Department of MACS, NITK Surathkal

MA855: Big data and Analytics
Jan-May 2023
Project   Details   MA648

Course Evaluation plan:		Seminar 10%	Report: 2.5
			Presentation: 05
			Teamwork: 2.5
Mid Sem:	25%	Project 20%	Report: 05
End Sem :	35%		Demonstartion: 10
Quiz :	10 %		Teamwork:05
Seminar :	10 %		
Project :	20%		

MA855 Big data and Analytics Mini Project:

Note:1. This project is for the whole class, i.e., all Teams.
*   2. Implement the task using Python Programming.*

Qn.1. The task is to create and analyze Co-Authorship network using Scopus or Web of science databases.

For a selected keyword or a combination of key words (List of keywords are given below), find:

a) Highest cited author and his h-index (from the world)

b) Highest publication author

c) Highest cited authors avg. citations, and the country name.

d)   Total number of publications of the highest cited author

e) Total publication in year

f)   Total citation per year

g) Author(country) having highest co-authorship with indian authors.

h) Highest cited author from India and the university.

i)   Comparative year wise article publication analysis of india, china and usa.

j)   Total number of grants given to the field
k) Country wise total number of publication

Qn. 2. Co-Author Relationship Prediction and Citation prediction using known machine learning   techniques.

Note:
Co-Author: If Authors A and B said to be co-author, then they written research article together.
Citation: If author A refers author B's journal, then A cites B' s work, and in a graph a directed link between node A to node B indicates the co-authorship relation.

Steps to download data(may vary between the databases):

STEP 1. Go to the homepage of SCOPUS https://www.scopus.com or https://apps.webofknowledge.com/
STEP 2. Type your relevant topic (KEYWORD) in Search box
STEP 3. Click source type- check to JOURNALS.
STEP 4- Document type- ARTICLE
STEP 5- Click Subject Area (i.e Computer Science , Information systems etc)
STEP 6. After all the steps- Click "LIMIT TO" or refine in web of science
STEP 7. find out best articles by click sort based on Number of citations
STEP 8. Click Export button after selecting articles based on method adopted,
        Choose CSV (comma separated value) Excel to export the data

STEP 9.From the above data Create new Excel file with required columns to analyze given set of attributes.

List of keywords (put with in double quotes " "):
Not exhaustive

1. Wireless Sensor networks
2.   bio geography-based optimization
3.   block-chain
4. quantum computing
5. game theory
6. internet of things
7. deep learning
8.   big data analytics
9. fuzzy logic
10. reinforcement learning
11. India
12. USA
13. China
14. artificial intelligence
15. computer science
16. meta heuristic
17. Evolutionary
18. bio-inspired
19. graph theory
20. graph coloring

21. ill posed problems
22. robotics
23. corona virus
24. cancer

Links:

Web of science and Scopus record:

1.https://www.youtube.com/watch?v=dQgmIcVXqu8&feature=youtube

2. http://networksciencebook.com/translations/en/resources/data.html

Instructions for Project:
1. The duration of Demonstration for each team (4 members) is 30 mins. The weightage for project components is given in table above.
2. The group members should work together to prepare the report and presentation slides.

Demonstration Schedule:

Sl.No	Team No.	Date and Time
1	Team 1, 2,3,4	April 12, 2023, 2:00 to 4:00 pm
2	Team 5,6,7, 8	April 13, 2023, 2:00 to 4:00 pm
3	Team 9,10,11,12	April 14, 2023, 2:00 to 4:00 pm
4	Team 13,14,15,16	April 15, 2023, 2:00 to 4:00 pm

Course Instructor: Dr. Pushparaj Shetty D.