



Akash Chauhan

Roll No.: 23035010563

BSc (Hons.) Data Science and AI

Indian Institute of Technology, Guwahati

+91-7017320226

akash.chauhan@op.iitg.ac.in

GitHub | Portfolio

linkedin.com/in/akashconnects

EDUCATION

| Degree Discipline | Institute/Board | CGPA/Percentage | Year |
|-----------------------------------|--|-----------------|--------------|
| BSc (Hons.) Data Science and AI | Indian Institute of Technology, Guwahati | 8.41 (Current) | 2023-Present |
| Diploma in Electrical Engineering | Board of Technical Education, U.P. | First Division | 2021-2024 |
| Senior Secondary | Alpine Public School, Khurja (CBSE) | 72.6% | 2021 |
| Secondary | Alpine Public School, Khurja (CBSE) | 80.8% | 2019 |

EXPERIENCE

- Delphic Global | Python Intern** June 2026 – Present
Python Development Internship GitHub
 - Developing Python applications with emphasis on clean, modular, and maintainable code following industry-standard programming practices.
 - Implementing practical Python concepts including Object-Oriented Programming, Exception Handling, File Handling, JSON, CSV, Logging, Modules, Virtual Environments, and package management using pip.
 - Building and debugging real-world Python projects while strengthening software development, version control, and problem-solving skills.
 - Tools/Tech:** Python, Git, GitHub, VS Code, JSON, CSV, Logging, Virtual Environment, pip
- Sensing Perception & Intelligence (SPIN) Lab, IIT Guwahati | Research Intern** Nov 2025 – April 2026
Natural Scene Classification using Machine Learning and Computer Vision GitHub
 - Expanded dataset scale from **100→2000 images** (Binary Classification) and **600→6000 images** (Intel Multi-Class Classification) enabling more robust experimentation and evaluation.
 - Designed feature-engineering pipelines using statistical, texture, and colour descriptors including Entropy, Edge Density, GLCM features (Contrast, Homogeneity, Energy, Correlation), and HSV colour features.
 - Performed exploratory data analysis and feature-space interpretation using statistical significance testing (Mann-Whitney U, Kruskal-Wallis), correlation analysis, PCA, and t-SNE visualizations.
 - Implemented and compared Logistic Regression, SVM, and Random Forest classifiers using 5-Fold Cross Validation, feature importance analysis, and misclassification analysis.
 - Improved Intel multi-class classification accuracy from **52.5% to 64.7%** while obtaining **88.5% accuracy** with **89.05% F1-score** on binary classification using Random Forest.
 - Tools/Tech:** Python, Numpy, OpenCV, Scikit-learn, Machine Learning, Computer Vision, Feature Engineering, Statistical Analysis, PCA, t-SNE

PROJECTS

- Fake News Detection System** April 2026
Python, NLP, Scikit-learn, NLTK, TF-IDF, Streamlit GitHub | Live Demo
 - Built an end-to-end fake news detector using 6000+ real-world news articles
 - Applied NLP techniques (text cleaning, stopword removal, TF-IDF vectorization) for feature extraction
 - Evaluated and benchmarked 3 ML models (Logistic Regression, Naive Bayes, Random Forest)
 - Reached 94%+ classification accuracy and deployed the model
 - Deployed as a live interactive web app using Streamlit Cloud
- Secure Password Manager** July 2026
Python, JSON, Logging, File Handling GitHub
 - Developed a secure command-line password manager supporting user registration, authentication, and encrypted password storage.
 - Implemented password generation, search, update, deletion, and validation features using modular Python architecture.
 - Applied Object-Oriented Programming, JSON-based data management, exception handling, logging, and file handling to build a maintainable application.
 - Designed the project with reusable modules, clean code practices, and robust input validation to simulate real-world software development.
- E-commerce Sales Prediction** April 2026
Python, Pandas, Scikit-learn GitHub
 - Developed a sales forecasting pipeline using real-world transactional data
 - Performed data cleaning, feature engineering, and exploratory data analysis (EDA)
 - Implemented Linear Regression and Random Forest models; achieved improved performance using Random Forest for non-linear patterns
 - Identified key insights such as seasonal sales trends and country-wise revenue distribution

TECHNICAL SKILLS

- **Programming:** Python, SQL, R, Java, C
- **Python:** Object-Oriented Programming, Exception Handling, File Handling, JSON, CSV, Logging, Modules, Virtual Environment, pip
- **Machine Learning & Data Science:** Scikit-learn, Pandas, NumPy, OpenCV, EDA, Feature Engineering, Statistical Analysis, Model Evaluation, Cross Validation
- **Visualization:** Matplotlib, Power BI, Excel
- **Developer Tools:** Git, GitHub, VS Code, Jupyter Notebook, Google Colab, MySQL

KEY COURSES TAKEN

- **Mathematics & Statistics :** Linear Algebra, Calculus, Optimization, Probability, Statistics
- **Data Science & AI :** Machine Learning, Deep Learning, Computer Vision, Time Series Forecasting, Recommender Systems, Data Mining, Data Visualization, RDBMS

CERTIFICATES

- **Deloitte Australia – Data Analytics Job Simulation**
Forage Feb 2026
- **Workshop on AI with Machine Learning**
IIT Roorkee Feb 2025
- **Trustworthy Generative AI**
Vanderbilt University (Coursera) May 2024
- **Introduction to R Programming**
Great Learning May 2024

ACHIEVEMENTS

- **Academic Excellence,** Secured CPI of **9.5** in the First Trimester at IIT Guwahati *2023*
- **Grade AA (10/10),** Awarded top grade in C Programming, AI Basics, and R Programming *2023–2024*
- **Consistent Performance,** Maintaining an overall CPI of **8.41** across all completed trimesters *Ongoing*
- **Diploma with Distinction,** Completed Diploma in Electrical Engineering with First Division. *2024*

ADDITIONAL INFORMATION

- **Languages :** Hindi (Native), English (Professional Proficiency)
- **Interests :** Data-Driven Problem Solving, Data Analytics, Emerging Technologies, Artificial Intelligence, Machine Learning, Python Development, Open Source Projects