

# Afzal Mansury

---

+91-9039511460, 7999467430 | afzalmansury1516@gmail.com | [LinkedIn](#)

## About Me

---

Data-Driven Engineer with 5+ Years of Experience in Data Science, Analytics, and Visualization

- **Professional Expertise:** Proven track record in leading and executing industry-grade data projects across diverse sectors, including Oil & Gas, Pharma, Banking, Hospitality, and Telecommunications. Adept at transforming complex data into actionable insights that drive business growth.
- **Technical Proficiency:**
  - **Programming Languages:** Python and SQL, with a deep understanding of data manipulation and analysis
  - **Machine Learning:** Comprehensive knowledge of machine learning algorithms and techniques, including **Linear and Logistic Regression, Lasso and Ridge Regression, Decision Trees, Random Forests, SVM, KNN, K-Means, XGBoost, NLP, and Ensemble Methods.**
  - **Libraries & Frameworks:** Extensive hands-on experience with NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, NLTK, Keras, and Hyperparameter Tuning.
  - **Data Engineering:** Experience in building scalable data pipelines, including the implementation of Medallion Architecture (Bronze, Silver, Gold layers) for improved data processing and analytics.
  - **Data Visualization:** Skilled in creating insightful, interactive dashboards using Tableau, effectively communicating complex data stories to stakeholders.
- **Cloud Experience:** Applied Microsoft Azure in a specific project, leveraging services like Azure Data Lake Storage and Databricks to enhance data processing efficiency

## EXPERIENCE

---

**Centelon Solutions** | Bangalore, India

(Nov 2020 - Present)

Senior Consultant - Data Science

Working for the Data team, mainly responsible for suggesting Machine and Deep learning solutions by analysing data.

### Projects:

1. **Email Triage | Centelon Product | NLP Based Classification Project**
  - **Objective:** Developed an automated email triage system to classify emails using NLP and Machine Learning, optimizing workflow efficiency by determining whether action is required and organizing emails accordingly.
  - **Approach:**
    - **Data Processing:** Pre-processed email text data, including text **normalization, tokenization, and feature extraction** (TF-IDF, Bag of Words).
    - **Model Development:** Implemented various binary classification models (**Logistic Regression, Decision Tree, Random Forest, SVM, XGBoost**) to analyse the email content and classify emails into actionable and non-actionable categories.
    - **Model Evaluation:** Performed cross-validation and hyperparameter tuning to select the best-performing model based on accuracy, precision, recall, and F1-score.

- **Automation:** Integrated the final model into an automated pipeline to move classified emails to the respective subfolders in real-time.
- **Impact:** Improved email management efficiency by automating the classification and organization process, reducing manual sorting time by **98%**.

## 2. Data Migration & Visualization | Oil & Gas Company in Australia

**Tool:** Microsoft Azure | Tableau

- **Objective:** Designed and implemented a data migration and visualization solution using the Medallion Architecture to optimize data processing and enhance data-driven decision-making.
- **Key Contributions:**
  - **Data Pipeline Implementation:** Developed a robust data pipeline using Microsoft Azure, incorporating Medallion Architecture (Bronze, Silver, Gold layers) to ensure data integrity, scalability, and efficient processing.
  - **Interactive Dashboards:** Created dynamic Tableau dashboards that transformed complex data sets into visually compelling insights, enabling stakeholders to effortlessly identify key trends and make informed decisions.
  - **Advanced Tableau Features:** Leveraged Tableau's advanced features such as calculated fields, parameters, and Level of Detail (LOD) expressions to enhance dashboard functionality and user experience.
  - **Optimized Visualization:** Implemented responsive containers in Tableau, ensuring optimal visualization across various screen sizes and devices.
  - **Client Collaboration:** Engaged closely with clients to understand their requirements, translating them into actionable solutions that exceeded expectations.
  - **Cross-Functional Collaboration:** Worked alongside data engineers and business analysts to gather, integrate, and validate accurate data sources, ensuring seamless data migration and visualization

## 3. Document Parsing for a Leading Pharma Firm

- **Requirement:** To develop a solution that can extract data from PDF invoices and store it in the database within 15 minutes after receiving the mail from vendors.
- **Solution:** We suggested and developed a solution that can extract text from PDF invoices and extract tabular data from PDF using Tabula & Camelot. Used custom-trained Tesseract for text detection and text recognition. Converted different PDF patterns into text using Apache Tika and extracted desired output. Also performed mail log creation to extract the information from mail like sender's email ID, email's subject, date, name of attachment & count of attachment.

**Oravel Stays Private Limited-OYO** | Pune, India

(June 2019 - Feb. 2020)

Junior Analyst

**Project:** Hotel Booking Cancellation | **Domain:** Hotel Industry

- Built the Machine learning classification models to predict which customer is likely to cancel the booking on which factors and help the customer support team devise policies and minimize the cancellation.
- Working with the Central Operation team, Business Development team and vendor for smooth operation. Also sharing important analysis from past booking data to minimize cancellations and improve overall sales.

**Reliance SMSL | Jaipur, India**  
Quality Engineer-Data

(June 2018 - June 2019)

**Project: Revenue growth-4G | Domain: Network & Digital Business**

- Analyse data for setup of the new major and minor towers in remote areas for expanding business and revenue in PAN Rajasthan
- Responsible for data pre-processing of unstructured data.
- Used data mining techniques to reveal patterns in the data that have implications for business decisions.

**Classroom Training– Detail**

---

**IMARTICUS**

(Nov. 2019 – June 2020)

Classroom Training– Data Science

**Projects:**

1. **Prediction of Employee Attrition (Language: Python)** - Built the classification models to predict which employee is likely to churn and help the business to devise policies and attract back the right talent, by using multiple classification algorithms like Logistic Regression, Decision Tree, Random Forest and XGBoost and compared results.
2. **Predict The Flight Ticket Price (Hackathon | Language: Python)** - Built the regression model to predict ticket price using multiple regression algorithms like Linear Regression, Lasso, and Ridge compared results.

**CERTIFICATIONS**

- 
- Pro Degree in Data Science (Internship) Imarticus Learning, Pune, Certified by Genpact.
  - Python for Data Science and Machine Learning Boot Camp Source - Udemy.
  - Python for Beginners Source - Udemy.
  - Tableau 2022 A-Z: Hands-On Tableau Training for Data Science - Udemy

**ACHIEVEMENTS**

- 
- Three times Nominated for **Spot Award** in Centelon while working on a client project.
  - Recognized as **Employee of the Quarter** in Oyo for improvising the Overall sales in remote areas.
  - Certificate of Leadership for significant contribution to developing a vibrant ecosystem of entrepreneurship on camp.

**ACADEMIC PROFILE**

- 
- **Mandsaur Institute of Technology, RGPV-Bhopal** (Aug 2012 - June 2016)  
B.E (Civil Engineering)
  - **Dashpur Vidhyalaya** (April 2011 - May 2012)  
Intermediate (M.P Board, Percentage – 85.0 %)