




NIRMAL AMIRTHALINGAM

Chicago, IL | +1 (540) 824-8595 | nirmalrak@gmail.com |  |  | 

EXPERIENCE

Data Engineer

RxLive

July 2023 - Present

St. Petersburg, FL

Infrastructure/Automation

- Automated ETL processes for office visit, hospitalization, ADT reports, population refresh, demographics, Medicare claims, and medication history files for **14 healthcare organizations** across 25 offices, covering over **125K patients**
- Developed daily list processing tasks using AWS services (S3, Glue, Lambda, EventBridge) and implemented workflows using AWS CodePipeline triggered by GitHub updates, file uploads, and cron-jobs

Analytics and Reporting

- Developed and maintained **16** key dashboards in **Domo**, including Drug Cost, Medication Adherence, HEDIS, Patient Engagement, Pharmacist Performance, Longitudinal Analysis, Operations, and Clinical
- Performed **ADR** and **risk analysis** using a **DID** approach, identifying trends impacting patient hospitalization rates, and presented findings to stakeholders using visualizations
- Facilitated collaboration with Data Science, Operations, and Engineering teams for continuous software development in an agile environment using Jira, Confluence, Miro, for data sourced from HealthGorilla, Surescripts, Acuity, and CallRail

Data Modeling

- Worked on **XGBoost** models for hospitalization risk and hospitalization risk reduction probability, selecting patients based on **admission risk** and other criteria for consultation
- Improved patient outcomes using **predictive diagnosis** models and specialized client-specific selection criteria, such as **HEDIS** Adherence metrics, and patient selection optimization based on office visit type

Research Assistant

Virginia Tech

Aug 2022 - May 2023

Blacksburg, VA

- Implemented an end-to-end object detection model for information extraction from Virginia Tech's ETD repository
- Extracted text and image objects from ETDs using object detection and stored in **PostgreSQL** database
- Implemented post-processing rules to filter detections and parse information to structured data (XML)
- Integrated to a **retrieval** and **search** system that supports **50K ETDs** by providing **Dockerized** services with **CI/CD**

Data Science Intern

AreaProbe

May 2022 - Aug 2022

Washington, DC

- Developed an **object tracking** model and deployed for 35 cameras installed in parking lots and housing communities
- Integrated the model with an **audio detection** pipeline to detect gunshots in real-time for a surveillance system
- Deployed the model using AWS **EC2** instances, and stored detected objects in **S3** bucket and structured data in **RDS**
- Queried data from RDS using **Athena** and connected to the surveillance dashboard along with the saved objects in S3

SKILLS

Programming: Python, SQL, JavaScript, Shell scripting

Databases/Tools: PostgreSQL, MySQL, MongoDB, Docker, Git, GitLab, PySpark, Kafka, Airflow, Postman

Libraries/Frameworks: PyTorch, TensorFlow, Pandas, Keras, Scikit-learn, XGBoost, OpenCV, NLTK, NumPy, OR-Tools

Cloud: AWS (S3, EC2, SageMaker, Glue, Lambda, RDS, DynamoDB, Athena, SQS, SNS), Azure (Databricks, ML Studio)

Data Visualization: Domo, Tableau, D3.js, Power BI, Excel Charts, Matplotlib, Seaborn

Modeling: Exploratory Data Analysis, Regression, Classification, Random Forest, Clustering, Time-series, DID, A/B Testing

PROJECTS

Trending YouTube Video Analytics using Sentiment Analysis

Aug 2022 - Nov 2022

- Scraped YouTube API for trending videos in the US to analyze key performance indicators using sentiment analysis
- Created interactive visualizations and dashboards in **D3.js** and **Tableau** to help understand general viewer sentiment
- Extended to trending videos in multiple countries by pre-processing raw JSON files using **AWS Lambda** and **Spark**
- Performed ETL using **Glue** to load analytical data into S3 bucket for creating visualizations in **QuickSight**

EDUCATION

Virginia Tech, Blacksburg, VA

Master's in Computer Engineering

May 2023

GPA: 3.96/4

Coursework: Data Analytics, Data Visualization, Information Storage and Retrieval, Applied Machine Learning