Sergey Litvinov

Machine Learning Engineer | Generative AI Specialist

Address: Moscow, Moscow City, Russia

LinkedIn: https://www.linkedin.com/in/sergey-litvinov/

Email: ltvn.sergey.work@gmail.com

Machine Learning Engineer with 6+ years of experience delivering high-impact AI systems across domains like generative media, geophysics, industrial automation, and NLP. Specialized in Generative AI, LLMs, and multi-modal model development, including text-to-image, image-to-video, and TTS. Proven leadership in R&D teams (6+ people), with expertise spanning model training, optimization, and deployment at scale. Adept in both research and engineering execution — bridging advanced AI theory with production-level implementation.

EXPERIENCE

NVI Solutions
April 2023 - Present

Data Scientist

- Led team of 6+ in AI R&D and commercial projects
- Implemented research-based algorithms for real-world applications
- Managed full project lifecycle: scoping, roadmap, delivery
- Integrated LLMs and generative models in client systems

GeoPrime, Moscow
May 2017 - March 2023

Senior Geophysicist

- Managed and executed large-scale seismic data analysis
- Applied ML/DL for signal processing in geophysical datasets
- Reduced processing time by 10x via automation pipelines

PROJECTS

LLM-based Tech Support Automation
July 2025 - present

- Created Telegram-based support system using LLMs, RAG, and agent logic
- Reduced specialist workload by 95%

Industrial Danger Detection with YOLO
October 2024 - December 2024

- Built object detection pipeline for real-time industrial monitoring
- Achieved 96–98% precision/recall

Visual Location Search System
April 2024 - October 2024

- Developed mobile image-retrieval app with segmentation, scene reconstruction
- Deployed vector search + segmentation on Android (Kotlin + TFLite)

Inappropriate Behavior Detection (Audio)
February 2024 - April 2024

- Built spectrogram-based deep learning pipeline for speech monitoring

AI Document Parsing Pipeline
May 2023 - February 2024

- Created microservices to analyze document structure and content
- Used YOLO + PaddleOCR; optimized via TensorRT
- Built GUI validator in PyQT

Seismic Noise Detection
June 2021 - August 2021

- Combined deterministic + ML methods for 10x faster noise detection

SKILLS

- **Languages**: Python, C++, SQL, Kotlin
- **ML Frameworks**: PyTorch, TensorFlow Lite, PaddleOCR, FastAPI, Gradio
- **Deployment Tools**: TensorRT, ONNX Runtime, Docker, Git, DVC, PyQT
- **Data/Infra**: PostgreSQL, Clickhouse, Qdrant, Minio, RabbitMQ, CVAT
- **Generative AI & LLMs:**
  - Trained and fine-tuned multi-modal generative models: Text-to-Image, Image-to-Image, Image-to-Video, Text-to-Audio
  - Tools used: Stable Diffusion, Flux, Lumina, Hidream, Sana, and ControlNets
  - Built video generation workflows with WAN, LTX
  - Fine-tuned open-source TTS models: Orpheus, Kokoro
  - Developed LoRA models for fine-grained control over outputs
  - Created RAG systems with custom embeddings and retrieval logic
  - Used GPT-family and other LLMs for automation, agents, and integration tasks

EDUCATION

Yandex Praktikum
May 2020 - March 2021

Data Science Specialist

Lomonosov Moscow State University (MSU)
September 2016 - June 2018

Master of Geophysics and Seismology

Lomonosov Moscow State University (MSU)
September 2013 - June 2016

Bachelor of Engineering Geology and Geophysics

LANGUAGES

English (Full professional proficiency), **Russian** (Full professional proficiency), **German** (Elementary proficiency)

CERTIFICATIONS

Certificate of completion 'Introduction to Machine Learning for Geophysical Applications'
September 2021 - Present

EAGE (European Association of Geoscientists and Engineers)

Certificate for prize-winner in hackaton held during Intellectual Data Analysis in Oil and Gas conference
August 2021 - Present

EAGE (European Association of Geoscientists and Engineers)

Statistics 101 provided by IBM
February 2023 - February 2033

Cognitive Class

Credential ID: c5f2673a900f4c8fa6842c9bcb34de0a

Show Credential

PUBLICATIONS

Machine learning for classification of seismic data
August 2021

EarthDoc

- This paper discusses the possibility of using neural networks to classify seismic data in order to increase the efficiency of data processing, reduce the time for a geophysicist to perform routine tasks and have a positive impact on the economic efficiency of the project. The result of using deep learning for the classification of seismograms in the presence of non-stationary man-made noise in space is presented. The approach made it possible to achieve high classification accuracy. As a result of the work, an important conclusion was made about the possibility of using this approach to search for man-made noise in seismic records.