

# Abhishek Singh

Email: [abhishek.singhbit11@gmail.com](mailto:abhishek.singhbit11@gmail.com), Mob: 9513142884

GitHub: - [mike2k11 \(github.com\)](https://github.com/mike2k11) LinkedIn: - [www.linkedin.com/in/abhishek-singh10](https://www.linkedin.com/in/abhishek-singh10)

Azure DP-900 Certified - [Microsoft Certified: Azure Data Fundamentals - Credly](#)

**To be a well-versed techie and to fulfill my responsibilities through sheer commitment and hard work that helps in my development as well as organization through perpetual learning.**

An Ambitious Problem Solver and Result-oriented API **Developer, Data Engineer and Gen-AI Enthusiast** with considerable training and enterprise experience in API design, Data Engineering and Machine Learning Models. Layout in all states of design, coding, and testing of the web services, and in the formation, planning, and implementation of the online business ventures. Working gives me a lot of skills and experience that I believe make me more confident. Analytical thinking and Problem-solving attitude. Always try to Learn and Update myself with latest skills and Technology.

EDUCATION QUALIFICATION			
Degree	Percentage	Institute/University	Year Of Passing
<b>B.E (ELECTRONIC &amp; COMMUNICATIONS)</b>	7 CGPA (Degree)	BIRLA INSTITUTE OF TECHNOLOGY, MESRA	May-2015
<b>H.S.C. (CBSE)</b>	82.20%	D.A.V Public School	May-2010
<b>S.S.C. (CBSE)</b>	86.60%	St Xavier School	May-2008

## Professional Summary And Technical Skills

I have a total **8 years** of professional IT experience starting carrier in **Infosys, PwC and currently in EPAM Systems**. where I have worked on Multiple stacks right from **Python, DataStructures (Stack,Queues,LinkedList), Flask, SQLAlchemy ORM, Rest API, PySpark, Databricks, Spark, Airflow, Azure ADLS, Azure Openai Api, GenAI Deeplake VectorStore langchain library, Web Scraping Library Scrapy and BeautifullSoup, Wrangling/Processing/visualization, Machine learning Algorithm, Git, Linux Environment, MS-SQL and SAP.**

I have an understanding of **Docker, AWS Lambda, Glue, S3, Step Functions, Azure Blob Storage Scikit Learn Library, NLP, Spacy and other Libraries , Jenkins CI/CD, Gensim-TFIDF, Sentiment Analysis, Word Embedding, Deep learning(RNNLSTM), HTML, Excel, Networking, SAP ABAP, SAP-HCM (SF), ADF, and few services on AZURE.**

I enjoyed learning new things and implementing it in project.

I have been a part of Organizations Internal society working for Fun activities for the kids and Employees at different occasions as well as volunteering in CSR in Tree plantation as well as social awareness programs.

## WORK HISTORY

---

### EPAM Systems | Jan-2022 -Present | Senior Software Engineer

#### Project2: LeMod || CTCO || Backend Engineer and GenAI

This project was POC which eventually becomes project. We developed chatgpt like private application for client.

**Business objectives:** To deliver a private AI-powered Solution to assist with code analysis, understanding architecture, and generating user stories. The proposal suggests running a Proof of Concept (PoC) for one CTC MF Application to validate the efficiency of the proposed AI-powered knowledge base in analyzing PL/1 code bases. The results of the PoC will be compared with the manual analysis captured by the App team. POC helps to figure out whether Gen AI can efficiently analyze data sources such as video recordings, images, HTML pages.

**Technology used:** Python, SQLAlchemy, Flask, Azure OPENAI API key, OpenAI Embedding, Deep lake Vector store, Redis search, Azure Blob storage, psycopg2.

---

### EPAM Systems

#### Project1: Phoenix || Shell || Data Engineer

Development of a OIM(Open Items) Report on Databricks NOTEBOOK using PySpark. The report was supposed to fetch the delta tables from ADLS mounted on it. Once the tables are fetched, we apply filters based on business needs. For eg: Filter based on year, excluding unnecessary columns, renaming of columns so that it could be used to join with other tables etc. we used around 8 tables from FDH and 5 tables from workflow table for all the 9 ERPs(SERP etc). Once the data is processed on Spark Cluster we are supposed to load the data on SQL server spark connection. PBI team are now informed to use the table from SQL table and then Report creation which is the end result of this whole process.

**Technology used:** PySpark, Databricks Notebook(spark cluster), ADLS, Delta Tables ADF, MS-SQL.

---

### PwC Kolkata February 2020 – Dec 2021

#### Senior Associate

#### Project1: GDH || Client: PwC Global || Developer

Developed a Rest API using Flask Framework for Data consumption from MS-SQL database. Created Multiple Python Script solution for data Processing. Understood the Business Part and Performed troubleshooting by changing the existing Code and adding new functionality written in Django Framework. I was Fully responsible for the deliverables. I was also involved Month Cycle activities with Data Engineering Team where our Task was to Run SSIS Packages written in SQL to consume the data file Upload by 143 Territories and Inserting them into Corresponding Tables in Different databases by transforming them using ETL tool such SSIS.

**Technology used:** Python, OOPs Concept, Flask, Multithreading, SSIS ETL Tool.

### **Project Description2: CIF, Cloud Infra || Data Engineer**

We had consumed the data from s3 which was coming from multiple sources. Created ETL Pipeline using AWS Glue and after processing needs to load in multiple target like RDS, Redshift.

Docker Compose file creation for the webservices deployment and managing them, working on AWS services. **Technology used: Python, AWS Services, Lambda, Glue, Redshift, DynamoDB, EMR, Kafka, Docker, Kubernetes.**

### **Power BI Report creation for Bruhat Bengaluru Mahanagara Palike || Client: BBMP**

Developed Multiple KPIs for the BBMP Bengaluru. BBMP wanted us to create Dashboard about all the People who were coming into Bangalore Post Lockdown and taking Quarantine in Public and Private Hotels. For Example Person Name, where they are staying (Facility Name), Total Quarantined people, Total Discharged, Covid Positive etc Datewise. Created Reports on PowerBI. Consumption of data from Multiple Sources as CDS, SQL etc and Created Dashboard and Published **Tools and Technology used:** Microsoft PowerBI, SQL, Excel CDS PowerApps.

---

### **Infosys Pune Sept'2015– Feb 2020**

#### **Technology Analyst**

#### **Project#1: – MFGADM || Python Backend Engineer**

Development of web applications using Flask Framework and other Backend tools and libraries. Consumption of Data and other resources was the main the main goal for our client and we had developed multiple endpoints for the same.

**Tools and Technology used:** Python, Flask, SQL Alchemy, Alembic, Jupyter Notebook, SQL, pyodbc.

#### **Project#2: – MFGADM || Client: Cisco MFGADM || Ticket Sentiments Analysis || CoE**

#### **Data Engineer and ML.**

Our Project was a part of the Journey Automation for Identifying Potential Customers and actual customers for a particular Hardware and Software products and sending them the marketing emails. The content of each emails depends on the product and common customer details. For all these works we were sourcing the data for each product campaign, Creating and Running Python Scripts for different Campaigns, filtering out Customer data based on Business Rules. QA Automation. Once the data is sourced and processed in Hadoop, the data is activated and deployed on a marketing tool names Eloqua through which Scheduled Marketing emails are sent to the customers.

Worked on Multiple Projects where I developed Machine Learning Models based Data wrangling, processing, analysis and data visualization. Creating Predictive Model using Machine Learning Algorithm for e.g:- Linear regression, Logistic, Decision Tree, KMeans, KNN classifier, PCA etc. Hyper-parameter Tuning Using Grid Search and Random Search, Confusion matrix, Cross Validation

**Tools and Technology used:** Python ML libraries, Pandas, Numpy, Scikit-learn, Nltk etc, Jupyter Notebook, SQL, Spark Hadoop, Airflow, kafka.

### **IoT and WebScraping || Client: Infosys POC Project**

In this Team I worked on Multiple PoCs based on IOT, web scraping etc. Successfully Connected Raspberry pi with Heat sensor, RFID Module along with other sensors. Collected data for processing. Designed and implemented a framework for web scraping.

**Technology used:** PowerBI, Azure Machine Learning.

### **Graduate Trainee, Sept15-Mar16 Infosys Ltd Mysore**

Training braced my skills both professionally as well as technically. To work as an engineer in Infosys, one has to clear the tests taken by Infosys and later a band is provided according to the cumulative performance in all tests. I was given HPF band (for 91%) in their training.

**Technology used:** Python, SQL, SAP ABAP, SAP HCM(SF).

Hobbies:

Swimming, Badminton, Running, Pool, Trekking, Dancing.    Personal

Details: Abhishek Singh (D.O.B – 10<sup>th</sup> June 1993)

Language – Hindi, English