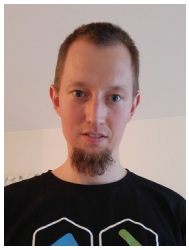
BARTCZAK TOMASZ



mobile: +48 693 302 884
email: kretes@gazeta.pl
born: 12.06.1984
city of residence: Warsaw, Poland

/tomaszbartczak /kretes @bartczakt

Official Education

10.2003 – 07.2008	Msc. in Computer Science at Technical University Of Łódź, The Faculty of Technical Physics, Computer Science and Applied Mathematics, Artificial Intelligence and Software Engineering

Brief Employment history

10.2019 - currently	Allegro, Team Leader, Machine Learning Research & Development
04.2018 – 09.2019	Allegro, Senior Research Engineer
06.2014 – 04.2018	Allegro, Senior Software Engineer
12.2009 – 05.2014	Pragmatists, Agile Java Developer
10.2008 – 11.2009	Altkom Software & Consulting, Java Developer
09.2007 – 09.2008	Transition Technologies S.A., Poland Solution Center Łódź, Java Developer

Professional profile

I consider myself an experienced software engineer turned machine learning. What I value the most is creating machine learning solutions that solve the right problem with the right approach and make an impact in the real world. I want to work in an environment with plenty of challenges of this kind.

My role in machine learning projects may be of multiple flavours. I appreciate doing research, literature review, system designing and architecture, implementation, deployment, analysis and tuning for better results. I feel like a technical person with a product perspective and people in mind.

I am a co-author of the allRank package, which is an easy tool for research of learning-to-rank algorithms. I worked on some research efforts within learning-to-rank that are yet to be published.

I have experience both on the engineering side (developing data pipelines for training data and predictions as well as microservices) and on the machine learning side (data analysis, creating custom architectures and loss functions to train models). I don't want to be specialized in any particular domain – I have worked CV, NLP, Learning to rank and I see the advantage in diversity of experience.

I am currently leading a machine learning research team that is working on multimodal search relevancy system – both on the research side creating new methods for handling the problem and on the production side applying current state-of-the-art to our domain problem.

I am an organised, open-minded person that always aims for improvement.
I have an Agile mindset, believing in tight feedback loops and close cooperation with client and team. In terms of machine learning I usually want to start small but think big – i.e. verify the simplest possible method that could have an impact and iterate to a complex solution.
I put strong emphasis on keeping and making code understandable and clean.

Technologies

I have experience with PyTorch, Spark(both Python and Scala), Python, XGBoost, Scikit-learn, Airflow, listing just the most important. I believe one should pick the right tool for the job and so it's more important to have a general intuition about what a technology offer and be able to learn it when needed.
I have a deep understaning of current state of the learning-to-rank research progress
I try to be up-to-date with novelty in general Machine Learning, NLP and CV.

Projects

Allegro 2017 – currently: **Allegro Relevancy Sorting**. We are developing a relevancy sorting algorithm for Allegro full-text search. This is an example of 'Learning to Rank' problem where using historical data consifering queries and offers we learn a ranking function that will maximise a utility of a list of offers to the user. From the initial work on a linear regression algorithm we moved to GBDT algorithm and now are on our way to deploying a deep neural network ranker.
I am the main contributor of this project and was very happy to observe an significant gains in business and monetary metrics when this model got deployed. We worked on data pipeline, feature engineering, model training up to a microservice implementation and deployment serving up to 2K ranking RPS. I am currently

leading a team with which we are expanding this model further and we are experimenting towards a next step which will be a neural ranking model – improving our performance substantially, including text, image, categorical and user data in a single end-to-end manner (PyTorch/allRank).

Allegro 2016.06 – 2016.12: Image quality assessment model. We were creating a model to assess whether a particular product photo was made according to a company standards (i.e. it didn't have any artificial text added, artificial frame around the product and was done on a bright background). We pushed the system from a handcrafted features to a fine-tuned CNN model trained in a multiclass classification setting. My contribution was threefold: 1. I insisted on and coordinated getting a human-labelled dataset for training and evaluation 2. I worked on an implementation of the CNN model (Keras/TensorFlow) 3. I contributed to the pipeline that was scoring every new image on the platform in a streaming fashion.

Allegro 06.2014 – 01.2016: Allegro Recommendation Engine. We built a service that is providing user with contextual and personalized recommendations for offers in the marketplace (those can be found on the allegro site under slogans 'you may like this', 'users similiar to you found this interesting' and similiar). We spend most of the time adapting industry-standard algorithms like Collaborative Filtering for a specific context of our business.
Using Apache Spark with Scala for processing and Scala/Spray! on top of Cassandra and ElasticSearch as a webservice – we are able to constantly provide recommendations for many different placements on the site.

Conferences/events I contributed to:
2015 Codepot- 'Data Processing with Apache Spark' workshop for 10 persons (code)
2015 Scalar.mini Introduction to Apache Spark
2015 Scala Forty-fives Batch processing with Apache Spark (video)
2017 PyData Warsaw "Data and test driven product development with Airflow, Jupyter and (Py)Spark" (video)
2018 PLinML with Ireneusz Gawlik "Learning to rank at scale" (video)
2018 PyData Warsaw with Ireneusz Gawlik "Learning to rank @ allegro.pl" (video)
2018 Allegro tech with Ireneusz Gawlik "A case-study of CNNs and transfer learning for image quality assessment at Allegro" (video)
2019 PyData Warsaw with Radosław Białobrzeski „Learning to rank with the Transformer" (video)

Professional attitude

I participated in a few courses on Coursera and Udacity, with the important ones being:
- Regression Models
- Machine Learning
- Mining Massive Datasets
- Introduction to Recommender Systems
- Principles of Reactive Programming
- Functional Programming Principles in Scala

In the recent years I read dozens of research publications in the field of machine learning.