|  |  |  |
| --- | --- | --- |
| Immagine che contiene testo, logo, emblema, simbolo  Descrizione generata automaticamente | Erasmus Plus: approvato il programma di lavoro 2023 | Progettare InEuropa | logo_tdtu |

Seminar series

Prof. TRAN TRANH

**Ton Duc Thang University (TDTU) Vietnam**

Data Mining

October 7-11, 2024

Monday 7 October 2023 11:00-13:00 Room TBA

Tuesday 8 October 2023 14:00-16:00 Room TBA

Thursday 10 October 2024 14:00-16:00 Room TBA

Friday 11 October 2024 11:00-13:00 Room TBA

Lecture will be also available via Cisco Webex at the address https://unicam.webex.com/meet/renato.deleone

Contents

1. Introduction of data mining: Data mining is the process of discovering patterns,  
   correlations, and anomalies within large sets of data to predict outcomes. By using a  
   combination of statistics and database systems, data mining transforms raw data into  
   valuable insights. This interdisciplinary field, also known as Knowledge Discovery in  
   Databases (KDD), involves several steps including data cleaning, data integration,  
   data selection, data transformation, pattern discovery, and knowledge presentation.
2. Definitions: Data set, transactions, power item set, support threshold, support  
   calculation, high frequent item sets, occupancy threshold, occupancy calculation,  
   high occupancy item set, etc.
3. Mining of high frequency item sets: Given a progress of mining high frequent item  
   sets, develop a program for mining frequent item set.
4. Mining of high occupancy item sets: Given a progress of mining high occupancy  
   item sets; develop a program for mining high occupancy item set.
5. Data mining applications: Using Python to develop a program for mining high  
   frequent item sets and high occupancy item sets.