



Mikhail Vasiliev

Deep Learning Engineer

Experience

2023—**Senior Machine Learning Specialist, Makves**
present Project: **Development and implementation of RAG system**
Tools: LangChain, Ollama, Saiga, GigaChat, Python, Hugging-Face, PyTorch, FastAPI, Ragas

- Developed and implemented a RAG system for automating customer request processing
- Optimized system hyperparameters using Ragas library and GigaChat LLM

Project: **Creation of a comprehensive security solution for corporate networks based on unstructured data**
Tools: python, transformers, EfficientNet, MobileNet, YOLO, PIL, PyOD, pandas, sklearn, pytorch, lightning, numpy, matplotlib, plotly, huggingface, onnx, fastapi, uvicorn, pyinstaller, pywin32, optimum, airflow, mlflow, cvat, natasha

- Implemented a neural network module for detecting violations of personal data laws, increasing detectable classes from 14 to 36 with top 1 accuracy reaching 98.9
- Developed a module for analyzing scanned document content: text, tables, stamps, signatures and corporate forms detection, increasing classes from 5 to 19 with mAP@.5 improved from .89 to .94
- Implemented sensitive data detection in text files with added NER module
- Created an ensemble of algorithms for anomaly detection in tabular data, including time series
- Developed sensitive data detection in audio files
- Collected and organized labeling for 8 datasets for classification and object detection tasks

* April 25, 1987

☎ +7 (916) 198 81 83 • ✉ gnu.xinm@gmail.com

🌐 onixlas.github.io • **in** [michael-vasiliev-ds](#)

🔗 [onixlas](#) • 📍 [LaHundo](#)

Personal Projects

2024 **Team Lead and Technical Expert, CheckDocAI**

Project: Telegram bot with AI module for document quality control for Gulfstream LLC, significantly reducing verification time and improving accuracy.

Tools: aiogram, YOLO, ONNX, Albumentations, CVAT

- Led a team of two data scientists and a backend developer, responsible for project development and implementation
- Successfully deployed for commercial use with monthly savings of 40 man-hours

Hackathons

2024 **VK HSE Data Hack, 1st place**

Hackathon for news article classification into 21 categories. Our solution combined results from a small transformer-based classifier and LLM predictions

Tools: transformers, Saiga3 8b, taiga dataset, streamlit

- Enriched the dataset
- Selected zero-shot classification model
- Trained classifier model
- Coordinated team work
- Presented results

Talks

24.04.2025 **Anomaly Detection in Data: iForest and PCA Algorithms**, *Moscow Python Meetup № 101*

24.02.2025 **Anomaly Detection in Data: HBOS and ECOD Algorithms**, *Moscow Python Meetup № 99*

29.01.2025 **NLP Neural Networks in Data Protection: Makves DCAP Experience**, *Moscow Python Meetup № 98*

25.06.2024 **Training and Applying Neural Networks as a Module of Russian DCAP System**, *Moscow Python Meetup № 91*

* April 25, 1987

☎ +7 (916) 198 81 83 • ✉ gnu.xinm@gmail.com

🌐 onixlas.github.io • in [michael-vasiliev-ds](#)

🔄 onixlas • 📍 LaHundo

Education

- 2024 **Data Analysis with SQL**, *Training Center "Specialist"*, professional development
- 2022—2023 **Computer Vision Engineer**, *Deep Learning School, MIPT*, professional retraining
- 2022 **Data Science Specialist**, *Yandex Practicum*, professional retraining
- 2021—2022 **Introduction to AI and Neural Networks for Aviation Applications**, *MAI*, professional development
- 2005—2008 **Translation and Translation Studies**, *MAI*, specialist degree
- 2003—2009 **Aviation and Space Thermal Engineering**, *MAI*, specialist degree

Languages

Russian	■■■■■	native
English	■■■■■	B2
German	■■■■■	B2
Esperanto	■■■■■	B2

Skills and Technologies

- | | | |
|----------------------|---------------|--------------|
| ○ Deep Learning | ○ Python | ○ Pytorch |
| ○ LLM, RAG | ○ SQL | ○ Lightning |
| ○ NLP, NER | ○ Linux | ○ Pandas |
| ○ Computer Vision | ○ Docker | ○ NumPy |
| ○ Speech Recognition | ○ YOLO | ○ Sklearn |
| ○ Machine learning | ○ Natasha | ○ FastAPI |
| ○ Anomaly Detection | ○ ONNX | ○ uvicorn |
| ○ Data analysis | ○ HuggingFace | ○ PyOD |
| ○ Data visualisation | ○ Ragas | ○ PySAD |
| ○ Statistics | ○ Ollama | ○ Optimum |
| | ○ U-Net | ○ pywin32 |
| | ○ AirFlow | ○ CatBoost |
| | ○ MLFlow | ○ XGBoost |
| | ○ CVAT | ○ PostgreSQL |
| | ○ Plotly | ○ MySQL |

* April 25, 1987