**Data Mining &Machine Learning**

### Installation instructions

Some links for installation of Data Mining tools.

R, RStudio

Python , Anaconda, Spyder

These installations depends on the operating system

###### **R and RStudio**

R is the engine, and RStudio is only a GUI.

1. Python

Spyder

Anaconda distribution ,

Notebook , jupyter notebook command

2 ) Practical work subjects

On ISTL Book

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Contenu** | **Notions** | **TP associés** | **Commentaires, Références** |
| Cours 1 (3h) | Introduction / Motivations Use-Cases , Vocabulaire, Modèles linaires de régression,  Classification | Framework de l’apprentissage artificiel, apprentissage statistique, Motivations, Droite de régression. Framework statistique (proba / stat, intervalle de confiance), Modèle de régression linéaire (une puis plusieurs variables)  Modèle de régression logistique | Installation des outils et bibliothèques. Anaconda python, R+RStudio. TP régression linéaire .  TP Regression logistique | ISTL Chap 1 et 2 et 3,  ISTL Chap 4 |
| Cours 2 (3h) | Revision des notions cours 1, Sampling and model selection  Méthode d’arbre de décision | Sélection de model, introduction au feature engineering. Bootstraping.  Arbre de décision, Ensembling | TP1 , prévision des survivants du Titanic | ISTL Chap 5 et 6, ISTL Chap 8 |
| Quizz num 1 avant TP1 |  |  |  |  |
| TP 1 (4h) | Qq TP simple sur ISTL, puis long TP :  Prediction des survivants du Titanic |  |  | Utilisation de R+RStudio, plus intégré. |
| Cours 3 (3h) | Méthode SVM.  Méthodes de clustering | Apprentissage non supervisé |  | ISTL Chap 9 |
| Quizz num 2 avant TP2 |  |  |  |  |
| TP2 (4h) | Sur SVM et Algo k-means, |  |  |  |
| Cours 4 (3h) | Systèmes de recommandation.  Wrap-up du cours.  Suite, état de l’art : deep learning,  Stacking de modele, blending | Filtrage collaboratif. Item based ou user based content filtering. |  | ISTL Chap 10,  MMS Chap 9 |
| Quizz num 3 avant TP3 |  |  |  |  |
|  |  |  |  |  |
| TP 3 (4h) | Filtrage collaboratif, algo sur les graphes |  |  |  |