



Success story

Automatization of the end-to-end Machine Learning pipelines

Our client, a global retailer, aimed to automate their Machine Learning (ML) pipeline to leverage user behavior data and gain a competitive edge by predicting customer needs more accurately. To do so, the retailer lacked the expertise in Big Data, software development, and data transformation. They turned to VirtusLab. They created a new end-to-end ML process using Python language with TensorFlow and PySpark frameworks. This new system helped our client to deliver a more personalized experience on their website while saving time and resources.



The Challenge

Our client aimed to automate their Machine Learning pipeline. It would allow them to take full advantage of existing customer behavior data and deliver a more personalized experience on their website. The client's Data Science team was already developing ML models but lacked the engineering expertise and manpower necessary to scale up their production.

In order to maintain the project timeline of delivering new solutions to the end-user the existing pipeline had to be fully automated. This called for a holistic approach in multiple areas:

- Data ingestion
- Feature generation
- Model training
- Model deployment and serving

Having previously collaborated with the client, VirtusLab assembled a dedicated team proficient in Big Data, software development, and data transformation.



The solution

VirtusLab built a fully automated end-to-end Machine Learning process that delivers new models on demand. They created small, manageable code pipelines, using decoupled component methodology. VirtusLab designed the pipelines, making it easier to understand individual data transformations, features, or models. They leveraged PySpark and TensorFlow frameworks only where necessary.

VirtusLab introduced a composable configuration for ease of production and scalability. They also developed common building blocks to extract complex logic and ensure consistency across modules. Overall, they built a fully automated end-to-end Machine Learning process that delivers new models on demand.

Working in collaboration with the client's Data Science team, rigorous engineering practices were integrated, including:

- Unit and acceptance testing
- Static type checking
- Linting
- Code reviews
- Continuous integration and deployment (CI/CD)





★ The results

VirtusLab's solution helped the client's Data Science team successfully productionise 25 Machine Learning models and over 200 pipelines.

Automated ML pipelines enabled our client to better personalize the online shopping experience. Our client introduced tailored product recommendations, customized promotions, and dynamic pricing that increased customer satisfaction and loyalty.

VirtusLab's small code pipelines enhanced the client's ability to manage complex data pipelines and Machine Learning processes. Using decoupled component methodology ensured that the system is:

- 1 Robust
- 2 Adaptable
- 3 Capable of handling the demands of large-scale data processing and model deployment

VirtusLab also streamlined the process of monitoring the performance of ML models, making it easier to address issues as they arose. For example, the client is now effectively counteracting problems like data drift — an unexpected change in input data that can affect the quality of the model's result.



The tech stack

/

Programming language

- Python

/

Libraries

- Pandas

/

Big Data Processing Framework

- PySpark
- Hadoop
- Hive

/

Machine Learning Framework

- TensorFlow

/

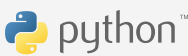
Cloud storage

- S3

/

Other tools

- Jenkins
- Conda
- mypy
- Jinja
- Oozie



About VirtusLab

At VirtusLab, we aim to lead in software technology, working consistently to enhance efficiency. Our profound commitment to research and development and a dedicated focus on emerging trends and inspirations fuels an innovative culture. This ethos precisely guides advancing our cutting-edge solutions, inviting collaboration to expand the boundaries of software technology collectively. We welcome you to be a part of this transformative journey.

Let's connect

Contact Details

info@virtuslab.com

/ POLAND

Kraków Headquarters

Virtus Lab Sp. z o.o.
ul. Szlak 49
31-153 Kraków

/ GERMANY

Berlin Office

+49 30 52014256
VirtusLab GmbH
Potsdamer Platz 10
10785 Berlin

/ UNITED KINGDOM

London Office

+44 (0)20 4577 1051
Virtuslab Ltd.
40 Bank Street HQ3
London E14 5NR

WWW.VIRTUSLAB.COM

