# Vishal Paliwal

Noida, UP | (+91) 989 7901 332 | vishal.paliwal003@gmail.com | LinkedIn

## SUMMARY

**5+ years** experienced **Data Engineer**, proficient in desiging efficient data pipelines using **PySpark** and ensuring data quality. Skilled in **SQL Data Warehousing** and dedicated to driving data-driven decisions and effective problem solving.

## EDUCATION

**UIET, CSJM University**                                                                                                        **Kanpur, UP**

*B.Tech in Information & Technology (First Division)*                                          *Graduation Date: May 2017*

## TECHNICAL SKILLS

**Key Areas:** PySpark Development, ETL Development

**Frameworks:** Spark, Hadoop, Python, SQL Datawarehousing, RDBMS, Shell Scripting

**Version Control Systems:** Git, SVN

**Scheduling / FTP Tools:** Stonebranch, Control-M, MoveIT

**Databases:** Postgres, Netezza, Oracle, SQL Server, Hive, ADLS, Synapse

**Others:** SQL Tuning, Performance optmization

## WORK EXPERIENCE

**Harman Connected Services**                                                                              **Bangalore, Karnataka**

*Senior Data Engineer*                                                                                                *Jun 2022 - Present*

- Working in a Retail domain to create a stable architechure which can be effectively utilised to analyze the campaigns running for multiple merchants and to gather the accurate analytics from it.
- Collaborated with client to successfully migrate data system from traditional **Hadoop Hive Architecture** to **Spark Architecture**, resulting in **3 to 4 times faster data load** and processing speed, improving overall system efficiency.
- Proactively engaged with stakeholders to gather and prioritize client requirements, translating them into actionable changes and enhancements in the data pipelines for 5+ different retail application.

**Capgemini**                                                                                                        **Navi Mumabi, Maharashtra**

*Consultant*                                                                                                                *Dec 2017 - May 2022*

- Worked as a Big Data Engineer on Microsoft Azure Platform, responsible to develop design solutions in Databricks using PySpark pipelines.
- Migrated on-premises Netezza Data Warehouse data with over 150 tables and 200+ previously developed Datastage jobs into the Azure system, resulting in a more efficient and scalable data infrastructure.
- Streamlined ETL workflows by automating manual processes, resulting in a 30% reduction in processing time for data extraction, transformation, and loading.
- Implemented automated error monitoring system for Stonebranch and MoveIT jobs, reducing troubleshooting time by 50%
- Mentored 20+ developers from different teams to maintain the quality check while migrating the code from on-prem Datastage jobs to Spark jobs.

## AWARDS & CERTIFICATIONS

**HackerRank:** Golden Badge - 5 Star Coder in Python and SQL (check here)

## SOFT SKILLS

**Skills:** Expertise in Microsoft Office, with a focus on Excel

**Languages:** Fluent in English and Hindi

**Public Speaking:** Experience and ability to speak with passion and confidence in public speaking and stand-up comedy.