NAMAN AGRAWAL

DATA ANALYST | MACHINE LEARNING ENTHUSIAST | MLOPS PRACTITIONER

Jaipur, Rajasthan, India | +91 8439410326 | dr.namanagrawal@gmail.com | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Passionate about transforming complex data into actionable insights using **Python, ML, DL** & **GenAI**; adept at **end-to-end ML deployment (Docker, AWS, MLFlow)** and leading teams to deliver measurable business impact.

TECHNICAL SKILLS

Languages	:	Python, R, MySQL, PostgreSQL
Tools	:	Power BI, Excel, Streamlit, Docker, GitHub Actions, AWS (S3, EC2, ECR), MongoDB, Kubernetes
ML/AI/DL	:	Linear Regression, Logistic Regression, Random Forest, XGBoost, SVM, ANN, CNN, RNN, LSTM, Transformers
Libraries	:	Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, TensorFlow, Keras, PySpark, SciPy, Spacy, nltk
Concepts	:	Model Deployment, Data Ingestion, Data Validation, Feature Engineering, NLP, Hypothesis Testing
NLP	:	Tokenization, Stemming, Lemmatization & Stop words, TF-IDF, Word Embedding, Sentiment Analysis
Productive	:	Linux, Windows, Excel, PowerPoint, markdown
MLOps	:	Docker, Kubernetes, GitHub Actions, AWS (S3, EC2, ECR), MLFlow, CI/CD

WORK EXPERIENCE *(Total: 5+ years)*

Samatrix Consulting Pvt. Ltd. (Gurugram) : *Data Analyst & Team Lead* (Sep 2022 – Present)

- **Leading a 6-member team** delivering data science tools and training to 800+ students, **achieving 4.8/5** satisfaction scores.
- **Deploy automated ML pipelines** for detecting real-time fraudulent transactions, improving accuracy by 15%.
- **Developed Recommender system** that analyze student feedback comments and filled forms, reducing attrition rates by 20%.
- Hands-on Experience in data analysis, Statistical Modeling, and Predictive Modeling using Python (Numpy, Pandas), SQL, and Excel.

MAT Holdings Inc. (Sonipat) : *Mechanical Engineer* (Feb 2021 – Sep 2022)

- **Automated data workflows** by developing VBA scripts to transfer and categorize production data across **Excel** sheets, reducing 70% manual efforts and **saving 10+ hours/day**.
- **Developing Pad Assembly** and part drawings, Printing drawings on CAD as per customer data for PPAP approval as well as Production.
- Managing complaints and issues from the production line

DB Engineering Pvt. Ltd. (Noida) : *Mechanical Design Engineer* (Aug 2020 – Jan 2021)

- Develop the Turning CAD Drawing of parts e.g., Knives, Holders, Cutters, Blades etc.
- Make packaging part drawings from OEM for Pad e.g., Shim drawings, Wire Indicators, etc.
- Creating and managing Job Cards according to Production Unit and maintain up-to-date documentation of process.

Preet Machines Ltd. (Ghaziabad) : *Asst. Design Engineer* (Jul 2019 – Aug 2020)

- Develop the detail drawings of Rolling Mill equipments e.g., Roller Table, Pinch Roll, Chain Transfer etc. using AutoCAD.
- Drafting Erection drawings of General Arrangement as per Layout and Bill of Materials (BOM) for purchase.
- Collect and analyze the Layout and carry out calculations for the preparation of Drawings.

PROJECTS

YouTube Comment Sentiment Analysis (Link) *(python, transformers, streamlit)*

- **89% accuracy** achieved by Boosting pre-trained model through custom preprocessing Hinglish text.
- **Designed text preprocessing pipeline** that handle null values, emojis, URLs, code-mixing, and Indian English dialects.
- **(0.91 F1-score)** Validated on 500 manually labeled comments.
- **Interactive dashboard** deployed for processing 700+ comments/min, showcasing **analysis and analytics** using graph.

Vehicle Insurance Claim Prediction – MLOps Project *(python, MLFlow, fastapi)*

- **Developed end-to-end ML pipeline** for predicting insurance claims. **(Recall: 1.0, F1 Score: 0.93, and Accuracy: 0.88)**
- Model deployed using **FastAPI on aws EC2 via Docker** with CI/CD using GitHub Actions.
- Implemented data ingestion, validation, transformation, training, and model registry on AWS S3.
- Enabled **real-time predictions** with a Fast API and used MongoDB Atlas for data storage.

Credit Default Risk Analysis (Link) *(python, catboost, streamlit)*

- **0.88 F1-score** achieved by developing ML model (CatBoost) on imbalanced dataset. (40k+ rows, 60+ features)
- **Feature Engineering** by filling nan values, VIF (>6), chi-square test, ANOVA and t-tests improving precision by 15% versus baseline.
- **SMOTE Implementation** for class balancing and hyperparameter tuning.
- **Interactive front-end** deployed for prediction for analyzing and categorizing risk into 4 buckets (1 to 4).

Real-Time Payment Fraud Detection System (Link) *(python, xgboost)*

- **0.92 F1-score** achieved using XGBoost on 6 million row dataset.)
- **Solved severe class imabalance** by using SMOTETomek and random undersampling **hybrid techniques**.
- **Feature Engineering** and created new features that improved precision by 18%.
- **Automated Preprocessing and training Pipeline** with (OneHotEncoder, StandardScaler etc.) via scikit-learn Pipeline.

EDUCATION

- **Chandigarh University ;** Master of Science (MS) - Data Science (Aug 2024 – Jul 2026)
- **Kanpur Institute of Technology, Kanpur;** Bachelors of Technology (B.Tech) (Aug 2015 – Jul 2019)

ACHIEVEMENTS

- **Secured Process Automation** | Developed VBA scripts to automate cross-Excel data transfer and categorization, reducing manual errors by 70% and saving 10+ hours/day
- **Kaggle Competitions Contributor** | Participated in Kaggle competition with code, dataset, and discussion contributions.
- **Published educational blogs and Jupyter notebooks** on machine learning and deep learning topics (e.g., RNNs, Gradient Descent, Backpropagation), combining mathematical intuition and analogies. e.g., RNN, Optimizers, BackPropagation, Logistic Regression