Agnieszka Mikołajczyk-Bareła

NLP Team Leader at Voicelab.AI, actively organizing "AI for Good" projects such as HearAI or detect waste. Author of scientific papers and publications interested in deep learning. Experienced in both NLP and Computer Vision.

LINKS & CONTACT

agnieszka.mikolajczyk.pg@gmail.com | 737 417 715
linkedIn.com/in/**agnieszkamikolajczyk** | twitter.com/ **agnmikolajczyk** | github.com/ **AgaMiko** |
google.scholar.com/ | **personal website**

EDUCATION

GDAŃSK UNIVERSITY OF TECHNOLOGY | Research Doctorate
September 2017 - 2022
Successfully defended with special award in January 2023. Thesis title: *Data augmentation and explainability for bias discovery and mitigation in deep learning.* Member of PhD students council for four years.

GDAŃSK UNIVERSITY OF TECHNOLOGY | Master's Degree
March 2016 - September 2017
Thesis title: *Utilization of the computational intelligence methods for diagnosis and detection of the abnormal states selected applications in medicine*
Awarded with scholarship for the best students. Best Master's thesis winner (Young Innovators 2018, Best Graduating Thesis in Innovative Technologies 2018, Best department thesis 2018).

GDAŃSK UNIVERSITY OF TECHNOLOGY | Bachelor's Degree
September 2012 - March 2016
Thesis title: *Analysis of skin lesions with the usage of image processing and computational intelligence methods and algorithms*
Awarded with scholarship for the best students. Best Bachelor's thesis winner (Young Innovators 2016).

EXPERIENCE

VOICELAB.AI - CONVERSATIONAL INTELLIGENCE | NLP Team Leader
December 2021 – Present | Gdańsk, Poland

Leading NLP team in R&D Lab Voicelab.ai. Working mainly with speech transcripts and natural conversations. Designed and prepared training pipelines, data, trained and tested models such as transformers, T5, BART, GPT, creating proof of concepts (PoC) and demos. Main person responsible for training and developing vlT5 model for keyword generation (currently TOP1 model for keyword generation on HuggingFace). Lead the team that created the first ChatGPT alternative in Poland called TRURL.AI. Taking care of models and spaces on huggingface.co/Voicelab (over 40 private models).

| Machine Learning Researcher
March 2020 – December 2021 | Gdańsk, Poland

For the first few months I have worked in ASR team on voice style transfer and was moved to NLP team to work on multimodal sentiment analysis. Later on I was working in the NLP team focusing on telephone conversations. I have worked with punctuation restoration, sentiment analysis, topic modelling, intent recognition, NER, keyword generation, and many others.

DESIGNBOTIC | ML Expert
March 2023 – Present | Sopot, Poland

Assisting the process development of an AI application for architectural studios in the role of Machine Learning Expert. Working on image classification with custom CNNs (PyTorch, ONNX).

GDAŃSK UNIVERSITY OF TECHNOLOGY | Research Assistant
June 2018 – Present | Gdańsk, Poland
Currently finishing my research project on bias ("systematic errors") in models and data, using the example of skin lesion classification.

ZATOICHI | ML Expert
March 2022 – September 2022 | Warsaw, Poland

Assisting the process development of an AI application for detecting and reading loud product's expiry date with mobile camera to help blind and visually impaired people. I was responsible for the algorithm design, consulting with training and experiments. Working on detection and OCR (PyTorch, OpenCV, various OCRs).

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY | Visiting Researcher at Explainable Artificial Intelligence Center
May 2019 – June 2019 | Ulsan, South Korea

Working on Explainable AI methods for local and global explainability.

INSTITUTE OF ELECTRICAL ENGINEERING | Research Assistant and C++ Software Developer
July - August 2013 & July 2014 – November 2016 | Gdańsk, Poland
Working in Electromagnetic Compatibility Laboratory. Developing software for automatic experiments inference.

RECENT PAPERS

1. TRANSFERABLE KEYWORD EXTRACTION AND GENERATION WITH TEXT-TO-TEXT LANGUAGE MODELS Pęzik, P., Mikołajczyk, A., Wawrzyński, A., Żarnecki, F., Nitoń, B., Ogrodniczuk, M. (2023, June). In International Conference on Computational Science (pp. 398-405). Cham: Springer Nature Switzerland.

2. THE (DE)BIASING EFFECT OF GAN-BASED AUGMENTATION METHODS ON SKIN LESION IMAGES, Agnieszka Mikolajczyk, Sylwia Majchrowska, Sandra Carrasco (2022) – MICCAI 2022, Singapur (paper was among the top 13% that received a provisional accept recommendation)

3. DEEP LEARNING-BASED WASTE DETECTION IN NATURAL AND URBAN ENVIRONMENTS, Majchrowska, S., Mikołajczyk, A., Ferlin, M., Klawikowska, Z., Plantykow, M. A., Kwasigroch, A., Majek, K. (2022). Waste Management, 138, 274-284.

AI FOR GOOD (VOLUNTEER)

HEARAI - SIGN LANGUAGE RECOGNITION | Principal ML Researcher
May 2021 – Present | Women in AI
The project involved automatic translation of sign language from video to text using the HamNoSys notation. I co-organized the project, proposed a research plan, and lead the team in the project *HearAI* for recognizing the Polish Sign Language. The research team included 12 people of which seven were recruited out of 75 applications. We worked directly on video using multimodal custom models with Vision Transformers, CNNs, and pose estimation (MediaPipe/OpenPose). The pipeline was breifly explained at: HearAI.pl blog - Pipeline

DETECT WASTE | ML Lead
September 2020 – March 2021 | Women in Machine Learning and Data Science
Organized and started an AI for Good project for detecting waste in natural and urban environments. Led a team of nine women during a six-month project. Organized a hackathon Hack4Environment, published results on our blog, extended existing datasets, and finally, prepared a research paper (published in Waste Management IF = 6.178). We used a two-stage detector to detect waste and classify it into seven waste categories inspired by waste categories in Gdańsk. See more at: detectwaste.netlify.app/

BIRD SONG RECOGNITION | ML Researcher
June 2019 – December 2019 | Women in Machine Learning and Data Science
Joined a team in Gdańsk. Prepared a deep learning model to classify bird songs of the most common Polish birds. Each recording was first filtered and then saved as a Melspectrogram. Used CNN's to classify melspectrograms. Prepared an open repository of existing papers, projects, and datasets related to bird song and call recognition.

OTHER RESEARCH PROJECTS

DETECTING AND OVERCOMING BIAS IN DATA WITH EXPLAINABLE ARTIFICIAL INTELLIGENCE | Principal Investigator
June 2020 – Present | Gdańsk University of Technology (Preludium Grant)
The project is divided into three research tasks: Developing globally aware local explanations for prediction justification; Developing global explanations for detecting undesirable bias in data; Eliminating influences of undesirable bias in data on the model.

OTHER

2023	Team Leader of the first team that developed the first ChatGPT alternative in Poland (TRURL.AI)
2023	Over 1500 citations on Google Scholar
2022	Co-organizer of PolEval 2022
2022	Co-created the Polish Corpus for Punctuation Prediction
2021	Team member of organizers of the Task 1 form PolEval 2021 Competition
2021	Co-created the first Polish Corpus for Punctuation Restoration
2020	Co-organizer of the Hackathon *Hack4Environment*
2019	First award for the best paper in 2018 in *Scientific Papers of the Faculty of Electrical and Control Engineering at GUT*
2018	Member of competition committee IAV Cup 2018
2018	Second award for the best graduating thesis in innovative technologies
2018	*Young Author Best Paper Award* for an oral presentation
2018	First award in the contest *MŁODZI INNOWACYJNI 2018* for the best master's thesis
2018	First award in the contest *MŁODZI INNOWACYJNI 2016* for the best graduation bachelor's thesis