

AKASH KUMAR NAYAK

Data Scientist

📞 +91-9668741796 — 📩 akashsiku2004@gmail.com — 💻 linkedin.com/in/akash-kumarnayak — 🌐 github — Portfolio

Summary

Entry-level Data Scientist with strong foundations in Python, SQL, statistics, and machine learning. Hands-on experience across data cleaning, exploratory data analysis, feature engineering, predictive modeling, NLP, and clustering through academic and real-world projects. Skilled in translating complex datasets into actionable insights using analytical and problem-solving approaches. Motivated to grow and contribute in data-driven teams.

Skills

Programming & Querying: Python (Pandas, NumPy, Scikit-learn), SQL

Data Analysis & Statistics: Data Cleaning, Exploratory Data Analysis (EDA), Descriptive Statistics, A/B Testing

Machine Learning & Deep Learning: Regression, Classification, Clustering, Random Forest, XGBoost, Neural Networks, RNN, LSTM (Foundational)

Natural Language Processing (NLP): Text Preprocessing, TF-IDF, Feature Engineering

Time Series Forecasting: Trend & Seasonality Analysis (Foundational)

Visualization & Tools: Power BI, Tableau, Advanced Excel, Jupyter Notebook, GitHub, MySQL, PostgreSQL

Experience

Data Science Intern

Mentormind (*Uber Dataset*)

Sept 2025 – Nov 2025

Python, SQL, Machine Learning, Power BI — [Link](#)

- Built a **regression model** on **200K+ Uber ride records** to predict fares, achieving an **R² score of 0.86**.
- Performed **feature engineering** (trip duration, pickup zones, time-of-day) and model evaluation to improve prediction quality and interpretability.
- Created **pricing and demand insights** and automated Power BI reporting, reducing analytical turnaround time by **30%**.

Projects

Employee Attrition Analytics Dashboard | Python, Power BI, Streamlit, XGBoost

Sept 2025 — [Link](#)

- Analyzed **15K employee records** to uncover attrition drivers by role, tenure, and pay band.
- Balanced data using **under-sampling** techniques and improved model accuracy from **95% to 97%**.
- Designed an **interactive Power BI dashboard** with **8+ KPIs** tracking attrition risk by department, enabling HR teams to design retention strategies that reduced attrition by **15%**.

Customer Segmentation Analysis | Python, Machine Learning

Oct 2025 — [Link](#)

- Performed **customer segmentation** on transactional data using **K-Means clustering** to identify distinct customer groups based on purchasing behavior.
- Analyzed segment characteristics to generate insights supporting **targeted marketing and personalized engagement strategies**.

Email Classification System (NLP) | Python, NLP, Machine Learning

Nov 2025 — [Link](#)

- Performed comprehensive **text preprocessing** including tokenization, stopword removal, and **TF-IDF feature extraction** to transform raw text into high-dimensional feature vectors.
- Evaluated multiple ML models, achieving **96% precision** in spam detection with minimal false positives.
- Automated the categorization pipeline, reducing manual sorting time by **40%** and increasing efficiency.

Education

Bachelor of Technology (B.Tech) in Computer Science

Graduated: 2025

Gandhi Engineering College, Bhubaneswar

CGPA: 7.6 / 10

Awards & Certifications

- **Advanced Certificate Programme in Data Science & AI-ML – upGrad (2025):** Completed an industry-oriented program covering data analysis, machine learning, SQL, Python, statistics, and end-to-end projects under expert mentorship. [Link](#)
- **IBM Machine Learning Certificate – Coursera (2025)** [Link](#)