WEEK -2

Lecture -14 Introduction to Datasets

With the advent of the internet, computing, and storage facilities, we now have a lot of data to analyze.

The Zachary Karate network was a popular social network dataset consisting of 32 nodes and some edges.

The module requires a good understanding of Python, networkx, and other APIs taught in the previous week.

The social network datasets are like Play-Doh, where one can mold them to ask a variety of questions.

The first question to ask on a friendship network dataset is whether the graph is connected.

The second question is about the degree of nodes and the average and standard deviation of friendships.

The datasets provide the ability to ask random questions and arrive at fantastic conclusions.

Title: Example of a Network of Dish Ingredients

Lecture -15 Network Dish Ingredients

With the advent of the internet, there are a plethora of datasets available.

One interesting dataset is a network of dish ingredients.

To create the network, all possible ingredients in the world are taken, and edges are placed between two ingredients if they are part of a dish.

Only popular combinations of ingredients are considered for the network.

There will be strong communities in this network, where ingredients are mostly connected to each other and not connected to other ingredients outside of the community.

Studying this network is interesting and provides hints on the overall behavior of the system.

The community structure in this network has properties that may be observed in other networks.

Lecture -16 Synonymy Network

An example of how network theory can be applied to the English language. If we create a network of English words based on their synonyms, the network is surprisingly connected, meaning there is a path from any given word to any other word through synonyms. This implies that some seemingly opposite words, such as love and hate, are actually synonyms.

This example shows how network theory can provide unexpected and eyebrow-raising results when applied to various entities or phenomena.

Lecture -17 Web Graph

The lecture discusses the importance of network datasets, particularly the web graph or the network of the World Wide Web. The example shows how the web graph, which is a collection of web pages connected by hyperlinks, became the basis of Google's search algorithm. By analyzing the connections between web pages, Google was able to provide more relevant search results. The lecture emphasizes the importance of mining network datasets to extract valuable information and insights.

Lecture – 18 :- Handling Real-world Network Datasets

Notes:

- ➔ Social networks are represented as graphs, with nodes representing people and edges representing friendships.
- ➔ Facebook and other social networking sites use this data to recommend new friends to users.
- ➔ Link prediction is the process of analyzing the network structure to determine the most probable next friend for a user.
- ➔ Recommending good friends to users can increase user engagement and the popularity of social networking sites.