

Olga Seleznova

MACHINE LEARNING ENGINEER

Professional Summary

Machine learning engineer with 3 years of experience in startup environments and research laboratories. Specializing in deep learning modeling and Python programming to address key challenges in natural language processing and speech recognition. Advanced degree in Data Science, complemented by AI and machine learning technologies certifications. Excited to ____ .

Strengths

- Passion for learning and exploring new technologies, showcased in my GitHub portfolio.
- Strong problem-solving and analytical skills.
- Excellent communication and presentation skills.

Technical skills

- Proficient in Python programming.
- Extensive experience with large sets of structured and unstructured data for machine learning, NLP, and speech recognition.
- Skilled in using ML frameworks such as PyTorch, HuggingFace, and Weights & Biases.
- Knowledge of standard ML techniques and algorithms.
- Experience building, training, fine-tuning, and evaluating deep learning models.
- Experienced with version control system Git, and GitHub service.
- Proficient in building and tuning natural language processing (NLP) systems using transformer architectures and classical tools like SpaCy, NLTK, Flair, and Gensim.
- Proficient in building and tuning speech recognition systems using TorchAudio, Wave2Vec, HuBERT, and Librosa.
- Hands-on experience working with large language models (LLMs) and AI frameworks like Llama, LangChain, LlamaIndex, and OpenAI.
- Experience using cloud technologies like AWS, container application Docker, and managing Linux systems.
- Ability to initiate, drive, and lead projects to completion with minimal supervision.
- Participation in Kaggle competitions and knowledge sharing with Medium.

Work experience

03/2023 - 05/2024. Research Engineer, Technion - Israel Institute of Technology

- Worked at "Speech, Language, and Deep Learning" Lab, leading modeling and research of deep learning projects in the speech and language domains, using Python, PyTorch, Whisper, Torch audio, NumPy, matplotlib, and other relevant technology.
- Developed over 4 clean and reusable packages for academic purposes by creating templates and streamlining submission processes.
- Guided 5 students through their technical challenges, optimizing deep learning project outcomes and enhancing research efficiency.

09/2021 - 09/2022. Data Scientist, Skai

- Mapped clients' needs by mining insights from eCommerce reviews, enhancing customer understanding and product development.
- Led research and implementation of an NLP system, including modeling, fine-tuning, and error analysis.
- Trained text classification models (BERT) and question-answering models (T5) to extract knowledge from consumer reviews.
- Built semantic similarity and keyword extraction modules to empower insights mining.

02/2021 - 07/2021. Data science intern, TRG Research and Development

- Performed Named Entity Recognition and built models to predict sentiments for social media posts: Twitter and Facebook.
- Improved baseline model, which predicted sentiment for three classes (positive, negative, and neutral) from an accuracy of 47% to 86.20% by fine-tuning the DistilBERT model from HuggingFace.

02/2021 - 07/2021. Researcher, University of Haifa

- Achieved an accuracy of around 85% in the classification of Hebrew posts from medical forums for low-scale, unbalanced data.
- Performed EDA, text preprocessing, and classification with Python libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, and transformers with PyTorch.

Projects

VoyageVocab. Language Learning assistant for travelers.

An application that uses Llama and LangChain to generate phrases useful for traveling abroad, given specific tasks such as ordering food or buying a ticket. The Gradio package serves as a simple user interface.

Link: github.com/OlgaSeleznova/VoyageVocab

Publications

Kim, S.-E., Chernyak, B. R., Seleznova, O., Keshet, J., Goldrick, M., & Bradlow, A. R. (2024). Automatic recognition of Second language speech-in-noise. JASA Express Letters, 4(2). <https://doi.org/10.1121/10.0024877>.

Education

Career advancement program in Data Science, Ydata.

Tel Aviv, Israel. 2020 - 2021.

Master's degree in Data Science, University of Haifa.

Haifa, Israel. 2018 - 2022.

Certifications

- Generative AI with Large Language. Coursera, DeepLearning AI.
- Introduction to Machine Learning in Production. Coursera, DeepLearning AI.
- Mathematics for Machine Learning: Multivariate Calculus and Linear Algebra.
- Coursera, Imperial London College. Data Structures and Algorithms in Python. Udacity.
- Using Python to Access Web Data. Coursera, University of Michigan.
- Python Data Structures. Coursera, University of Michigan.