

NITESH PANDEY

AWS | | PySpark | | Machine Learning Engineer

CONTACT: 9555195182

EMAIL: 25decniteshpandey@gmail.com

PROFILE

Passionate AWS | | PySpark | | Machine Learning Engineer with 3+ years of experience in quality data management and accuracy by leveraging Spark, Python, AWS, Machine Learning. Skilled in Continuous Improvement, Agile and Business Process Improvement.

Self-driven, highly motivated individual who loves data puzzles and problems around it.

GITHUB:

<https://github.com/nitesh251>

SKILLS

Programming Languages: Python

Technologies: Spark

Cloud: AWS (Glue, Lambda, step function, DynamoDB, CloudWatch, SNS, S3, Athena, API gateway, IAM etc.)

Database: SQL, NoSQL (DynamoDB)

Python libraries: NumPy, Pandas

Tools: Jira, Bitbucket, PyCharm

AWARDS AND ACHIEVEMENTS

2019: Spot Winner- MLEU (Cognizant)

2018: Gold Medal – B-tech in Computer Science

2017: College Internship Coordinator

2015: First Prize - Inter college script writing competition.

2011: District topper in class 10th in CBSE

HOBBIES

Reading and Writing stories, poems
Blogger at
itsniteshpandey.blogspot.com

EDUCATION

Galgotias University

2014 - 2018

B. Tech in computer science & engineering **8.55**

Foundation School

2011 - 2013

Senior Secondary from CBSE board **84.6%**.

Sri Saty Sai School, Rishikesh

2010-2011

Higher Secondary from CBSE board **95%**

WORK EXPERIENCE

Sunlife Financial - Analyst Development

Aug 20*

Project 1 :- IFRS17

Developing a data pipeline that will incorporate the transformation logic as per the latest standard using AWS Services - Glue, S3, SNS, Lambda, Redshift, PySpark, Sql, Python.

- Developing the PySpark transformation script to migrate data from s3.
- Developing Glue job to transform and load data into RDS.
- Created a Lambda to trigger glue jobs.

Project 2 :- HR Resume screening And Ranking Model

Developing a Machine Learning Model to Select Resume and Rank them As per the JD Using Naive Bayes Classification Algorithm and Natural Language Processing (NLP), Python

Cognizant Technology Solutions – Associate Project

Mar 18 – Aug 20

Project 3 :- Smart Grid Integration System

Developed metadata driven data pipeline using aws services and Pyspark which was scalable, easy to extend, reusable and easily maintainable due to its generic nature.

- Developed the PySpark transformation script to manage and merge six Meter reads files through Glue and Lambda.
- Created a Python Script in Lambda to monitor the various Glue Jobs , query the DynamoDb table and send notification as required.
- Updated the already implemented Lambda function to execute the New Glue Jobs and create their metadata entry in DynamoDb.
- Developed a PySpark Script to upload data to DynamoDb after some transformations and execute an ECS task in Aws Glue.

Project 4:- Transformer Failure Prediction

Worked On A Machine Learning POC For Transformer Failure Prediction Using Various Machine Learning Classification Algorithm Like Support Vector (SVM), Random Forest, Python