

PRANAV SENTHILKUMARAN

NJ 08901 | 732-322-4983 | senthilkumaranpranav@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

- Azure-Certified Data Engineering and Business Intelligence Strategy Analyst Intern with close to 2 years of experience in transforming raw data into actionable insights using tools such as Power BI, SQL, and Azure. Demonstrated ability to automate workflows, implement AI-driven data processes, and drive process optimization. Strong foundation in data analysis and reporting from academic and real-world projects, with a focus on problem-solving and effective decision-making across diverse settings.

EDUCATION

Rutgers University, New Brunswick, New Jersey

Master of Science in Data Science, GPA-3.667/4.0

2024 – E.May 2026

- **Relevant Coursework:** Data Mining, Data Wrangling, Regression and Time Series Analysis, Data Structures and Algorithms, Statistical Models and Computing, AI and LLM in Data Science, Probability, Database Management Systems

Indian Institute of Technology, Madras, India

Diploma in Programming and Data Science (Online Degree)

2022 – 2026

Government College of Technology, Coimbatore, India

Bachelor of Engineering in Electronics and Communications, GPA- 3.48/4.0

2020 – 2024

PROFESSIONAL EXPERIENCE

Space Ace Inc.

Feb 2026 - Present

Backend & Data Engineering Intern

Chicopee, Massachusetts

- Working on designing an end-to-end backend for a streaming platform, focusing on REST API design, PostgreSQL schema modeling, and MongoDB collections, with OAuth-based authentication using Python, FastAPI, and SQL.
- Containerizing 5+ backend services with Docker and deploying on AWS (ECS, RDS, S3, Secrets Manager, SES) using Terraform modules to support authentication flows, data storage, and API performance.

Kentuckiana Curb Company

Dec 2024 - April 2025

Business Intelligence Strategy Analyst Intern

Louisville, Kentucky

- Automated data workflows and supported the end-to-end development of an Azure-based chatbot, including data ingestion from 5 sources with Azure Data Factory and testing in Azure Chat playgrounds on AI Foundry within a private VNet.
- Contributed to self-service BI tools like Dotnet, Izenda, and PowerBI, and generated reports and printable labels using SSMS and Report Builder and involved in Python projects, such as automating email generation.

Spark Minda Technical Centre

Jan 2024 - April 2024

Machine Learning Intern

Bangalore, India

- Engineered a precise Automatic Passenger Counting System utilizing Computer Vision algorithms to accurately count individuals, achieving a detection accuracy of 98% using the YOLOv8, object detection model (bounding boxes).
- Optimized the system to process 30 frames per second on Raspberry Pi (CM4), enabling real-time passenger detection and enhancing passenger count accuracy in crowded environments by 25%.

National University of Singapore

June 2023 - Aug 2023

Project Intern - "Data Analytics and Deep Learning"

- Mastered advanced data manipulation techniques and machine learning algorithms through 15+ hands-on projects, utilizing Python libraries, AWS, R Studio, NLP, and text mining for diverse analytical challenges.
- Directed the development of a predictive modeling framework for analyzing near fault non-pulse ground motions, leading to the creation of a robust database consisting of 10,000+ seismic events.
- Enhanced model efficiency by 30% by optimizing ANN training using advanced gradient descent algorithms.

RESEARCH PROJECTS

Bristol Myers Squibb AI-Hackathon

2026

Dynamic ML Pipeline for mAb Fucosylation Prediction

- Built a full-stack ML dashboard predicting mAb fucosylation from 10 bioprocess variables, training 7 models (Ridge, PLSR, Random Forest, XGBoost, GPR, ANN, Hybrid), achieving $R^2=0.83$ with XGBoost+SHAP on 10,000 samples and demonstrating 1,440% ANN recovery by scaling from 500 samples, proving neural network incompatibility with small pharma datasets; integrated Gemini 2.5 Flash for ICH Q8/Q9-aligned regulatory report generation.
- Engineered passwordless authentication (Google OAuth 2.0, Gmail SMTP OTP with signed token sessions) and a full support ticketing system as part of a production-grade React + FastAPI deployment.

Offline Computation Tool to Estimate Charge Variants of a Biopharmaceutical Protein

- Estimated charge variant distributions using an offline Python tool with strict data validation, applying convolution, polynomial powering, and FFT-based convolution, with moment-matching approximation for scalability, and stress-tested on over 100 profiles, evaluating 10–50 charge states per run with stable runtime and memory.

- Generated PMF plots, charge state distributions, and summary statistics, enforcing no renormalization to preserve 100% probability mass within the truncation window.

Glycan Clearance Analysis & Visualization for Therapeutic Optimization

- Designed a dynamic Power BI dashboard with 6+ analytical visualizations (stacked bar plots, line charts, card visuals) to analyze glycan clearance trends across 5 production sites (A–E), enhancing clearance prediction accuracy by 20%.
- Applied K-means clustering (k=3, Elbow Method) on 50+ glycan profiles, classifying them based on AUC and half-life.

News Headline Classification: Comparison of Traditional ML and Transformer Models 2025

- Built a multiclass NLP pipeline on AG News (120K+ headlines, 4 categories) using TF-IDF with Logistic Regression and Linear SVM, achieving 92% accuracy as baseline performance.
- Fine-tuned DistilBERT and RoBERTa with Hugging Face Transformers and PyTorch, leveraging contextual embeddings and few-shot training frameworks to reach 95% accuracy and higher macro F1 versus traditional models.
- Applied LoRA parameter-efficient fine-tuning (Q+V projections) with dropout and hyperparameter tuning, training only 1.4% of parameters while improving generalization and efficiency.

Conference based Agenda Optimization: Identifying Similar Abstracts 2025

- Detected duplicate/highly similar conference abstracts (2014–2025) using text tokenization, TF-IDF & cosine similarity (>0.85) on 254K+ entries (30 features), leveraging Depth First Search graph traversal for node transitivity to cluster related abstracts into connected components, including isolated entries.
- Built a Python pipeline and Tableau dashboard for automated preprocessing, similarity scoring, and multi-year clustering to optimize agenda planning.

HLS based Signal Modulation Classification System using Deep Learning 2024

- Used CGRDNN, a hybrid neural network for recognition of signal modulation through high-level synthesis, facilitating practical FPGA implementation using high-level synthesis and demonstrating ASIC development potential.
- Reduced processing time by 20% compared to conventional models for complex signal classification.
- Trained the neural network to attain 98% accuracy in classifying modulated signals, including those corrupted by random noise, in challenging non-cooperative channel scenarios.

TECHNICAL AND OTHER SKILLS

Programming and Tools: Python, C, R, Azure, AWS, n8n, SQL, Power BI, Tableau, PostgreSQL, MongoDB, SSRS, Dotnet Report, MATLAB, Verilog, Postman, Docker, Git, CI/CD, REST APIs, WebSocket, Flask, Streamlit

Libraries and others: Pandas, Numpy, Scikit-learn, PyTorch, Matplotlib, Seaborn, TensorFlow, ETL, LangChain, Ad-hoc Reporting, Optical Character Recognition (OCR), Exploratory Data Analysis (EDA), Statistical Analysis, Data Modeling, Data Validation, Data Quality, Data Warehousing, Big Data, Feature Engineering, Model Evaluation, Cross Validation, Hyperparameter Tuning, Object-Oriented Programming (OOP), Graph Algorithms

Area of Research: Machine Learning, IOT, NLP, Computer Vision, LLM, Agentic AI

Communication and Interests: Team Collaboration, Fluent in English and Tamil; Interests: Art, Badminton, Piano

PROFESSIONAL DEVELOPMENT AND ACHIEVEMENTS

- Grading Assistant** - “Applied Statistical Learning”, “Machine learning Principles”, “Introductory Statistics for Business”, “Basic Statistics for Research” Rutgers University
- Bloomberg Finance Fundamentals, n8n Certification** 2025
- Microsoft Certified: Azure AI Engineer Associate (AI-102) Certification** 2025
- Graduate Student Mentor-** Summer STEM Camp, Bristol Myers Squibb - (AI and Smart Pills) 2025
- Tata’s Data Visualization:** Empowering Business with Effective Insights virtual experience on Forage 2025
- Research Paper** - “Improving Security in Cyber-Physical Systems: Utilizing AI Optimization for Intrusion Detection in Blockchain Environments”
- Awards-** Presented a paper on “Intelligent IOT Home Solutions” at Anna University - *consolidated 1st place* 2023
- Leadership-** TEDxGCT Curator, Head Cartoonist at **Student Journalist Council**, Head of **Badminton Club**