

Description du module

Machine Learning

Généralités
Nombres de crédits ECTS

3

Sigle du module

TSM_MachLe

Responsable du module

Dr. Thilo Stadelmann, ZHAW

Langue
Explications concernant les langues définies par site :

- L'enseignement se déroule dans la langue définie ci-dessous selon le site / la réalisation.
- La documentation est disponible dans les langues définies ci-dessous. En présence de plusieurs langues, un pourcentage par langue est indiqué (100% = documentation complète)
- L'examen est entièrement (100 %) disponible dans la langue cochée selon le site / la réalisation.

	Berne	Lausanne		Lugano	Zurich	
Enseignement	<input type="checkbox"/> E 100%	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> F 100%	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> E 100%	<input type="checkbox"/> D 100%
Documentation	<input type="checkbox"/> E 100%	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> E 100% <input type="checkbox"/> F %	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> E 100%	<input type="checkbox"/> E % <input type="checkbox"/> D %
Examen	<input type="checkbox"/> E 100%	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> E 100% <input checked="" type="checkbox"/> F 100%	<input type="checkbox"/> E 100%	<input checked="" type="checkbox"/> E 100%	<input type="checkbox"/> E 100% <input type="checkbox"/> D 100%

Catégorie du module

- FTP Bases théoriques élargies
- TSM Approfondissement technique et scientifique
- CM Modules de savoirs contextuels

Périodes

2 périodes d'enseignement frontal et 1 période d'exercice par semaine

Conditions préalables
Connaissances et compétences prérequis

- **Mathématiques** : calcul de base / algèbre linéaire / calcul des probabilités (c.-à-d. dérivées, multiplication de matrices, distribution normale, théorème de Bayes)
- **Statistiques** : statistiques descriptives de base (c.-à-d. moyennes, variances et covariances, histogrammes, boîtes à moustaches)
- **Programmation** : bonne maîtrise d'un langage de programmation structuré (c.-à-d. Python, Matlab, R, Java, C, C++)
- **Analyse** : méthodes d'analyse des données élémentaires (prétraitement des données, arbres de décision, clustering par k-means, régression linéaire & logistique)

Brève description / Explication des objectifs et du contenu du module en quelques phrases

L'apprentissage automatique ou *Machine Learning* (ML) est le fruit de l'intelligence artificielle et de la science informatique, et se définit comme une discipline universitaire visant à « donner aux ordinateurs la faculté d'apprendre sans avoir été explicitement programmés pour le faire ». (A. Samuel, 1959). A l'heure actuelle, l'apprentissage automatique est le moteur méthodologique des grandes tendances de fond du *Big Data* et du *Data Science*. Les experts en ML sont très recherchés aussi bien dans le milieu universitaire que dans le monde industriel.

Ce cours s'appuie sur des connaissances de base en mathématiques, programmation et analyses/statistiques telles qu'elles sont dispensées dans les différents cursus bachelor d'ingénieurs. Partant de là, il enseigne les bases des techniques modernes de l'apprentissage automatique par une méthode centrée sur la mise en pratique et l'application à des problèmes de la vie réelle. Le processus de création d'un système d'apprentissage est abordé dans sa globalité :

- formulation de la tâche à l'étude sous forme de problème d'apprentissage ;
- extraction de caractéristiques intéressantes à partir des données disponibles ;
- sélection et paramétrage d'un algorithme d'apprentissage convenant à la situation.

Les sujets abordés englobent des concepts transversaux tels que la conception et le débogage de système ML (comment interpréter les modèles appris et leurs résultats) ainsi que l'ingénierie des caractéristiques (*feature engineering*) ; les algorithmes abordés comprennent (parmi d'autres) les machines à vecteurs de support (*Support Vector Machines*) ainsi que les méthodes d'apprentissage ensemblistes.

Objectifs, contenu et méthodes**Objectifs d'apprentissage et compétences visées**

- Les étudiants **connaissent le contexte et la taxonomie** des méthodes d'apprentissage automatique ;
- Sur cette base, ils **traduisent** des problèmes donnés en **tâches d'apprentissage** et **sélectionnent une méthode d'apprentissage adaptée** ;
- Les étudiants **sont capables de convertir** un set de données en un **set de caractéristiques** adapté à la tâche à l'étude ;
- Ils **évaluent** l'**approche choisie** de façon structurée en faisant appel à un design d'expérience adapté ;
- Les étudiants **savent comment** sélectionner des modèles, et « **déboguer** » des caractéristiques et des algorithmes d'apprentissage si les résultats ne répondent pas aux attentes ;
- Les étudiants sont capables d'utiliser les résultats d'évaluation pour **optimiser les paramètres** d'un système donné et **maximiser** ses performances ;
- Les étudiants **ont pris connaissance d'exemples de différentes sources de données** / types de problèmes et sont **capables de consulter la littérature scientifique pour acquérir des connaissances spécialisées additionnelles**.

Contenu du module avec détail des contenus d'enseignement

- **Introduction** (2 semaines) : convergence des participants ayant suivi diverses formations
- **Apprentissage supervisé** (7 semaines) : méthodes de modélisation à partir de données labellisées Apprentissage non supervisé

Sujets transversaux : *feature engineering* ; *ensemble Learning* ; débogage des systèmes ML

Algorithmes : p. ex., SVM, apprentissage ensembliste, modèles graphiques (réseaux bayésiens)

- **Apprentissage non supervisé** (3 semaines) : méthodes de modélisation à partir de données non-labellisées

Algorithmes : p.ex., réduction de dimensionnalité, détection d'anomalie, analyse de prototypes

- **Chapitres spéciaux** (2 semaines) :

Algorithmes : p.ex., apprentissage par renforcement, systèmes de recommandation, modèles de mélange gaussien

Méthodes d'enseignement et d'apprentissage

Cours en classe ; exercices de programmation (p.ex., Python 3)

Bibliographie

- T. Mitchell, « *Machine Learning* », 1997
 C. M. Bishop, « *Pattern Recognition and Machine Learning* », 2006
 G. James et al., « *An Introduction to Statistical Learning* », 2014
 K. Murphy, « *Machine Learning – A Probabilistic Perspective* », 2012

Mode d'évaluation**Conditions d'admission aux examens de fin de module (tests exigés)**

Aucune (il est fortement recommandé d'effectuer toutes les heures de laboratoire et d'assister aux cours)

Principe pour les examens :

**Tous les examens finaux des modules sont organisés sous forme écrite.
 Les examens de rattrapage peuvent se dérouler sous forme écrite ou orale.**

Examen de fin de module régulier et examen de rattrapage écrit

Type d'examen	Ecrit
Durée de l'examen	120 minutes
Moyens autorisés	<input type="checkbox"/> Pas de moyen auxiliaire <input checked="" type="checkbox"/> Moyens autorisés
	<input type="checkbox"/> Moyens électroniques : _____ <input type="checkbox"/> Format papier : _____ <input checked="" type="checkbox"/> 1 page A4 (recto et verso) de notes manuscrites (pas de livre, pas de diapositive, pas d'autres notes)

Cas particulier : examen de rattrapage oral

S'il est procédé à un examen oral (uniquement si ≤ 4 étudiants concernés), la règle suivante s'applique :

Type d'examen	oral
Durée de l'examen	30 minutes
Moyens autorisés	Pas de moyen auxiliaire