

Viktor Andriichuk
Data Engineer with hands-on experience in Python Backend Development /
ML / Data Analyst / DevOps
4 years of experience

Contacts:

Personal web-site: <https://vandriichuk.com/data-engineer/>
GitHub: <https://github.com/vvandriichuk>
LinkedIn: <www.linkedin.com/in/viktor-andriichuk-data-expert>

E-mail: v.andriichuk@gmail.com
Telegram: @vandriichuk

About me:

I worked a lot with different types of data - with "regular" and with what can be called "big data" (I did a lot of things with my hands in Data Engineering, ML, analytics and backend) using Python. I used to strive to learn tools for the sake of the tools themselves. And since 2023, I realised it's not about what tools I use but what value I bring to the business. I realised the main thing is the goal and task; the tools are secondary. I also realised that I can not only "do tasks" - I can also bring additional value to a business in the form of creating effective solutions for finding insights for business growth from start to finish (from collecting data from different sources, proper storage, creation of necessary pipelines for transforming data, finding information in data, or creating dashboards). I help businesses grow by making data-driven decisions and implementing data-driven solutions, from design to implementation.

I have invested and continue to invest a lot of time and money in studying the fundamental principles of processing and storing data and algorithms, allowing me to look at problems in a broader, more profound way and with greater understanding.

I prefer to work in small teams - startups or growing start-up businesses. If necessary, I can be the only developer on the team. I wouldn't say I like bureaucracy. This is my element if it is required to complete tasks from different directions. At the same time, I can be a team player. I have had mentoring experience.

Technical skills:

- Programming Languages: Python (primary language), Java (secondary language)
- Frameworks: Flask, FastAPI
- Databases: PostgreSQL, Redis, MongoDB, Snowflake, RedShift, Athena, Aurora
- Cloud Providers: AWS (RedShift, Lambda, Glue, Athena, Quicksight, S3, Aurora, StepFunctions)
- Big Data: Hadoop, PySpark
- ETL tools: dbt, AirFlow
- Data Science knowledge and tools: Classic ML and DL algorithms, pandas, numpy
- Delivery: CI/CD
- Other tools: docker, celery, pytest

Professional Experience:

Project:	Consulting Services for Fabiosa Media - the top #4 Facebook media publisher in the world
Period:	10.2023 - 12.2023 (Short-Time Contract)
Role:	Data Engineer
Achievement, Pride	<ul style="list-style-type: none">• I worked closely with Data Scientists and Data Analytics to gather requirements, provide technical guidance, and ensure the successful delivery of data engineering projects.• I designed all pipelines, from collecting data from Facebook API to storing it in the AWS S3 from scratch.• I created a platform where businesses can find insights into their data and build dashboards and recommendations from Gen AI. It helps to decrease the time for making decisions.
Responsibilities included:	<ul style="list-style-type: none">• Solution Design• ETL• Building Insights Visualisation Platform
Environment:	Data Science, Data Analyse, AWS AirFlow, AWS S3, DuckDB, pandas, boto3, Python, Docker, Streamlit, LLM, OpenAI API, Nginx, CI/CD by GitLab, Tableau
Project:	Deloitte V360 – a big analytics platform for one of the biggest consultant groups in the world

Period: 02.2023 – 10.2023
Role: Big Data Engineer, Machine Learning Expert
Achievement, Pride

- Using the fuzzy matching method, I created a system that found similar companies based on their characteristics to eliminate duplicates and tidy up the client base of a massive company such as Deloitte.
- I proposed and implemented a blue-green deployment strategy in AWS, eliminating site downtimes during data updates. This ensured uninterrupted user access, showcasing innovation and user-centric focus on the deployment process.
- I enhanced data handling efficiency by accelerating the loading time of raw data into Spark DataFrames by 4X by converting CSV files into Parquet format, thereby improving the overall data processing speed.
- I successfully decreased the size of data stored in S3 by 8X, optimising storage usage and cost and demonstrating keen resource management skills and cost-consciousness.

Responsibilities included:

- Created jobs for AWS Glue for data transformations
- Built a system for matching data by AWS Fuzzy Matching, SQL and Machine Learning algorithms

Environment: AWS Glue, Athena, Aurora, Step Functions, RedShift, OpenSearch, Python, Spark

Project: Clip Translator – a startup, pet project for automatising translation video to different languages, reached hundreds of users in the first month after launch
Period: 04.2023 – Current
Role: Python Backend Software Engineer, ML engineer, DevOps
Environment: Python, Deep Learning, PostgreSQL, Flask and FastAPI, Celery, Docker, Vue, Nginx, CI/CD by GitLab, Alembic, pytest

Project: NDA – a network of sites for health care (from GlobalLogic)
Period: 11.2021 – 02.2023
Role: Data Engineer, Python Software Engineer
Responsibilities included:

- AWS AirFlow
- ETL processes for Snowflake and dbt
- Data Quality pipelines
- Extras for custom Python frameworks for data engineering.

Environment: Snowflake, dbt, Great Expectation, Python, FlaskAPI, pytest, numpy

Project: PepsiCo – Southern Africa, Australia branches (from Provectus)
Period: 03.2021 – 11.2021
Role: Data Engineer
Responsibilities included: DWH from scratch
Environment: Snowflake, dbt, Airflow, Python

Project: Appen – Quality AI training data for the AI lifecycle (from Provectus)
Period: 06.2020 – 03.2021
Role: Data Engineer
Responsibilities included: DWH and Data Marts from scratch
Environment: AWS Airflow, AWS RedShift, dbt, Great Expectation

Project: Analytcs Dashboards (from Provectus)
Period: 03.2020 – 06.2020
Role: Python Developer
Environment: AWS S3, AWS Glue, AWS Athena, AWS Quicksight, AWS S3, AWS Lambda Functions (Python for collect data and save to S3), Terraform

Project: BSG – automatic marketing campaigns' global company

Period: 01.2017 – 03.2020
Role: Python Backend Developer, Machine Learning Developer
Responsibilities included:

- AntiSpam Machine
- Data Lake and DWH
- Recommendation System
- Chat-bot

Environment: Python, FastAPI, MongoDB, PostgreSQL, Redis, Kafka, docker, ML

Education:

Institution: Kharkiv National University of Radio Electronics
Date(s) Attended: 2022 – 2023
Qualification(s): Data Science, Master's degree
Institution: EPAM Master's Program
Date(s) Attended: 2022 – 2023
Qualification(s): Big Data Course, Masters Program
Institution: National Technical University of Ukraine 'Kyiv Polytechnic Institute'
Date(s) Attended: 2020 – 2022
Qualification(s): Applied Mathematics, Master's degree
And a lot of different professionals courses.

Domains, Directions: Telecom (SMS-, Email-marketing), Social Media, Health Care, E-commerce, Gaming, Agro and some others

Languages: English – Upper-Intermediate; Turkish – A2; Ukrainian, Russian – Native.