

RESUME

Pankaj Kumar Patra

Berhampur, Odisha, India 

+91 7008853870 | +91 8280626637 

pankajpatra1998@gmail.com 

[linkedin.com/in/pankaj-kumar-patra-363754151](https://www.linkedin.com/in/pankaj-kumar-patra-363754151) 

<https://pankajkumarpatra.github.io/> 

My goals are to accomplish a sound position in an organization and work enthusiastically in a team, contributing to the development of self and the organization.

Professional Summary

- Setting up Cloud environment and working with Azure Databricks.
- Perform data integration and manage ETL workflows using Informatica and Azure Data Factory.
- Manage and optimize Data Warehouse, Data Lake, and Data Marts.
- Implement Data Governance and Automate tasks for daily scheduled jobs using PySpark.
- Understand Multi-Hybrid cloud infrastructure, migrate on-premise applications to the cloud, maintain Data Integrity, availability, and confidentiality.
- Prepare, optimize, and deploy, data preparation processes
- Having good problem-solving skills with excellent interpersonal skills and good relationship qualities
- Good team member, positive attitude and result oriented, self-motivated, quick learner, willing to adapt to new challenges and new technologies

Work Experience

Data Engineer at Tata Consultancy Services Limited

Oct 2020 - Present

- With over 3 years of experience, I have experience in big data engineering and Microsoft Azure and worked on data pipeline using Azure Databricks.
- I have over 2 years of experience in designing REST APIs and working on Data Lake or Lakehouse projects and building scalable and metadata-driven data ingestion pipelines using CDC (Changed Data Capture) retention policy.
- As there was no other Python developer, I take the credit here for developing PySpark script on Databricks for 3 years to reduce manual data validation post all kind of ingestion process which reduces manual effort by 60 percent.
- Integrated Production data pipeline with Control-M (Job Scheduler) for 1 year to complete scheduling, orchestrating, and validating pipelines.

- Performed ETL operations using Azure Data Factory, Azure Logic Apps, Azure SQL database, Azure Data Lake Storage (ADLS) for different file formats like Delta Lake, Parquet, CSV, JSON.
- Working closely with Admin team for 1 year to create suitable databricks clusters for executing high-performance data processing for structured and unstructured data.
- Communicate effectively with vendors and Data customers to fine-tune Data Platform. Participated in Solutions design and Data modelling for application enhancements.
- Managed a 24 * 7 Support team and solve most of the ingestion and ETL related issues.
- Developed PySpark script to load data from Microsoft Azure to SQL Server Management Studio(SSMS). The script works 30 percent faster than previous framework.
- On daily basis, I interact and collaborate with multiple teams and various stakeholders to meet deadlines. to stay updated.

Data Analyst Intern at Aptus Data Labs

May 2018 - July 2018

- Text Mining - Transformation, Cleansing & term-document matrix extraction
 - Worked on Data Analysis projects - Collected data from various sources, performed ETL operations, analyze data using various models
 - Worked on POSTMAN API tool to build and Testing API
 - Prepared Python for intermediate level and learn the use of it in Big Data
 - Participated in Gartner meet - discussed on ways to evaluate Product Demand variability
-

Experience Profile

- Internship May 2018 to July 2018 – Data Science and Analyst Intern, Aptus Data Labs
 - Microsoft certified: Azure Fundamentals ([AZ - 900](#))
 - Microsoft certified: Azure Data Fundamentals ([DP - 900](#))
 - Microsoft certified: Power BI Data Analyst Associate ([PL - 300](#))
 - I am working on preparing myself for Data Engineering on Microsoft Azure (DP-203) certification
-

Academic Qualifications

- BTech in Computer Science – National Institute of Science and Technology (N.I.S.T.) with C.G.P.A. of 8.1

Co-Curricular Activities

- (July 2018 – Present) Working as a Student Guide for N.I.S.T. Counselling Service (N.C.S.), understanding the concept of counseling and the role of the student counselor, guiding students – understanding their issues, helping them find solutions, supporting them as a career mentor, motivating them to pursue excellence in their field of interest, and participating in various team-building activities.
 - (Oct 2018 – Present) Worked as an Active Club Member, Data Science Club, N.I.S.T., organized workshops on Python, Linear Regression, Logistic Regression, basics of Data Science, and Analytics.
-

Technical Skills Summary

Machine Language Skills	Python, Scala, SQL, MongoDB
Data Analysis	PySpark, Shell Scripting, Spark, Hadoop, REST API implementation
Web Development	JavaScript, HTML, SEO
Tools	Microsoft Office, Control – M, Azure Portal, Databricks, Azure Data Factory (ADF), Azure Logic Apps, Microsoft Power BI
Operating System	Windows, Linux

Personal Details

D.O.B.: 28th January, 1999
Nationality: Indian
Gender: Male
Languages Known: English, Hindi, Odia
Strength: Hard-working, Flexible, Eager to learn

Declaration

I hereby declare that all the above-mentioned information and the particulars furnished are true and unambiguous, to the best of my knowledge.

Mr. Pankaj Kumar Patra
Berhampur, Odisha, India
Ph.- +91 7008853870