Akash Kannan

akash2462003@gmail.com | linkedin.com/in/ak2467 | https://github.com/Akash2467 | +1 (585) 415-3531 | Rochester, NY

EDUCATION

Rochester Institute of Technology   Rochester, USA
Master of Science - Data Science   **Aug 2025 - Dec 2026**
Vellore Institute of Technology   Bangalore, India
Post Graduate Program – Artificial Intelligence: CGPA: 10/10   **Jul 2024 - Jun 2025**
Vellore Institute of Technology   Chennai, India
Bachelor of Science – Mathematics and Computing: CGPA: 8.4/10   **Sep 2021 - May2024**

SKILLS

- **Programming & Scripting**: Python, R, Java, C, C++
- **Machine Learning & AI**: Machine Learning, Deep Learning, Reinforcement Learning, Natural Language Processing (NLP), Computer Vision, Feature Engineering, Model Evaluation & Validation
- **Libraries & Frameworks**: Scikit-Learn, TensorFlow, PyTorch, Pandas, NumPy, Matplotlib, Seaborn, NLTK, SpaCy, OpenCV
- **Data Analytics & Visualization**: Power BI, Tableau, Excel, Exploratory Data Analysis, Data Mining, Statistical Modelling,Minitab
- **Databases**: MySQL, SQLite, MongoDB, SQL
- **Platforms & IDEs**: Jupyter Notebook, Visual Studio Code, Google Colab, Kaggle
- **Soft Skills**: Adaptability, Teamwork, People Management, Problem Solving, Communication, Analytical Thinking

---

Internships & Experience

Project Intern – IGCAR Kalpakkam   **August 2023-December 2023**

- **Tools and Techniques**: Python, PyTorch, NumPy, MCTS, Minimax, Machine Learning
- Worked on a project focused on inventing faster algorithms for matrix multiplication using AI and reinforcement learning.
- Designed a game-playing agent inspired by **AlphaTensor** to explore optimized tensor decompositions.
- Implemented and analyzed search strategies such as **Monte Carlo Tree Search (MCTS)** and **Minimax** on simplified environments like Tic-Tac-Toe before extending to tensor operations.
- Collaborated with researchers in the AI division and presented findings on algorithmic efficiency.

PROJECTS

AutoML Web App with Streamlit |Link

- **Tools & Techniques**: Streamlit, Scikit-learn, Pandas, Matplotlib, NumPy, Machine Learning
- Developed an interactive AutoML web app for seamless dataset upload and **automatic problem type detection** (classification, regression, clustering).
- Implemented end-to-end preprocessing, model training, and result visualization within a user-friendly interface.
- Integrated advanced features including model saving and **prediction on new data**, **cluster visualizations.**

Automatic Text Summarization using NLP and LLMs | Link

- **Tools and Techniques**: Python, NLTK, Pytorch, TF-IDF, T5, LLMs, NLP
- Developed an **extractive summarizer** to generate concise summaries from paragraph-level input text.
- Built a complete NLP pipeline including tokenization, lemmatization, TF-IDF vectorization, and supervised learning.
- Trained on 50,000 samples from the WikiHow dataset using the T5 Large Language Model.
- Evaluated the model using **ROUGE** metrics to measure summary quality and effectiveness.

Smart Billing System using Object Detection (YOLOv8) | Link

- **Tools & Techniques**: YOLOv8, Ultralytics, PyTorch, OpenCV, Streamlit, Computer Vision
- Built an **end-to-end automated billing system** that detects grocery items in real-time and computes the total bill.
- **Fine-tuned YOLOv8** on the Freiburg Groceries Dataset, significantly improving detection accuracy and reducing loss.
- Developed an interactive **Streamlit interface** to visualize detected items, display itemized lists, and show the final bill.
- Enhanced system performance and usability by optimizing real-time item detection and seamless interface interaction.

Cricket Playing XI Prediction using Clustering and Ranking | Link

- **Tools and Techniques:** Pandas, Scikit-learn, KMeans Clustering, Random Forest, Matplotlib, Machine Learning.
- Predicted the optimal playing XI for different cricket formats (ODI, T20, Test) using comprehensive player performance data.
- Applied **clustering** techniques to group players by roles and skills for better team composition.
- Used a **Random Forest** model to generate performance-based player rankings.
- Selected the final XI based on overall ranking, maintaining role balance and considering current player form.

Writing Preference Prediction using Ensemble Learning | Link

- **Tools and Techniques:** Python, Pandas, NumPy, Scikit-learn, Statsmodels, Matplotlib, Seaborn
- Predicted writing preference (Left/Right/Ambidextrous) using 100+ anthropometric features from the ANSUR II Male dataset.
- Applied **GLM-based feature selection** to identify the most relevant predictors for handedness.
- Trained and evaluated **ensemble models** including Gradient Boosting, Bagging, Stacking, and Random Forest.
- Achieved **87.7% accuracy** and highlighted key predictors such as arm length, shoulder breadth, and hand size.

CERTIFICATES

Basics of Data Science for Metallurgy and Manufacturing – Indian Institute of Metals   **May,2023**