GIMAS9AI - INFORMATION THEORY

GIMAS9AI - Mines Nancy **Information Theory**	**Crédits : 2 ECTS** **Durée : 21 heures**	**Semestre : S9**

Responsable(s) : PEYRE Rémi

Mots clés : Information, Kolmogorov complexity, Shannon entropy, Data compression, Kullback-Leibler divergence, Cramér-Rao bound, model selection

Pré requis : Intermediate-level knowlegde in probability theory and statistics ; general knowledge in mathematics ; general programming skills

Objectif général : Getting acquainted with the concepts of information theory which are useful for an engineer in mathematics, especially in data science

Programmes et contenus : This course offers a panorama on various topics around information theory :

- **How can one measure an amount of information? Link with data compression. The case of Kolmogorov complexity. The case of the Shannon entropy.**
- **Main results on Shannon information: chain rule, data treatment inequality; &c.**
- **Lossy data compression: what is the maximum compression rate that you can achieve for a signal up to a certain tolerable distorsion?**
- **Kullback-Leibler divergence and large deviation theory: how surprising is a result with respect to a given belief?**
- **The Cramér-Rao bound: in statistics, this is a fundamental limit on how much information you can get about a hidden parameter.**
- **Information theory as a tool for model selection: justification for the AIC and BIC criteria.**

Compétences :

Niveaux	Description et verbes opérationnels
Connaître	To know the definitions of Kolmogorov complexity, Shannon entropy, Kullback-Leibler divergence, Fisher information; together with their main mathematical properties.
Comprendre	To understand what “measuring an amount of information” means, and in which sense compressing, describing and predicting are equivalent.
Appliquer	To implement some basic data-compression and decompression algorithms. To compute and compare AIC and BIC criteria.
Analyser	To compute how much information is fundamentally contained in a partly random signal, or how surprising is a signal w.r.t. a given model.
Synthétiser	To use the tools of information theory to give a precise meaning to how much a signal is “complex”, or “blurry”.
Évaluer	To compare the respective relevances of two models in statistical data analysis. To compare a statistical technique with the Cramér–Rao benchmark.

Évaluations : (*) The main exam shall be a classical 3-hour written test (maybe with a small programming part). In case of failure, the second-chance exam shall be a homework followed by an interview about the student’s work and some other questions.