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**)

Relevant Courses: Advanced Mathematical Statistics | Database Design | Applied Business Analytics and Information Visualization | Business Intelligence and Knowledge Management | Big Data Analytics | Advanced Data Mining | Data Visualization | High Performance Scientific Computing

APJ Abdul Kalam Technological University *Kerala, India* **05/2015 - 07/2019**

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

Professional Experience

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 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.

Associate Data Analyst **Medfanum Pharmaceuticals** *Remote, India* **10/2021 - 7/2022**

- Automated Inventory Data Pipelines: Improved stock accuracy by 30% and reduced stock outs by 25% through the development of ETL processes using AWS Glue and Kafka for real-time data updates.
- Created Real-Time KPI Dashboards: Cut report generation time by 40% by building interactive dashboards in Tableau that integrated data from Amazon RDS and ERP systems.
- Executed Predictive Analytics Models: Increased forecasting accuracy by 20% and boosted conversion rates by 15% through the application of machine learning techniques with AWS S3, Lambda and SageMaker for demand forecasting and customer segmentation.

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.
- Modernized Frontend Stack: Updated legacy jQuery code to a modern Node.js and React stack. Improved performance with the virtual DOM, used Webpack for better module management, and implemented Redux for state management to enhance code reusability and reduce technical debt.

Academic Projects

Note: *Some of the open projects are hosted on a base-tier plan, which may result in occasional downtime due to inactivity. Initial load times may be up to 50 seconds. I am aware of these limitations and am working on improving the hosting to ensure a better user experience.*

- **dbLinkPro:** A Python package on PyPI for MySQL, PostgreSQL, MongoDB, and Cassandra databases with Docker integration. Link
- **British Airways Review Dashboard:** Visualizing customer feedback by aggregating and analyzing reviews including numerous metrics like monthly ratings and comfort scores.
- **Developed a Large Language GPT model from scratch using Python and PyTorch,** incorporating tokenization, data compression, and fine-tuning with transformers for NLP tasks. Link
- **Adaptive Distillation for Student Model Compression Based on Hardware Constraints,** compresses large teacher models (utilizing HuggingFace) into optimized and fine-tuned student models using adaptive knowledge distillation and validation. Link
- **Reddit Data Pipeline Engineering**: Integration of Reddit, Airflow, Celery, Postgres, S3, AWS Glue, Athena, and Redshift to create a seamless ETL process. Link
- **Personalized Recipe Recommendation System:** Selenium for data extraction, BERT for semantic analysis and EnsTM for advanced topic modeling. Link
- **Real-Time Facial Emotion Detection and Audio Feedback System for the Visually Impaired:** A computer vision project to detect facial emotions using YOLOv5, converts emotions to speech providing real-time audio feedback trained using T4 GPU. Link

SNEH PILLAI

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

- **RAG QA**: Demonstrates deploying a Retrieval-Augmented Generation (RAG) application using AWS services such as ECR, and Langchain, Hugging Face, and Docker. Link

Research Experience

- **(Ongoing) Enhancing Demand and Risk Forecasts with Macroeconomic Data:** Developing a Bayesian forecasting model combining macroeconomic data (GDP growth, inflation) with historical sales to improve demand forecasting. Utilizing **Mistral 7B, gpt-j-6B, LlaMa2-8B** to analyze trends, patterns, and enhance predictive accuracy from datasets and reports.
- **(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.
- **(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
- **Machine Learning Course,** Udemy
- **Introduction to Probability and Data with R,** Coursera
- **Python Mega course,** Udemy