

Anshika Gupta

Email: anshikagupta379@gmail.com | Mobile: +91 8824623961 | LinkedIn: [My LinkedIn](#)

PROFESSIONAL SUMMARY

A dedicated and quick learner with a passion for Data Science, Machine Learning, Deep Learning, and Business Intelligence. Possessing hands-on experience in developing and deploying machine learning models, data visualization, and analytics.

SKILL SUMMARY

Languages: C++, JAVA, Python, SQL

Technologies & Frameworks: Scikit-Learn, TensorFlow, Keras, Pandas, Numpy, PyTorch, MySQL, MongoDB, Tableau, Power BI, Statistics, HTML, CSS, KNIME, Wireshark, Hadoop, Pig, Spark, FastAPI, Docker, MS Excel

WORK EXPERIENCE

Mazars India

Gurugram, India

Data Analytics Intern

Feb 2023 – Apr 2023

- Developed more than 100 queries to reduce audit related tasks by 50 % using Tableau and Excel and **MySQL**.
- Prepared customized **macros, pivots using Excel, used filters, prepared dashboards and reports** on **Tableau** and **Power BI**.
- Led a team in Designing, optimizing and deploying a **Machine Learning Model** for **Anomaly Detection** in audit data.
- Executed web scraping tasks for data collection and performed data visualization using forecasting through charts, graphs and insights. Handled large data files (over 100 Mb) and performed **data cleansing**.
- Prepared custom POVs for more than 10 clients' various business consultancy solutions.

Upskillz

Gurugram, India

ML Research Intern

Jul 2022 – Sep 2022

- Gained hands-on experience with frameworks such as Scikit-learn, TensorFlow and Keras.
- Applied optimization techniques such as L1/L2 regularization, dropout, batch normalization, Adam and hyperparameter tuning to real-time projects.
- Conducted thorough Exploratory Data Analysis (EDA) to uncover insights and prepare data for modeling.
- Carried out projects including :
Image Classification using Transfer Learning by fine tuning models like AlexNet, VGG16 and ResNet.
Heart Attack Prediction using supervised learning techniques like Logistic Regression, SVM, Random Forest.

EDUCATION

The NorthCap University

Gurugram, India

Bachelor of Technology in Computer Science; **CGPA: 8.75/10**

Oct 2020 – Jun 2024

Courses: Operating Systems, Data Structures, Analysis of Algorithms, Artificial Intelligence, Machine Learning, Deep Learning, Networking, Database Management Systems, Computational Statistics, OOPs, Software Engineering etc.

Modern Public School

Bhiwadi, India

RELEVANT PROJECTS

1. Crop Disease Prediction using Deep Learning (CNN):

Objective: Performed prescriptive analysis to develop a custom CNN model for predicting crop diseases using the new-plant dataset (87,000 RGB images, 38 classes) and a diagnosis and treatment recommendation system : research oriented.

Key Responsibilities:

- Led** a team of 3 in designing and training a custom Convolutional Neural Network (CNN) model on **Google Colab (T4 GPU)**, achieving **98.7% train accuracy** and **96.8% validation accuracy** over 15 epochs using **TensorFlow/Keras**.
- Coordinated performance comparisons with popular architectures, including **AlexNet, InceptionV3, ResNet, GoogleNet**, and **MobileNet** using **Transfer Learning** techniques.
- Built a real-time disease diagnosis and treatment recommendation system, deploying a web interface using **FastAPI, Postman, HTML, CSS**, and **JavaScript**.

2. Customer Segmentation and Recommendation System:

Objective: Analyzed a transactional dataset from a UK-based online retailer (541,908 transactions) to understand customer behaviour and purchasing patterns and then built a recommendation system using unsupervised learning.

Key Responsibilities:

- Performed EDA, scaled the features , did **PCA (Principal Component Analysis)** for dimensionality reduction.
- Applied **K-means clustering** algorithm to segment customers based on **RFM (Recency, Frequency, Monetary) analysis** and evaluated the performance using **silhouette scores, elbow method and radar charts**.
- Implemented **collaborative filtering** and **content-based filtering** techniques to build a recommendation engine.

3. Health Insurance Claim Analysis:

Objective: Designed a predictive model to estimate health insurance claim amounts using supervised learning.

Key Responsibilities:

- Handled missing data, encoded categorical variables, plotted trend patterns using pyplot, scaled the features to prepare dataset for modelling and plotted a correlation heatmap.
- Used **Linear Regression, Decision Tree, Random Forest, Gradient Boosting, XG Boost, Light GBM** and evaluated model performance using **MSE and R2 score**.
- Implemented a method to predict insurance claim amounts for custom output using the trained models.

4. Credit Card Fraud Detection:

Objective: Developed a robust credit card fraud detection using a PCA transformed dataset (284,807 transactions) addressing the challenge of extreme class imbalance.

Key Responsibilities:

- Addressed class imbalance where fraudulent class was significantly underrepresented using **Borderline SMOTE**, used Standard Scaler to normalize features.
- Employed algorithms such as **Logistic Regression, Naïve Bayes, K-Nearest Neighbours (KNN) and Random Forest**, evaluated the model performance using **accuracy, precision, recall, F1 and AUC score**.
- Tuned model parameters to maximize detection accuracy while maintaining computational efficiency.

5. Power BI Analysis of IPL Performance and Trends:

Objective: Analysed IPL (Indian Premier League) match data to uncover insights and trends using descriptive analysis.

Key Responsibilities:

- Imported and transformed IPL match data using Power BI's data modelling capabilities.
- Developed interactive dashboard to visualize match statistics, player performance, and team trends.
- Utilized various Power BI features including **charts, slicers, and filters** to enhance analysis, generated reports and presentations to effectively communicate findings.

6. Flight Delay Analysis and Visualization using Tableau:

Objective: Analyzed factors contributing to flight delays, cancellations, and diversions across various airports and airlines.

Key Responsibilities:

- Developed interactive visualizations to assess flight delays due to weather, air carriers, and NAS across different time periods.
- Identified key insights on airport security, flight cancellations, and airline performance regarding delays
- Created dashboards with filtering and drill-down capabilities for detailed data exploration.

CERTIFICATIONS

- Data Structures and algorithms certification from Coding Ninjas.
- Top Performer in Data Structures at Coding Ninjas.
- Data Analytics Certification from ICT Honeywell.
- Introduction to Cisco Packet Tracer.
- AWS Certification in Machine Learning Foundations and Cloud Foundations.