THEJUS KANNOTH

Chicago, USA | +91 9074837971 | kannoththejus@gmail.com | Linkedin | GitHub

EDUCATION

Depaul University, Chicago, USA **2023-2025**

MS in Data Science *3.93/4 GPA*

APJ Abdul Kalam Technological University, MACE, Kothamangalam, India **2017-2021**

B.Tech in Electronics and Communication *8.2/10 GPA*

WORK EXPERIENCE

Cognizant Technology Solutions **Nov 2021- Aug 2023**

Programmer Analyst

- Coordinated with eight members for the development of CRM-based enterprise applications using Microsoft Dynamics 365, Power Apps, JavaScript.
- Contributing to the implementation of modifications, testing and maintenance of the sales module.
- Handled testing and debugging customisations and creating custom plugins, web resources, and JavaScript functions.

PROJECTS

Time Series Analysis and Forecasting – Traffic Volume **Jun 2025**

Independent Project

- Forecasted 24-month traffic volume on I-94 (Minnesota) using ARIMA, SARIMA and SARIMAX models.
- Conducted seasonal decomposition, trend analysis, and cross-validation to ensure model robustness.
- Engineered time-based, weather, and holiday features for multivariate SARIMAX modeling.
- Achieved optimal results with SARIMA(0,1,1)(0,1,1)[12] based on AIC and residual diagnostics.

Predicting Injury Risk in NBA Players **March 2025**

Prof. Ilyas Ustun - DePaul University, Chicago

- Utilized machine learning models (XGBoost, Random Forest, Logistic Regression) to predict NBA player injury risk, achieving 97.97% accuracy and optimizing performance using SMOTE and PCA.
- Evaluated model performance using precision, recall, F1-score, and confusion matrices to ensure reliability in injury prediction.
- Developed a scalable data pipeline, applied hyperparameter tuning, and implemented effective techniques for class balancing and dimensionality reduction.

Sleep Quality Prediction **March 2025**

Prof. David Hubbard - DePaul University, Chicago

- Developed a machine learning pipeline to predict sleep quality using wearable technology data, achieving 62% accuracy with a Voting Ensemble model, optimizing performance across multiple classes (Poor, Moderate, Good).
- Performed data preprocessing including handling missing values, outliers using Z-scores, and class imbalance with SMOTE, ensuring balanced and effective model training.
- Implemented and evaluated multiple models (Logistic Regression, KNN, Random Forest, Gradient Boosting, XGBoost) using performance metrics like accuracy, precision, recall, and F1-score, with the Voting Ensemble model delivering the best-balanced results.

E-Commerce Fraud Detection **Nov 2024**

Prof. Ahmed Abid - DePaul University, Chicago

- Built a scalable data processing pipeline using PySpark on AWS EMR to clean, preprocess, and transform a 1.5M-row e-commerce transaction dataset stored in AWS S3.
- Performed advanced data cleaning, aggregation, and exploratory data analysis using PySpark and staged trans formed data in AWS Athena for efficient querying.
- Built a machine learning model using Spark MLlib for fraud detection, leveraging algorithms like Logistic Regression and Random Forest, achieving 95% accuracy.
- Automated pipeline execution and model inference workflows using AWS SageMaker, with end-to-end monitoring and reporting.

Insurance Cost Prediction **Jun 2024**

Prof. Ilyas Ustun - DePaul University, Chicago

- Developed a machine learning pipeline to predict medical insurance costs, achieving an accuracy of 88%.
- Optimized Random Forest, Gradient Boosting, and Stacking Regressor models using hyperparameter tuning and ensemble methods.
- Preprocessed data by encoding categorical variables, normalizing numerical features, and handling missing values.
- Conducted exploratory data analysis with visualizations to identify key factors influencing insurance costs.

TECHNICAL SKILLS

Languages: Python, R, SQL

Tools & Frame Works: Tableau, Excel, MySQL, AWS (S3, EMR, SageMaker), Oracle, Jupyter Notebook, PySpark Hadoop, Visual Studio, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Prophet, Statsmodels, PyTorch, TensorFlow

Machine Learning, Forecasting & Predictive Analytics: Decision Tree, KNN, SVM, XGBoost, Linear Regression, K-Mean, Random Forest, Logistic Regression, Predictive Modeling, Time Series analysis and Forecasting, ARIMA, SARIMA, SARIMAX, Seasonal Decomposition Classification, Regression, Cross-Validation, Model Evaluation, Feature Engineering, Big Data

CERTIFICATIONS

- Generative AI with Large Language Models, DeepLearning.AI & AWS
- ChatGPT Prompt Engineering for Developers, DeepLearning.AI & OpenAI

EXTRA-CURRICULAR ACTIVITIES

- Led the Planning and execution of student led events under the international student office, managing timelines, resources.
- Elected as a Senate Member during undergraduate studies and actively worked on initiatives to enhance student well-being and campus life.