Last Updated on February 14th, 2025

Anastasiia Semina

t.me/sad_bkt github.com/sad-bkt linkedin.com/in/anastasiia-semina semina.anastasy@gmail.com

Data Scientist and Machine Learning Engineer with experience in deep generative models, NLP, and computer vision. Specialized in AI-driven automation, mathematical optimization, and predictive modeling. Conducted experiments with txt2img and img2img models (Flux, Stable Diffusion), as well as txt2video models (Kling, Luma, Gen-2). Experienced in mentorship, statistical analysis, and model evaluation, with a strong academic and research background.

Experience

Mentored a course on image processing and generation | *Mentor* Dec. 2024 - Jan. 2025

I help the lecturer with course materials and check students' homework.

X5 Tech | *NLP Engineer* Oct. 2024 - Feb. 2025

- **Diploma thesis Product attitude**

  I am developing a project for review classification and clustering to identify growth opportunities and improve retail store performance. Compared multiple Hugging Face models on a custom dataset. Achieved the best performance with DeepPavlov/rubert-base-cased for sequence classification, reaching 63% accuracy. The result is notable given the dataset's poor labeling quality.

Gazprom Neft | *Data Scientist, Engineer* Feb. 2024 - Present

- **Automatic scheduling**

  I engaged in a project focused on forecasting production volumes and identifying critical production plant metrics to optimize operational efficiency.
- **Mentorship of 2 interns**

  Mentored two interns in conducting statistical analysis for three technical installations, leading to the proposal of several data-driven production forecasting methods. Guided them in Python and statistical modeling.

EveryPixel | *Computer Vision and Prompt Engineer* Aug. 2024 - Sep. 2024

- **Neuroproduction project**

  Conducted comparative analysis of multiple models for generating realistic group images with interactions between people, using webui-forge, webuiapi. The following models were analyzed: SDXL Juggernaut, RealisticStockPhoto v2.0 (SDXL) and v3.0 (SD 1.5), SDXL RealvisXL 4, combinations of SDXL models with custom LoRA, MidJourney, BRIA.AI, DALLE-3, Adobe Firefly Image 3, and Flux. Identified issues such as hallucinations, facial distortions, unnatural proportions, hand inaccuracies, and token truncation.

Wonderslide | *Prompt Engineer for Stable Diffusion* Feb. 2024 – Oct. 2024

Generated images on various topics (realistic portraits, abstract figures) using models from Civitai for automated presentation project.

NymphLens telegram bot | *Prompt Engineer for Stable Diffusion* Oct. 2023 – Jul. 2024

Prompt engineering and comparison of models for a project that helps individuals create social media photos by automatically generating images with their faces.

BIA Technologies | *Machine Learning Engineer* Aug. 2022 - Aug. 2023

- **Automation of planning process**

  I automated task planning in Jira by gathering requirements, assessing tools and algorithms for optimal task allocation, coding the solution, and collaborating with a user interface designer. The planning process was comprehensive, it considered factors like deadlines, priorities, sequencing, people competencies, and capabilities, using technologies like OptaPlanner and DoCPLEX.
- **Layoff Prediction and Employee Attrition Analysis**

  I developed a predictive model to forecast layoffs and analyze employee attrition reasons, aiming to proactively address retention issues and reduce turnover.

Skills

Languages: Python, SQL, LaTeX
Tools: Git/GitHub, Bash, PyCharm, VS Code, MS Office, Atlassian Jira, Atlassian Confluence, Docker
Libraries: matplotlib, seaborn, sklearn, opencv, pytorch, ultralytics, diffusers, nltk, gensim, spacy, cplex, pyomo
Frameworks: AUTOMATIC1111, Django

Education

Master's in Artificial Intelligence Sep. 2023 – Aug. 2025
ITMO University *Current GPA: 5.0/5.0*

Bachelor's in Informational Systems and Technologies Sep. 2019 – Aug. 2023
ITMO University *GPA: 4.0/5.0*

Project and Product Management 2022
Tinkoff courses

Projects and courses

Image manipulation detection | *Grad-CAM, ResNet50* 2024

- Developed a ResNet50-based model to classify original and photoshoped images from PS-Battles dataset, implemented Grad-CAM for interpretability. Reached 0.89 accuracy and 0.84 F1-score weighted.

Deep generative models course | *diffusers, VAE, GAN* 2024

- Developed Bayesian models and implemented autoencoders.
- Implemented GAN and sampled images in the StyleGAN latent space.
- Trained Stable Diffusion v1.5 using Dreambooth method and fine-tuned LoRA models.

Computer Vision course | *Multi-class classification, Anomaly detection, Self-supervised learning* 2023

- ResNet realisation in pytorch.
- Experiments with variational autoencoder.
- Multi-class classification using a dataset with missing labels during training.

RoadSignsDetector | *YOLOv8, RTMDet* 2023
Web application for real-time road signs detection.

Achievements

MegaOlympiad ITMO 2023

- Artificial Intelligence - Top 5 out of 200+ participants.
- Game-theoretic modeling - Top 3 out of 90+ participants.