

Yerkezhan Abdullayeva

Berlin, Germany

NLP Data Analyst | Machine Learning Engineer | AI Researcher

 +4915752668911

 yerkezhan.abdullayeva@gmail.com

 [Portfolio](#)

 [Yerkesoul](#)

 [Yerkezhan](#)

Education

Master of Science in Cognitive Systems Universität Potsdam <i>Major: Advanced NLP, Machine Learning, Data Science, Artificial intelligence, Deep learning</i>	<i>October 2020 - September 2024</i> Germany
Bachelor of Computational Linguistics Al-Farabi Kazakh National University	<i>September 2016 - June 2020</i> Kazakhstan

Skills

- **Programming Languages:** Python, SQL
- **Data Analysis:** Pandas, NumPy, SciPy, Matplotlib, Seaborn, Power BI
- **Machine Learning & AI:** PyTorch, TensorFlow, Keras, Scikit-learn, HuggingFace, NetworkX, LangChain
- **Natural Language Processing (NLP):** Named Entity Recognition (NER), Text Classification, Data Mining, Event Extraction, Generative AI, Spacy, NLTK, RAG
- **Cloud & Data Engineering:** AWS, Apache Kafka, PySpark
- **CI/CD Tools:** , Git, JIRA Atlassian, Visual Studio Code, Flask, Django

Work Experience

Natural Language Processing Data Analyst UNICEPTA Corporate Intelligence (Berlin)	<i>April 2022 - July 2023</i>
<ul style="list-style-type: none">• Developed data mining and machine learning modules for named entity recognition (NER), event extraction, and relationship extraction from unstructured texts using Python.• A team of 2 people (myself included) replaced a department with a model• Enhanced NLP frameworks and libraries, improving accuracy and efficiency in text processing tasks.• Processed and analyzed large-scale unstructured textual data and data visualizations to inform business decisions.	

Certificates

[AWS Certified Cloud Practitioner](#)
Principles of Computing (Part 1 and 2): Coursera

Projects

Master Thesis in Generative AI: Assessment of Map Navigation and Spatial Reasoning Abilities of Large Language Models	2024
<ul style="list-style-type: none">• Developed five Python-based map traversal games to assess the spatial reasoning capabilities of LLMs.• Tested 15 large language models (GPT-4, Llama-3, Claude-3) through API calls and conducted quantitative and qualitative analysis.• Results were published as part of a conference paper. DOI link: https://doi.org/10.48550/arXiv.2406.14035.	
Individual Research: Environmental Agenda Detection	2023
<ul style="list-style-type: none">• Built text classification models using political data from the Manifesto Corpus, developed classification models using BERT, RoBERTa, and XLM-RoBERTa for accurate text detection.	
Computer Vision + NLP: GuessWhat game	2022
<ul style="list-style-type: none">• Building a ensemble model with ResNET and LSTM for the dialogue between agents of the game	
Chat Bot for Charity project Arman	
<ul style="list-style-type: none">• Telegram chat bot development with deployment on AWS cloud server with an SQL database	2024

Languages

English – Advanced (C1/C2), Russian – Advanced (C1/C2), German – Intermediate (B1/B2), Kazakh – Native, Chinese – Beginner (A1)