Tatyana Pichugina PhD

linkedin.com/tatyana-pichugina   github.com/tapichugina
ttatyana.pichugina@gmail.com
New Jersey, USA (Open to Relocation)

SUMMARY

Data scientist with a solid background in Computer Vision and a proven track record of driving projects from inception to deployment. Proficient in utilizing a wide range of technical skills to extract insights from complex data sets, build efficient machine learning pipelines, and deploy models for real-world applications. I have worked in a fast-paced, cross-functional setting, both independently and as part of a team. I am passionate about learning and using new technologies.

SKILLS

Toolsets: | Python | PyTorch | OpenCV | SQL | Tableau | Linux | AWS | Flask | Docker | Git | Jira

Technical: Computer Vision | Deep Learning | Machine Learning | ML Experiments Tracking | Exploratory Data Analysis | Data Cleaning | Data Mining | Feature Engineering | Data visualization | Summary Reports | Reproducible Pipelines | Principal Component Analysis | Anomaly Detection | Clustering | Semantic segmentation | ML Model Deployment

MLops: Containerization (Docker) | REST API (Flask), | ML Lifecycle (MLflow, WandB)

General: Agile Project Management | Problem-Solving | Team Work

WORK EXPERIENCE

Machine Learning Engineer (Computer vision) | Invonto, USA | Feb 2023 - present
- Orchestrated end-to-end machine learning projects: established business oriented performance metrics, initiated image dataset collection, performed data labeling, executed iterative model experimentation, deployed and monitor model.
- Significantly enhanced model performance in semantic segmentation by optimizing deep neural networks (U-Net), achieving an exceptional increase in Intersection over Union (IOU) from 0.76 to 0.90. Employed **PyTorch Lightning** as a dynamic training framework to drive these advancements effectively.
- Successfully deployed a trained model utilizing **Flask** as a web application on an **EC2** instance in **AWS**, ensuring efficient and user-friendly access to the model's capabilities.
- Gathered, curated and refined image datasets, resulting in improved model generalization and accuracy.
- Employed advanced computer vision techniques to accurately measure objects within images utilizing OpenCV
- Collaborated closely with cross-functional teams and stakeholders to integrate machine learning solutions into practical applications
- Applied best practices in version control and documentation, providing transparent and well-structured project documentation that facilitates knowledge sharing and ensures project reproducibility.

Senior Researcher | Max Planck Institute for Evolutionary Biology, Germany | 2018 - 2022
- Mined, analyzed and interpreted over 100 Tb unstructured data (microscopy videorecording) for 3 decision-making research projects with multiple stakeholders.

- Checked data quality and data collection problems using **SQL,** and **Python API** for Omero Server Database.
- Designed **Python (Jupyter)** pipeline to check data quality, clean, and segment objects for thousands of microscopy images, to facilitate data analysis and reduce hundreds of hours of manual work.
- Detected anomalies in instrumental data collection, identified the problem and developed an automated **Python** script to check the experimental setup.
- Identified meaningful clusters using **Scikit-learn** Gaussian Mixture Model combined with **PCA**.
- Tracked version changes and shared code for 3 projects using **Git** and **GitLab**.
- Led Image processing discussion group for institute members, organized talks, invited guest speakers, and managed surveys.

Researcher | the University of Auckland, New Zealand | 2012 - 2016
- Designed from scratch a 3D predictive model using the Monte Carlo simulation method for whole yeast genomes using **C**.
- Performed over 3000 extensive calculations using Auckland University Supercomputer (NESI) using SLURM, and **Linux**.
- Identified meaningful clusters of genetic elements in 3D genomes data set using **R**.
- Conducted statistical analysis for over 1000 3D genome structures.
- Designed 3D visualizations of genes, epigenetic marks, and replication origins positions using **R**.
- Published 3 papers in peer-reviewed journals (NAR, Scientific Reports, Nucleus) in collaboration with interdisciplinary cross-country research teams.
- Performed literature search and summarized findings for over 300 scientific papers.
- Gave 3 presentations at international conferences.

EDUCATION
PhD in Molecular Medicine, the University of Auckland, New Zealand
Awards/Activities: The University of Auckland Liggins institute PhD scholarship. NAR journal breakthrough article. Top-ranked PhD thesis (8 out of 71 completed in 2016 in the Faculty of Medical and Health Sciences of the University of Auckland). The Maurice and Phyllis Paykel Trust travel grant.
Master of Science, Chemical and Biological Physics, Novosibirsk State University, Russia
Bachelor of Science, Physics, Novosibirsk State University, Russia