Michał Bień

✉ michal@mbien.pl   ⌂ Glorf

EXPERIENCE

Allegro.pl   **Poznań, PL**
Research Engineer - Machine Learning   *September 2021 -*
- Developed recommender systems for over 400M products catalogue increasing CTR and conversion
- Gave speeches and represented my company on international conferences
- Closely operating with business, led the project for complementary product recommendation
- **Technologies: Python, SQL, Tensorflow, Kubeflow, Google Cloud Platform, Vertex AI, BigQuery, Recommender Systems, Information Retrieval, Pandas, Dask, NVIDIA Merlin, Kubernetes, Docker, PyTorch, ONNX**

EPFL Data Science Laboratory (DLAB)   **Lausanne, CH**
Data Scientist - Student Assistant   *February 2021 - August 2021*
- Developed automated solution for building rich social surveys from text data
- Performed succesful data analysis which resulted in a scientific publication
- **Technologies: Python, SQL, Pandas, Dask, NumPy, Matplotlib, scikit-learn, Docker**

Applica.ai (acquired by Snowflake)   **Poznań, PL**
Cloud Engineer   *September 2019 - August 2020*
- Managed product's critical infrastructure reliability and scalability
- Created cloud-native installer for one-click deployments, including on-premises.
- **Technologies: Python, Bash, Kubernetes, Helm, Terraform, Ansible, Docker, AWS**

CERN   **Geneva, CH**
Openlab Summer Student   *July 2019 - August 2019*
- Participated in Openlab Summer Student Programme as one of 40 selected participants
- Deployed ML pipelines at PB scale with Spark and Kubernetes
- **Technologies: Python, Scala, Spark, Hadoop, Kubernetes, Kafka, TensorFlow**

Poznań Supercomputing and Networking Center - GÉANT   **Poznań, PL**
Backend/Cloud Developer   *May 2018 - June 2019*
- Developed infrastructure-management platform for large European research project
- Managed Kubernetes deployments automation in distributed cluster system
- **Technologies: Java, Spring, Hibernate, PostgreSQL, Docker, Kubernetes, Helm**

EDUCATION

École polytechnique fédérale de Lausanne   **Lausanne, CH**
MSc in Digital Humanities   *September 2020 - July 2022*
- Studied DH with regard to ML & NLP applications
- Thesis: "NLGRank: A discriminative evaluation for natural language generation metrics" with A. Bosselut
- Extracurricular: ALPS 2021 (Advanced Language Processing School) organized by LIG Grenoble

Poznań University of Technology   **Poznań, PL**
BSc in Computer Science   *October 2016 - February 2020*
- Studied CS with regard to ML & NLP
- Thesis: "Cooking recipe generator utilizing deep learning-based language model" with A. Ławrynowicz
- Extracurricular: CLARIN workshops on NLP resources for polish language processing

SELECTED PROJECTS

Epicer   *2021 - 2023*
- Served as a R&D head of a startup dedicated to NLG applications
- Trained custom LLM solution based on GPT-3 architecture
- **Technologies: Python, Rust, Dart, Huggingface Transformers, Google Cloud Platform, Firebase, PyTorch, Cloud TPU, Cloud Run, Pandas, NumPy, Docker, Github, Flutter, JIRA**

FDH Wikibio   *September 2020 - December 2020*
- Built PoC disinformation tool which generates fake biographies uploaded to Wikipedia using GPT-2
- Project details: `http://fdh.epfl.ch/index.php/WikiBio`
- **Technologies: Python, SPARQL, GPT-2, Huggingface Transformers, PyTorch**

PUBLICATIONS

On the Context-Free Ambiguity of Emoji	**ICWSM**
J Częstochowska,K Gligorić,M Peyrard,Y Mentha,M Bień,A Grutter,A Auer,A Xanthos,R West	*2022*
RecipeNLG: A cooking recipes dataset for semi-structured text generation	**INLG**
M Bień, M Gilski, M Maciejewska, W Taisner, D Wiśniewski, A Ławrynowicz	*2020*
Big Data Analysis and Machine Learning at Scale with Oracle Cloud	**CERN**
M Bień, R Castellotti, L Canali	*2019*

INVITED TALKS

Multilingual Translation at Scale: Large Language Models in Allegro	**Poznan, PL**
GHOST Day - Applied Machine Learning Conference	*April 2023*
Information Retrieval at Scale in Allegro	**Prague, CZ**
sMallTalk - Mall Group Tech Seminar	*November 2022*
RecipeNLG: A dataset for semi-structured text generation	**Grenoble, FR**
Advanced Language Processing School 2021	*January 2021*
AI at CERN: Big Data and Machine Learning in the Cloud	**Warsaw, PL**
ML in PL Conference	*November 2019*
Big Data Analysis and ML at Scale with Oracle Cloud Infrastructure	**Geneva, CH**
CMS Machine Learning Forum	*August 2019*

LANGUAGES

- **Polish:** Mother tongue
- **English:** C2
- **French:** B1 (bilingual high-school certificate)