Evgeny Krivosheev

Email: e.krivoshe@gmail.com   Google Scholar   GitHub   Personal page: evgeneus.github.io

Lead and Senior AI Engineer with 8+ years of Industrial and Research experience aspiring to build a career in the technology sector with a focus on AI, applied research, and data-driven solutions.

Current stack: Python, PyTorch & Lightning, Huggingface, Scikit-learn, AWS Sagemaker, Airflow, Jenkins, Docker, SQL, Statistics, Transformers, Llama2, LayoutLM, ELECTRA, ViT, Graph CNN, ResNet etc.

ML/AI Expertise: Natural Language Processing (NLP), LLMs, Embeddings Learning, Computer Vision, Graph Neural Networks, Self-Supervised Learning, Domain Adaptation, Data Fusion and Crowdsourcing.

WORK EXPERIENCE

Jumio Corporation (Vienna, Austria, Remote)   2021 - present
Tech Lead and R&D AI Engineer

- Leading Document Classification and Retrieval Team for supporting ID verification for customers such as Airbnb, UBER, TikTok. My scope is design, tune, productionalize multimodal (text, vision transformers) solutions, define AI roadmap for the team, mentoring engineers, cross-team communications.

- Research and development of first-class citizen Multimodal ID Retrieval Engine for the RAG (Retrieval-Augmented Generation) system, aimed at enhancing document classification accuracy, reducing inference time and delivering fast scope extension without retraining models.

- Innovating and proofing new product concepts using cutting-edge technologies. Tuning LLAMAv2 for data extraction from ID documents (Lora, Haggingface, PEFT), and prompt-tuning for achieving quality-inference time trade-off.

- Developed Data Extraction NLP models and services for processing PII documents, such as Bank Statements. Reduced reliance on the human data extraction workforce by 45% (in France, Germany, Italy), surpassing target KPIs (XLM, ELECTRA, RoBerta, PyTorch, Huggingface, ONNX, AWS, GCP OCR).

- Implemented a multi-lingual NLP document classification model taking ownership for the entire feature life-cycle. Increased Accuracy by +10% of the Document Verification product, resulting in significant reduction of high-severity errors experienced by customers.

- Designed Automated Testing solution for ML services, enabling continuous integration of model updates and code changes at Pull Requests (Airflow, Jenkins, Docker, AWS SageMaker).

University of Trento, Multimedia group (Trento, Italy)   2020 - 2021 (1.5 years)
Post-Doc in Computer Vision, Representation Learning

- Developed a deep learning framework for Unsupervised Domain Adaptation pivoted around Graph Embeddings Aggregation, GANs, and Curriculum learning, achieving +5% improvement in accuracy with respect to the state-of-the-art algorithm (accepted at CVPR 2021, see project page).

- Implemented several Self-Supervised visual embedding models (SimCLR, Moco, Byol), customised Byol with Ensemble training and Data augmentations (MixUp, CutMix), improving the accuracy by 2%.

- Co-supervised PhD students in technical and methodological aspects throughout research projects.

IBM Research (Zurich, Switzerland)   2019
Visiting Researcher (Information Retrieval, NLP)

- Invented and programmed Self-Supervised Graph Convolutional Net for linking business entities from text to relational database (Dun & Bradstreet) resulting in +7% linkage improvement for Swiss business entities with respect to the current IBM linkage system (accepted as AAAI 2021 paper).

- Deployed the Graph Convolutional Net as a web service with PyTorch, Flask, Docker (see demo).

University of Trento (Trento, Italy)   2016-2020 (4 years)
PhD in Human-AI, NLP (Cum Laude, advisor Prof. Fabio Casati)

- Developed a number of cost-aware Machine Learning, Active Learning and Hybrid crowd-machine learning algorithms for document classification for low-risk decision making.

- Main author of 9 papers published on top-tier conferences (WWW, AAAI, NeurlPS, VLDB, HCOMP).

- Co-supervised students and interns in research projects.

IBM Benelux & TU Delft (Amsterdam, Netherlands) 2018

Visiting Researcher (Machine Learning)

- Developed Active Learning and Adaptive Human-ML classification algorithms for balancing between training ML models and acquiring reliable training data.

- Reduced document classification cost by 10% while improved accuracy by 5% over several classification tasks with respect to state-of-the-art approaches (two NeurlPS publications).

iTechArt (Minsk, Belarus) 2015

Python Software Engineer

- Developed RESTful endpoints and features in admin interface for classpass.com company (Django, JS).
- Troubleshooted backend, database problems that speeded up SQL queries and email services.

EDUCATION

University of Trento (Trento, Italy) 2016 - 2020

PhD (Cum Laude), advisor Prof. Fabio Casati
Machine Learning, Hybrid crowd-machine algorithms for document classification.

Belarusian State University (Minsk, Belarus) 2011 - 2016

Specialist in Information Systems and Radiophysics (equivalent to Master degree)

PROFESSIONAL ACTIVITIES

Program committee member at NurlPS 2020 workshop, poster session chair at ICPR 2020, invited speaker at PiSchool 2020, reviewer at WSDM 2020, WWW 2018, WWW 2017, ICWE 2017, BPM 2017.

SELECTED PUBLICATIONS

GIPSO: Geometrically Informed Propagation for Online Adaptation in 3D LiDAR Segmentation **ECCV 2022**, Cristiano Saltori, Evgeny Krivosheev, Stéphane Lathuilière, Nicu Sebe, Fabio Galasso, Giuseppe Fiameni, Elisa Ricci, Fabio Poies.

Curriculum Graph Co-Teaching for Multi-Target Domain Adaptation **CVPR 2021**, Subhankar Roy*, Evgeny Krivosheev*, Zhun Zhong, Nicu Sebe, Elisa Ricci, * - equal contribution.

Business Entity Matching with Siamese Graph Convolutional Networks **AAAI 2021,**
Evgeny Krivosheev, Mattia Atzeni, Katsiaryna Mirylenka, Paolo Scotton, Christoph Miksovic, Anton Zorin.

A review and experimental analysis of active learning over crowdsourced data, **Artificial Intelligence Review 2021 (Journal)**, Burcu Sayin, Evgeny Krivosheev, Jie Yang, Andrea Passerini, Fabio Casati.

Active Hybrid Classification, **NeurlPS 2020** workshop, Evgeny Krivosheev, Fabio Casati, Alessandro Bozzon

LANGUAGIES

English (Fluent), Russian (Native)

HONORS:

- The 1st place in the AI Jumio Corporation Hackathon with the Embeddings Learning for Document Retrieval project. The solution is recognised to be implemented on production.

- Won GoBoiler Research Scholarship, Purdue University, USA.

- Won PhD Scholarship, the University of Trento, Italy, 2016 – 2020.

- The 1st place in the Gomel Olympiad in Astronomy, Belarus.

- Won a University grant and founded the "Mutual assistance" project. For several years, I was leading my team of 10 senior students. We have been organising workshops and lectures on mathematical disciplines for Physics, Biology, and Humanities departments.

DOCUMENTS: Italian/EU Permanent Work Permit