

## Ashish Nagdev

Email: [ashunagdev@gmail.com](mailto:ashunagdev@gmail.com)

Phone: +919975475153

### Profile Summary

#### Overview

- A Results-oriented **Big Data Lead/Architect** with **10+ years of practical experience**.
- Proven ability and effectiveness in **Design Data Platforms and Lead development of high quality code**.

#### Career Objective

- Work on a **diverse range of new technologies** in **challenging and dynamic environment**, Contribute to **development and innovation of organization**, upgrading my **skills and knowledge**.

#### Technical Knowledge

- **Proficient:** Scala, Python, Java, Spark, Kafka, DataBricks, Snowflake, AWS, GCP, Data Science tools.
- **Working Knowledge:** Cassandra, BigQuery, Redshift, NeuralNetwork, Scalding, Docker, Kubernetes.
- **Tools:** IntelliJ, Eclipse, SBT, Git, Jenkins, Teamcity, udeploy, Cloudera, Hue, Airflow, Jupyter, Rstudio.
- **Operating Systems:** Linux, Windows, Solaris, AIX.

#### Other

- Strong design and analytical skills.
- Designing **scalable, performant, resilient, cost-effective** product architecture.
- Experience in designing Data Platforms, Data Warehouse and Data Lake on Big Data Systems.
- Product demonstrations and training sessions on Machine Learning, Big Data, Spark, BDD, TDD.
- Have a good grasp of finance, network, retail, healthcare domain fundamentals.

### Employment History

- **Pythian (Nov 2021 – Present):** Lead Cloud Data Engineer – Building big data platforms on cloud.
- **TIAA/Nuveen Asset Management (Jan 2019 – Nov 2021):** Design, Lead Development for Multi Asset Solutions Portfolio & Risk Analytics for Nuveen Innovation team and Developed Asset Retention model for prevention of Funds Redemptions and Funds Recommendation Engine for cross sell.
- **Citi Corp Services India(July 2017 – Jan 2019):** Design and Development of Equities Trade Data Services Dataview Model for real time analytics, pre-post trade and one analytics etc.
- **DataMetica Solutions Private Limited (June 2016 – July 2017):** Worked as Senior BigData Engineer, Building recommendation engine for very popular US based retailer Kohls.
- **Persistent Systems Ltd (Jan 2013 – May 2016):** Worked as a Module Lead, designing and developing on IBM NPI product. Worked with a team which migrated a product from oracle to db2.

### Professional Project Summary

#### Employer: Pythian Pvt Ltd. (Nov 2021 – Present)

- **Overview:** Working as a Lead Cloud Data Engineer; Responsible for Designing and building Data platform architectures on Cloud for multiple clients; Practice Lead for DataOps, Data Quality practices.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design, Developed a real time Data acquisition platform using AWS MSK, Kafka Streaming, Kafka connect for UK based top travel client with 10 TB of data in real time.
  - Build a model to migrate data from SAP to Google Cloud using CDF, Bigquery, Composer for client. Building migration module from oracle to Cloud SQL using DMS and ora2pg tool.
  - Build Sales offering for Data Quality and Data Ops by ingesting data from all clouds AWS, GCP.

#### Employer: TIAA; Asset Retention and Cross Sell Fund Recommendations (Jan 2020 – Nov 2021)

- **Overview:** Build a propensity framework which can calculate the likelihood that Financial Advisor will redeem in next quarter from a particular fund Asset class using historical transactions, demographics, activities, fund's performance and statistical data. Developed Fund Recommender using associativity and collaborative filtering algorithms to prevent redemption and improve sales activities.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design, Developed model using Logistic Regression, Apriori and Collaborative filtering algorithms.
  - Developed end to end model by data collection from multiple sources, data cleaning, Data Analysis, model implementation, model deployment, output ingestion and presentation using python, pandas, numpy, scikit, tensorflow, keras, apriori, hive, SQL Server.
  - Developed Time Series sales forecasting model for Funds daily subscriptions data using SARIMA, LSTM.

#### Employer: TIAA; Multi-Asset Portfolio and Risk Analytics (Jan 2019 – Jan 2020)

- **Overview:** Design and Development of Multi Asset Solutions Portfolio & Risk Analytics for Nuveen and TIAA portfolios. Developed Python/R based Analytics framework using libraries like QuantMod, Rshiny, Keras, Pandas etc. using Bloomberg Port by Upload data and Download reports.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design and Lead Development for Multi Asset portfolio Solutions Analytics using Python, R, Pyspark, Bloomberg Port etc. and generating reports like Value-At-Risk, TrackingError factors, Characteristics, Attribution, Scenarios reports etc. for portfolio managers covering Nuveen's \$9 billion multi-asset funds.
  - Design and developed Data Lake for Analytics, Risk module using Python, Spark, hive/impala and AWS.
  - Participated and Won hackathon in generating Nuveen Sentiment Analysis using VADER and GDELT database data and developed framework for Positive, Negative, Neutral sentiment.

**Employer: CitiGroup; Data Fabric (July 2017 – Jan 2019)**

- **Overview:** Worked as a Data Technical Lead and design and developed real time streaming framework for trade data. Which helped to gain business for sending reports in real time that is T+0 interval. Streaming data on 26 nodes production cluster for 100 million/day, Handling deduplication in streaming data, doing analytics.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design HLD, LLD and Development from scratch for real time reporting of trade data to clients.
  - Real time Trade data joining with Reference data in streaming with seconds latency.
  - Developed framework using Java, Spark Streaming, Kafka, Impala, Hive, Cloudera.
  - Lead COE on various initiatives related to Kubernetes, Machine Learning.
  - Developed Automated testing framework using Cucumber, Team City, Udeploy and Autosys.

**Employer: DataMetica Solutions Pvt Ltd.; Recommendation Engine (June, 2016 – July, 2017)**

- **Overview:** Development and enhancement of Recommendation Engine for popular eCommerce vendor, It extracts valuable insights, generating intelligence provides recommendations of products from data based on algorithms like Contextual, View Based, Cart Based, Frequently Bought time series, Best Seller, New Arrivals etc. Developed an Informatica to Hadoop Migration tool using Neo4j, Spark, Sqoop, Groovy.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design and Develop various data science algorithms in Scala, Pig, Hbase, Scalding etc
  - Running Hadoop Jobs in Azkaban in Production and Dev cluster.
  - Working directly with Clients and with remote US team in Agile Scrum manner.
  - Used IntelliJ, Gradle, Ant, Maven tools to develop and build the code.
  - Individually build a framework for Informatica to Hadoop migration using Neo4j, Spark, Sqoop, AWS.

**Employer: Persistent Systems Ltd.; NPI 1.0, 1.1, 1.2 (May 2014 – May 2016)**

- **Overview:** Netflow is the network data passing through router. TNPM Next will capture this data at 300k flows per second, does analytics on it, capture alarms, alerts user and visualize dashboards, charts on top network utilization and streaming and capturing various attacks like DDOS.
- **Role & Responsibilities:** Following is brief on role in this project:
  - Design and Development of ongoing analytics module which does analytics like various kinds of aggregation of captured data at rate of 300k fps in technologies like scala, java, h2, akka.
  - LLD for Analytics and Storage Module matching Solution team requirements.
  - Streaming real time data to visualize charts in diagnostic view and to visualize DDOS attack.
  - Good hands-on on apache spark streaming, Kafka and apache calcite for streaming live data.
  - POC on h2 in memory store, column store db like Cassandra, kairosdb, parquet, off-heap storage.
  - Used tuning tools like Java Heap Analyzer, Jconsole to monitor speed, memory, GC utilization.

**Employer: Persistent Systems Ltd.; IBM TNPM 1.3.3, TNPM 1.4.0, 1.4.1 (Jan 2013 – Apr 2014)**

- **Overview:** TNPM 1.3.2 was supporting just Oracle as a backend. Our task was to support DB2 as a backend for TNPM by which we can save that oracle license cost and as DB2 is product of IBM. New functionalities in our product like Multiple collector on single host, 16 datachannel for load balancing, support to higher version of Oracle, DB2, Dash & Java.
- **Role & Responsibilities:** Following is brief on role in this project:
  - I was involved in initial analysis of required code changes. Then I worked on changes in Java and CPP code to support DB2 as backend. Also on module which deal with loading files to database, and stored procedure code.
  - Development work to support DB2 as backend in java, cpp, Worked on various enhancements.

**Accolades & Certifications**

- ☐ **Google Cloud Architect Certified.**
- ☐ **Google Cloud Data Engineer Certified.**
- ☐ **Databricks Data Engineer Certified.**
- ☐ **Snowflake Data Engineer Certified.**
- ☐ **"Innovator of the quarter", "Star Performer Award", "You made a Difference" awards** throughout 10 years with various companies.

**Educational Qualification**

Course	Board/University	Class Obtained	Marks	Month Year
BE (Computer Science)	Shivaji University, Kolhapur	First class with Distinction	67.25%	Jun 2012
H.S.C	Maharashtra State Board	First class with Distinction	74.25%	Mar 2008
S.S.C.	Maharashtra State Board	First class with Distinction	77.06%	Mar 2006

**Personal Details**

Date of birth : 10<sup>th</sup> April, 1991  
 Sex : Male  
 Nationality : Indian  
 Marital status : Single  
 Language known : English, Hindi, Sindhi, Marathi

I hereby declare that all the information furnished above is true to the best of my knowledge and belief.

**Date:**

Ashish Nagdev