HSLU Hochschule Luzern

Certificate of Advanced Studies (CAS) Machine Learning

Curriculum

Program Directors

Prof. Dr. Umberto Michelucci
umberto.michelucci@hslu.ch

Dr. Aygul Zagidullina
aygul.zagidullina@hslu.ch

Website hslu.ch/cas-ml

The CAS in Machine Learning at the HSLU is a hands-on, practice-oriented program for professionals looking to deepen their understanding of machine learning. Using real-world data, participants develop practical skills in building, training, and evaluating machine learning models. The course covers key concepts, algorithms and tools. Combining theory with practical exercises and providing ready-to-use knowledge, it is ideal for people aiming to integrate machine learning into their work.

The program boasts a strong hands-on approach, integrating practical exercises into every topic discussed to ensure participants can directly apply their knowledge in practice. Each module includes guided labs, coding challenges and case studies using industry-relevant tools and datasets. Participants work with popular frameworks such as TensorFlow, and Scikit-learn, learning to implement machine learning models from scratch, optimize algorithms, and interpret results effectively. These practical sessions help participants understand not only the technical aspects of machine learning but also how to handle challenges such as overfitting, data preprocessing, and model evaluation.

The CAS in Machine Learning comprises 11 modules.

Module 1 – Introduction to Python – Anaconda installation, Introduction to Jupyter notebooks / Python Basics / Pandas: Reading and writing tables, indices, groupby, join, etc. / Plotting: Intro to Matplotlib and basic plots / Numpy: Creating and working with arrays, little linear algebra, etc. / Project work / Project presentation on Saturday afternoon.

Module 2 – Supervised Machine Learning – ML context in data science / Gradient descent / Linear regression / Loss functions to describe model performance / Dealing with missing data through imputation / Transforming data through scaling / Ttrain / test / validation set and cross-validation / Feature selection / Bias/variance trade-off, diagnosing under- fitting and over-fitting / Regularisation / Model selection principles including hyperparameter search / Difference between regression and classification / Basics of logistic regression and transformation of problem into prediction of class probability / Losses functions for classification / Classification model assessment / Multiclass classification / Imbalanced class data and strategies to resolve / Introduction to ensemble methods / Data leakage and strategies to overcome it.

Module 3 – Deep Learning – The different types of perceptrons, activation functions / Backpropagation / Data and model learning pipeline / Hands-on: First fully-connected neural network with keras / Choosing the right number of hidden layers, neurons / Choosing the activation function / First example of a CNN / Inspect architecture, role of each layer / Visualize training: training/validation loss over epochs / Overfitting vs Underfitting / Hyperparameters tuning: dropout and learning rate variations and observations / Improving model accuracy: convolutional filters and dense layers sizes / Project on Google Street Dataset.

Last udpated 16.01.2024

Participants conclude the CAS in Machine Learning program with an in-depth understanding of machine learning concepts, methods and tools reinforced by practical, hands-on experience.

The modules build on each other and offer both foundational knowledge and advanced technical skills to apply across a variety of domains. From Python programming to neural networks, computer vision, NLP, and MLOps,

Participants gain the skills necessary to handle complex data challenges and deploy robust machine learning solutions. This curriculum highlights the program's goal of blending theory with real-world practice. This empowers graduates to tackle challenges in data science, AI, and beyond with confidence, and to apply their expertise with a view to driving innovation in their respective fields.

Access and pre-process the data, normalization / CNN design and implementation / Training and evaluation, analyze the classification performances / Visualize predictions.

Module 4 – Computer Vision – Digital Image formats introduction: rRaster vs vector images, / Image features and formats / CNN recap / Conv, Pool, Dense Layers / Metrics for Image Classification / Confusion matrices / Metrics for Image Localization : IoU / Transfer learning and an introduction to pre-trained models (VGGNet and Resnet) / Hands-on exercises: Introduction to Image Processing / Introduction to Convolutions

Module 5 – Generative AI – Autoencoders / Variational autoencoder / Generative Adversarial Networks (GANs) / Taxonomy, models definition, capabilities, the generative story / Training and limitation details, practice with code representation / Hands-on: Colab notebooks with the different models implemented and trained.

Module 6 – Natural Language Processing – Text mining basics / Text similarity / Text generation with AI / Word embeddings / Presentation of day 1 project / NLP + ML grand challenge / Hands-on challenge: Train a supervised classifier to distinguish between real and fake news articles

Module 7 – Unsupervised Learning – Clustering: K-means, hierarchical clustering / Dimension reduction: principal component analysis (PCA) / t-distributed stochastic neighboring.

Module 8 – Transformers Architectures and Large Language Models – Prompt Engineering / Transformers / Transformer-Based Contextual Sentence Classification / Basic Retrieval Augmented Generation (RAG) with LangChain and Claude.ai Model / Fine-tune Gemma on Medical Dataset / Hands-on: HuggingFace account creation; LLM playground for prompt engineering; Notebooks with Contextual Sentence Classification, Retrieval Augmented Generation, Fine-tuning.

Module 9 – Keras Introduction – Quick keras overview / Training workflow, software and hardware requirements, evolution / Built-in functions and keras models, compilation / Layers and customization, APIs. / Custom training loops / Model, loss and optimizer / Gradient descent definition and remarks. tf.GradientTape presentation.

Module 10 – Reinforcement Learning – Objectives and origins, interdisciplinarity / Process : Observation, action, reward / Terminology, tasks and methods / Applications in industry / Reinforcement learning vs supervised learning / Model-based, model- free, policy-based, value-based approaches / Data-efficient solutions / Contextual bandit : agents and regret metrics / Hands-on: Tensorflow Agents : usage and application; Contextual bandits agent for movie recommender.

Module 11 – Machine Learning Operations (MLOPS) – MLOps definition and principles, challenges in ML models deployment / Codespaces and Docker. Work environment for the course, practical deployment. / Model platforms, registry and Mlflow models concept, tracking and registration. / Processing: request-response, batch and stream / ML System pipeline/ Model Drift, data orchestration, data validation / Feature stores and platforms / Hands-on: Exercices on the different pipelines presented, MlFlow introduction, monitoring.