Alexander V. Belikov

+33 (0) 787848423
a.belikov@gmail.com
alexander-belikov.github.io
Paris, France

EDUCATION

Ph.D. in Physics, University of Chicago, IL, 2011
M.S. in Applied Mathematics and Physics (summa cum laude),
Moscow Institute of Physics and Technology, Russia, 2005
B.S. in Applied Mathematics and Physics (summa cum laude),
Moscow Institute of Physics and Technology, Russia, 2003

EXPERIENCE

Head of R&D Sept. 2023 - present
Qantev, Paris, France

- Leading research initiatives within the medical insurance domain, focusing on LLMs and embedding techniques.

Lead Data Scientist Apr. 2023 - Sept. 2023
Qantev, Paris, France

- Directed and supervised a team of four data scientists, overseeing the design and prototyping of data schema and anomaly modeling in the medical insurance domain.
- Developed a model of mixed-curvature embeddings of ICD/CPT codes used in anomaly detection (medical code / price inconsistency) and medical code inference.
- Achieved a significant improvement in character error rate, reducing it from 55 % to 6 % by fine-tuning a handwritten Optical Character Recognition (OCR) model for Spanish.
- Prototyped a medical claim anomaly model using pre-trained LLMs.

Head of Data Science Nov. 2019 - Apr. 2023
Hello Watt, Paris, France

- Orchestrated and led a team of 2 to 5 data scientists, fostering collaboration between product, operational, and development teams to optimize dataflows and address data-centric and modeling requirements.
- Developed novel machine learning models of signal disaggregation and anomaly detection using probabilistic graphical models, anomaly detection, fairness.
- Contributed to knowledge dissemination through two publications in high-impact peer-reviewed journals and multiple blog publications.

Postdoctoral fellow Jan. 2016 - Oct. 2019
University of Chicago, Knowledge Lab

- Prototyped a novel model of agent evolution on a graph using Seq2Seq methods used to predict the state of the graph, the evolution of individual agents and to identify clusters of agents and events.
- Developed a model of validity of scientific claims in biological literature (with AUC reaching 0.8), and subsequently a model for prediction of the gene-gene interaction with AUC reaching 0.76 based on original network features.

Quantitative researcher Aug. 2015 - Jan. 2016
Barclays Capital, Equity Derivatives Group, New York

- Introduced an effective method for estimating portfolio sensitivities between trading days that accounts for the change of the volatility surface (C++, deployed in production).
- Implemented new types of contracts: options on volatility control indexes.

Quantitative researcher Jun. 2014 - Aug. 2015
JP Morgan Chase, Model Review and Development, New York

- Developed models of mortgage defaults using regularized logistic regression and decision trees (python, scikit-learn).
- Implemented the model of rating migration (loan default estimation) used for the Comprehensive Capital Analysis and Review (CCAR) of the wholesale portfolio (python, deployed in production).

Postdoctoral researcher Oct. 2011 - Nov. 2013
Institut d'Astrophysique de Paris

- Predicted the cosmological annihilation signal for a contracted (due to supermassive black holes) dark matter density. Demonstrated that the spectral properties of the annihilation signal can be used to differentiate dark matter from astrophysical signals.

PhD candidate/Research Assistant Oct. 2005 - Sept. 2011
University of Chicago

- Discovered the connection between the winding angle of random curves appearing in the scaling limit of critical two-dimensional systems and the properties of local operators of conformal field theory.
- Predicted the diffuse gamma-ray background from annihilating leptohilic dark matter and estimated the impact of annihilating dark matter during the reonization epoch (developed a C++ library for estimating cosmological dark matter signals).
- Found semi-analytical solutions for a non-linear PDE in the DGP modified gravity theory.

SIDE PROJECTS

- ETL python package for transformation and loading of tables and json-like structures into graph databases (ArangoDB, Neo4j).
- Academic knowledge graph (KG), relation extraction, entity linking (spacy, SciBERT).
- Signal generation from market analyst reports using network measures (IBES).
- Project management tool (sparql, schema.org, cypher, neo4j).
- Optimization of scientific projects based on KG dynamics (portfolio optimization).
- State-of-the-art model of career transitions (state space with partial order) using graph neural networks (GNN).
- Additive mixture modeling (signal/background separation) using Bayesian neural networks (Pyro).

RELEVANT SKILLS	**ML**: Linear models, graphical models, ensemble methods, SVM, regularization, optimal transport, neural networks (VAE, GAN, RL), graph neural networks (GNN, MPNN) **NLP**: entity linking, co-reference, relation extraction, embeddings, topic modeling (spacy, nltk, BERT, GPT), LangChain. **Languages**: Python (pandas, scikit-learn, pytorch, networkx, igraph, pyro, pymc3), C++, R, Haskell, Java, Cypher, AQL, SPARQL. **Semantic Technologies**: RDF, SHACL, OWL. **Technology**: SQL, mongoDB, ArangoDB, neo4j, git, bash.
PUBLIC SPEAKING	More than 30 presentations at conferences and seminars. Organizer of journal clubs at the Institute of Astrophysics in Paris, Knowledge Lab at the University of Chicago and Hello Watt.
PROFESSIONAL INTERESTS	Natural language processing: NER, relation extraction, embedding, summarization. Graph neural networks, variational methods. Knowledge graphs.
LANGUAGES	English, Russian (native), French (advanced), Italian (elementary)
PUBLICATIONS	Author to 20+ publications in refereed journals, including "Prediction of robust scientific facts from literature" featured in Nature Machine Intelligence.