

UFR de Sciences Appliquées et de Technologie Mention : Informatique

Niveau : Master 2 GDIL

EC : Data Mining

UE : Extraction de connaissances

**Projet de TP** : Etude comparative des modèles de classification

**Présenté par : Doctorant :**

* Mr KOUA
* Faback Dieng (P26 830)
* Cheikhna Diaw (P31 5381)
* El hadji Lèye ()

Année académique : 2022-2023



[I. Introduction 3](#_Toc151817484)

[II. Etude Comparative des modèles de classification 3](#_Toc151817485)

[1. Régression Logistique 3](#_Toc151817486)

[2. Arbre de Classification 3](#_Toc151817487)

[3. Naïve Bayes 3](#_Toc151817488)

[4. SVM (Support Vector Machine) 3](#_Toc151817489)

[III. Conclusion 3](#_Toc151817490)

# Introduction

# Etude Comparative des modèles de classification

### Régression Logistique

La régression logistique est un algorithme supervisé de classification. Elle mesure la relation entre la variable dépendante catégorielle et une ou plusieurs variables indépendantes en donnant une estimation à la probabilité d’occurrence d’un évènement à travers l’usage de sa fonction logistique.

### Arbre de Classification

Un arbre de décision ou de classification permet d’expliquer une variable cible à partir d’autres variables dites explicatives. L’algorithme permet de déterminer la meilleure caractéristique dans l’ensemble de données, diviser les données en sous-ensembles contenant les valeurs possibles de la meilleure caractéristique, générer de manière récursive de nouveaux arbres de décision en utilisant les sous-ensembles de données crées et lorsqu’on ne peut plus classifier les données, on s’arrête.

### Naïve Bayes

### SVM (Support Vector Machine)

Le SVM est l’un des modèles le plus populaire de l’apprentissage automatique. Il appartient à la famille d’algorithme d’apprentissage supervisé. Son but est de séparer les données en classes

# Conclusion