy in workplace.
- Findings revealed that 67% struggles with inflexible workplace policies and lack of adjustments.

SKILLS

Language: Python (NumPy, pandas, scikit-learn, matplotlib, ggplot), SQL, R, COBOL, JCL, VSAM

Database: DB2, IMS DB, Teradata, MySQL

Project Management: Agile, Jira, Kanban, Git, Scrum, Rational Clear Case, ServiceNow, Smart sheet, Power Point

Tools: Power BI (Data Model, DAX, KPI), ETL, SSIS, SSMS, SSRS, Tableau, AWS (Hadoop, spark, Redshift), Excel

Machine Learning: Regression, Time series, Kmeans, DBSCAN, Umap, Statistical analysis, Casual inference, A/B testing

ACCOMPLISHMENTS

- Recognized as a [Trailblazer in Analytics](#) on the Cal State East Bay student testimonials portal.
- Selected as a [Women in Leadership](#) fellow from over 100+ applicants gaining insights into Diversity, Equity & Inclusion.
- Vice President of Sky Mediation club at Cal state East Bay

CERTIFICATION

- AWS Certified Cloud Practitioner