Valentin Malykh

Research Scientist

📞 +7 (903) 966 8133
✉ valentin@maly.hk
🖳 val.maly.hk
GitHub: madrugado
Google Scholar: https://goo.gl/b6UPb5

Education

2019 (expected)   **PhD**, *Institute for Systems Programming, Russian Academy of Sciences*, Moscow.
Thesis: Noise Robustness in Various NLP Tasks

2009   **M.Sc.**, *Moscow Institute of Physics and Technology*, Moscow.
Major: Computer Science

2007   **B.Sc.**, *Moscow Institute of Physics and Technology*, Moscow.
Major: Computer Engineering

Experience

Oct 2018–p.t.   **Applied Research Scientist**, *VK.com*, Saint-Petersburg.
Work on research problems related to social network, like text classification, summarization, etc. Accepted paper at ECIR'2019 conference. Shared task on headline generation at Dialogue'2019 conference.

2016–p.t.   **Research Scientist**, *Laboratory of Neural Systems and Deep Learning, Moscow Institute of Physics and Technology*, Dolgoprudny, Moscow Region.
Work on `iPavlov.ai` project. Research on noise robustness in NLP tasks.
Detailed achievements:
- Published 15 papers,
- including ACL Demo Paper on DeepPavlov library,
- & 4 papers in NIPS Proceedings on ConvAI Challenge series.

2016–p.t.   **Lecturer**, *Moscow Institute of Physics and Technology*, Dolgoprudny, Moscow Region.
I was teaching Deep Learning in NLP course in 2016-2018. The cumulative audience of that course is estimated to 3000 people. Also I'm teaching Neural Networks and Reinforcement Learning modules in Machine Learning course.

2015–2016   **Developer**, *Yandex*, Moscow.
I was working in Yandex.News on the whole ML stack for ranking and also participated in clusterization development.
Detailed achievements:
- New ranking formula for news clusters.
- Improved news agency ranking.

2014–2015 **Research Engineer**, *Sputnik*, Moscow.

Sputnik is a Russian government-sponsored search engine, which was developed from scratch by a small team of engineers. I was in the Search Quality Department.
Detailed achievements:
○ Web pages classifier.
○ Malicious documents ranking.

2012–2014 **Research Engineer**, *Cognitive Technologies*, Moscow.

Cognitive Technologies is a company with long history of work in computer vision domain. It created second most common OCR solution in Russia. Nowadays company's main interest is self-driving cars.
Detailed achievements:
○ Computer vision & control for a robot car.
○ Research at Astarta project – high-load document classifier.

Miscellaneous

2017 **Certified Instructor**, *NVIDIA Deep Learning Institute*.

In addition to being a Certified Instructor, I have authored whole NLP Workshop in NVIDIA DLI and presented it at GTC EU in 2017.

2017&2018 **Organizer**, *ConvAI Challenge series*.

ConvAI Challenge is devoted to creation of human-level conversational intelligence. I was co-organizing in both scientific and technical areas.

2015 **Winner**, *DeepHack.Game*.

DeepHack.Game was a hackathon on Reinforcement Learning applied to Atari games. URL: `http://game.deephack.me`. Our team is a winner of qualification round, and has overall 2nd place.

Publications

○ Generalizable Architecture for Robust Word Vectors Tested by Noisy Paraphrases, V. Malykh, The 6th International Conference On Analysis Of Images, Social Networks, And Texts (AIST), 2017
○ The first conversational intelligence challenge. M. Burtsev et al. NIPS 2018 Proceedings Challenge Chapter.
○ ConvAI dataset of topic-oriented human-to-bot dialogues. V. Logacheva, M. Burtsev, V. Malykh. NIPS 2018 Proceedings Challenge Chapter.
○ DeepPavlov: Open-Source Library for Dialogue Systems. M. Burtsev et al. Proceedings of ACL 2018 Demo Track, Melbourne, Australia, 2018.
○ Robust Word Vectors: Context-Informed Embeddings for Noisy Texts. V.Malykh et al. Proceedings of 4th Workshop on Noisy User Texts, Empirical Methods in Natural Language Processing, Brussels, Belgium, 2018.
○ Alekseev, A., Malykh, V., Nikolenko, S., Shenbin, I., and Tutubalina, E. AspeRa: Aspect-based rating prediction model. 41st European Conference on Information Retrieval, 2019.
○ Gavrilov, D., Kalaidin, P., Malykh, V. Self-attentinve model for headline generation. 41st European Conference on Information Retrieval, 2019.

Languages

English	Fluent	*IELTS Academic 7.0*
Russian	Native speaker	