B V Nandhan

bvnandhan@gmail.com | *+91 9110222329* | github | *LinkedIn* Leetcode Portfolio

Work Experience

AtkinsRealis , Bangalore Aug 2023 - Present

Data Engineer

- **Merck Project - Construction Cost Intelligence Platform :**
  - Built a robust **Python**-based **ETL automation** script to extract **unstructured** data from Excel files, transform it using **Pandas** and PyODBC, and load it into **Azure SQL** Database—ensuring **end-to-end data validation**, **integrity**, and auditability.
  - Stored input files in **Azure Data Lake Storage** and used **Azure Functions** to **schedule** and **trigger** the ETL script, enabling **serverless**, **event-driven processing** without manual intervention.
  - **Azure Function** served as the **compute** layer, **triggered** by a **new file upload**, to run the **ETL** process: read files from **Data Lake**, **transform** data, and **load** it into **SQL**—all without **managing infrastructure**.
  - **Orchestrated** the entire flow using **Azure Data Factory**, enabling dependency handling and future scalability.
  - Achieved a **95% reduction** in **manual processing** time and enhanced **scalability**, with **secure access** handled through **Azure Key Vault**.
  - *Tech Stack :*Python,ETL,Pandas,Numpy,SQL,Azure Functions,Azure Data Factory,Azure Data Lake,PyoDBC.
- **GM BRB Project - Cost Visualization Tool :**
  - Developed an **automated Python ETL solution** to extract data from **complex Excel** files, transform **hierarchical** structures using **Pandas**, and **load clean data** into a **SQL** database.
  - Designed and **normalized database schemas** to handle **deeply nested**, **hierarchical data** efficiently, ensuring **data consistency** and **scalability**.
  - Built **interactive React-based dashboards** on top of the transformed data to deliver **real-time insights**, trend analysis, and streamline **data verification and reporting**.
  - **Reduced manual effort by 98%** through end-to-end automation using **Azure Data Factory, Azure Functions, and Azure SQL, with secure access via Azure Key Vault**.
  - *Tech Stack :* Python, Pandas, ETL, NumPy, SQL, React.js, Azure Data Factory, Azure Functions, Azure SQL, Azure Data Lake Storage, Azure Key Vault.
- **Digital Twin Project - Metro Asset Tracking System :**
  - Designed and implemented **data pipelines** using **Azure Data Factory (ADF)** and **Azure Data Lake Storage**, reducing overall data **processing time by 85%.**
  - Developed **complex backend logics,Stored Procedures,Views,schemas** in **SQL** to optimize **large-scale queries**, resulting in a **75% improvement** in **performance** and **responsiveness**.
  - **Managed cloud infrastructure** and **automated workflows** using **Azure Logic Apps**, improving operational efficiency and **automation by 90%.**
  - Integrated **React.js dashboards** to **visualize** processed data and system insights, enhancing **real-time monitoring** and usability for stakeholders.
- *Tech Stack :* Advanced SQL, Azure Data Factory,Azure Data Lake Storage Azure Logic Apps, React.js, ETL,Python.
- **Thames Water Project - Storm Drain Analytics :**
  - **Automated complex data processing tasks using Python**, leveraging **Pandas** and **NumPy** for efficient data transformation and cleaning.
  - **Visualized insights** using **Matplotlib** and **Seaborn**, enabling faster decision-making through clear **trend analysis** and **exploratory data** views.
  - **Reduced manual effort by 98%** and significantly **cut processing costs** by **replacing manual workflows** with **fully automated pipelines.**
  - Used **SQ**L for **data integration and validation**, ensuring consistency across processed datasets in the **ETL workflow**.
- *Tech Stack* : Python , Pandas , NumPy , Matplotlib , Seaborn, SQL,ETL,Azure.

Additional Responsibilities & Initiatives :

- **Developed** and **deployed REST APIs** (**Node.js,FastAPI**) and **React.js** frontends for **internal tools**, built **Dockerized applications**, and implemented **CI/CD pipelines** using **Azure DevOps** ,Designed **Data warehouses**,**monitored apps** .
- Contributed to **big data** processing with **Databricks** and **PySpark**, led **data migrations** to Azure Data Lake and SQL, managed **database backups**, **encryption**, **monitoring**, and conducted **internal training** sessions on Python and Pandas.

AtkinsRealis , Bangalore Jan 2023 - May 2023

Software Development Intern

- Built a **document management system** using **Python** and **React.js** to centralize file access, **implement role-based permissions**, and streamline collaboration across teams, **boosting productivity by 85%.**
- Designed a **responsive frontend UI** for **document upload and retrieval**, reducing search time by 70%, and implemented **efficient backend logic** for secure, organized file handling.
- *Tech Stack :* Python, React.js,Redis,SQL.

Skills

Languages: Python, C++,Java,C,HTML,CSS,JavaScript,TypeScript,SQL .

Technologies & Tools: SQL, PySpark ,Microsoft Azure , Azure Databricks , Azure Data Factory , Azure Synapse Analytics , Numpy , Pandas, Git , Github ,React.js , Redux, MongoDB,Express , Node.js , Next Js,PostgreSQL, Prisma, Docker, TailWind CSS , Recoil , Context API,Redis,WebRTC,PubSubs,WebSockets,CI/CD,Prometheus, Message Queue, Grafana,Kubernetes ,Kafka, Grpc,Jest , Data Structures and Algorithms , Styled Components , SASS,AWS.

Project Work

• **AI Podcast Clipper SAAS :** Live Source Code Demo

Built a full-stack AI-powered SaaS platform that automatically converts long-form podcast videos into short, viral clips using LLMs (Gemini 2.5), WhisperX, computer vision, and serverless GPU compute (Modal)—optimized for YouTube Shorts, Reels, and TikTok. Implemented a scalable video pipeline with speaker face tracking (LR-ASD), OpenCV-based cropping, FFMPEG recomposition, and integrated Stripe credits for monetization; deployed using FastAPI, Next.js, AWS S3, and Inngest queues.

Tech Stack : **TypeScript,Next.js,Stripe,Python,WhisperX, Gemini 2.5 API, LR-ASD, OpenCV, FFMPEG, AWS S3, Modal.**

• **FootBall Stadiums Data Engineering Pipeline and Visualization :** Source Code Demo

Built an end-to-end data engineering pipeline that scrapes data from Wikipedia on the Top 300 Football Stadiums, cleans and transforms it using Apache Airflow, and stores the processed data in Azure Data Lake. Further data processing and orchestration are handled through Azure Data Factory and Azure Synapse, with insightful visualizations delivered via PowerBI.

Tech Stack : **Python,Airflow,Azure Data Factory,Azure Synapse,Docker,Azure Data Lake,Azure SQL, Postgres.**

• **Ride Booking Application :** Source Code Demo

Built a full-stack ride booking application using the MERN stack with secure, role-based authentication for Users and Captains, OTP-based ride verification, and token blacklisting for logout security. Implemented real-time communication with Socket.io for ride requests and status updates, and integrated the Maps API for location suggestions and live GPS tracking with auto-refresh. Enhanced user experience with GSAP animations and containerized the app using Docker for consistent deployment.

Tech Stack : **React.js,Node.js, MongoDB,Express.js, GSAP , Maps API,Docker.**

• **Zapier – Automation Platform :** Source Code Demo

Built a custom Zapier-style automation platform that allows users to create workflows, trigger actions via webhooks, and automate tasks like sending emails or interacting with Solana. Implemented the Transactional Outbox Pattern to ensure atomicity and data consistency across services using PostgreSQL and Apache Kafka. Developed background workers for reliable task execution with automatic retries, ensuring no event loss or duplication. The system was architected with Next.js, Node.js, Prisma, and Docker, delivering a scalable, fault-tolerant, event-driven platform for workflow automation.

Tech Stack : **Next.js,Node.js,Apache Kafka,Postgres,Docker.**

• **AI Website Builder :** Live Source Code Demo

Built an AI-powered website builder that generates fully customizable websites from natural language prompts using the Claude API. Integrated Web Containers to simulate a live development environment in the browser, offering real-time previews and downloadable code. Developed a responsive, interactive frontend with React.js and TypeScript, enhancing user experience and reducing runtime errors.

Tech Stack : **React.js, Express.js, Node.js, Artificial Intelligence,Web Containers ,TypeScript ,Claude API.**

• **Imagify - AI Powered Image Generator :** Live Source Code Demo

Developed Imagify, an AI-powered image generation platform that transforms user prompts into high-quality, customizable images using the Clipdrop API. Implemented real-time previews, fast image generation, and efficient handling of concurrent requests for a seamless user experience. Integrated user authentication and Razorpay for secure payments, and built a responsive React.js UI for accessibility across devices. The platform was optimized end-to-end for performance, scalability.

Tech Stack : **React.js, Express.js, Node.js, Artificial Intelligence,MongoDB,TypeScript ,ClipDrop API,Cloudinary.**

• **Real Estate App :** Source Code Demo

Designed and built a Real Estate application with secure user authentication, allowing users to browse listings, upload properties, save favorites, and manage profiles. Integrated a messaging system for direct communication between users and property owners, boosting engagement and interaction across the platform.

Tech Stack : **React.js, Express.js, Node.js, Zustand,Docker,Socket.io,MongoDB.**

• **SQL Data Warehouse :** Source Code

Designed and implemented a modern data warehouse using SQL Server, following the Medallion Architecture (Bronze, Silver, Gold). Built ETL pipelines to extract, transform, and load data from ERP and CRM sources, enabling analytical reporting and business insights. Developed optimized fact and dimension tables for efficient querying and reporting .

Tech Stack : **SQL Server, SSMS, SQL, Medallion Architecture, Star Schema, Fact & Dimension Tables.**

Education

MSRUAS Bangalore Aug 2019 - June 2023

B.Tech. in Computer Science and Engineering ***CGPA:9.5/10***