

**Summary:** A Junior Data Scientist specializing in multimodal Deep Learning, designing advanced architectures like Foundation Models, Transformers, Self-Supervised learning (VICReg), and Machine Learning (CatBoost and LightGBM), to handle complex, large-scale datasets including signal, image, tabular, and textual data.

# SeyedHasan MirHosseini

[LinkedIn](#) | [Email](#) | [Github](#) |  
+358449325220 | Tampere, Finland

## Work Experience

### August 2025 – October, Data scientist (internship) at University of Eastern Finland

**Responsibilities:** Developing a foundation model from EEG recordings using Transformer and self-supervised learning (VICReg) pretraining, then fine-tune the pretrain model to predict behavioral performance.

**Source:** EEG Foundation Challenge 2025, <https://eeg2025.github.io/>

**Technologies and tools:** Python, PyTorch

**Supervisor:** Mastaneh Torkamani Azar ([mastaneh.torkamani@uef.fi](mailto:mastaneh.torkamani@uef.fi))

### January 2024 – 2025, Data Analyst at Gamification Group of Tampere University

**Responsibilities:** Experimental data organization, cleansing, preprocessing, statistical analysis, visualization of facereader data, physiological signal processing (PPG and EDA), co-authoring in publications, [GitHub Repository](#)

**Tools & Technologies:** Python: Pandas, NumPy, SciPy, Matplotlib, R

**Supervisor:** Dr. Wilk Oliveira Dos Santos ([wilk.oliveira@tuni.fi](mailto:wilk.oliveira@tuni.fi))

### Mater's Thesis: Seizure Detection from EEG Signals Using Multimodal Deep

#### Learning , GitHub repository (Grade: 5/5)

**Unimodal Model:** LSTM-Transformer: 79% | CNN-Transformer: 90 %

**Multimodal Model:** Fusion, LSTM&CNN -Transformer: 82%

**Supervisors:** Prof. Jaakko Peltonen [jaakko.peltonen@tuni.fi](mailto:jaakko.peltonen@tuni.fi),

Doctoral researcher Saana Seppälä [saana.seppala@tuni.fi](mailto:saana.seppala@tuni.fi)

## projects

| Project Title                              | Model                                | Performance                         | Github Repository                 |
|--|--------------------------------------|-------------------------------------|-----------------------------------|
| Predicting Loan Payback                    | LightGBM, CatBoost                   | ROC: 0.92135%,                      | <a href="#">GitHub Repository</a> |
| Predicting the Beats-per-Minute of Songs   | Deep Learning                        | RSME: 26.41                         | <a href="#">GitHub Repository</a> |
| Binary-Classification-with-a-Bank-Dataset  | LightGBM Autoencoder                 | ROC AUC: 97% 96,4%.                 | <a href="#">GitHub Repository</a> |
| Predicting Road Accident Risk              | Autoencoder                          | RMSE: 0.056                         | <a href="#">GitHub Repository</a> |
| Predict the Introverts from the Extroverts | CatBoost Random Forest LightGBM      | Accuracy: 0.974                     | <a href="#">GitHub Repository</a> |
| Detect Behaviour with Sensor Data          | GRU + CNN                            | F1-Score: 81%                       | <a href="#">GitHub Repository</a> |
| Image2Biomass Prediction                   | U-Net, Resnet                        | R <sup>2</sup> : 0.60 (In progress) | <a href="#">GitHub Repository</a> |
| Digitization of ECG Images                 | Self-supervised learning-Transformer | In progress                         | In progress                       |
| Protein Function Prediction                | Transformer                          | In progress                         | In progress                       |

## Education

| Specialization                       | Year, University                                    |
|--------------------------------------|---|
| Master of Statistical Data Analytics | 2023 - 2025<br>Tampere University, Tampere, Finland |
| Master of Business Administration    | 2013 - 2016<br>Tehran University, Tehran, Iran      |
| Bachelor of Mathematics              | 2008 – 2013<br>Damghan University, Damghan, Iran    |

## Courses

Data science course with Python and R  
Web Development with Django, HTML, CSS

## Technical Skills

| Skills  | Tools                              |
|---|------------------------------------|
| Programing Language   | Python and R                       |
| Machine Learning, Data Mining   | Pandas, NumPy, SciPy, scikit-learn |
| Deep Learning: Transformers, LSTMs, CNNs, RNN, Autoencoders, Embedding, Foundation Model, Transfer Learning | PyTorch                            |
| Image and Signal Processing: EEG, EDA, PPG and spectrogram analysis   | SciPy and Librosa, MNE             |
| Data Visualization  | Matplotlib, Seaborn                |
| Statistical Analysis  | R                                  |
| Web Development   | Django, HTML, css                  |

## Soft Skills

Organized and Punctual  
Self-Directed Research  
Collaborative Teamwork  
Communication and Engagement  
Project Management & Organization  
Problem-Solving Mindset

## Languages

Persian: Native  
English: C1  
Finnish: A1