

Contact

vivek.nayeeni@gmail.com

www.linkedin.com/in/vivek-nayeeni
(LinkedIn)

Top Skills

Generative AI

LangChain

OpenAI API

Languages

English (Native or Bilingual)

Hindi

Telugu

Certifications

Big Data Hadoop Foundations

AMCAT Certified Business Analyst

Hadoop Data Access - Level 1

Talend Big Data V6 Certified Developer

Databricks Certified Associate Developer for Apache Spark 3.0

Vivekananda Nayeeni

Data Engineering Lead at LexisNexis

Raleigh-Durham-Chapel Hill Area

Summary

- Certified Cloud Big Data Engineer with a demonstrated history of working in the health care and Financial industry.
- Architected and implemented Data Lakehouse using Delta Lake and Apache Iceberg
- Experienced in building Big Data Lakes, Data Pipelines in Hadoop and AWS Cloud environments, and data migrations using various Big Data tools and custom frameworks.
- Highly skilled in AWS (EMR, Glue, Redshift, Athena, Cloud Formation, Lambda, Kinesis, DynamoDB), Big Data, Hadoop, Spark, Python, Airflow, Snowflake, Databricks, Hive, Presto, Terraform, SQL, Talend, Sqoop, Kafka, HBase, Unix.
- Experience working with Palantir Technologies and its proprietary Big Data tools such as Palantir Gotham, Foundry, Contour, Authoring, Slate, Mesa and Magritte.

Experience

LexisNexis

5 years 4 months

Data Engineering Lead

February 2025 - Present (1 year 1 month)

Raleigh-Durham-Chapel Hill Area

Senior Data Engineer

November 2020 - February 2025 (4 years 4 months)

Raleigh-Durham-Chapel Hill Area

Genworth

Application Development Manager (Data Engineer)

March 2020 - November 2020 (9 months)

Raleigh, North Carolina, United States

- Contributed to the development of a Unified enterprise data lake platform from scratch on AWS, designed to act as the central repository for critical data across various organizational teams, including Underwriting Data Science, Risk, EDMO, and Finance.
- Collaborated with various teams across the enterprise to collect data requirements essential for the platform construction including capacity planning.
- Leveraged a suite of services like Snowflake, EMR, Glue data catalog, Athena, and Apache Spark to architect and build the platform
- Implemented infrastructure configuration using Terraform.
- Utilized snowflake as the data serving layer, facilitating seamless integration with Tableau and Data Robot for advanced analytics.

Vanguard

AWS Big Data Developer

February 2019 - March 2020 (1 year 2 months)

Charlotte, North Carolina

- Migrated data from on-premise Cloudera to AWS Cloud for the Global Investment and Financial Division (GIFS), ensuring a seamless transition to a more scalable environment.
- Developed ETL pipelines to consolidate data from over 20 sources, including Charles River Development (CRD), Bloomberg, Yahoo Finance, MorningStar, OneTick, Forex rates, Blackrock etc.
- Extracted data from third-party external APIs, shared locations, and relational databases using Python, Attunity, and Sqoop. Transformed data with EMR, PySpark, and Hive, and conducted data analysis using Presto. Managed job scheduling with Control-M.
- Improved spark job performance, achieving a runtime reduction of over 50%, significantly improving processing speed.
- Designed visualizations in Tableau by integrating with Presto tables.
- Worked closely with the Transaction Cost Analysis data science team in analyzing market and trade data.

Molina Healthcare

Big Data Engineer

October 2016 - February 2019 (2 years 5 months)

Long Beach, California

Worked with Palantir Technologies in building member, provider, and claims related use cases using its proprietary Big Data tools such as Palantir Gotham, Foundry, Authoring, Slate, Mesa, Magritte, and Contour.

Built an Enterprise Data Platform (EDP) from scratch on Cloudera Hadoop distribution using Big Data tools such as Spark, Talend, Sqoop, Attunity, Hive, Impala, HBase, Python and Scala.

Also, Worked on several other projects such as:

- MOSAIC: Effectively identifying the Unable to Contact (UTC) patients across the organization;
- MEMBERSPHERE: Creating a Multidisciplinary view of patients medical and pharmaceutical care;
- SEMANTIC LAYER: Building the member, provider, claims, pharmacy, and authorizations 360-view data using tools such as Apache Spark for complex data processing.

Cognizant

Big Data Developer (Associate - Projects)

March 2016 - February 2019 (3 years)

UnitedHealth Group

Big Data / Talend ETL developer

March 2016 - October 2016 (8 months)

Greater Boston Area

1. Worked for Military and Veterans (M&V) Project as a Big Data developer for TRICARE 2017 contract.
2. Used MAPR distribution to build the Hadoop Big Data Lake.
3. Built Ingestion, Enrichment and Provisioning engines in Data Lake to ingest the data, curate the data and export it to various downstream applications using Talend, Sqoop, Hive, HBase, Pig, Spark, Scala.
4. Written Hive and Pig scripts to perform complex transformations.

West Virginia University

Graduate Teaching Assistant

August 2014 - December 2015 (1 year 5 months)

Morgantown, West Virginia

- Worked as a Graduate Teaching Assistant (GTA) at the Freshman Engineering campus of West Virginia University.
 - Mentored around 400+ students on-campus in Big Data, MATLAB, Data Science, and Statistics.
-

Education

West Virginia University

Masters, Computer Science and Electrical Engineering (GPA:
4.0/4.0) · (2014 - 2015)

CVR College of Engineering, Hyderabad

Bachelors, Engineering · (2010 - 2014)

Sri Chaitanya Junior College