

Anton BELY

anton.belyy@gmail.com · avbelyy.github.io

LinkedIn · GitHub · Scholar

EDUCATION

- | | |
|-------------------|--|
| AUG '19 — MAY '22 | M.Sc.Eng. in Computer Science (NLP) , Johns Hopkins University, USA GPA: 4.0. Advisors: Benjamin Van Durme (Primary), Vladimir Braverman (Secondary) Research projects: knowledge graph completion, controlled text generation, schema induction |
| SEP '13 — MAY '18 | B.Sc. in Informatics and Applied Mathematics , ITMO University, Russia GPA: 3.6. Advisors: Andrey Filchenkov (Primary), Konstantin Vorontsov (Secondary) Research projects: plagiarism detection, hierarchical topic modeling Thesis: Construction and Quality Evaluation of Heterogeneous Hierarchical Topic Models <ul style="list-style-type: none">• Distinguished thesis award (given to 2 out of approx. 50 undergraduates) |

EMPLOYMENT

- | | |
|-------------------|---|
| JUN '24 — PRESENT | Senior Machine Learning Engineer in Credit Risk Detection , Stripe, USA <ul style="list-style-type: none">• Default prediction: I lead project Voyager, which scores millions of Stripe merchants to predict whether they might go default in the near future and lead to Stripe-liable losses. Voyager is a PyTorch DNN-based model, which saves tens of millions in credit losses yearly.• Model explainability: I implemented and deployed scalable post-hoc explanation pipelines for several ML models at Stripe, leading to higher model trust during review-based actioning. |
| JUN '22 — MAY '24 | Machine Learning Engineer in Knowledge Graphs and LLM Datasets , Apple, USA <ul style="list-style-type: none">• KG: I am a founding engineer of the ODKE project, which transforms unstructured Web data into millions of facts every week to continuously improve coverage and freshness of Apple KG. I implemented scalable fact extractors using in-house and OpenAI LLMs, built fact verification model using Spark MLlib, built CI/CD pipelines and oversaw integrating ODKE with Apple KG.• LLM Datasets: I used MapReduce, Spark and ML to turn petabytes of semi-structured Web data into billion-scale, high-quality datasets used to train LLMs for Apple Intelligence. |
| JUN '21 — AUG '21 | ML Research Intern in Conversational AI , Microsoft Research, USA <ul style="list-style-type: none">• Used explainable k-means to speed up labeling data for semantic parsing by 35%. [paper] |
| SEP '17 — AUG '19 | Machine Learning Engineer in Compliance Risks , Tochka Bank, Russia <ul style="list-style-type: none">• Client scoring: using Hadoop and Spark, enriched internal data for 200K+ bank clients and 50M+ transactions with Web-scraped public data to improve client scoring accuracy by 10%.• Intent recognition: built ML models to detect intent for the top-90% client questions in chat.• Call center planning: using ILP, automated CC scheduling while improving accuracy by 10%. |
| MAR '17 — AUG '17 | Machine Learning Engineer , Antirutina, Russia <ul style="list-style-type: none">• Tender anomaly detection: developed clustering algorithms to identify bidding anomalies, allowing to discover bid-rigging behavior on auctions with contract amount exceeding \$3B.• Precise IE: designed information extraction pipelines for precise identification of vendor codes, volumes and quantities of goods from unstructured and diverse vendors' price lists. |
| OCT '15 — OCT '16 | Software Engineer in Ads , VK.com, Russia <ul style="list-style-type: none">• URL fraud: built service to periodically detect malicious URL redirect changes in VK ads.• Click fraud: built ML models to detect users that generate fraudulent clicks in VK ad network. Model was deployed semi-automatically and helped recover up to 3% monthly ad revenue.• Ads search: launched moderator search interface (incl. full-text search) over 30M+ VK ads.• Ads scoring: implemented advertiser ranking for faster moderation of top-10% clients. |

LANGUAGES AND TECHNOLOGIES

| | |
|--------------|---|
| LANGUAGES | Python (advanced); Golang, Scala, JavaScript, SQL, C++, bash (proficient) |
| TECHNOLOGIES | Airflow, Spark, Iceberg (advanced); XGBoost, PyTorch, AWS, Docker, K8s (proficient) |