

Vaibhav Rustagi

(Data Engineer/ ML Engineer /Senior Data Analyst)

Microsoft Certified: Azure Data Engineer Associate

Microsoft Certified: Azure Data Scientist Associate

Email: vaibhrustagi@gmail.com

Krishna Nagar, Delhi- 51

Mobile: 9013257864; 8983245705



[View my Microsoft Badges](#)

Work Experience

6.2 Years

Professional Summary

Seasoned professional having an overall 6+years of experience in cross-functional domains (Life Sciences and Automobiles), primarily focusing on Advanced Data Analytics, Reporting & Data Modernization and deployment on Azure Cloud. Have expertise in Azure cloud computing and Machine learning & predictive analytics.

Fractal Analytics, Gurgaon

Dec'20 - Present

Designation: Senior Cloud Engineer

Project Details: An **American** multinational consumer goods corporation headquartered in Cincinnati, Ohio, founded in 1837. It specializes in a wide range of personal health/consumer health, and personal care and hygiene products; these products are organized into several segments including Health Care; Fabric & Home Care; and Baby & Family Care.

- **Application:** Building Inventory tracker (North America region) from scratch on Azure platform
- **Technology Used:** Azure Data Factory, Databricks, DataLakes/Blob Storage, Azure SQL DB, Logic App, PySpark, SparkSQL, Power BI etc

Roles & Responsibilities

- **Implemented Data Modernization** - Data integration, data cleaning and data transformation using **MS Azure** services.
- **Developed SharePoint – Blob sync connection** using **Logic app** connector which pulls all the vendor's weekly sales data into blob container once uploaded by User to a specific SharePoint folder, scheduled for a weekly run.
- **Designed** production **data pipelines** for data ingestion, storing data to a **hierarchical namespace** format, using **ADF** and data featurizing, transformation, aggregation and merger with different data sources (CSV, xlsx and delta lake) into a final dataset using **Databrick**.
- **Built** newly Azure **SQL DB** for storing Inventory tracking database for all the vendors across North regions with essential **KPI metrics**.
- **Developed** integration solutions between SQL DB to Power BI for building **descriptive** and **predictive analytics reports** using potential visualizations and multiple KPIs and further publish the report to **Power BI**.
- Developed **modeling techniques** to discover and preparing **visual dashboards** showcasing essential business KPIs as for **MOM/QOQ/YOY** comparison/target using **Power BI** for leadership decision-making calls.

Cognizant Technology Solution, Noida

Feb'15 – Dec'20

Designation: Associate

#4 Project Details:

Sep'19 – Dec'20, Noida

Client: An **American** worldwide manufacturer, marketer and distributor of consumer and commercial products with a portfolio of brands including: Rubbermaid storage and trash containers, home organization and reusable products etc.

- **Application:** Data Modernization on Azure platform
- **Technology Used:** Azure Data Factory, Databricks, DataLakes/Blob Storage, Azure Synapse, PySpark, Power BI

Roles & Responsibilities

- **Implemented Data Orchestration** -Data Transformation and Data integration using **MS Azure services**.
- **Developed** production **data pipelines** for data ingestion, storage, processing & transformation and visualization for building **descriptive** and **predictive analytics reports** using Azure analytics services and further send to **Power Bi** for viewing reports.
- **Developed batch processing** and integration solutions across region **heterogeneous source** of data (**SAP**, flat file from **SFTP location- Salesforce/Avaya**) into Azure platform using **ADF**, **Logic app** for file ingestion etc.
- Developed **notebooks** using **Python/SQL/PySpark** as part of **data extraction, data cleaning, transformation** and processing.

#3 Project Details: Cognizant IT Data and Analytics

June'18 – Sep'19, Noida

As a part of the IT DnA, my responsibilities include to develop deep/machine learning models for the firm's internal clients, designing and implementing Data Science POCs to optimize various processes, deploying ML model on Azure Cloud and integrating the same with other live projects for improving AI adoption and enablement within the organization.

Skill: Python, Machine Learning, Data Visualization – Matplotlib, Power BI, OpenCV, Flask, Text-Mining, TF-IDF, development of ML model using Azure Cognitive services.

- 1. Concur Global:** Concur team is responsible for building global expenditure reports for the employee's travel and accommodation (International & Domestic) providing both cashless and reimbursement facilities.
Objective: To develop **automated data processing** software using **Azure Form Recognizer**. To identify and extract text, useful details from the documents- hotel invoice, cab receipt, meals invoice etc for **auditing** and to ease the expense reimbursement processing.
Analytic Approach:
 - Identified and extracted **key-value pairs** and **table data form** documents which is then outputs structured data that includes the **relationships in the original** file.
 - Validated** the **Json file** for each document (image/pdf) for better understanding of data insights at **cluster level**.
 - Models trained** comprised of dual approach for custom model **without label** for tabular extraction and custom model **with labeled** data which perform better and can work with more **complicated documents**.
- 2 CWS India:** Corporate Workspace services to ensure the safety and security for all employees/guest across all the premises of Cognizant Offices in India.
Objective: To prepare a web based contactless attendance application via **Facial Recognition** approach to mark the **time-in/time-out** of the employees/vendors of the organization and to deploy this **web application** on **Azure platform**, this is being used across Cognizant Offices India-wide for Clients/Vendors/Guest at entry/exit premises..
Analytical Approach:
 - Developed face recognition solution which uses OpenCV Haarcascade technique to mark the attendance by matching the facial landmarks of the employees with employee face images stored in our trained model.
 - Delivered with a latency of 1-2sec to recognize the employee identity using **LBPH face recognizer, Cascade Classifier** etc and integrated with **Flask** framework for web interface.
 - Deployed** the web application via **Azure Web application** while storing the employee images to **Azure Blob** storage and further loading files URLs to **Azure SQL Database** for maintaining logs.
- 3 ISpace:** Created Innovation idea **Search Engine**- cum- Recommendation system for Project manager/Delivery Lead.
Objective: User performs a search of specific business relevant/ technology keywords, the system must be able to pick the similar combination of the given keywords and generate the list of innovation idea entries more relevant to the search from a set of existing innovation ideas across the other Business units.
Analytical Approach:
 - Developed a custom **Search engine** using a Python library called "**Whoosh**" with multiple functionalities as mentioned below:- Sentence correction, Synonyms, Variation searched word, Blocking unwanted/irrelevant meaning of searched text using **TextBlob, thesaurus** library. Further searched results are fed to a **Wordcloud** for visualization of trending technologies keywords.
 - Built **User profiling** method to extract important similar keywords from a set of characteristics of project description as per given project ID by individual logger using **TF-IDF and Cosine distance** method to calculate similarity score.

#2 Project Details:

Dec'17 – May'18, Noida

Client: Pharma MNC, is an American publicly traded biopharmaceutical company founded in 2013. It is the world's sixth-largest independent biotech company by market cap.

- Application:** Marketing Analytics and Business Intelligence
- Technology Used:** Python, Teradata, SQL, Power BI

Roles & Responsibilities

- Analytics and Reporting** of Marketing data on how the brand promoted through various channel and analyzing the digital marketing channel and providing **insights** on the TRP.
- Analysis** on the **Survey data** of survey conducted among doctor and patients. Also analyzing the calls and activity data during rep visit with doctor for brand promotion.
- Extracting raw data from **Teradata** database via **Teradata SQL** and unstructured data from third party global data vendor using HiveQL.

#1 Projects Details:

Feb'15 – Nov '17, Pune

Client: Automobile MNC, US based was founded in 1837 and is the world's leading manufacturer of advanced products and services for agriculture, forestry and heavy construction vehicles.

- **Application:** E-commerce based Sales - Financial Management
- **Role:** Data Analyst
- **Technology Used:** SAS, SQL, Advanced Excel

Roles & Responsibilities

- Involved in entire life cycle of project– **Analysis, Enhancement and Client Interaction**. Application involves high impact critical process i.e. **Sales, Invoicing, Return/Credits, Inventory** with direct client interaction involves **15 months** of working experience at the client site **Technology Center of India (TCI)**, Pune.
- Prepared detailed **documentation/reports** that involve analysis of factors such as the operating environment, business strength, financial strength, management strategy, sales- market-implied parameters and distribution support.
- Analyze customer data with **analytic** programming **SAS**, to do descriptive analytics and creating **visualizations** to understand the customer data.

Soft Skills

- Good in initiating conceptual ideas and work on developing new ideas and approaches for solving problems.
- Exceptional ability to maintain and build client relationships, able to comfortably converse with all facets in the clients' organization.
- Strong ability to analyze and solve problems efficiently and take independent decisions.
- Keep Stake holders well informed about the progress of the tasks and complete tasks in a timely manner.

Education

Year	Qualification	Institute	CGPA/ %
2014	B.Tech - ECE	SRM University, Chennai	7.75/10
2010	XII, CBSE, Delhi	Greenfields Public School, Delhi	70.8%
2008	X, CBSE, Delhi	Greenfields Public School, Delhi	80.8%

Achievements

- Won "GEM", 2016 and "Rough in the Diamond", 2016 award for commendable contribution to MLEU (BU) Practice.
- Early Promoted to Associate in 2017 for commendable contribution to MLEU (BU) Practice.
- Member of Cognizant Social Event Management group: All India level Outreach Fest (conducting games and talent shows for children of all Govt. school, Pune) and participated in Go Green cultured activities.
- Received Employer award for Q1 FY2020 amidst pandemic to deliver social distancing solution apps with determined performance and commitment with timeline.

Key Interests

- Traveling, Long Driving, Playing Snooker
- Connecting with new people, which helps to improve my interpersonal skills.