

Paras Nainwal

Haldwani, Uttarakhand | +91 7466861580 | nainwalparas5@gmail.com | [Linkedin](#)

SUMMARY

Aspiring Data Engineer with strong analytical and organizational skills, experienced in academic group projects and data workflows. Skilled in Python, SQL, and pandas for data manipulation, with exposure to visualization and pipeline optimization. Seeking to apply technical expertise to build scalable data solutions in a dynamic, data-driven environment.

SKILLS

Programming & Scripting:

Python (pandas, NumPy, scikit-learn), SQL, PySpark, C++

Big Data & Engineering:

Apache Spark (Databricks), Kafka, Airflow, Hadoop, ETL/ELT pipelines, Delta Lake

Machine Learning:

scikit-learn (Random Forest Classifier, Logistic Regression), GridSearchCV, train-test split, ROC-AUC evaluation, feature engineering, ensemble methods

Databases:

MySQL, PostgreSQL, MongoDB

Tools & Platforms:

Git, GitHub, VS Code, Jupyter Notebook, matplotlib, seaborn, Excel

Work History & Projects

Lloyds Banking Group - Data Science on Forage

- Built a customer churn prediction model using random forest and ensemble methods, achieving ROC-AUC score of 0.82
- Performed end-to-end data preprocessing pipeline: handled missing values, encoded categorical variables, implemented feature scaling using pandas and scikit-learn
- Optimized model performance through hyperparameter tuning with GridSearchCV and conducted feature importance analysis to identify key churn drivers
- Delivered actionable insights by translating model results into business recommendations for customer retention strategies

Electronic Arts - Software Engineering on Forage

- Designed and documented a new feature for EA Sports College Football, creating detailed feature proposal and technical specifications for cross-functional stakeholders
- Developed object-oriented architecture with C++ class definitions and UML class diagrams for game mechanics implementation
- Optimized codebase performance by refactoring data structures, improving algorithm efficiency and reducing processing overhead

Automated CSV → Database ETL Pipeline

- **Developed** an automated ETL pipeline to process CSV/Excel files and **validated** schema consistency.
- **Cleaned** and **standardized** bad data before loading into MySQL/PostgreSQL databases.
- **Implemented** error recovery and **streamlined** workflows with Python (pandas) and Airflow.
- **Generated** automated reports and **optimized** data handling for reliability and scalability.

EDUCATION

Veer Madho Singh Bhandari Uttarakhand Technical University, Dehradun, India
Bachelor of Technology (Computer Science)

Uttarakhand
2023 - 2027

CERTIFICATE

[Oracle Data Platform Foundations Associate](#)