

SANJA SIMONOVIKJ

sanja_sss@yahoo.com ♦ sanjass.github.io ♦ github.com/sanjass

EDUCATION

Massachusetts Institute of Technology

Master of Engineering, Concentration in Artificial Intelligence (GPA: 5.0/5.0)

Jun 2020 - Jun 2021

Bachelor of Science, Computer Science and Engineering (GPA: 4.7/5.0)

Sep 2016 - May 2020

SELECTED INDUSTRY EXPERIENCE

Indeed

Austin, TX

Senior Data Scientist, Metadata and Taxonomy Team

Jan 2023 - present

Data Scientist, Metadata and Taxonomy Team

Jul 2021 - Jan 2023

- Building and improving production-level NLP deep learning models for metadata extraction from jobs and resumes, currently powering core Indeed features such as matching and ranking, both in the US and internationally
- Leading a key initiative to transition 10+ core Indeed markets to a fully machine learning powered approach to job metadata extraction, providing more than 15% increase in key coverage and quality metrics (OKRs)
- Conceptualized and improved upon numerous human-in-the-loop ML model-based workflows, facilitating ongoing model improvement and expansion of the occupation Taxonomy in over 12 new international markets
- Built and deployed custom and reusable ML components and metrics, enabling faster iteration of Data Science experiments, improved analysis, and more informed strategic product decisions
- Designed and implemented modular and extensible internal tooling that is used to enable crucial Data Science and Taxonomy processes, including performance monitoring, category investigations, and model development

Microsoft

Cambridge, MA

Data Science Intern, Docs ML team, Microsoft Fluid Office

Jan 2021

- Created a prototype of an intelligent component in Office predicting when text can be represented as a table
- Engineered features and optimized lightweight models that perform at least 20% better than a heuristic baseline

Qualcomm

Cambridge, MA

Software Engineering Intern

Jun-Aug 2019

- Performed automatic triage of build and test failures by detecting failure signatures using NLP techniques
- Built and deployed an end-to-end framework, reducing team's triaging and debugging time from hours to seconds

SELECTED RESEARCH AND TEACHING EXPERIENCE

MIT EECS Department Graduate Teaching Assistant for a popular ML course (6.036)

Feb - May 2021

Improbable AI Group, MIT CSAIL Graduate Research Assistant - Prof. Pulkit Agrawal

Feb 2020 - Jun 2021

ALFA Group, MIT CSAIL Undergraduate Researcher - Prof. Una-May O'Reilly

Feb 2019 - May 2020

SELECTED AWARDS AND PROJECTS

- **Indeed Hackathon:** project Indeed Rhizome (Social Graph) - #1 place overall, Global winner Aug 2021
- **HackMIT:** project LabelLearn, #1 place overall, Grand Prize and track hackathon winner Sept 2019

SKILLS

Programming Languages: Python, Java, Javascript, Bash, R, SQL, Julia

Technologies: Git, Docker, AWS (S3, EC2, Lambda, CloudWatch, EBS), Elasticsearch, Mongo, Django, Flask

ML/data tools: TensorFlow, TFX, Pytorch, Kubeflow, Streamlit, standard ML packages (pandas, dask, pyspark, numpy, scikit-learn, keras), Jupyter Notebooks, Apache Beam, Apache Spark, Optuna

Human languages: English, Macedonian (native), Serbian (proficient), Spanish (intermediate), French (limited)