

VIKTOR ANDRIICHUK

Summary of Qualifications

3+ years of IT development (data engineering + python backend developer + machine learning engineering),
Experience in creating ETL processes (AirFlow or pure Python; PySpark; AWS Glue, AWS Kinesis),
Experience in creating Data Lakes (AWS S3, Google BigQuery) and Datawarehouses (AWS RedShift),
Experience in developing web applications with Flask and FastAPI frameworks,
Data acquisition experience with API and parsers,
Knowledge and experience in Data Science (NLP, Recommendation Systems),
1+ years in Java Core, Scala (for understanding Spark's RDD etc.),
Understanding async in Python.

Technical Skills

Languages: Python, Java Core, Scala

Web frameworks: Flask, FastAPI

NoSQL: MongoDB, Redis

SQL: PostgreSQL, Greenplum, RedShift

Env: Docker

OS: Linux, OS X

Clouds: AWS, GCP

SVN: git

OS: Windows, macOS

Big Data: Spark, Airflow, HDFS, AWS Glue, AWS Kinesis

Professional Experience

Provectus

2021 – Present

Position: Data Engineer

Project: Appen

Responsibilities included:

Developing Python scripts for AWS AirFlow and SQL queries for RedShift DB

Creating ETL processes for DWH

AWS

Environment: Apache Airflow, RedShift

Project: NDA

Responsibilities included:

Developing reverse proxy microservice with Flask and Nginx.

Environment: Flask

Project: NDA

Responsibilities included:

VIKTOR ANDRIICHUK

Create ETL processes for Data Lakes

Environment: AWS S3, AWS Glue, AWS Kinesis

BSG.world (telecommunication)

2019-2021

Position: Python Developer, Data Engineer, Machine Learning Developer

Project: AntiSpam Machine

Responsibilities included:

Created NLP models

Created AntiSpam microservice with Flask, MongoDB, Redis and Machine learning algorithms

Created admin microservice for control AntiSpam microservice with Grafana, Prometheus

Environment: Flask, PostgreSQL, MongoDB, Redis, Docker, ML

Project: Recommendation System

Responsibilities included:

Created recommendation system mechanism, ML development

Created microservice with Flask

Collect data from users behaviors and than load to Data Lake

Environment: Docker, Flask, PostgreSQL, Redis

Project: DataLake and DWH

Responsibilities included:

Responsible for building scalable distributed data solutions using Spark

Developed Spark scripts by using Python

Experienced in performance tuning of Spark Applications

Experienced in handling large datasets using Partitions

ETL for save data from e-shops to Data Lake in real-time

Environment: Apache Spark, Apache HDFS, AWS EMR

Education

Taras Shevchenko National University of Kyiv (Bachelor)

2002

Faculty: Finance

National Technical University of Ukraine (Master)

2021

Faculty: Applied Mathematic

Also I like to run different pet-projects and learn new technologies.