Resume

Table of contents

Basics

Name	**Milena Šošić**
Label	Mathematician & computer scientist
Email	
Url	
Summary	Machine Learning expert with eighteen years of experience in data mining, text mining, business intelligence and general software algorithms development and implementation. Worked on high-visibility projects in the commercial, banking and telecommunications areas. PhD candidate in computer science, NLP specialization.

Work

2021.01 - PRESENT

Senior Data Scientist

GridDynamics

Design and implement new solutions with focus on predictive analytics. Run projects in NLP domain using newest machine learning technologies.

- Data Modeling · Analytics · Artificial Neural Networks · PyTorch · Keras · TensorFlow

2020.02 - 2022.01

Advisory Board Member

Serbian AI Society

Serbian AI Society is being launched in Serbia by the most prominent people in artificial intelligence, with one goal - networking and strengthening the domestic AI scene. The aim of this organization is to bring together all stakeholders dealing with this subject which will contribute to the presence and growth of artificial intelligence in Serbia.

- AI · Growth · Networking · Strength · Serbia

2019.02 - 2021.01

AI Software Developer

Devtech

Lead AI team, mentor and drive AI efforts. Develop PoC solutions for the projects by adding AI components.

- AI · Data · Mentoring · PoC · Project Management

2016.01 - 2019.02

Machine Learning Team Lead

Selerant

Manage state-of-the-art machine learning tasks for text segmentation and classification. Shape product features according to the business requirements.

- Text classification · Semantic search · PoC · Project Management

2013.12 - 2015.12

Data Miner

Telenor

Develop different up-sell, cross-sell and churn models. Support risk models development. Accountable for segmentation, customer lifetime value calculation.

- Customer classification · Credit risk · NPS scoring

Education

2020.10 - PRESENT

PhD Candidate

Mathematics Faculty, University of Belgrade, Serbia

Computer Science, NLP

- Text classification · Multilingual (English, Serbian) · Conversation · (Deep) Machine Learning

2004.10 - 2010.12

Magister

Mathematics Faculty, University of Belgrade, Serbia

Computer Science

- Bioinformatics

1997.10 - 2004.04

Diploma

Mathematics Faculty, University of Belgrade, Serbia

Computer Science

- Programming Languages (Pascal, C, C++, Java, Prolog, Lisp), Algorithms

Architecture, Computer System Architecture, Mathematical Logic in Computer Science, Automata Theory, Databases, Microprocessors, Microcomputers, Assembly Language, Compilers and Interpreters, Programmable Systems Architecture

- Algebra (I and II), Analytical, Constructive and Euclidian Geometry, Calculus (I and II), Real and Complex Functions, Numerical Methods(I and II), Probability and Statistics

Certificates

DSC Europe 21 Certificate of Attendance
Data Science Conference

Understanding and Visualizing Data with Python
Coursera

Learn Python Programming Masterclass
Udemy

Machine Learning
Coursera

Deep Learning Specialization
Coursera

Natural Language Processing Specialization
Coursera

Publications

2022.12.01

SRPOL--A LEXICON BASED FRAMEWORK FOR SENTIMENT STRENGTH OF SERBIAN TEXTS
Review of the National Center for Digitization

2022.06.01

Effective methods for email classification: Is it a business or personal email?
Computer Science and Information Systems

Skills

Programming languages
Python, C#, C++, C, Java, JavaScript, SQL

Databases
MS SQL Server, Oracle, DB2

Machine learning tools
Scikit-learn, TensorFlow, PyTorch, SPSS, STATISTICA, Weka, IBM Intelligent Miner, scipy, numpy, pandas, Jupyter, matplotlib/seaborn

NLP tools & model architectures
BERT, GPT-3/4, gensim, fastText, word2vec, nltk, spacy, CoreNLP, HuggingFace Transformers, Simple Transformers

Languages

Serbian

Native speaker

English

Fluent

Interests

Natural Language Processing (NLP)

Computational linguistics, Semantic text analysis, Named Entity Recognition (NER), PoS tagging, Topics modeling, Language modeling, Knowledge construction from richly formatted textual data, Summarization

Serbian Language Resources

Member of the JeRTeh group dedicated to the development of resources and technologies for the Serbian language.

Hobbies

Hiking, reading, playing flute

Projects

2008.03 - 2010.12

Applying classification methods on N-gram genome analysis

Scientific thesis in the field of bioinformatics using machine learning and data mining techniques - identification of genomic islands using the N-gram statistical behavior on the particular regions of the genome sequence.

- Magister thesis · Research project

2020.01 - 2022.06

Email classification into Business and Personal categories

Classifying email messages with mixed content to Business/Personal categories.

- Text classification · BERT embeddings · Tf-Idf · Tensorflow · Research project

2017.05 - 2019.12

Hierarchical Multilabel Document Classification

Assigning textual data to categories with complex hierarchical structure and distribution.

- Document classification · Product attributes detection · Multilabel stratified sampling

2021.03 - 2021.09

Health-Care Search & Answer

Applying advanced search algorithms to retrieve relevant search results. Experiment with different Q&A modeling techniques to provide relevant answer.

- Text similarity · Extreme classification · Q&A · Transformers · Transfer learning · BERT modeling · BioASQ

2022.03 - 2022.12

HR - Analytics

Extracting candidate attributes from the resume using machine learning techniques. Find similarity between candidate resume and position description. Compute demand time for opened positions.

- Named Entity Recognition · Text similarity · Transformers · BERT modeling · Regression

---

2022.07 - 2022.12

SRPOL

Create a word sentiment lexicon and complete framework to calculate sentiment strength for a given text writen in Serbian language.

- Word lexicon · Contextual sentiment triggers · Text segmentation · Sentiment strength