

SDS 3120: Data Mining

Dr. Mutua Kilai

May 2024

Department of Pure and Applied Sciences

Purpose of the Course

Important

The purpose of the course is to equip learners with necessary skills for developing, manipulating, recovering and maintaining databases.

Learning Outcomes

- By the end of the Course the learner should be able to:
 - Explain concepts of data mining
 - Apply mining on different kinds of data
 - Apply mining on different kind of knowledge
 - Discuss principles of pattern analysis and machine learning
 - Demonstrate representation discovery

Course Content

- In this course we will consider the following problems:
 - Classification, cluster and outlier analysis
 - Mining time series and sequence data
 - Text mining, web mining and pattern analysis

Course Content Cont'd

- We cover fundamental data mining ideas:
 - Clustering, SVM
 - Semi-supervised learning
 - Information retrieval
 - Most important algorithms in data mining

Logistics of the Course

- **Course Website:** <https://sam-mutua.github.io/datamin23/>
- All Lecture Notes will be uploaded per week.
- All Assignments will also be posted in the course website.
- The website also has useful resources on data mining.

Dr. Mutua Kilai, PhD

Programming Languages

- We will use **R** and **Python** programming languages which are **Free and Open Source**.
- We will also use SQL for databases lectures.
- Details on Installing the softwares can be found [here](#)