

# Neelraj Khawale

Dallas, TX, 75252 | +14695790604  
neelkhawale@gmail.com | [www.linkedin.com/in/neelraj-khawale](http://www.linkedin.com/in/neelraj-khawale)

## EDUCATION

**Master of Science (M.S.) - Business Analytics**  
University of Texas at Dallas, TX

Jan 2022 - Dec 2023  
GPA - 3.70

**Bachelor of Engineering (B.E.) - Computer Engineering**  
Savitribai Phule Pune University, India

Jun 2016 - May 2020  
GPA - 3.20

## PROFESSIONAL EXPERIENCE

**Data Analytics Intern (Remote)**

Sept 2023 - Dec 2023

SoftClouds LLC, San Diego, CA

- Designed, implemented, and maintained ETL pipelines to extract, transform, and load large amounts of data from various sources.
- Optimized dashboard performance by enhancing the data model and streamlining SQL queries resulting in a 25% improvement in dashboard performance.
- Enhanced functionality of "Kapture" Knowledge Management app, integrating Tableau dashboard API for real-time analytics and app health monitoring.

**Analytics Officer**

Jun 2022 - Jan 2023

Data Science Club, UTD, Richardson, TX

- Led 5 analysts in data collection and analysis, creating data-driven reports, boosting club engagement and participation by 30%.
- Coordinated kick-off events and workshops for 80+ members.
- Mentored club members to build projects, offering guidance and support for successful project outcomes.

## ACADEMIC PROJECT EXPERIENCE

**Unicorn Companies Analysis**

Jun 2023 - Aug 2023

- Analyzed unicorn companies' data using SQL and Tableau.
- Identified key metrics: ROI, time to unicorn status, industries, countries, and top investors.
- Utilized advanced SQL techniques for data cleaning and exploration (joins, CTEs, window functions).
- Created a Tableau dashboard, incorporating features such as parameters, top N filters, reference lines, and calculated fields for enhanced visualization.

**Integrated HR Data Analytics**

Jan 2023 - May 2023

- Conducted Python-based analysis on Employee Development, Recruitment, and Attrition Data.
- Enhanced training effectiveness and recruitment processes through data-driven insights.
- Utilized feature selection and classification models for attrition analysis, informing retention strategies.
- Developed visualizations and reports to support decision-making in Tableau.

**S3 Redshift Batch ETL Pipeline**

Sept 2022 - Dec 2022

- Designed and implemented scalable ETL pipelines using PySpark and AWS services (S3, Redshift).
- Developed data models and dimensional schemas to support analytical queries and reporting requirements.
- Automated pipeline execution and monitoring using Apache Airflow DAGs, ensuring reliability and efficiency.
- Conducted data quality checks and implemented error-handling mechanisms to maintain data integrity.

## SKILLS AND CERTIFICATIONS

**Certifications:** Google Data Analytics, Alteryx Foundation, R for Data Analytics.

**Tools:** R, SAS, AWS, Hadoop, Hive, Spark, MS Excel, VBA, SQL – MSSQL, MySQL, Azure Databricks, Oracle, NoSQL - MongoDB, Tableau, Power BI, Alteryx.

**Data Engineering:** ETL Development, Data Modeling, Database Management, Data Warehouse.

**Machine Learning:** Linear and Logistic Regression, Statistics, Statistical modeling, Naive Bayes, Support Vector Machine, Decision Tree, Random Forest, XG Boosting, KNN, K-mean clustering, Neural Networks, Sentiment Analysis in NLP, Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn.

**Programming Language:** C, C++, Java, Python, HTML, CSS, PHP, JavaScript.