

SNEH PILLAI

<https://www.linkedin.com/in/snehpillai/> | <https://github.com/snehsuresh>
<https://snehpillai.vercel.app> | snehpillai02@gmail.com | Boston, MA | 774-704-0786

Education

University of Massachusetts, Dartmouth

Dartmouth, MA

Expected December 2024

- Master of Science in **Data Science** (Currently Enrolled, **3.95 GPA**)

APJ Abdul Kalam Technological University

Kerala, India

05/2015 - 07/2019

- Bachelor of Technology, **Computer Science and Engineering**

Professional Experience

Research and Industry
Support (Data Science)

University of Massachusetts,
Dartmouth

Dartmouth, MA

10/2024 - Present

- **Developed a Bayesian forecasting model** with large language models like Mistral 7B, gpt-j-6B, LLaMa2-8B to enhance demand projections, processing 1 million+ data points, resulting in a **95% credible interval** for accuracy.
- Utilized PyMC3 and TensorFlow for advanced analytics techniques like Bayesian hierarchical frameworks and MCMC techniques, achieving robust hyperparameter optimization through cross-validation.
- **Analyzed 12 key macroeconomic indicators** and 5 years of historical sales data using time series analysis, identifying consumer behavior drivers and improving forecast accuracy through various **KPIs**, evaluated with MAE, RMSE, and R-squared metrics.
- Simulated real-time sales data using Apache Kafka and Apache Spark to test the forecasting model's responsiveness and accuracy, enabling dynamic adjustments to demand projections based on live market fluctuations.

Data Scientist I

Playerzpot

Navi Mumbai, MH, India

08/2022 - 08/2023

- **Built a hierarchical BERT-based model** that outperformed baseline models by improving transactional data analysis through masked language modeling and random token detection strategies, supported by CI/CD best practices.
- **Increased player retention** by 76% F1 score in classifying multi-turn player dialogues using a RoBERTa-based model, deployed on AWS SageMaker to provide real-time insights for player support and engagement improvements.
- **Identified critical reasons for disengagement** by analyzing player conversations with unsupervised clustering techniques and LLaMA2-13b, uncovering underlying sentiments driving player churn.
- **Boosted transaction volume** by ₹5 million through building recommendation models that assisted players in targeting campaigns, optimizing player engagement and winnings.
- **Conducted Root Cause Analysis** on player behavior patterns using Apache Spark for distributed data processing and Databricks for scalable data engineering and advanced analytics, identifying factors such as session duration, in-game purchases, and engagement frequency that contributed to player dissatisfaction and churn.

Associate Data Scientist

Medfanum Pharmaceuticals

Remote, India

10/2021 - 7/2022

- **Developed and deployed a predictive model** using Python and scikit-learn to analyze historical sales trends and seasonality, achieving a 25% reduction in inventory holding costs by accurately forecasting demand for seasonal medications, utilizing Tableau for data visualization of insights.
- **Conducted time series analysis** in R with ARIMA, leveraging historical sales data to identify patterns and anomalies, which improved sales prediction accuracy by 15%, enabling the supply chain team to optimize inventory levels for critical drugs, while using Microsoft Excel for preliminary data analysis.
- **Automated data extraction and transformation processes** by designing ETL workflows in Apache Airflow and using SQL for data normalization and aggregation, resulting in a 30% increase in reporting efficiency and faster insights for the marketing team, with data stored in AWS S3 for scalability.

Software Engineer

Organiza Tech Pvt. Ltd.

Navi Mumbai, MH, India

10/2019 - 08/2021

- **Boosted Mapping Accuracy:** Improved XBRL mapping accuracy from 75% to 98% by developing an auto-mapping feature in Unifeye utilizing Python for parsing, Node.js for backend operations, and MongoDB for data storage, integrating Apache Kafka for real-time data streaming and Docker for containerization.
- **Enhanced Server Performance:** Reduced server response times by 35% in high-concurrency tasks by integrating Golang, using goroutines and channels for efficient non-blocking I/O, and implementing Redis for in-memory caching and optimized load balancing with nginx.
- **Transformed KYC System:** Led the migration of the KYC system from PHP to Node.js. Designed a microservices architecture with Express.js and GraphQL to speed up identity verification and improve code maintainability, using Kubernetes for managing services and AWS Lambda for serverless functions

Research Experience

- **(Ongoing) Optimizing Knowledge Distillation for Balanced Cloud and On-Device Inference Workloads:** Developing a novel pipeline using knowledge distillation to balance cloud and on-device compute for deep learning inference. This work focuses on minimizing computational load, preserving performance, and enabling efficient handling of specialized tasks while maintaining privacy and energy efficiency across distributed environments.

SNEH PILLAI

<https://www.linkedin.com/in/snehpillai/> | <https://github.com/snehsuresh>
<https://snehpillai.vercel.app> | snehpillai02@gmail.com | Boston, MA | 774-704-0786

- **(Ongoing) Keyword Masking for LLM Privacy Protection:** to develop keyword masking techniques aimed at preventing large language models (LLMs) from inferring sensitive personal attributes (e.g., location, gender, income) from synthetic user text.
- **Enhancing Supply Chain Management through Business Intelligence and Knowledge Management:** Integrating Business Intelligence (BI), Supply Chain Management (SCM), and Knowledge Management (KM) to enhance transparency, visibility, and cost optimization.

Skills

- **Language:** Python | JavaScript | Typescript | R Language | C/C++
- **Artificial Intelligence / Machine Learning :** Machine Learning | Deep Learning | PyTorch | GPT | LangChain | Few-shot classification | Retrieval Augmented Generation and Fine Tuning | Transformers | YOLO Object Detection | Topic Modelling (LDA, NMF, BERTopic) | Autoregressive Models | Diffusion Models | Generative Adversarial Networks (GANs) | Prompt Engineering
- **Cloud Computing and Big Data Technologies:** AWS | Kubernetes | Terraform | Celery | Azure | GCP
- **Databases / DBMS:** MySQL | PostgreSQL | MongoDB | Cassandra | CosmosDB | Redis
- **Frameworks and Libraries:** Node | Express | React | Numpy | Pandas | Firebase
- **DevOps and Automation:** Docker | Airflow | CI/CD | MLFlow | DVC | Github Actions | Jenkins
- **Data Visualization:** Tableau | Amazon Quicksight | D3.js
- Stochastic Processes and Forecasting | CUDA | OpenMP | Linear Algebra | Data Structures & Algorithms | OOP | Full-stack Development | BI Tools | Knowledge Management | Shell Scripting

Certifications

- | | |
|--|---|
| • Inferential Statistics , Coursera | • Introduction to Probability and Data with R , Coursera |
| • Machine Learning Course , Udemy | • Python Mega course , Udemy |