

VAIBHAV BODHE

(+91) 9881564473

bodhevaibhav007@gmail.com

https://www.linkedin.com/in/vaibhav-bodhe/

https://github.com/VaibhavB98.

DATA ENGINEER

Inteliment Technologies Pvt. Ltd.

WORK EXPERIENCE

Inteliment Technologies Pvt. Ltd. (Jan 2023- Present)

Project 1: Data Services Portal (Development):

Implemented a **Data Services Portal** to modernize and streamline data flow, replacing inefficient email-based processes with a **centralized web application**, enhancing transparency and efficiency across Business Units.

- **Roles and Responsibilities**
  - As a **Business Analyst**:
    - Orchestrated discussions and meetings to comprehend the firm's workflow objectives, subsequently **crafting logical workflows and wireframes** to guide developers and establish frontend design requirements.
    - Helped client get the **cost estimation of cloud resources** for their analysis, coordinating with Microsoft to obtain accurate quotations for planned deployment.
    - Facilitated a subject matter expert meeting with a **Microsoft MVP** to explore custom configuration options for **Azure Purview**.
    - Developed and **maintained Entity-Relationship Diagrams (ERDs)** for multiple schemas within the foundational systems at the ground level, serving as primary data sources for the Data Warehouse (DWH).
  - As a **Cloud Data Engineer**:
    - Led the setup of essential Azure resources for the client, including **ADLS Gen2, Azure Purview, Azure Synapse Analytics, and Azure SQL Database**, forming the foundational architecture of the portal.
    - Established and configured **Azure Purview to conduct periodic catalog scans**, ensuring the maintenance of the **Business Glossary and extraction of lineage from existing entities**.
    - Designed the **backend database schema** for the web application, ensuring efficient data capture with minimal loss and redundancy.
    - Supported the backend team in connecting the application to cloud resources through **APIs**, enabling seamless integration and functionality.
- **Achievements**
  - Successfully executed each milestone aligning seamlessly with planned sprints, showcasing adept project management and efficient delivery.
  - Appreciation from the client directly for being “**extremely flexible and diligent as we worked through several challenges and scope/requirement changes**”

Project 2: Data Quality Framework Implementation :

Implemented **single-handedly a robust framework** to quantify and **assess data quality** emanating from the Data Warehouse, empowering users to make informed decisions based on **standardized data quality metrics**. This initiative streamlined data analysis processes, ensuring clearer insights and enhanced data utilization.

- **Roles and Responsibilities:**
  - Developed a **framework** using **Python and PySpark** to compute **data quality scores** for data transmitted outside of the Data Warehouse (DWH) and consumed by Business Units.
  - Retrieved data **from Azure Data Lake Storage Gen2 (ADLS Gen2) location and integrated with Synapse Spark pool along with Azure SQL Database** to fetch computation rules.
  - Implemented computation logic based on rules stored in **Azure SQL Database and generated metadata files for reporting purposes**.
  - Collaborated with stakeholders **to refine project objectives and ensured adherence to best practices** in data processing and quality assessment.
- **Achievements:**
  - Initiated reporting of **Quality scores**, leading to pinpointing of **critical issues in the Data Warehouse** which expedited the resolution process, significantly enhanced **data reliability and efficiency**.
  - Implementation of framework led to the decision to extend its application to staging area, **ensuring erroneous data is identified and addressed before integration into the Data Warehouse (DWH)**.
  - This optimization resulted in a **more reliable reporting system and significantly enhanced the effectiveness of data analysis processes**.

Project 3: Data Services Portal (Support):

As **Technical Lead for Data Services Support**, I manage **data extraction and report creation from the Data Warehouse**. This includes scripting **SQL queries, stored procedures**, and documenting reports. I also analyze requirements, locate data sources, and utilize tools like **SQL Server and Synapse Analytics** for data extraction and loading.

- **Roles and Responsibilities:**
  - Lead Data Services Support team**, overseeing **data extraction and report creation from the Data Warehouse**. This includes scripting **SQL queries, creating pipelines ,stored procedures**, and meticulously documenting reports.
  - Analyze requirements from diverse business units to tailor customized reports, while identifying and **sourcing relevant data from the Data Warehouse**.
  - Utilize tools like SQL Server and **Synapse Analytics for efficient data extraction and loading**, collaborating with stakeholders to ensure successful delivery of data solutions aligned with business objectives.
- **Achievements:**
  - Successfully **delivered over 125 data** requests within a single year, earning **recognition from customers in the form of a Customer Appreciation Award**.

Cognizant Technology Solutions. (Nov 2020-Jan 2023):

- **Roles and Responsibilities**
  - The role demand day-to-day activities of **creating connections to various data lake storages (Amazon S3, ADLS gen 2)**.
  - Creating **metadata store, analyzing huge volumes of data (Millions of records)**, specific to the medical domain using **Databricks, SQL, and Pyspark**.
  - Suggest **attribute-specific Data quality rules to stewards and then Implement a Data Quality evaluation framework** using Ataccama One Web
  - Coordinating **with domain-specific Data scientists from various Medical fields of expertise** for client-specific needs.
  - Documenting the **evaluation results for Data quality rules and providing analysis for publishing the data downstream**.
- **Achievements**
  - Delivered the MVP comprising of **Implementing Data Quality Framework for five datasets in less than the stipulated time**.

ACADEMIC BACKGROUND

<ul style="list-style-type: none"><li>Post-Graduation and Diploma in Advanced Business Analytics(ePGD-ABA) Indian Institute of Management, Ahmedabad</li></ul>	Mar 2023 – Current
<ul style="list-style-type: none"><li>Bachelor Of Engineering (B.E) Smt.Kashibai Navale College of Engineering, Pune (M.H)</li></ul>	Jul 2016 – May 2020 B. Tech CGPA – 8.11/10.0
<ul style="list-style-type: none"><li>Senior Secondary (Class XII) Ira Internatioal School,Butibori, Nagpur</li></ul>	Mar 2015 – Mar 2016 Percentage – 75.60%
<ul style="list-style-type: none"><li>Secondary (Class X) CBSE Stella Maris Convent School, Rajura</li></ul>	Mar 2013 – Mar 2014 CGPA – 9.8/10.0

SKILLS AND ACHIEVEMENTS

Technical Skills
RDBMS   SQL   Azure - Azure Datafactory (ADF) Azure Synapse Analytics(Synapse)  Azure DataBricks (ADB)  Data Warehousing  Spark-Architecture PySpark   Ataccama OneWeb   Machine Learning   Python  MongoDB  Hadoop
Core Competencies
Adaptability and Quick Learner   Multitasking   Disciplined   Communication and Interpersonal Skills   Business Analysis   Team Player   Agile Framework
Certifications
<ul style="list-style-type: none"><li>Microsoft Certified: Azure Data Engineer Associate (DP-203).</li><li>Microsoft Certified: Azure Data Fundamentals (DP-900).</li><li>Microsoft Certified: Azure Fundamentals (AZ-900).</li><li>IBM Data Science Professional Certificate.</li></ul>

PROJECTS

<p><b><u>Analysis of Brand Preferences Among Top 2W Companies in Indian Auto</u></b></p> <ul style="list-style-type: none"><li>The project <b>focused on</b> analyzing <b>User Generated Content (UGC) from multiple sources, including comments on YouTube videos reviewing electric 2W vehicles and reviews on automobile sites.</b></li><li>Utilized text analysis techniques, such as <b>sentiment analysis, to quantify customer perceptions of electric scooters.</b></li><li>Leveraged a <b>pre-trained Roberta Model to perform sentiment analysis on the collected reviews</b>, providing insights into customer sentiment.</li><li><b>Conducted topic modeling to identify prevailing topics among customers</b>, helping to understand the most discussed aspects of electric scooters.</li><li>Evaluated three firms, with <b>one acting as the focal firm and the other two as competitors.</b></li><li>Analyzed the strengths and weaknesses of the focal firm and competitors, aiming to understand areas for improvement and competitive advantages.</li><li><b>Provided actionable recommendations to the focal firm based on the analysis</b>, aimed at addressing shortcomings and leveraging strengths to enhance competitiveness in the market.</li><li><b>GitHub link:</b> <a href="https://github.com/VaibhavB98/Analysis-Of-Brand-Preference">https://github.com/VaibhavB98/Analysis-Of-Brand-Preference</a></li></ul>
<p><b><u>Area Segmentation using Data Mining and Machine learning for new enterprises, to maximize customer turnout.</u></b></p> <ul style="list-style-type: none"><li>The Project is focused on segmenting the area into clusters (High competition, Low competition, No Competition ) for new enterprises (Here Coffee shops), so that they have a clear idea to target areas that would bring out maximum customer turnout.</li><li>The Project uses Foursquare API for Data collection, Pandas for Data Transformation, Scikit-learn for ML implementation, and Jupyter notebook.</li><li><b>GitHub link:</b><a href="https://github.com/VaibhavB98/Area-Segmentation-for-Coffee-shop-Enterprise">https://github.com/VaibhavB98/Area-Segmentation-for-Coffee-shop-Enterprise</a>.</li></ul>
<p><b><u>Plagiarism Detection using Data mining.</u></b></p> <ul style="list-style-type: none"><li>The Project was set to create a web application to detect and quantify the similarity score between the database and query document</li><li>Based on that score the evaluator could decide whether to accept the piece of work or not.</li><li>Clustering was done using datagrams, and the model was deployed locally but was provisioned to deploy over commercial servers too.</li></ul>

INTERESTS

**Sports:** Following and watching Cricket, and MMA.  
**Reading:** Reading Self-help, Finance, Business Blogs, and Reports.

SOCIAL RESPONSIBILITY

Inteliment Technologies Pvt. Ltd.
<ul style="list-style-type: none"><li>Active Member of <b>Learning and Research Department</b> to create learning material for <b>freshers in leading technology.</b></li><li>Conducted <b>numerous session and hands on training</b> on technologies and tools for aspirants through <b>Intern program</b> at the organization.</li></ul>
Cognizant Outreach
<ul style="list-style-type: none"><li>Active Member of <b>Cognizant Outreach.</b></li><li>Served as <b>a Program Manager</b> for the <b>Technology for Everyone Program.</b></li></ul>