Maksim Shmalts

• +49 176 32626044   • schmaltzmax@gmail.com
• linkedin.com/in/maksim-shmalts   • github.com/maxschmaltz

• NLP / LLM engineer   • 5 years theoretical background plus practical expertise in NLP
• 3 years professional experience in the area   • Fast learner with advanced analytical skills

Work Experience

2023-2024 — **Working Student for LLMs**, IAV GmbH
- Implemented and integrated a tool for evaluation of LLMs and RAG pipelines (`Python`, `LangChain`, `Ragas`, `OpenAI`)
- Currently, connecting OpenAI function calling to our internal on-premise models (`Python`, `TGI`, `LangChain`)

2021-2023 — **Data Scientist / NLP Engineer**, MKSKOM
- Implemented backend for a custom Llama-2-based Retrieval-Augmented Generation Engine for a large customer (`Python`, `LangChain`, `Llama.cpp`, `HF Transformers`)
- Increased the number of responses from potential customers on a freelance platform by 5 times by independently designing and implementing an LLM-based tool for automatic search and contacting potential customers (`Python`, `LangChain`, `OpenAI`, `Puzzle` [internal tool])
- Implemented tools for Natural Language Processing, Time Series Prediction, Optimization Modeling, Data Analysis (`Python`, `PyTorch`, `spaCy`, `scikit-learn`, `statsmodels`, `Pandas`, `NumPy`, `streamlit`)
- Participated in hackathons for RL and Anomaly Prediction (`Python`, `scikit-learn`, `statsmodels`)
- Communicated with potential customers

2022-2023 — **Tutor HiWi**, Eberhards Karls Universität Tübingen
- Held tutorials and contributed into evaluating assignments and projects for the courses "String Algorithms" (WiSe22/23), "Python for Beginners" (WiSe22/23), "Statistical Language Processing II" (SoSe23), "Reinforcement Learning for LLMs" (SoSe24)
- Created Finite-States Transducers for measuring distance between German dialects for the course "String Algorithms" (`Python`, `pynini`)

2021-2022 — **Assessor**, Yandex
- Evaluated search engine results on IT-themed queries, evaluated machine translations

2021-2022 — **Course Instructor (Linguistics for Olympiades)**, Lomonosov Moscow State University Gymnasium
- Held lectures, prepared and evaluated assignments for the Course "Linguistics for Olympiads"

Education

2022-2024 — **Eberhard Karls Universität Tübingen**, Bachelor (in progress): Computational Linguistics
- 2024 Bachelor thesis on applying LLMs to German compounds (`Python`, `LangChain`, `PyTorch`, `HF transformers`)
- 2024 Group project on investigating influence of RL fine-tuning data on biasing LLMs (`Python`, `HF transformers`)
- 2023 Participated at the SemEval 2023 and published a paper at the ACL Anthology (`Python`, `HF transformers`)

2019-2022 — **Lomonosov Moscow State University**, Bachelor (incomplete): Theoretical and Applied Linguistics
- 2022 Personal project DERBI: a tool for automatic inflection of German words (`Python`)
- 2022 Conference diploma for DERBI at Lomonosov-2022 (Moscow), Science Sessions-2022 (Kaliningrad), XXII International Conference of Young Slavists (Tallinn)
- 2021 Term paper on predicting ablaut class of German strong verbs (`Python`, `PyTorch`)

2017-2019 — **Lomonosov Moscow State University Gymnasium**, High School: History and Philology
- 2019 Prize-winner of All-Russian Olympiad for Linguistics (gives no-exam admission to the university of choice)
- 2019 Graduation with a Gold Medal

Publications

2023 — **John Boy Walton at SemEval-2023 Task 5: An Ensemble Approach to Spoiler Classification and Retrieval for Clickbait Spoiling**, ACL Anthology, Toronto

The paper summarizes my approach to the SemEval-2023 Task 5 that uses an ensemble of encoder and seq-to-seq Transformer-based LMs.

2022 — **Derbi: Deutscher regelbasierter Inflektor**, Studia Slavica XX, Tallinn (issued 2024) **&**
DERBI: DEutscher RegelBasierter Inflektor, Faktor14, Tübingen

The papers give a brief overview on DERBI, my model for automatic inflection in German.

Facts

Natural languages	Russian (native), English (C1), German (B2), Spanish (A2)
Programming languages	`Python`, `R`, `C#`, `Java`
Tech stack	`LangChain`, `HF Transformers`, `spaCy`, `PyTorch`, `scikit-learn`, `Pandas`, `NumPy`; `Docker`, `GitLab` (CI/CD), `Unix` tools
Misc achievements	Prize winner for various national and international music competitions, driving license categories A and B, startup founder (now closed)
Interests	Music, sports (in particular, kickboxing, table tennis, volleyball), history, philosophy, traveling