



NEMANJA JANJIĆ

MACHINE LEARNING ENGINEER (6+ YEARS)

CONTACT

 nemanja1995ng@gmail.com

 [linkedin.com/in/nemanja-janjić-kvant-politika](https://www.linkedin.com/in/nemanja-janjić-kvant-politika)

 github.com/nemanja1995

 Belgrade, Serbia

LANGUAGES

Serbian, English

EDUCATION

- **Faculty of Electronic Engineering, University of Nis (2014-2018) Bachelor degree in Computer Science**
- **Faculty of Electronic Engineering, University of Nis (2018- 2020) Masters degree in Computer Science [drop out]**

SKILLS

- **Machine Learning:** TensorFlow, Keras, NumPy, NLTK, Spacy, OpenCV, OpenAI Gym, Gensim, Rasa, Scikit-Learn, Hugging Face, Tesseract, Celery, Dask, Pandas, Polars
- **DevOps (MLOps):** AWS, SageMaker (aws), Docker, MLFlow, Kubeflow, Feast, Argo, CircleCI, Jenkins, Grafana, Kibana, Sentry
- **Databases:**
 - SQL - MySQL, PostgreSQL
 - NoSQL - Cassandra, MongoDB, Neo4j, Redis, RabbitMQ, Kafka, Milvus, Postgres
 - Other: Snowflake
- **LLM:**
 - Models - Llama, GPT (finetuned, chat, vision, embeddings), Prompt Engineering
 - Frameworks - LangChain (LangGraph),
- **Programming languages:** Python, C#
- **Other programming languages:** C++, ASM (assembler), Java, Javascript (Vanilla, React), PHP, Octave, MatLab

WORK EXPERIENCE

Compstak (<https://compstak.com/>), Tech Lead, DS Team

DEC 2024 - / (0 YEAR, 10 MONTHS)

Compstak is an American Prop-tech platform specializing in commercial real estate

As Tech Lead of the Data Science team, I design and deploy scalable AI platforms, bridging data science, engineering, and infrastructure. I build and mentor the team, hiring and onboarding new engineers to deliver production-ready solutions.

- Driving the design, development, and deployment of scalable platforms and AI-driven solutions
- Building and the Data Science team, including hiring and onboarding new engineers
- Designing CI/CD pipelines and automation (CircleCI) for applications, jobs, and services
- Managing cloud infrastructure (AWS EKS, ALB, Vault, Kafka, Argo Workflows, Milvus)
- Developing and scaling APIs, with monitoring, tracing, and robust error handling
- GPU-based model serving, and orchestration
- Designing Document AI pipelines (information extraction, LLM integration with GPT & LLaMA)
- Driving experiments with multimodal AI, information retrieval, and hybrid search systems

MLOps/ Machine Learning Engineer

OCT 2022 - DEC 2024 / (2 YEAR, 2 MONTHS)

- Creating continuous machine learning (CML) infrastructure (MLFlow, Kubeflow, Feast)
- Deploying models to production (Circle CI, Argo/ Argoworkflows, AWS)
- Leading in-house project for processing documents (NLP, OCR)
- Developing, Training and deploying LLMs (Llama, GPT), and lot more "traditional" smaller models or heuristic algorithms.

Synechron (<https://www.synechron.com/>), Machine Learning Engineer

JUN 2021 - JUN 2022 / (1 YEAR)

- **Client ABN (Dutch Banck)**
- Synechron is a consulting firm focused on the financial services industry.
- Developing an OCR (OMR) system for checkmark and signature detection (tesseract, OpenCV). Working with RASA
- Creating MLflow and Sagemaker infrastructure on AWS.
- Working on the engineering part of the email classification system (celery)
- Working on job candidates' recommender system for Synechron recruitment.

Zuccharbar (<https://www.zuccharbar.com>), Machine Learning Engineer

JUL 2019 - JUN 2021 (1 YEAR, 11 MONTHS)

- **Zuccharbar was Silicon Valley based ed-tech start-up**
- Developing a Machine Learning product. That includes reading SOTA papers, testing architectures, searching/ adapting datasets, model training/ testing, building data pipelines, model productization, monitoring, and maintenance.
- Models: BERT, BioBERT, NER BERT, QAA BERT, QGen BERT
- Using NoSQL (MongoDB, Neo4j) and SQL (MySQL) databases
- Working with Docker for deploying databases and services (limited experience with AWS)
- ***Working on Market Analysis and pitch decks***

ERS (European Respiratory Society) project, Machine Learning Engineer (part time)

JUN 2020 - JUN 2021 / (1 YEAR)

- Developing and deploying data processing pipelines used for clusterization and visualization of data
- Models: Doc2Vec, Universal Sentence Encoder, BERT Multi-Label Classifier

Student club Cognitive AI, Co-founder

SEP 2018 - NOV 2019 (1 YEAR, 3 MONTHS)

- Group studying Machine Learning courses (Courses from Coursera, Stanford and MIT mathematics course)
- Working on Machine Learning project
- Projects: Neural Style Transfer model, Visualize CNN filters, visualizing CNN heatmaps, visualizing intermediate activations, Linear regression, and logistic regression visualization
- Teaching younger club members

ADDITIONAL INFO

TRAININGS

- Belgrade Big Data Analytics Summer School 2019 – LAMBDA
- Belgrade Big Data Analytics Summer School 2020 – LAMBDA

ACKNOWLEDGEMENTS

- **Award for best graduate paper in my generation. The thesis of that paper is “Use of blockchain technology in the electoral process”.**

CERTIFICATES

- **Name: Machine Learning**
Issuing Organization: Coursera
- **Name: Neural Networks and Deep Learning**
Issuing Organization: Coursera
- **Name: Structuring Machine Learning Projects**
Issuing Organization: Coursera
- **Name: Improving Deep Neural Networks: Hyperparameter tuning, Regularization, and Optimization**
Issuing Organization: Coursera
- **Name: Convolutional Neural Networks**
Issuing Organization: Coursera
- **Name: Sequence Models**
Issuing Organization: Coursera
- **Name: [Official] Rasa Certification Workshop**
Issuing Organization: Udemy
- **Name: Getting Started with AWS Machine Learning**
Issuing Organization: Coursera

Courses without certificates:

- Name: Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning
- Name: Convolutional Neural Networks in TensorFlow
- Name: Natural Language Processing in TensorFlow
- Name: Sequences, Time Series and Prediction
- Name: Stanford - CS231n: Convolutional Neural Networks for Visual Recognition
- Name: Coursera, University of Toronto - Introduction to Self-Driving Cars

PROJECTS

- Faculty project: Implementation DQN (reinforcement learning) algorithm in OpenAI Gym environment.
- Faculty project: Neural Style Transfer in Tensorflow 2.x
- Project: NaissCoin - As part of Master paper named 'Blockchain technology in tourism industry'
- EDTech Market Research

INTERESTS

• Traveling:



- Organising exhibitions: Fake Banksy, true message
- Writing: Utičnica (short story)
- Reading
- Running
- Guitar
- Dancing (Bachata, Salsa)