

## Presentation

**Ahmad Aghaebrahimian**

# Hyperparameter Tuning for Deep Learning in Natural Language Processing

### **Abstract:**

Although Deep Learning has advanced a lot through past several years, it still seems like a black art for many people mostly due to the fact that obtaining consistent and good results from a deep architecture requires optimizing many parameters. Hyperparameter tuning is an essential task in deep learning which can lead to significant changes in the network performance. This talk is the essence of approximately 3000 GPU hours on optimizing a network for a text classification task on a wide array of hyperparameters. Word embedding types, word embedding sizes, word embedding updating, character embedding, character embedding sizes, deep architectures (CNN, LSTM, GRU), optimizers, gradient control, classifiers (Softmax, Sigmoid, CRF), dropout, deep vs. wide networks, pooling, and batch sizes are the hyperparameters studied in this work using a grid search scheme. I will talk about the most critical parameters and the insight behind them that researchers can modify or prioritize in a deep architecture to get the best performance with the least effort on the part of humans and the machine.

**Biography:** Dr. Ahmad Aghaebrahimian is a Research Associate at the Zurich University of Applied Sciences (ZHAW) with backgrounds in computer science and linguistic. He completed his Ph.D. in Mathematical Linguistics at the Charles University in Prague where he integrated Deep Neural Networks and linguistics to improve the state of the art in large-scale open-domain Question Answering. His areas of interest include Artificial Intelligence in general and Deep Neural Networks in particular, as well as Corpus Linguistics, Question Answering, and Information Retrieval.

**Organization:** Zurich University of Applied Sciences

**Contact:** agha@zhaw.ch