Tom Bukic

+385 95 855 5546 | tom@bukic.net | tom.bukic.net
linkedin.com/in/tombukic | github.com/tbukic | gitlab.com/tbukic

Creator of LangTree. Data Scientist building AI solutions from research to production with 6+ years of experience.

Experience

Data Scientist, Owner
accelerate

Jan 2024 – Present
Zagreb, Croatia

- Created LangTree (ac8.ai/langtree) - a tool for complex LLM workflow orchestration.
- Implemented AI-based solutions (RAG, semantic clustering, document processing, synthetic data generation, LLM-as-a-judge) for various industries.
- Executed research in AI for healthcare (voice analysis for detecting cognitive impairment).
- Executed time series analysis and prediction, advanced statistical analysis, and created reinforcement learning solutions.
- Built and managed on-premises GPU workstation running Debian for ML infrastructure. Deployed local LLMs, VertexAI, AWS, Databricks, HPC.
- Built ontologies and data models. Created data ingestion and cleaning pipelines.
- *Technologies:* Python, LangChain, LangGraph, HuggingFace, transformers, BERTopic, torch, PyTorch, sklearn, CatBoost, MLFlow, rllib, Ray, VertexAI, AWS, Databricks, Dask, Prefect, Docker, Kubernetes, AAS, OWL

Data Scientist
AVL

Nov 2021 – Dec 2023
Zagreb, Croatia

- Designed and built end-to-end ML solutions and digital twin software. Established reinforcement learning effort resulting in 2 conference publications. Mentored junior colleagues.
- Set up cloud data ingestion pipelines from vehicle to cloud. Executed data science projects for leading automotive OEMs.
- *Technologies:* Python, RLlib, MLFlow, Prefect, Dask, Docker, Kubernetes, Azure

Machine Learning Engineer
BID

Feb 2020 – Jun 2020
Remote

- Fine-tuned early BERT model with touristic sector data. Integrated with RASA before default inclusion.

Alpha tester
deeplearning.ai

Jan 2020
Remote

- Reviewed and improved two courses in NLP specialization.

Machine Learning Engineer
Bricsys

Jul 2019 – Sep 2019
Gent, Belgium

- Created NLP prototype for automated processing of support requests.

Machine Learning Intern
Ericsson Nikola Tesla

Jul 2018 – Sep 2018
Zagreb, Croatia

- Implemented reinforcement learning agent for Ludo game in team of 3. Created game engine and visualization.

Data Science Intern
Multicom

Jul 2014 – Jan 2015
Zagreb, Croatia

- Created prototype for text searching algorithms on Oracle database.

Education

MsC Applied Cognitive Science (ongoing), University of Zagreb
MsC Applied Mathematics (thesis remains), University of Rijeka
BSc Mathematics, 2017, University of Rijeka

Technical Skills

ML & AI: LLMs, NLP, Deep Learning, RL, Time Series, Model Discovery | PyTorch, scikit-learn, transformers, BERTopic, LangChain, LangGraph, HuggingFace, Ray, RLlib
Data Engineering: Python, Pandas, NumPy, MLFlow, Dask, Prefect, Jupyter
DevOps & Infrastructure: Docker, Kubernetes, Apptainer, Linux/Debian, GitLab CI, GitHub Actions, AWS, Azure, HPC