+386 40914122
luka.krsnik@outsmartify.com
LinkedIn
Github
Website

Luka Krsnik

Data Scientist

Summary

A data scientist, with 6 years of experience working on various natural language processing tasks. Seeking a practical challenge where I can use and expand my knowledge and experience.

Education

Master's degree
Computer science
University of Ljubljana
2014 - 2017
Ljubljana, Slovenia

Bachelor of science
Computer science
University of Ljubljana
2012 - 2014
Ljubljana, Slovenia

Skills

Python (PyTorch, Pandas, Scikit-learn, NumPy, Flask)
SQL (SQLite, Postgres)
Data Science
Machine Learning
Natural Language Processing
Neural Networks
BERT

Publications

- Enhancing deep neural networks with morphological information

Work experience

Data Scientist Mar 2017 – Sep 2023

Centre for Language Resources and Technologies

Collaborated closely with linguist researchers, to develop tools for text analysis and automated annotations.

- Enhanced and developed multiple pipelines actively utilized for research projects (STARK, cordex and Classla)
- Designed and implemented several LSTM neural networks, improving annotation quality across various languages (Classla, stress assignment)
- Led the development of a BERT-based tweet selection tool, instrumental in selecting tweets for subsequent manual annotation (standardness)

Data Scientist Aug 2018 – Oct 2021

University of Ljubljana, Faculty of Computer and Information Science

Participated in research involving cross-lingual embeddings and multilingual models (multilingual BERT)

- Designed experiments to assess the impact of adding morphological data on the performance of neural networks (BERT, Fasttext+LSTMs).
- Conducted end-to-end experiments across multiple languages, applying this approach to Named Entity Recognition (NER) and subsequently published significant findings in the Natural Language Engineering journal.
- Explored cross-lingual model transfer through experiments involving embeddings and anchor points (anchor points).