Masdar City
Abu Dhabi, UAE

Aleksandr Matsun

+971567818597
matsun.av@gmail.com
github.com
linkedin.com

Employment

Research Engineer — **MBZUAI** — **June 2024 – November 2024; March 2025 – Current**

- Developed a pipeline for in-silico drug generation, filtering, docking and retro-synthesis
- Discovered multiple SIRT1 allosteric activators with estimated binding affinity higher than that of existing ones by 15% on average
- Devised a multi-omics-based model for aging-related gene and drug discovery

Researcher — **Technology Innovation Institute** — **December 2024 – February 2025**

- Implemented deep learning algorithms for lesion-presence classification and lesion localization in chest X-Ray images with over 0.95 F1 score and over 0.60 AP50 on out-of-distribution data samples
- Worked on a residual imaging-based pipeline for automated chest X-Ray report generation task

ML R&D Engineer, Intern — **Insilico Medicine** — **June 2023 – July 2023**

- Developed a deep learning-based pipeline for disease hypotheses extraction form a bioentity knowledge graph
- Introduced a performance metric for hypothesis generation, based on Levenshtein Distance
- Created a pre-training technique for hypothesis generation models, based on link prediction. Engineered an algorithm for link prediction in the knowledge graph with over 98% accuracy

ML Engineer — **CITaS** — **August 2021 – May 2022**

- Prepared a deep learning-based model for person re-identification for CCTV system, improving the previous mAP score by 7%
- Introduced and implemented the instance segmentation module into the pipeline. Established an embedding based on the segmentation results for further improvement of re-identification
- Supervised a group of junior MIPT students and evaluated research projects

Software Developer — **MIPT - Starkits Lab** — **November 2019 – January 2022**

- Implemented techniques for the image-processing module of the soccer-playing robots. Reduced the error of conversion between 2D and 3D coordinates by 13%
- Prepared the strategy and decision-making modules for the robots based on the particle filter algorithm
- Established an overhaul of the firmware code that resulted in 8% faster work

ML Engineer, Intern — **Sberbank** — **August 2020 – November 2020**

- Developed an ML model for crop harvest prediction, using weather-related data
- Devised an RNN-based method for filling missing data cells, that improved the accuracy by more than 10%

Education

Abu Dhabi, UAE — **MBZUAI** — **August 2022 – June 2024**

- M.S.E. in Computer Vision, June 2024

Moscow, Russia — **MIPT** — **September 2015 – June 2021**

- M.S.E. in Applied Mathematics, June 2021
- B.S.E. in Applied Mathematics, June 2019

Publications

- **ConDiSR: Contrastive Disentanglement and Style Regularization for Single Domain Generalization** WACV 2025 – arxiv.org

- **CoReEcho: Continuous Representation Learning for 2D+time Echocardiography Analysis** MICCAI 2024 – arxiv.org
- **DGM-DR: Domain Generalization with Mutual Information Regularized Diabetic Retinopathy Classification** MICCAI-DART 2023 – arxiv.org

Skills

- Python; OOP; Algorithms; Docker; NumPy; Pandas; Git
- Deep Learning; PyTorch; Optimization; Statistics; GNN; LLM; ONNX; AWS
- Object Detection; OpenCV; CNN; ViT; YOLO; Medical Imaging
- Multi-omics; Drug discovery; Target identification; Docking; CASP

Projects

Computer Vision

- **Remote object detection** (2023). Algorithm for object detection in satellite images, utilizing Large Selective Kernel technique with Diffusion-based detection. github.com
- **Surgical tools detection** (2022). Framework for weakly supervised two-stage surgical tool detection in endoscopic videos. github.com
- **Image processing in robo-soccer** (2021). Modules for image processing and 2D/3D coordinate conversion for soccer-playing robot. github.com

Natural Language Processing

- **Text moderation model** (2024). Large Language Model fine-tuned for text moderation on presence of offensive content, deployed on Hugging Face. huggingface.co
- **NER for Knowledge Graph** (2023). Model for Named Entity Recognition-based composition of the bioentity knowledge graph. github.com
- **Machine Learning Q&A bot** (2022). Open-book LLM-based Q&A bot for ML-related queries with flexible knowledge base. github.com