

## *Certificate of Open Studies*

*in Applied Data Science Machine Learning*

### RÈGLEMENT D'ÉTUDES

#### **Article 1. Objet**

L'École polytechnique fédérale de Lausanne (EPFL), par son unité « Extension School », délivre un *Certificate of Open Studies* en Applied Data Science Machine Learning (COS ADSML).

#### **Article 2. Objectifs de la formation**

Le COS ADSML a pour objectif de fournir une expérience pratique des dernières technologies dans le domaine de la science des données. L'apprentissage est axé sur les outils de *predictive modeling and analytics*, la puissance des *neural networks* et des techniques de *deep learning* sur des ensembles de données de différents types.

#### **Article 3. Organisation générale de la formation**

- 3.1 Le COS ADSML est un programme de formation ouvert au public, à distance et au rythme du participant.
- 3.2 Le COS ADSML se compose d'une partie de cours prérequis, puis d'une partie projet (Capstone Project). La réussite des cours prérequis permet l'admission au Capstone Project, lequel permet la réussite du COS ADSML.
- 3.3 Les cours prérequis sont un ensemble de cours en ligne. Le Capstone Project consiste en la mise en pratique de ces cours par l'accomplissement d'un projet individuel.
- 3.4 Le COS ADSML vaut 15 crédits ECTS, indivisibles et décernés en bloc.
- 3.5 Le plan d'études, en annexe au présent règlement, présente la liste et le contenu des cours pour les cours prérequis, ainsi que les exigences pour le Capstone Project.

- 3.6 Les modalités d'inscription et d'organisation pour les cours prérequis sont communiquées en ligne.
- 3.7 La formation se déroule entièrement en anglais.
- 3.8 Le COS ADSML requiert une infrastructure informatique d'une qualité adéquate pour suivre les cours en ligne, communiquer avec l'école et, le cas échéant, interagir en direct par vidéoconférence dans des conditions satisfaisantes pour l'EPFL.

#### **Article 4. Durée de la formation**

- 4.1 Le COS ADSML s'étend sur une durée minimum de trois mois et maximum de dix-huit mois, selon le rythme du participant.
- 4.2 La formation peut être suspendue pour une durée d'un mois ou plus, par période mensuelle et sans prolongation de la durée maximum du COS ADSML, aux conditions prévues dans les Conditions Générales (Terms and Conditions) du COS ADSML (disponibles sur la page internet dédiée à la formation). Durant la période de suspension, le montant des frais de formation est réduit.
- 4.3 La formation ne peut être ni interrompue, ni prolongée. L'abandon en cours de formation vaut échec au COS ADSML.

#### **Article 5. Inscription à la formation**

- 5.1 La procédure d'inscription, le montant des frais et les modalités de paiement sont indiqués dans le document de conditions générales de la formation (Terms and Conditions), disponible sur la page internet dédiée à la formation.
- 5.2 L'inscription nécessite la communication d'informations bancaires valables pour le prélèvement des frais pour la formation, ainsi que la transmission des documents requis pour la vérification de l'identité du participant.
- 5.3 Au premier jour de la formation, l'inscription devient définitive et les frais sont dus. Le participant a dès lors accès à l'ensemble des contenus de cours du COS ADSML pour une période initiale d'un mois. Il peut prolonger son inscription de mois en mois selon un système d'abonnement, sous condition de paiement des frais et dans la limite de la durée maximale pour réussir la formation.

#### **Article 6. Contrôle des connaissances**

##### **Cours prérequis**

- 6.1. Les cours prérequis font chacun l'objet d'un contrôle des connaissances, sous la forme d'un ou de plusieurs projets soumis en ligne.
- 6.2 Chaque cours prérequis est évalué comme réussi (« pass ») ou échoué (« fail »).

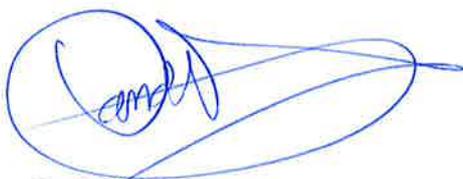
- 6.3 En cas d'échec au cours, le participant peut le tenter à nouveau, dans la limite de la durée maximale pour réussir le COS ADSML.
- 6.4 La réussite des cours prérequis permet d'obtenir une attestation de cours. Elle ne donne lieu à aucun crédit ECTS.
- 6.5 Lorsque les conditions ne sont pas réunies pour que soit garanti un contrôle des connaissances de qualité et exempt de fraude, l'EPFL y renonce. Le cas échéant, le participant peut être invité à se déplacer, à ses frais, pour un contrôle des connaissances à l'EPFL.

### Capstone Project

- 6.6 Le Capstone Project fait l'objet d'une évaluation. Dans le cadre de l'évaluation, le participant est soumis, par vidéoconférence, à une défense orale de son projet et à un nouveau contrôle des connaissances sur le contenu des cours prérequis.
- 6.7 Le participant peut soumettre son projet dans des versions intermédiaires, jusqu'au délai de dépôt de la version finale. Le projet est finalement apprécié comme réussi (« pass ») ou échoué (« fail »).
- 6.8 En cas d'échec, le participant dispose d'une seconde et dernière tentative, dans la limite de la durée maximale pour réussir le COS ADSML. Le cas échéant, il effectue son projet sur un nouveau sujet. L'échec en seconde tentative vaut échec définitif au COS ADSML.
- 6.9 En cas de réussite, le COS ADSML est décerné.
- 6.10 Lorsque l'EPFL considère que les conditions ne sont pas réunies pour que soit garanti un contrôle des connaissances de qualité et exempt de fraude, elle y renonce. Le cas échéant, le participant peut être invité à se déplacer, à ses frais, pour un contrôle des connaissances à l'EPFL.

### Article 7. Entrée en vigueur

Le présent règlement d'études entre en vigueur le 1<sup>er</sup> juin 2019.



Prof. P. Vanderghyest  
Vice-Président pour l'éducation



Prof. M. Salathé  
Directeur académique de l'Extension School

## Annexe au règlement d'application

Plan d'études (juin 2019) de la formation COS in Applied Data Science Machine Learning

<u>Titre des cours</u>	<u>Crédits</u>
Course 1. Introduction to Data Analysis with Python:	0 ECTS
<ol style="list-style-type: none"><li>1. Welcome</li><li>2. Getting started</li><li>3. The JupyterLab and Jupyter notebook</li><li>4. Exploring our first data set</li><li>5. A first look at NumPy</li><li>6. A first look at Pandas</li><li>7. The basics of data visualization</li><li>8. Probability for data science</li><li>9. Linear algebra for data science</li><li>10. Course project</li></ol>	
Course 2. Applied Data Analysis	0 ECTS
<ol style="list-style-type: none"><li>1. Welcome</li><li>2. Getting the data</li><li>3. Cleaning the data</li><li>4. Manipulating the data</li><li>5. Working with text data</li><li>6. Working with time-series data</li><li>7. Databases in Python</li><li>8. Statistical data analysis</li><li>9. Course Project</li></ol>	
3. Applied Machine Learning 1	0 ECTS
<ol style="list-style-type: none"><li>1. Introduction to Machine Learning</li><li>2. Fitting a first model</li><li>3. Cost functions and outliers</li><li>4. Linear regressions</li><li>5. Gradient descent</li><li>6. Feature engineering</li><li>7. Regularization</li><li>8. Advanced Scikit-learn</li><li>9. Course project</li></ol>	
4. Applied Machine Learning 2	0 ECTS
<ol style="list-style-type: none"><li>1. Course introduction</li><li>2. K-nearest neighbors</li><li>3. Bias-variance trade-off</li><li>4. Logistic regressions</li><li>5. Decision trees and SVMs</li><li>6. Clustering and dimensionality reduction</li><li>7. Introduction to deep learning</li><li>8. Multilayer and convolutional neural networks</li><li>9. TensorFlow ecosystem</li><li>10. Chatbots overview by Swisscom</li><li>11. Course project</li></ol>	

Travail personnel

Crédits

Capstone Project

15 ECTS

The capstone project must demonstrate the learner's acquisition of the following:

1. Data Preparation
  - Collect and import the required data sets
  - Demonstrate good understanding of the data and its structure (values, encoding, format, etc.)
  - Prepare the data appropriately for the analysis and modeling (data cleaning, manipulation, feature engineering and encoding)
2. Exploratory data analysis
  - Demonstrate a deep understanding of the data specific to the project goals by performing detailed exploratory data analysis including all necessary descriptive statistics and visualization methods
  - Discuss insights and potential difficulties to justify the proposed implementation
  - Discuss data quality and completeness with respect to the intended project goals
3. Machine learning models
  - Identify and discuss suitable machine learning models, baselines, metrics and evaluation strategies
  - Tune the different models, analyze their performance
  - Discuss the results and potential trade-offs (complexity, interpretability, computational resources) using the appropriate terminology
4. Communication
  - Use data to tell a story and discuss the value of the results. Highlight interesting insights through discussion and visualizations
  - Document and discuss each step of the analysis to support the approach and implementation
  - Present and evaluate findings, discuss results in the context of the project goals