DANG NHAT

DATA ENGINEER

12/04/1995

Go Vap District, HCMC

+84 0354883673

dangnhatsimon@gmail.com

Profile

LinkedIn

Education

Obninsk Institute for Nuclear Power Engineering
B.Sc in Nuclear Power Engineering and Thermal Physics (2014-2019).
GPA: 4.6/5.0

IBM Skills Network
Data Engineering (2021-2022).

DataCamp
Data Engineer (2021-2022).

Google Career
IT Automation with Python (2021-2022).

Google Career
Data Analytics (2021-2022).

Certificate

IBM - Big Data 101.
IBM - Hadoop 101.
DataCamp - Building Data Engineering Pipelines in Python.
freeCodeCamp - Data Analysis with Python.
Hackerrank - Problem Solving Basic.
Hackerrank - Python Basic.
Hackerrank - SQL Advanced.
Sololearn - Python Intermediate.
Sololearn - SQL Intermediate.

Working Experience

Apr 2024 – Present

Back End Engineer
Definer Inc.
-Definer designs and develops solutions, builds DevOps, automates cloud IT infrastructure construction, and builds analytical platforms for big data. -Using Prisma for interfacing to interact with databases using a type-safe and auto-generated query builder. Leveraging psycopg2 to create OOP and executes query to enhance and report to customers.
-Creating task definition on AWS, building configuration file CircleCI, docker compose file.
-Run jobs on cloud: Google Cloud, AWS.
- Build config CircleCI to automate the build, test, and deployment processes of application.
- Docker is utilized to containerize application components, package application and its dependencies into containers, which can then be deployed consistently across different environments.
-Using REST API with Flask to build website.
-Core technologies: Python, Flask, GCP, AWS, CircleCI, DevOps, Prisma ORM, Docker.

Oct 2022 – Present

Data Engineer
Power Engineering Consulting Joint Stock Company 2 (PECC2)
-The joint stock consultancy company in electricity energy industry. Data from any monitoring system are ingested to database server at Company from invested power plants. Processing data with 1GB/day, ensure available data for Data Scientist and ML Teams.
-Contribute to Big Data Development Plan for Company. Writing automation script with Python and GUI, shell scripting on Linux.
-Contributing to interactive reports sent streaming to the end users and improve customer experience.
-Core technologies: Big Data Analytics, ETL, Data Engineering, Python, Shell, Git, Linux.

Oct 2019 – Oct 2022

Mechanical Piping Engineer
BIM Manager and Mechanical Lead. Thermal-Mechanical and piping engineering.

Study Project

IBM Data Engineering

-Web scraping and using APIs to extract data with BeautifulSoup, Selenium, Scrapy.
-Creating, designing, & managing relational databases & applying database administration to MySQL, PostgreSQL, & IBM Db2. Composing advanced SQL techniques like views, transactions, stored procedures and joins. Monitoring and optimizing database performance.
-Developing shell scripts using Linux commands, environment variables, pipes, and filters. Schedule cron jobs.
-Developing working knowledge of NoSQL & Big Data using MongoDB, Cassandra, Cloudant, Hadoop, Apache Spark, Spark SQL, Spark ML, and Spark Streaming.
-Implementing ETL & Data Pipelines with Bash, Airflow & Kafka; architecting, populating, deploying Data Warehouses; creating BI reports.
-Core technologies: Data Science, ETL & Data Pipelines, RDBMS, NoSQL and Big Data, Apache Spark, Python Programming, Data Analysis, SQL, Shell Scripting.

py-frame

-Creating a Telegram Bot with python-telegram-bot that can do API call to coingecko and get real time price of coin.
-Creating search function for searching specific keywords across multiple files combined with simple GUI using tkinter.
-Core technologies: Python with frameworks.

DANG NHAT

DATA ENGINEER

Languages

- Vietnamese
- English
- Russian
- Chinese

Study Project

IBM DevOps Software

-Create and refine a product backlog using the sprint planning process. Create burndown charts to forecast the ability to meet a sprint goal. Assign and track stories using a kanban board. Incorporate Scrum artifacts, events, and benefits.
-Develop shell scripts using Linux commands. Schedule cron jobs. Collaborate on a team project with Git.
-Create and deploy an AI-based application onto a web server using IBM Watson AI Libraries and Flask.
-Build cloud native applications, deploy microservices using Docker, Kubernetes. Create and leverage a YAML file to configure and create resources.
-REST API endpoints and invoke them using cURL and Postman; Use SwaggerUI to document and test APIs. Improve unit testing through advanced test-driven development methods.
-Create IaC scripts using Terraform, automate CI/CD tasks using Jenkins and GitHub actions.
-Perform defensive coding following OWASP principles, troubleshooting using logging, stack trace, and log analytics, scanning and pen testing with Kali Linux. Monitoring and troubleshooting; and test monitoring with Prometheus and Grafana.
-Core technologies: Python, Flask, AI, CI/CD, Microservices, Scrum, Agile, REST, Serverless, Git, Monitoring, Shell Scripting, ETL, Kanban, Zenhub, Scrum, Agile, Kubernetes, DevOps, IaC, Shell Bash, OWASP.

DataCamp

-Ingesting data from external sources with many of different file types. SQL Database Design.
-Cleaning data, using record linkage technique to merge multiple datasets together, used when values have typos or different spellings. Pandas optimizations. Streamlined data ingestion with pandas and data visualization.
-Resilent Distributed Datasets, using pyspark to create a data transformation pipeline. Spark SQL for Big Data Analytics.
-Testing data pipeline. Manage and Orchestrate Workflows with Airflow. Software Engineering Principles in Python.
-Core technologies: Python, Spark, ETL, Airflow, SQL, Git.

Big Data with Spark

-From cleaning data to creating features and implementing machine learning models, execute end-to-end workflows with Spark: Linear Regression, Logistic Regression/Classifiers, creating pipelines.
-Data wrangling and feature engineering with PySpark. Building recommendation engines using Alternating Least Squares in PySpark.
-Core technologies: Big Data, ETL & Data Pipelines, Apache Spark, SparkSQL, SparkMLlib, Python Programming, Java Programming, SQL, Shell Scripting.

Skills

Domain	Data Structures & Algorithms, Probability & Statistics.
Programming	Python, R, Java, LaTeX.
Framework	Flask, Django, Spring.
Scripting	Bash, Shell with Linux, Git.
Database	SQL(MySQL, PostgreSQL), NoSQL(MongoDB, Cassandra).
Big Data	Hadoop, Spark, Kafka.
Cloud	GCP, Azure, IBM.
Visualization	Tableau, Power BI.
Soft Skills	Detail-oriented, Structured & Analytical Thinking, Problem Solving, Teamwork, Diligent, Energentic, Resilient.