



CA-I

Name : Shankar Ankush Beralé.

Class : B-Tech -2

PRN : 1921321242077

subject : Social Networks.

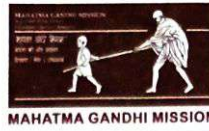
Q.1] A) Write What is link Prediction in social Networks explain with examples. Also write the usage of Link Prediction.

→
a) Link Prediction :-

i) Link Prediction is used to predict future possible links in a network.

ii) Link Prediction is the algorithm based on which facebook recommends people you may know, Amazon predicts items you're likely going to be interested in.

iii) Link Prediction is a task in social network analysis that involves predicting the likelihood of a connection between two nodes in a network.



iv) Given a network of nodes and edges, the goal is to predict which pair of nodes are likely to be connected in the future based on their current connectivity patterns.

b) Example:-

i) Let's say we have a social network of users and their friends. Each user is represented as a node in the network, and each friendship is represented as an edge between two nodes.

We can use link prediction to predict which users are likely to become friends in the future based on their current social network.

e.g. Facebook, LinkedIn, Snap, Instagram, etc.

c) Usage:-

i) Recommender Systems:-

i) Link prediction is often used in recommender systems, which are designed to recommend new connections or products to users based on their past behaviour.

ii) By predicting which pair of users are likely to become friends or collaborate in the future, recommender systems can suggest new connections or



PAN: 1921321242077

Name: Shankar Bevale

products to users that they may be interested in.

2) Fraud Detection:

i) Link prediction can also be used to detect fraud or other malicious behaviour in social network.

ii) e.g. if a network of users is being used for illegal activities, link prediction algorithm can be used to identify suspicious connections and alert authorities.

3) Marketing:

i) Link prediction can also be used in marketing to identify potential customers or influencers.

ii) By analyzing social network data, marketers can identify which users are most likely to become advocates for a brand or product and target their marketing efforts accordingly.



PAN: 1921321242077.

Name: Shankar Bevale.

4) Disease spread:

i) Link prediction can be used in public health to predict the spread of diseases.

ii) By analyzing social network data, public health officials can identify which individuals are most likely to come into contact with infected individuals and take preventive measures to limit the spread of the disease.



PRN: 1921321242077.

Name: Shankar Bevale

Q.1 / B]

What is web graph? Explain google page ranking using web graph?

→
a) Web graph.

i) A web graph is a mathematical model that represents the structure of the world wide web as a directed graph.

ii) In this model, web pages are represented as nodes, and links between pages are represented as edges.

iii) The web page graph is a useful tool for studying the properties of the web and understanding the behavior of users who navigate it.

b) Google page ranking using web graph:

i) Google PageRank is an algorithm that uses the web graph to rank web pages based on their importance or relevance to particular search query.

ii) The pageRank algorithm works by assigning a score to each web page based on the number and quality of links pointing to it. The more links a page has from other pages with high PageRank scores, the higher its own PageRank score will be.



PRN: 1921321242077.

Name: Shankar Bera.

iii) The quality of a link is determined by the PageRank score of the page that is linking to it. Pages with high PageRank scores are considered to be more authoritative and valuable, and therefore have a greater influence on the PageRank scores of the pages they link to.

iv) To calculate the page rank score for a web page, Google uses a recursive algorithm that iteratively calculates the PageRank Score for all the pages in the web page.

v) The algorithm starts by assigning an initial PageRank scores for all the pages in the web graph.

vi) The algorithm starts by assigning an initial PageRank score of $1/N$ to each page, where N is the total number of pages in the web graph.

vii) It then iteratively updates the pageRank score for each page based on the pageRank scores of the pages that link to it.

viii) This process continues until the PageRank scores converge to a stable value.



PRN: 1921321242077.

Name: Shankar Berale

ix] The web graph is useful tool for studying the properties of web.

x] Google PageRank algorithm was developed by Larry Page and Sergey Brin, the founders of Google, and is based on the idea that a ~~we~~ web page is more important if it is linked to by other important pages.

xi] The PageRank algorithm has become an important tool for search engine optimization (SEO) and web marketing, as it can help webmasters understand how their pages are perceived by search engines and users.

xii] By analysing the web graph and optimizing their pages to increase their PageRank scores, ~~web~~ web masters can improve their visibility and attract more traffic to their sites.



PRN: 1921321242077,
Name: Shankar Berale

Q. 2] A) Write a short note on.

a) Clustering Coefficients:

- i) clustering coefficient is a measure of the degree to which nodes in a social network tend to cluster together.
- ii) It measured the proportion of nodes neighbors that are also neighbors of each other.
- iii) High clustering coefficients are typically observed in social network where individuals tend to form closed-knit communities or groups.
- iv) low clustering coefficients, on the other hand, indicate that nodes tend to be more isolated from each other and may be more likely to form random connections.
- v) clustering coefficients can be used to identify important nodes in a social network, such as leaders or influencers who are highly connected to other nodes in their community. Tl
- vi) they can also be used to analyse the structure



PRN: 1921321242077.

Name: Shankar Bevale.

and dynamics of social networks and to develop models for predicting the behavior of nodes in a network.

b) Neighbourhood Overlap:

i) Neighbourhood overlap is a measure of the similarity between the neighbors of two nodes in a social network.

ii) It measures the proportion of a node's neighbors that ~~can~~ are also neighbors of another node.

iii) High neighborhood overlap is typically observed in social networks where individuals tend to form tightly interconnected communities or groups.

iv) Low neighborhood overlap, on the other hand, indicates that nodes tend to be more isolated from each other and may be more likely to form random connections.

v) Neighborhood overlap can be used to identify nodes that are likely to form connections in the future.

vi) It can also be used to analyse the structure and dynamics of social networks and to develop models for predicting the behavior of nodes in a network.



PAN: 1921321242077.
Name: Shankar Bevale.

Q.2/B

- i) Calculate neighbourhood overlap if no. of friends of $A = 20$, no. of friends of $B = 10$ and total no. of friends $= 18$.



Given Data,

no. of friends of $A = 20$

no. of friends of $B = 10$

Total no. of friends $= 18$.

neighbourhood overlap $= ?$

$$\text{Neighbourhood overlap}(A, B) = \frac{\text{no. of friends}(A \cap B)}{\text{no. of friends}(A \cup B)}$$

a) First we find no. of friends $(A \cap B)$

A has $= 20$ friends

B has $= 10$ friends.

Total $= 18$

\therefore The number of friends $(A \cap B)$

$$= \min(20, 10, 18)$$

$$= 10, 12,$$

~~$A \cap B$~~

$$\boxed{(A \cap B) = 12}$$



PRN! 1921321242077.

Name! Shankar Bevale

b) find no. of friends (A ∪ B)

$$\text{no. of friends (A ∪ B)} = \text{no. of friends (A)} + \text{no. of friends (B)} - (A ∩ B)$$

$$\therefore (A ∪ B) = 20 + 10 - 12$$

$$= 30 - 12$$

$$\text{no. of friends } \boxed{(A ∪ B) = 18}$$

c) find neighborhood overlap (A, B)

$$= \frac{\text{no. of friends (A ∩ B)}}{\text{no. of friends (A ∪ B)}}$$

$$= \frac{12}{20} = \frac{12}{20}$$

$$= \frac{6}{10}$$

$$= \underline{\underline{0.66}}$$

∴ The neighborhood overlap (A, B) is 0.66 OR 66%



PRN: 1921321242077.
Name: Shankar Bevale.

Q.2)

Q]

ii]

Explain Triadic closure, membership closure and foci closure.

→

a) Triadic closure.

i] Triadic closure refers to the tendency for people who share a common connection to become connected themselves. This phenomenon is also known as the "friend of a friend" effect.

ii] e.g. if person A is friends with B, and person B is friend with person C, there is greater chance that person A and person C will become friends as well.

iii] This effect is important in social network because it can lead to the formation of cluster or communities within the network.

b) Membership Closure:

i] Membership closure refers to the tendency for people who share a common attribute or characteristic to become connected. This phenomenon is also known as homophily.



PRN: 1921321242077,

Name: Shankar Berale.

ii) e.g. people work in the same industry are more likely to form connections with each other, even if they do not work for the same company.

iii) This effect can be particularly important in shaping the structure of social networks, as it can lead to the formation of subgroups or cliques.

c) Foci Closure:

i) foci closure refers to the tendency for people who share a common interest or activity to become connected. This phenomenon is also known as focal closure.

ii) e.g. people who enjoy playing chess are more likely to form connection with each other, even if they do not have any other shared attributes.

iii) This effect can also be important in shaping the structure of social network, as it can lead to the formation of interest-based groups or communities.