

# Apprentissage automatique pour les sciences et l'ingénierie de l'environnement

Master en sciences de l'environnement, FGSE, Université de Lausanne

Description du cours (pour le semestre de Printemps 2022, dernière mise à jour le 8 septembre 2021)

**Professeur:** Tom Beucler (Professeur Assistant à l'IDYST, Site web du groupe, tom.beucler@unil.ch)

**Assistant à l'enseignement:** Milton Gomez (Doctorant à l'IDYST, milton.gomez@unil.ch)

## 1 Résumé

Notre capacité à observer et simuler notre environnement s'améliore continuellement, et nous sommes désormais inondés de milliers de Tb de données chaque jour, qui sont difficiles à explorer avec la programmation traditionnelle. Les algorithmes d'apprentissage automatique (AA), qui par définition peuvent apprendre à exécuter une tâche à partir de données sans être programmés explicitement pour cette tâche, sont particulièrement efficaces pour extraire des informations de larges jeux de données géo-environnementales. Après leur entraînement, ces algorithmes sont peu coûteux à utiliser, ce qui fait d'eux des outils idéals pour remplacer des modèles réalistes/complexes lorsque le temps et les ressources sont limités. En outre, leur capacité à résumer efficacement de grandes quantités de données les rendent prometteurs pour faire de nouvelles découvertes scientifiques.

Ce cours appliquée de 10 semaines présentera les algorithmes d'AA les plus courants et comment les appliquer en sciences et ingénierie de l'environnement; l'objectif est qu'à la fin du cours, tu saches :

1. Nommer les algorithmes d'AA les plus courants et résumer leurs avantages et inconvénients, particulièrement dans le contexte des sciences environnementales,
2. Les implémenter avec Python (principalement avec les librairies Numpy/Scikit-Learn/Keras/Tensorflow dans des Google Collab notebooks),
3. Savoir d'expérience quels algorithmes sont les plus appropriés pour les applications environnementales qui te tiennent à cœur (comme par exemple ta thèse de Master).

Dans le cadre des ces trois objectifs, le cours sera structuré de la manière suivante (voire la version anglaise pour plus de détails) :

- **Cours magistraux** ( $\approx 2$  heures/semaine, 15% de la note);
- **Lecture obligatoires de manuels et d'articles entre les cours, accompagnée de questions de compréhension** ( $\approx 3$  heures/semaine, 15% de la note);
- **Travaux dirigés d'informatique** ( $\approx 3$  heures/semaine, 20% de la note);
- **Projet final** ( $\approx 2$  heures/semaine, 50% de la note).
- Il n'y aura **pas d'examen final** ou de travail à la maison autre que le projet final et les manuels et articles à lire entre les cours.

Pour la partie d'AA du cours, nous utiliseront principalement le livre “Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow” (code/pdf) de Géron (2019) et le livre “Deep Learning with Python” (code/pdf) de Chollet (2017), mais nous vous invitons à explorer les nombreuses ressources en ligne sur le sujet, par exemple en cliquant sur ce lien.

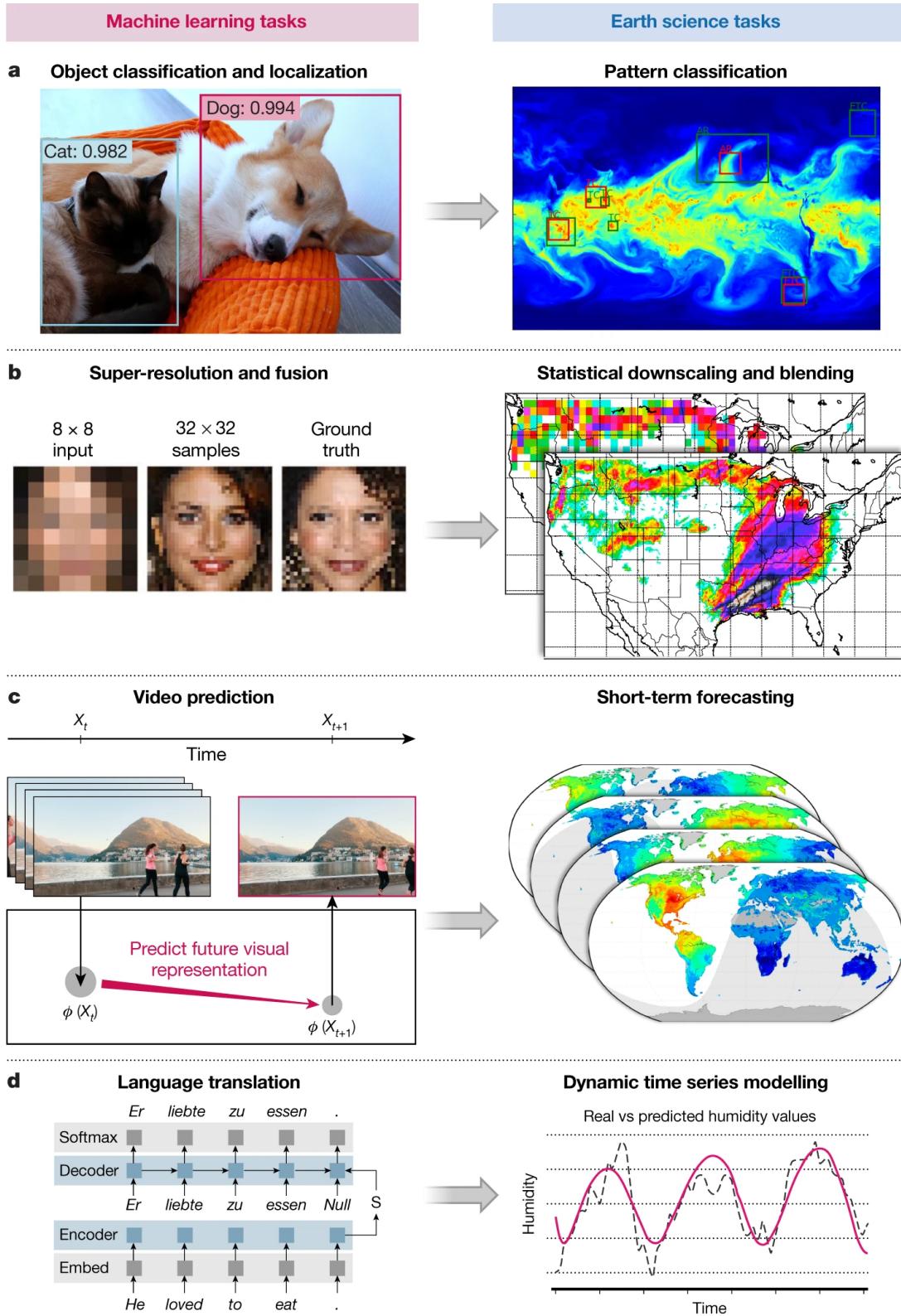


Figure 1: a) Object recognition in images links to classification of extreme weather patterns using a unified convolutional neural network on climate simulation data<sup>41</sup>. b) Super-resolution applications relate to statistical downscaling of climate model output<sup>72</sup>. c) Video prediction is similar to short-term forecasting of Earth system variables. d) Language translation links to modeling of dynamic time series.

Source: Figure 2 in <https://www.nature.com/articles/s41586-019-0912-1>.

## **2 Veuillez consulter la description anglaise du cours pour plus d'informations**

Pour rester à jour, nous te suggérons d'ajouter le calendrier Google suivant à ton propre calendrier : Lien Google Cal. Les horaires et lieux indiqués entre parenthèses ont été prévus pendant l'été 2021 et sont sujets à changement; n'hésite pas à consulter le calendrier Google pour suivre les dernières mises à jour.