



Stefan Seman

ML Engineer | Data Scientist | PhD Researcher

Phone: +38162262131

LinkedIn: linkedin.com/in/sseman

Website: manse404.github.io/webpage

Research: ResearchGate

Email: stefan.seman404@gmail.com

GitHub: github.com/ManSe404

Scholar: Google Scholar

PROFESSIONAL SUMMARY

ML Engineer, Data Scientist, and PhD Researcher with a strong background in higher education and published research. Specialize in machine learning, AI, generative AI, and software development, with unique expertise bridging data science and biomedicine. Analytical thinker with a pragmatic problem-solving approach and excellent communication skills.

TECHNICAL SKILLS

Languages

Python, SQL, Bash, LabVIEW

AWS Cloud

Lambda, S3, ECS, Step Functions, DynamoDB, Bedrock

Data & Storage

DynamoDB, Redis, PostgreSQL, Pinecone, Milvus, Data Lake

AI/ML Frameworks

PyTorch, LangChain, LlamaIndex, Transformers, PySpark

Azure Cloud

OpenAI, Data Factory, DevOps, Synapse, Functions, Search

Tools & DevOps

Docker, KubeFlow, FastAPI, Flask, Git, Grafana

PROFESSIONAL EXPERIENCE

Senior Machine Learning Engineer & Software Engineer

404 Solutions

Oct 2024 - Present

- Main AI Engineer for critical infrastructure and core systems
- Owned document intelligence services and cloud APIs
- Built serverless workflows for high-volume data processing

Senior Data Scientist & Software Engineer in AI, R&D

Publicis Sapient

Oct 2023 - Sept 2024

- Led development from POC to fully production-grade systems
- Built enterprise RAG systems with retrieval optimization
- Delivered client-facing solutions and technical propositions

Data Scientist & Developer -> Senior Specialist

Valcon Netherlands

May 2022 - Oct 2023

- Led and mentored a team of 6 engineers
- Owned design of document processing microservices
- Built data extraction and document classification services from scratch

Data Scientist

Oct 2021 - Apr 2022

Optimal Systems GmbH

- Systematic model evaluation and hyperparameter tuning for ML classification
- Data-driven insights and testing/validation of ML systems

Researcher

Mar 2020 - Mar 2022

BioIRC - Bioengineering R&D Center

- Built components for EU Horizon 2020 computational platform (SILICOFCM)
- ECG, CPET, and biomedical data processing; peer-reviewed publications

PhD Researcher

Oct 2018 - Mar 2022

University of Belgrade

- Developed research methodology and statistical frameworks
- Applied machine learning to biomedical datasets; published papers internationally

Sports Scientist

Oct 2014 - July 2018

Sport and Recreation Center Obliquus

- Built algorithms for signal processing and data collection
- Functional testing and performance analysis for athletes

KEY PROJECTS**AI Document Summarizer**

RAG-based summarization with LLMs
AWS Bedrock, LangChain, LlamaIndex

AI-Code Copilot

Enterprise AI for internal knowledge
Azure OpenAI, Milvus, Fine-Tuning

Document Chatbot

Enterprise chatbot with RAG
FastAPI, Azure Search, Docker

Document Extractor

Deep learning for text extraction
PyTorch, Azure Data Factory

Large-Scale ETL

Serverless data pipelines
AWS Lambda, Step Functions, S3

SILICOFCM (EU H2020)

Clinical trials platform
Python, Biostatistics, ECG

Doc Classification

ML-based categorization
Flask, Docker, Azure

Data Extraction Service

ETL with Power BI integration
Azure Data Factory, Synapse

Voice Support Chatbot

Speech-to-text telephony AI
Azure LUIS, Speech Services

SELECTED PUBLICATIONS (10+ TOTAL, 100+ CITATIONS)**A Machine Learning Approach to Gene Expression in Hypertrophic Cardiomyopathy**

Pavic, J., Zivanovic, M.N., Seman, S., et al.
Pharmaceuticals (MDPI), 2024

Ventilatory Efficiency Parameters Outperform Peak VO₂ in Monitoring Therapy Effects in HCM

Seman, S., Tesic, M., Popovic, D., et al. (SILICOFCM Investigators)
European Heart Journal, Vol. 45, 2024

Defining the Importance of Stress Reduction in Managing Cardiovascular Disease

Popovic, D., Bjelobrk, M., Tesic, M., Seman, S., et al.
Progress in Cardiovascular Diseases, Vol. 70, 2022 (60+ citations)

Physical activity and exercise as an essential medical strategy for COVID-19 and beyond

Seman, S., Srzentic Drazilov, S., et al. (HL-PIVOT Network)
Experimental Biology and Medicine, Vol. 246, 2021

EDUCATION

PHD CANDIDATE**Sport Science**

University of Belgrade
2018 - Present

MASTER'S DEGREE**Sports Medicine**

University of Novi Sad

BACHELOR'S DEGREE**Sport Science**

University of Belgrade

CERTIFICATIONS

- ✓ Microsoft Azure AI Associate
- ✓ Databricks Lakehouse Fundamentals
- ✓ Johns Hopkins COVID-19 Contact Tracing

- ✓ Microsoft Azure Data Fundamentals
- ✓ Databricks Generative AI Fundamentals

Languages: Serbian (Native) English (Fluent)

Member: ACSM SEBM