# NEHA SINGH

**Senior Data Analyst at Accenture**

@ nehasinghiiitm@gmail.com  7509941320  Mumbai  www.linkedin.com/in/neha-singh01/

## WORK EXPERIENCE

### Analytics Advisory Analyst

**Accenture**

May 2021- Present  Mumbai

- Designed and maintained ETL workflows using **Azure Databricks**, handling high-volume data with complex transformations.
- Developing **Dynamic use cases for claims and identifying fraud** using R programming, Machine Learning models.
- Working in the selection and extraction of features to make the data ready for modeling.
- Identifying trends, patterns, and anomalies in data.
- Analysed and optimized queries using **Spark** concepts such as Partitions, broadcast join, designed both managed and external tables.
- **Migration** of modules/engine from **R to Azure Databricks** and developed solution resulting increase in performance by 56%.

### Trainee Consultant

**Datamatics Global Service Ltd**

July 2019-May 2021  Mumbai

- Developed **Interactive and Dynamic R Shiny Dashboards** for financial data using R programming complex mathematical logic and calculations implemented. I.e., weighting and scoring logic
- Built **Configurable and auto code generators in R**, which allows code to be redevelop internally based on given inputs
- Predicted **Price Elasticity of product** using regression algorithm and created dashboard to forecast the analysis
- Worked on Large Datasets using R libraries and developed insights and business analysis, created dashboards using BI tools and R-shiny.

## EDUCATION

### Integrated Post Graduate(B.TECH+M.TECH), Information Technology

**IIITM Gwalior**

July2014- May2019

## CERTIFICATIONS

- Introduction to cloud computing on AWS for Beginners course completion certificate from Udemy.
- Machine Learning Hand on with R and Python course completion certificate from Udemy.
- Certified Data Science in Real Life from Coursera Certificate ID(BP38G39K8TBX)
- Certified Managing Data Analysis from Coursera. Certificate ID(HP322NMMLYVU)

## SKILLS

R | SQL | SparkR | Machine Learning | shiny | Python | Azure Databricks | Azure Data factory | Big Data Analysis | Statistical Modeling | Data Visualization

## PROJECTS

**Supplier Warranty Recovery**

- A machine learning model has been developed to determine whether a dealer claim will be accepted or rejected by a supplier
- Load 24 months data in databricks and deciding on the features required.
- Applied **XGBoost** to improve the accuracy of model.
- Used metrics such as confusion matrix, accuracy score to understand the performance of the models.
- Tasks: Data extraction, Data preprocessing, Data Exploration, Random Forest, Modeling, Model Evaluation, Python

**Warranty Claim Management**

- Based on defined rules, determine whether a claim raised by dealer for vehicle repair should be accepted or not.
- Responsible for developing complex use cases, troubleshooting issues and implementing solutions in R programming.
- Created **pipeline in Azure** to automate the task end-to-end.
- Used SQL queries to Fetch data from Azure data lake.
- Active monitoring of jobs and pipeline to ensure seamless functioning of the analytics flow.