



John Smith

Data Engineer

Profile

Dedicated Data Engineer with 5+ years' experience dealing with large datasets. Eager to build robust databases that lay the groundwork for game-changing insights at LionHeart Algorithm LLC. Implemented natural language processing tools to ensure machine-readable databases were ready for the team of data scientists.

Employment History

Data Engineer at FNB, Nong Phai

January 2018 — December 2020

Responsible for scaling machine learning models and making these models fit within banking environments with the sci-kit-learn, Tensorflow, and Keras, also adapting these models/architectures created by data scientists to fit in with the UX and the UI of the brand.

- Translated business propositions into quantitative queries and organized the necessary data.
- Developed scalable databases capable of ETL processes using SQL and Spark.
- Estimated the workflow and increase the efficiency of data pipelines that process over 60 TB of data daily.
- Utilized MongoDB to create NoSQL databases that collect data from a variety of sources.
- Involved in creating tables, join conditions, partitioning tables, correlated subqueries, views, sequences, nested queries, and synonyms for business application development.

Data Engineer at ABSA, Washington

January 2015 — December 2017

Responsible for developing database triggers, packages, functions, and stored procedures using PL/SQL and maintain the scripts for various data feeds across multiple regional and international offices of the company

- Co-develop a SQL server database system to maximize performance benefits for clientele.
- Assisted senior-level Data Scientists in the design of ETL processes, including SSIS packages.
- Developed coherent Logical Data Models that helped guide important client business decisions.
- Collaborate and coordinate with development teams to deploy data quality solutions and create and maintain standard operating procedure documentation.

Education

BS, Computer Science, Texas University, Austin

January 2021 — January 2021

- Dual concentration Machine Learning,
- Received a Business Foundations Certificate
- Part of the winning team for the 2016 Longhorn.

Details

1515 Pacific Ave
Los Angeles, CA 90291
United States
3868683442
email@email.com

Place of birth
San Antonio

Driving license
Full

Skills

SQL

Java

Apache Spark

Hadoop

Python

Coding

Languages

English

German

Hobbies

Cycling, Songwriting, Running

Courses

CCA Cloudera Certified Associate, Cloudera, Online.

January 2021 — January 2021

Achievements

- Developed a data pipeline with Delta Lake that led to process optimization and a corresponding revenue increase of 21%.
- Successfully figured out ETL issue while following PL/SQL best practices that resulted in an insight that increased the client's customer base by 37%.