Dagmawi Yohannes

Data Engineer

A data engineer and data science enthusiast with experience in using relational and non-relational databases, as well as Java, C++, and Python. Practical knowledge on data engineering tools like Kafka and Airflow, as well as the ability to build and manage SQL- based, high-performance, scalable, and reliable end-to-end data pipelines.

✉ dagmawi.yoha03@gmail.com

📍 Addis Ababa, Ethiopia

in linkedin.com/in/dagmawi-yohannes-b36488236

github.com/dagmawiii03

📱 +251-942-18-4341

sites.google.com/view/dagmawi-yohannes/home

🐦 @dag_mawi111

medium.com/@dagmawi.yoha03

WORK EXPERIENCE

Data Analyst
ATYM Engineering

05/2021 - 01/2022   *Addis Ababa, Ethiopia*
IT service and consulting company

Achievements/Tasks

- uncover trends, find patterns, test hypotheses, and solve issues, an extensive exploratory data analysis is performed by cleaning and utilizing an existing data collection.
- developing the parameters for a machine-learning algorithm to help it determine which categories specific data points belong in while working with various classification projects
- Working on a few regression tasks and examining the patterns in the datasets that have been collected

Front-end developer
Dejen Technologies

01/2021 - 03/2021   *Addis Ababa, Ethiopia*
IT service and consulting company

Achievements/Tasks

- Front-end development using Html /CSS and React
- Social Media marketer

EDUCATION

Data, ML & Web 3 Engineering Training
10 Academy

05/2022 - 08/2022   *Santa Clara, CA*

Courses

- Data cleaning
- Kafka streaming
- Dbt transformation
- Exploratory Data Analysis
- Time Series Analysis
- Big data techniques
- Spark transformation and Airflow orchestration
- Tech stack building and change automation
- Statistical and predictive Modeling
- Casual inference and Deep learining

BSc in Electrical and Computer Engineering
Addis Ababa University

09/2015 - 10/2021   *Addis Ababa, Ethiopia*

Courses

- Object Oriented Programming(java).
- Computer Archietecture and Organiztion.
- probablity and Random Process.
- C++.
- Applied Mathematics I, II & III.
- Computational Methods.

SKILLS

python · pandas · SQL · AWS · Looker · Snowflake · Kafka · Airflow · Spark · Dbt · Postgresql · Tensorflow · Pytorch · scikit-learn · MLflow · streamlit · Tabelau · Redash · superset

PERSONAL PROJECTS

Speech-to-text data collection with Kafka, Airflow, & Spark ⧉
- Using Apache kafka to stream data as soon as a user uploads the audio data captured and retain the data.
- Using Apache Spark to make data transformation and load it to S3 bucket and Apache airflow to orchestrate the tasks.

Pharmaceutical sales prediction ⧉
- Building model with deep learning of the Long Short Term Memory which is a type of Recurrent Neural Network.
- Serving predictions using Streamlit platform.

A/B Hypothesis Testing ⧉
- MLOps planning and set up a DVC and MLflow.
- A machine learning model using 5-fold cross validation using 3 different algorithms: Logistic Regression , Decision Trees, XGBoost, RandomForest.
- Define the loss function for the model using the validation data.

CERTIFICATES

Python and SQL (sololearn) (04/2021) ⧉
certificate of completion for courses given in python and SQL

certificate of completion (10 Academy) (08/2022) ⧉
Certificate of Completion for mastery of essential topics in Machine Learning, Data Engineering and Web3 Engineering by Distniction

Project Management (04/2021)
from Mated Management Institute

Certificate of Participation(03/2022)
Enterpreneurship and Employability skill trainging from center of African Studies

LANGUAGES

English
Full Professional Proficiency

Amharic
Full Professional Proficiency