



## **MASTER 2 DATA SCIENCE**

Cours : Knowledge Discovery in Graph Data.

Établissement : UNIVERSITE PARIS SACLAY

Enseignant : FATIHA SAIS

**Sujet** : Project Exam : Key Selection - 05/01/2021

Etudiant : Salma DAMAK.

Etudiant : Haytham AIT ETTAJER.

Etudiant : Panongbene Jean Mouhamed SAWADOGO.

## **SOMMAIRE :**

I. INTRODUCTION

II. DATASET.

III. GRAPH ANALYSIS.

IV. TO BUILD MINIHASHING MATRIX.

V. TO BUILD SIMILARITY MATRIX

VI. LOCALITY SENSITIVE HASHING

# I. INTRODUCTION

In this work, we analyzed the properties of a social network dataset. For this, we use the python language with the Jupyter tool. In this folder you will find the Jupyter file which has the name `network_analysis.ipynb` (this file contains the python code to analyze the network), you will find the `README.rd` file which explains how the code works and other explanations for the specifics of our code. The `Project.pdf` pdf file which explains the work we need to do. In the data directory, you will find the `wiki-Vote.txt` file which contains the data representing our graph. You will also find this report which explains in detail our work as a whole.

## II. DATASET