

Assignment submitted on 2024-01-30, 23:01 IST

1) Which of the following is/are anonymous social networking sites ? **1 point**

- Whisper
- Sarahah
- Facebook
- X (Formerly Twitter)

Yes, the answer is correct.

Score: 1

Accepted Answers:

Whisper

Sarahah

2) Which of the following statements is/are true about ephemeral content on social networks ? **1 point**

- Content on social networks disappears after a fixed period of time.
- Snaps on Snapchat are an example of ephemeral content.
- Stories on Instagram are an example of ephemeral content.
- None of the above.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Content on social networks disappears after a fixed period of time.

Snaps on Snapchat are an example of ephemeral content.

Stories on Instagram are an example of ephemeral content.

3) Which of the following 4 V's of Social Media addresses the correctness of information posted online ? **1 point**

- Velocity
- Veracity
- Volume
- Variety

Yes, the answer is correct.

Score: 1

Accepted Answers:

Veracity

4) Facebook Friends and Instagram Followers follow the principles of a unidirectional **1 point** network.

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

False

5) ____ helps users on social media to find posts related to a specific topic by clicking **1 point** or searching.

- Mentions
- Hashtags
- Shares
- Likes

Yes, the answer is correct.

Score: 1

Accepted Answers:

Hashtags

6) Degrees of Separation has reduced on Social Media due to:

1 point

- Facilitating Connections and Interactions across various locations
- Interconnected Networks
- Share Information and Communicate Globally
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Facilitating Connections and Interactions across various locations

Interconnected Networks

Share Information and Communicate Globally

7) Which of the following is the correct syntax for installing numpy version 1.21.5 **1 point**

- pip3 install numpy==1.21.5
- pip3 install package==numpy version==1.21.5
- pip3 install numpy --v=1.21.5
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

pip3 install numpy==1.21.5

8) Which of the following libraries can be used to send and receive requests to a server? **1 point**

- numpy
- requests
- math
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

requests

9) A self-contained directory containing its own Python installation (for a particular version of Python) and libraries, allowing it to manage and isolate project dependencies. **1 point**

- Python Standard Libraries
- PIP
- Virtual Environment
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Virtual Environment

10) Which of the following are the advantages of a version control system ? **1 point**

- Helps in keeping a record of all changes made to a file in its history.
- Allows for collaboration between various people working on the same project.
- Allows for backup and recovery in a project.
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Helps in keeping a record of all changes made to a file in its history.

Allows for collaboration between various people working on the same project.

Allows for backup and recovery in a project.

11) Social Media constitutes: **1 point**

- Social Networks (Facebook, LinkedIn etc.)
- Virtual Worlds (Habbo, Club Penguin etc.)
- Livecasts (Twitch, Yahoo! Live etc.)
- Content Sharing Platforms (YouTube, Flickr etc.)

Yes, the answer is correct.

Score: 1

Accepted Answers:

Social Networks (Facebook, LinkedIn etc.)

Virtual Worlds (Habbo, Club Penguin etc.)

Livecasts (Twitch, Yahoo! Live etc.)

Content Sharing Platforms (YouTube, Flickr etc.)

12) Which of the following platforms allow for managing/hosting git repositories ? **1 point**

- GitHub
- GitLab
- Bitbucket
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the above

13) Select the correct syntax for allowing git to track all files in a repository (assume you are in the root of the repository) **1 point**

- git add --all
- git commit -m "track"
- git status
- git log

Yes, the answer is correct.

Score: 1

Accepted Answers:

git add --all

14) Creating a new file in a git repository adds it to : **1 point**

- Tracked Files
- Untracked Files
- Remote Repository
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Untracked Files

15) Select the correct syntax for pushing an existing repository to a remote repository on GitHub **1 point**

- git push origin master
- git commit -m "Push to GitHub"
- git add remote origin
- git pull origin master

Yes, the answer is correct.

Score: 1

Accepted Answers:

git push origin master



16) Which of the following is the correct syntax for installing a package Ubuntu computer ? **1 point**

- sudo apt install <PACKAGE_NAME>
- sudo apt-<PACKAGE_NAME> install ubuntu
- sudo apt-install <PACKAGE_NAME>
- sudo get <PACKAGE_NAME>

Yes, the answer is correct.

Score: 1

Accepted Answers:

sudo apt install <PACKAGE_NAME>

1. Which of the following is NOT a social network?

-
- a) Facebook
 - b) Koo
 - c) Naukri
 - d) LinkedIn

2. StackOverflow and Quora belong to which class of online networks?

-
- a) Discussion forums and Community-based Question Answering
 - b) Bookmarking and content curation networks
 - c) Consumer review networks
 - d) Blogging and publishing networks

3. Which of the following is the verified twitter handle of Bengaluru City Police?

-
- a) @BlrCityPolice
 - b) @blrpolice
 - c) @bangalore_police
 - d) @bengaluru_police

4. Which of the following Twitter handle likely belongs to a bot account (i.e., likely a fake account)?

-
- a) @DelhiPolice
 - b) @delhipolice1
 - c) @hydcitypolice
 - d) @Navimumpolice

6. Trending is a feature of which of the following networks?

-
- a) Twitter
 - b) Facebook
 - c) Youtube
 - d) All the above

7. Which of the following is a Facebook feature?

-
- a) Charity Donations
 - b) Groups
 - c) Pages
 - d) All of the above

8. Which of the following are anonymous networks?

-
- a) Whisper
 - b) Blind
 - c) 4chan
 - d) Indeed

9. When a person blocks someone on Facebook, the other party gets a notification.

-
- a) True
 - b) False

10. Orkut was owned by:

- a) Google
- b) Microsoft
- c) Amazon
- d) Uber

1. Which is the correct online platform, as depicted in Figure 1?

- a) Koo
- b) Blind
- c) Reddit
- d) LinkedIn



2. Which of the following statement(s) is/are true about Anonymous social networks?

- a) Anonymous social networks help in providing enhanced privacy and anonymity.
- b) Anonymous social networks provide more freedom of expression.
- c) Anonymous social networks reduce the fear of judgment or social consequences.
- d) None of the above

3. Choose the correct option to create a virtual environment in Python.

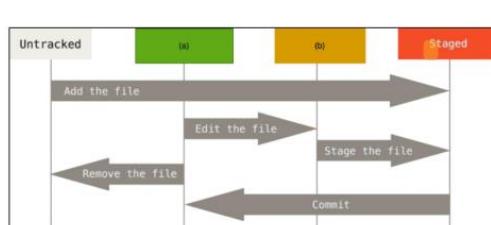
- a) `python3 -m venv <path_to_virtual_environment>`
- b) `python3 venv -m <path_to_virtual_environment>`
- c) `virtualpython3 -m venv`
- d) `virtualenv python3 -m venv myenv`

4. Why is it important to create a virtual environment in Python?

- a) Creating a virtual environment isolates project dependencies and prevents conflicts between different projects.
- b) Creating a virtual environment enhances the performance of Python programs.
- c) Creating a virtual environment to ensure compatibility with different operating systems.
- d) Creating a virtual environment optimizes memory usage in Python applications.

5. Fill in the correct option to stage files/directories.

- a) Modified, Unmodified
- b) **Unmodified, Modified**
- c) Modified, Unstaged
- d) Modified, Tracked



6. Which command lists all the remote addresses associated with a local repository?

- a) `git remote -v`
- b) `git pull origin master`
- c) `Git stash`
- d) `Git checkout`

7. Which of the following concept is closely related to the fact that infrequent, arms-length relationships are more beneficial for employment opportunities, promotions, and wages than strong ties?

-
- a) Six degrees of separation
 - b) Small world phenomenon
 - c) Strength of weak ties
 - d) None of the above

8. Which of the following is/are shortcomings of online social media platforms?

-
- a) Spread of misinformation
 - b) Cyberbullying and harassment
 - c) Addiction and time consumption
 - d) Limited control over privacy

9. When a piece of content, such as a video, meme, or news article, gains attention on social media platforms, it can quickly reach a large audience within a short period. Which V of social media does it refer to?

-
- a) Velocity
 - b) Veracity
 - c) Value
 - d) Volume

10. What is the primary purpose of the Gab platform?

-
- a) To promote free speech
 - b) To facilitate online shopping experience
 - c) It provides educational resources
 - d) Support for professional networking

1) Which of the following is NOT a social network?

- Facebook
- Gmail
- Twitter
- LinkedIn

Yes, the answer is correct.

Score: 1

Accepted Answers:

Gmail

2) Quora belongs to which class of social networks?

- Discussion forums
- Bookmarking and content curation networks
- Consumer review networks
- Blogging and publishing networks

Yes, the answer is correct.

Score: 1

Accepted Answers:

Discussion forums

7) Which of the following is a Facebook feature?

- Charity Donations
- Groups
- Pages
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the above

8) Which of the following are anonymous networks?

- Reddit
- Glassdoor
- Twitter
- Facebook

Yes, the answer is correct.

Score: 1

Accepted Answers:

Reddit

Glassdoor

- 1) Which of the following is the correct social platform, as depicted in Figure 1?



Figure 1

- Whisper
- Facebook
- Reddit
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Reddit

- 2) Which of the following statement(s) is/are true about Anonymous social networks?

- Content posted on Anonymous Social is not connected with particular online identities or profiles
- Anonymous social networks might ask for log-in details and sync your contact details to personalise the user experience
- Anonymous social networks help build a professional network and seek a job
- Anonymous networks are safe.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Content posted on Anonymous Social is not connected with particular online identities or profiles

Anonymous social networks might ask for log-in details and sync your contact details to personalise the user experience

3) What is the output of the response variable in the below code:

```
import requests  
response=requests.get("https://precog.iiit.ac.in/")
```

- < Response [200] >
- <Response [400]>
- <Response [800]>
- None of the above responses

Yes, the answer is correct.

Score: 1

Accepted Answers:

< Response [200] >

4) State whether True or False:

As per guidelines for creating an anonymous Twitter account, 'You can't make an anonymous Twitter account by using your real phone number.

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

5) What does Value specify in 5Vs on Social Media?

- Confirmation of the information posted as legitimate
- Data can be obtained from many different sources
- Ability to get value out of the platform
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Ability to get value out of the platform

6) Which one of the following is an example of shoulder surfing?

- Mail claiming that your system has been infected or the account has been breached
- Someone sees your password standing behind you while you type without letting you know.
- Emails that claim you've won a prize and require you to go on a link
- None of the above.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Someone sees your password standing behind you while you type without letting you know.

7) You have two files in the folder: "privacy.py" and "security.py". Now you want to command to stage the privacy.py file and commit. Which of the following option are correct?

- git add privacy.py git commit -m "Commit message"
- git add privacy.py git commit -m "privacy.py"
- git add security.py git commit -m "security.py"
- git add security.py git commit -m "privacy.py"

Yes, the answer is correct.

Score: 1

Accepted Answers:

git add privacy.py git commit -m "Commit message"

8) Which of the following statements is/are true?

- Python is a language that can be used to draw network graphs.
- It is an interpreted language.
- Python Functions do not have restrictions on the type of the argument and the type of its return value.
- None of the above

Partially Correct.

Score: 0.66

Accepted Answers:

Python is a language that can be used to draw network graphs.

It is an interpreted language.

Python Functions do not have restrictions on the type of the argument and the type of its return value.

9) Match the following. Choose the most appropriate option.

Social Networks	About
 i) Tinder	1. Help to stay anonymous
 ii) Flickr	2. Social news aggregation
 iii) Reddit	3. Organizing photos
 iv) Koo	4. Regional language microblogging website
 v) Whisper	5. Dating service

10) Choose the incorrect option about ephemeral content.

- (i) Instagram Stories are a type of ephemeral content.
- (ii) Ephemeral content helps in generating more user-generated content.

- Only i
- Only ii
- Both i and ii
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

None of the above

9) Which of the following is the social networking site is owned by Microsoft?

- Bing
- LinkedIn
- Gabb
- Flickr

No, the answer is incorrect.

Score: 0

Accepted Answers:

LinkedIn

1) Which of the following activities count as phishing?

- Using social media to resolve a crisis.
- Threatening and bullying users online with aggressive language
- Mails from fake websites asking for credit card information such as account number and password.
- None of the above.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Mails from fake websites asking for credit card information such as account number and password.

3) What is the difference between a bidirectional and unidirectional social media network?

- A bidirectional network allows both users to connect with each other at once, whereas a unidirectional users in the network.
- Unidirectional networks use messages for communication whereas bidirectional networks use emails.
- Bidirectional networks always involve pictorial data such as images and videos whereas unidirectional messages.
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

A bidirectional network allows both users to connect with each other at once, whereas a unidirectional network requires separate approval from both users in the network.

4) Which of the following options have only unidirectional networks?

- Twitter, Facebook, LinkedIn
- Twitter, Google+
- Twitter, LinkedIn, Google+
- Facebook, Google+

No, the answer is incorrect.

Score: 0

Accepted Answers:

Twitter, Google+

5) Which of the following incidents marks the first time Twitter was used to solve a crisis?

- Boston Marathon bombing incident
- Plane landing in Hudson River
- UK Riots
- Floods in Chennai

No, the answer is incorrect.

Score: 0

Accepted Answers:

Plane landing in Hudson River

6) What are anonymous social networks?

- Networks where the content is hidden and identity of user posting it can't be revealed
- Networks where the content is posted and get destroyed by itself
- Professional networks to connect with individuals having same occupation or career interests
- Multimedia networks which allow users to share video and images with friends

No, the answer is incorrect.

Score: 0

Accepted Answers:

Networks where the content is hidden and identity of user posting it can't be revealed

7) Which of the following are examples of microblogging social media networks?

- Twitter
- YouTube
- LinkedIn
- Facebook

No, the answer is incorrect.

Score: 0

Accepted Answers:

Twitter

9) Which of the following are location based applications?

- Tinder
- LinkedIn
- Snapchat
- Foursquare

No, the answer is incorrect.

Score: 0

Accepted Answers:

Tinder

Foursquare

10) Which of the following is/are ephemeral networks?

- Pinterest
- Snapchat
- Twitter
- Facebook

No, the answer is incorrect.

Score: 0

Accepted Answers:

Snapchat

11) What does Veracity in the 4/5 Vs of Social Media refer to?

- Confirmation whether the information posted is legitimate
- Value of the content being posted.
- Amount of content being posted on the platform
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Confirmation whether the information posted is legitimate

12) Which of the following social media networks uses “circles” as a building block of social media networks?

- Foursquare
- Twitter
- Facebook
- Google+

No, the answer is incorrect.

Score: 0

Accepted Answers:

Google+

13) What is the main difference between Periscope and YouTube?

- Periscope is mainly used to broadcast live videos whereas YouTube is used to upload and store content that can be viewed at any time.
- Periscope is a location-based network whereas YouTube is not.
- Periscope is an anonymous network whereas YouTube is not.
- Periscope is used for images, whereas YouTube is used for videos.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Periscope is mainly used to broadcast live videos whereas YouTube is used to upload and store content that can be viewed at any time.

Content may vary by preview only.

14) What is the difference between a hashtag and a mention?

1 pc

- A hashtag is used to identify a keyword or topic of interest and facilitate a search by other users while a mention is a Tweet that contains another user's @username to specifically tag him.
- A hashtag is used for objects and a mention is used for people.
- A hashtag can only be used with images, whereas mentions are used with text
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

A hashtag is used to identify a keyword or topic of interest and facilitate a search by other users while a mention is a Tweet that contains another user's @username to specifically tag him.

15) What is pip?

- Package manager for python modules
- Command for administrative rights in Ubuntu
- Command used to check reachability of networks
- Command used to uninstall packages in Ubuntu

No, the answer is incorrect.

Score: 0

Accepted Answers:

Package manager for python modules

16) What is sudo used for?

- To change user access permissions
- Grants administrative rights in Ubuntu when used with any command
- Used to install python in Ubuntu
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Grants administrative rights in Ubuntu when used with any command

17) Which of these commands can be used to make a file executable by everyone?

- Chmod +x filename
- chmod a+x filename
- Chmod 444 filename
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

*Chmod +x filename
chmod a+x filename*

18) Which of these commands are used to import a package in Python?

- Get
- Use
- Import
- Package

No, the answer is incorrect.

Score: 0

Accepted Answers:

Import

19) What does it mean if an account is verified on Twitter?

- The account was successfully created.
- The account has been verified to be a constructive bot.
- The account is of public interest and may belong to government agencies, celebrities etc.
- The account is linked with a Facebook page

No, the answer is incorrect.

Score: 0

Accepted Answers:

The account is of public interest and may belong to government agencies, celebrities etc.

20) Which of the following statements is true?

- Python is a high level programming language.
- Python is an interpreted language and object-oriented language.
- In Python identifiers are case sensitive .
- All of the above.

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above.

1) What does JSON stand for ?

1 point

- JavaScript Object Notation
- Java Standard Object Notation
- JSON Scripted Object Name
- Joint Syntax Object Network

Yes, the answer is correct.

Score: 1

Accepted Answers:

JavaScript Object Notation

2) In which of the following formats does an API usually return its response ?

1 point

- JSON
- XML
- TSV
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

JSON

XML

3) MongoDB stores data in a structure most similar to ____.

1 point

- JSON
- SQL
- XML
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

JSON

4) MySQL can be used for querying relational databases.

0 points

- True
- False

Yes, the answer is correct.

Score: 0

Accepted Answers:

True

5) Given a MongoDB collection named "words" with the following document structure: **1 point**

```
{  
    "_id":ObjectId("615243d88f0ae43f255874c1"),  
    "count": 300,  
    "department": "the"  
}
```

Which MongoDB query would you use to retrieve all words with count in descending order ?

- db.words.find().sort({ count: -1 })
- db.words.find().sort({ count: 1 })
- db.words.find().order_by("count", DESCENDING)
- db.words.find().sort({ descending: 1 })

Yes, the answer is correct.

Score: 1

Accepted Answers:

`db.words.find().sort({ count: -1 })`

6) Which of the following are the components of a graph ?

1 point

- Vertices
- Edges
- Unique Identifier/ID
- All of the Above

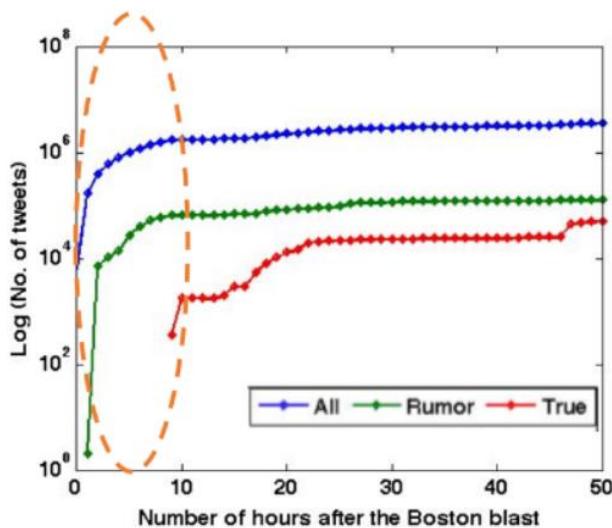
Yes, the answer is correct.

Score: 1

Accepted Answers:

`All of the Above`

7) Which of the following is true regarding the situation depicted in the graph below



- False information/rumors spreads faster than facts/true information on social media
- True information spreads faster than false information/rumors on social media
- Cannot be Determined

Yes, the answer is correct.

Score: 1

Accepted Answers:

`False information/rumors spreads faster than facts/true information on social media`

8) Select the correct sequence of steps taken towards evaluating social media data. **1 point**

- Data Collection/Filtering > Data Characterization > Data Classification > Data Evaluation
- Data Collection/Filtering > Data Characterization > Data Evaluation > Data Classification
- Data Characterization > Data Collection/Filtering > Data Evaluation > Data Classification
- Data Characterization > Data Collection/Filtering > Data Classification > Data Evaluation

Yes, the answer is correct.

Score: 1

Accepted Answers:

Data Collection/Filtering > Data Characterization > Data Classification > Data Evaluation

9) Which of the following attributes can be inferred from data collected via X (formerly **1 point** Twitter) ?

- Number of Likes of the Tweet
- Number of times the Tweet has been viewed
- Unique identifier of the user/author of the Tweet
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Number of Likes of the Tweet

Number of times the Tweet has been viewed

Unique identifier of the user/author of the Tweet

10) Which of the following processes can be used to receive annotated posts while **1 point** determining their legitimacy ?

- A group of friends annotating posts given to them
- Crowd-sourcing / Mechanical Turk
- Getting Data from Fact Checking / News Outlets
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the above

11) TweetCred is a chrome extension that ranks the credibility of each tweet on Twitter. **1 point**

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

12) Which Python library would you most likely use to get the top 10 hot posts from the **1 point** subreddit "r/politics" using the Reddit API?

- praw
- requests
- urllib
- json

Yes, the answer is correct.

Score: 1

Accepted Answers:

praw

13) An indicator similar to hashtags is called _____ on Reddit.

1 point

- Flair
- Karma
- Comment
- Upvote

Yes, the answer is correct.

Score: 1

Accepted Answers:

Flair

14) Communities on Reddit are called _____

1 point

- Subreddits
- Flairs
- Threads
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Subreddits

15) A higher Follower-Friend ratio on Twitter can be associated with the popularity of **1 point** the user

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

1. What is the primary data format utilized by the Reddit API when returning information?

- a) JSON
- b) XML
- c) HTML
- d) CSV

4. Suppose you want to retrieve the top 5 hottest posts from the subreddit "r/technology" using the Reddit API. Which Python library would you most likely use to achieve this?

- a)

```
import praw
reddit = praw.Reddit(client_id='MY_CLIENT_ID',
client_secret='MY_CLIENT_SECRET',
user_agent='MY_USER_AGENT')
posts = reddit.subreddit('worldnews').hot(limit=5)
for post in posts:
    print(post.title)
```
- b)

```
import requests
url = "https://www.reddit.com/r/worldnews/new.json?limit=5"
response = requests.get(url)
data = response.json()
for post in data['data']['children']:
    print(post['data']['title'])
```
- c)

```
import praw
reddit = praw.Reddit(client_id='MY_CLIENT_ID',
client_secret='MY_CLIENT_SECRET', user_agent='MY_USER_AGENT')
posts = reddit.subreddit('worldnews').new(limit=10)
for post in posts:
    print(post.title)
```
- d) None of the above

5. Which of the following tweet content falls in category of linguistic features?

- a) Number of followers
- b) Presence of negative emotion words
- c) Presence of smiley
- d) Presence of pronouns

6. Why do we do network analysis on social media platforms?

- a) To analyze the performance of computer networks and improve data transmission speeds.
- b) To understand the social structures and relationships between individuals or entities in a network.
- c) To study the behavior of network protocols and optimize their efficiency.
- d) To identify and analyze patterns of diseases spreading within a population.

7. Which of the following is a recent example of misinformation spread on social media platforms?

- a) Tweet claiming that a particular brand of COVID-19 vaccine causes severe side effects, with no evidence to support the claim.
- b) Facebook post sharing information from a well-known health organization about the effectiveness of vaccination.
- c) Instagram story containing accurate information about a recent political development with credible sources cited.
- d) YouTube video providing a detailed explanation of climate change supported by scientific evidence.

8. Given a MongoDB collection named "employees" with the following document structure:

```
{  
  "_id": ObjectId("615243d88f0ae43f255874c1"),  
  "name": "Ram",  
  "department": "CSE",  
  "age": 35,  
  "salary": 50000  
}
```

What MongoDB query would you use to retrieve all employees with a salary greater than or equal to 60000?

- a) db.employees.find({ "salary": { \$gte: 60000 } })
- b) db.employees.find({ "salary": { \$gt: 60000 } })
- c) db.employees.find({ "salary": { \$lte: 60000 } })
- d) db.employees.find({ "salary": { \$lt: 60000 } })

9. You are developing a social media platform, and you have a database table, "posts" to store user posts. The "posts" table schema is given below:

```
CREATE TABLE posts (
post_id INT PRIMARY KEY,
user_id INT,
post_content TEXT,
post_date DATETIME
);
```

Now, you want to retrieve the 5 most recent posts made by a specific user with the user ID "12345". Which SQL query would you use?

- a) **SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_date DESC LIMIT 5;**
- b) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_id DESC LIMIT 5;
- c) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_date ASC LIMIT 5;
- d) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_id ASC LIMIT 5;

10. What is the Facebook Graph API as discussed in lecture?

- a) An API that provides access to a user's Facebook friends list.
- b) An API that allows developers to create and manage Facebook pages.
- c) An API that enables developers to integrate Facebook login into their applications for user authentication.
- d) **An API that provides access to a wide range of user data, interactions, and connections on Facebook.**

Which of the following is/are true about APIs?

- 1. APIs ~~cannot~~ be designed for external use by third-party developers to access the functionality and data of an application.
- 2. The API is a set of rules and protocols for building and interacting with software applications.
- 3. An API allows different systems to communicate with each other over a network, such as through HTTP requests and responses.
- 4. APIs can access a wide range of functionality and data, such as retrieving data from a database, creating or updating resources, or performing operations on a remote system.

- Tweet about “Floods in McDonald’s at Virginia beach” is related to which Misinformation news spreads?
 - Boston’s explosion
 - Hurricane Sandy
 - Shootouts in the Mexican city
 - None of the above

Assignment submitted on 2023-02-06, 13:31 IST

1) Which of the following is/are false about RoboMongo (Robo3T)?

- RoboMongo is a free and open-source graphical user interface for MongoDB.
- It allows developers to interact with MongoDB databases in a visual and user-friendly way.
- It doesn't provide features such as in-place editing, SSL/TLS support, and support for MongoDB Atlas clusters.
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

It doesn't provide features such as in-place editing, SSL/TLS support, and support for MongoDB Atlas clusters.

2) Choose / Match the most appropriate option with respect to data collection:

Data	API
1. submissions data	i) Reddit
2. Other user's comments data	ii) Twitter
3. user profile data	iii) Facebook

- 1- i) 2. iii) 3. ii)
- 1- ii) 2. i) 3. iii)
- 1- iii) 2. ii) 3. i)
- 1- i) 2. i) 3. ii)

No, the answer is incorrect.

Score: 0

Accepted Answers:

1- i) 2. i) 3. ii)

3) Which of the following statements is/are true?

S1: Research on Hurricane Sandy and Boston blasts shows how Twitter users debunk falsehoods during disasters.

S2: Boston Blasts happened on 15 April 2013.

- S1
- S2
- Both
- None

No, the answer is incorrect.

Score: 0

Accepted Answers:

Both

4) Which of the following statements are true about Relational Databases?

- i. They provided a standard way of representing and querying data and are table-based.
 - ii. They are a good fit for static datasets.
 - iii. The relationships between tables in a relational database are defined by using keys.
- I, iii
 i, ii
 II, III
 All are correct

Yes, the answer is correct.

Score: 1

Accepted Answers:

All are correct

5) Which of the following is/are designed to handle large amounts of unstructured and semi-structured data?

- CouchDB
 Oracle SQL
 MongoDB
 JavaDB

Yes, the answer is correct.

Score: 1

6) Which of the following can be potentially used to collect the below data?



MySQL

Graph API

Facebook API

RoboMongo

Yes, the answer is correct.

Score: 1

Accepted Answers:

Graph API

7) Which of the following is/are a Visualization tool?

- PhpMyAdmin
- RoboMongo
- Compass
- Tableau

Partially Correct.

Score: 0.75

Accepted Answers:

PhpMyAdmin

RoboMongo

Compass

Tableau

8) Which features perform better to distinguish fake and legitimate tweets?

- tweet-based features
- Meta-data features
- user-based features
- All the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

tweet-based features

9) Tweet about "Floods in McDonald's at Virginia beach" is related to which Misinformation news spreads?

- Boston's explosion
- Hurricane Sandy
- Shootouts in the Mexican city
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Hurricane Sandy

10) Choose the correct statement (s):

- Inter-annotator agreement is used to calculate how good/bad annotations are.
- NoSQL databases have a flexible schema, which means that the structure of the data can change over time.
- Kohen's Kappa is one of the best ways to calculate inter-annotator agreement.
- It is not important to report how trustworthy the annotation is.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Inter-annotator agreement is used to calculate how good/bad annotations are.

NoSQL databases have a flexible schema, which means that the structure of the data can change over time.

Kohen's Kappa is one of the best ways to calculate inter-annotator agreement.

12) Choose the correct statement (s):

- S1:** The decision tree picks the best features from the available data than naive Bayes.
S2: Naive Bayes algorithm is harder to debug.

- Only S1 is true
- Only S2 is true
- Both S1 and S2 are true
- Both S1 and S2 are false

Yes, the answer is correct.

Score: 0

Accepted Answers:

Both S1 and S2 are true

13) Which of the following is/are python wrappers to collect data?

- PRAW
- reddit_api
- API_wrapper
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

PRAW

14) Which of the following option(s) is/are true about API?

1 point

- APIs cannot be designed for external use by third-party developers to access the functionality and data of an application.
- The API is a set of rules and protocols for building and interacting with software applications.
- An API allows different systems to communicate with each other over a network, such as through HTTP requests and responses.
- APIs can access a wide range of functionality and data, such as retrieving data from a database, creating or updating resources, or performing operations on a remote system.

Yes, the answer is correct.

Score: 1

Accepted Answers:

The API is a set of rules and protocols for building and interacting with software applications.

An API allows different systems to communicate with each other over a network, such as through HTTP requests and responses.

APIs can access a wide range of functionality and data, such as retrieving data from a database, creating or updating resources, or performing operations on a remote system.

3) MySQL is a kind of:

- Relational database
- NoSQL database
- Text based database
- In memory database

No, the answer is incorrect.

Score: 0

Accepted Answers:

Relational database

4) Which of the following are non-relational databases?

- MySQL
- Oracle SQL
- MongoDB
- JavaDB

No, the answer is incorrect.

Score: 0

Accepted Answers:

MongoDB

5) Why is Facebook API called the Graph API?

- It returns all data in a graph format
- It is easy to plot graphs directly from the data returned by the API
- It stores data in a graph format
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

It stores data in a graph format

6) With respect to Boston blast, how quickly fake content became viral?

- Under 30 mins
- 1-2 hours
- 3-4 hours
- 7-8 hours

No, the answer is incorrect.

Score: 0

Accepted Answers:

7-8 hours

7) PhpMyAdmin is a:

- API
- Programming language
- Database
- Visualization tool

No, the answer is incorrect.

Score: 0

Accepted Answers:

Visualization tool

9) After approximately how many hours of the Boston blast did the number of tweets spreading true information become equal to the rumour tweets? 1 point

- 50
- 95
- 150
- 9

No, the answer is incorrect.

Score: 0

Accepted Answers:

95

10) Hurricane Sandy occurred in the year: 1 point

- 2011
- 2012
- 2013
- 2014

No, the answer is incorrect.

Score: 0

Accepted Answers:

2012

11) Which of the following rumors were spread during hurricane sandy incident? 1 point

- Flood in McDonalds at Virginia beach
- Shark on waterlogged highway
- Fake pictures of cloud
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

13) Hoax on the death of 8 year old boy is associated with which of the following incidents? 1 point

- Hurricane Sandy
- Boston blasts
- UK riots
- Hurricane Irene

No, the answer is incorrect.

Score: 0

Accepted Answers:

Boston blasts

14) Which of the following is/are not user based features? 1 point

- Number of friends
- User has URL
- Number of URL's in tweet
- Number of times listed

No, the answer is incorrect.

Score: 0

Accepted Answers:

Number of URL's in tweet

15) Which of the following features was the most influential in determining the credibility of a tweet in the Boston blast dataset? 1 point

- The number of words in a tweet
- The number of retweets of a tweet
- The number of characters in a tweet
- The number of URLs in a tweet

No, the answer is incorrect.

Score: 0

Accepted Answers:

The number of characters in a tweet

16) Approximately how many new accounts were created that posted about Boston blasts?

- 30 thousand
- 40 thousand
- 50 thousand
- 20 thousand

No, the answer is incorrect.

Score: 0

Accepted Answers:

30 thousand

17) Which of the following is among top 10 features for real time tweet analysis?

- Number of followers
- Number of followees
- User has location in profile
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

User has location in profile

18) What is NDCG?

- Technique for collecting ground truth data for classification tasks
- Technique for calculating credibility of user generated content
- Metric for calculating the confidence in classification accuracy
- Metric for calculating the quality of ranking

No, the answer is incorrect.

Score: 0

Accepted Answers:

Metric for calculating the quality of ranking

19) In Boston blasts incident, fake contents were spread mostly from

- Web
- Mobile
- Others
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Mobile

20) WOT stands for

- Web Of Trust
- Web Of Test
- Web Organization
- None of the above

No, the answer is incorrect.

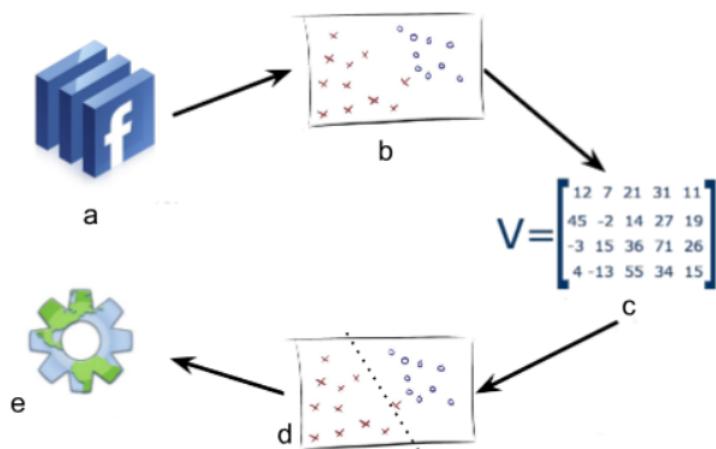
Score: 0

Accepted Answers:

Web Of Trust

1) Select the correct sequence of steps as used by Facebook Inspector (as Discussed in Lecture)

1 point



- Data Collection > Ground Truth Extraction > Feature Generation > Supervised Learning > Result via RESTful API
- Feature Generation > Data Collection > Supervised Learning > Ground Truth Extraction > Result via RESTful API
- Ground Truth Extraction > Data Collection > Supervised Learning > Feature Generation > Result via RESTful API
- Data Collection > Feature Generation > Ground Truth Extraction > Supervised Learning > Result via RESTful API

Yes, the answer is correct.

Score: 1



2) Which of the following is/are correct statements in context of Web of Trust Score? 1 point

- Web of Trust score outputs the reputation of an input website
- Domain Age can be used as a feature for a Web of Trust score
- Domain Owner can be used a feature for a Web of Trust score
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Web of Trust score outputs the reputation of an input website

Domain Age can be used as a feature for a Web of Trust score

Domain Owner can be used a feature for a Web of Trust score

3) "Shantanu never likes people taking his photo because he doesn't want it to be posted online". According to Westin's 3 Categories, Shantanu is a : **1 point**

- Fundamentalist
- Pragmatist
- Unconcerned
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Fundamentalist

4) "Geeta browses all social media websites on her laptop casually, but always opens her banking website in incognito for checking her balance" According to Westin's 3 Categories, Geeta is a :

- Fundamentalist
- Pragmatist
- Unconcerned
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Pragmatist

5) According to Westin's 3 Categories, the lowest percentage of US citizens fall into the category of (as discussed in the lecture): **1 point**

- Fundamentalist
- Pragmatist
- Unconcerned

Yes, the answer is correct.

Score: 1

Accepted Answers:

Unconcerned

6) According to Westin's 3 Categories, the highest percentage of US citizens fall into the category of (as discussed in the lecture) : **1 point**

- Fundamentalist
- Pragmatist
- Unconcerned

Yes, the answer is correct.

Score: 1

Accepted Answers:

Pragmatist

7) Which of the following can be used in authentication for the Twitter API ?

1 point

- Bearer Token
- Client ID
- Client Secret
- None of the above

Partially Correct.

Score: 0.33

Accepted Answers:

Bearer Token

Client ID

Client Secret

8) Which of the following can be achieved using Twitter API ?

1 point

- Retrieving reverse chronological home timeline
- Retrieving user Tweet timeline
- Retrieving user mention timeline
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Retrieving reverse chronological home timeline

Retrieving user Tweet timeline

Retrieving user mention timeline

9) Which of the following is the correct syntax for creating a tweets table with the specified attributes?

1 point

tweet_id : Primary Key, Not Null
text: Not Null

CREATE TABLE tweets (
 tweet_id VARCHAR(50) NOT NULL,
 text VARCHAR(500) NOT NULL,
 PRIMARY KEY (tweet_id)
);

CREATE tweets TABLE (
 tweet_id VARCHAR(50) NOT NULL,
 text VARCHAR(500) NOT NULL,
 PRIMARY KEY (tweet_id)
);

CREATE TABLE (
 tweet_id VARCHAR(50) NOT NULL,
 text VARCHAR(500) NOT NULL,
 PRIMARY KEY (tweet_id),
 TABLE NAME(tweets)
);

CREATE TABLE tweets (
 tweet_id VARCHAR(50),
 text VARCHAR(500),
 PRIMARY KEY (tweet_id),
 NOT NULL (text)
);



10) Consider a table named "users" containing user profiles from Twitter, which of the **1 point** following attributes can be used as a primary key in the table :

- Full Name of the User
- User ID of the User
- No. of Followers of the User
- No. of Tweets by the User

Yes, the answer is correct.

Score: 1

Accepted Answers:

User ID of the User

13) A table in SQL is called a _____ in MongoDB.

1 point

- Database
- Index
- Document
- Collection

Yes, the answer is correct.

Score: 1

Accepted Answers:

Collection

14) A row in SQL is called a _____ in MongoDB.

1 point

- Collection
- Field
- Index
- Document

Yes, the answer is correct.

Score: 1

Accepted Answers:

Document

15) Which of the following can be used to create a new database called "project" and **1 point** create a collection named "users" inside it in MongoDB ?

- 1. use project;
2. db.users.insert({});
- 1. use project;
2. db.users.insert();
- 1. create_db project;
2. db.users.insert({});
- 1. create project;
2. db.users.insert();

Yes, the answer is correct.

Score: 1

Accepted Answers:

*1. use project;
2. db.users.insert();*

1. Consider the below response:

"To this segment, consumer privacy is very important; they feel that they have been victims of privacy invasions, they are pessimistic about the future of privacy protection, and about a third of them favor creating a general federal regulatory agency on consumer privacy."

To which of the three Westin Privacy categories is "they" in the above sentence most likely to fall?

-
- a) Fundamentalists
 - b) Pragmatists
 - c) Unconcerned
 - d) None of the above

2. Which of the following features is used by the credibility model-TweetCred?

-
- a) Tweet links
 - b) Tweet Author
 - c) Tweet network
 - d) Tweet meta data

3. Cronbach's alpha

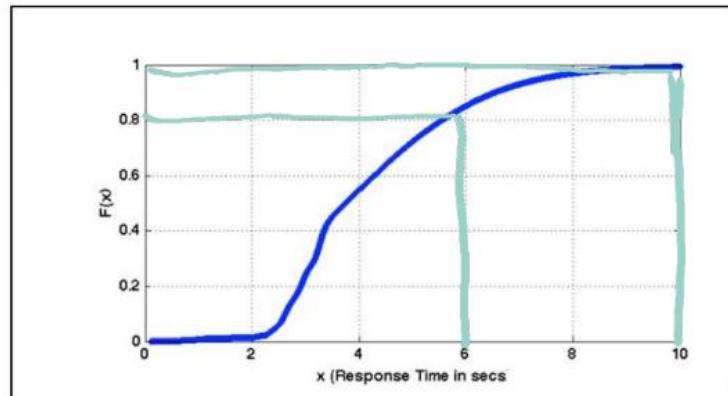
-
- a) determines the statistical significance of a research study.
 - b) **assess the reliability and internal consistency of a questionnaire.**
 - c) analyze the effect size of an intervention.
 - d) calculate the standard deviation of a sample population.

4. You developed a search engine that ranks websites based on their relevance to a user's query. Each website has a relevance score assigned by human experts (1: Not relevant, 2: Moderately relevant, 3: Highly relevant). You compare your ranking algorithm's performance with another using NDCG@5. The other algorithm's ranked list for a specific query includes two "Highly relevant" websites at positions 1 and 4, one "Moderately relevant" website at position 2, and two "Not relevant" websites at positions 3 and 5. Based on NDCG@5, which algorithm is likely to be better for user satisfaction at the top 5 results?

-
- a) **Your algorithm, as it prioritizes more "Highly relevant" websites at the top.**
 - b) The other algorithm, as it has more websites with non-zero relevance scores in the top 5.
 - c) It's impossible to say without knowing the specific relevance scores of your ranked websites.
 - d) Both algorithms are equally good for user satisfaction.

8. Consider the below graph and choose the correct option:

- a) For 82% of the users, the response time was less than 2 seconds, and for 99%, the response time was under 10 seconds.
- b) For 99% of the users, the response time was less than 6 seconds, and for 100%, the response time was under 10 seconds.
- c) For 82% of the users, the response time was less than 6 seconds, and for 99%, the response time was under 10 seconds.
- d) None of the above



9. What is the Likert scale?

- a) rating scale in surveys and questionnaires
- b) gives the score for any real-time tweets
- c) Interannotator agreement
- d) None of the above

10. TweetCred is

- a) a real-time web-based system to automatically evaluate the credibility of content on Twitter.
- b) a real-time app-based system to automatically evaluate the score of tweets on Twitter.
- c) a real-time app-based system to automatically assess the quality of content on Twitter.
- d) None of the above

- Which of the following is the most powerful tool to build real-time applications that can process and analyse tweets and can do sentiment analysis, trending events, etc.?
 - Twitter API → *select data for specific parameters* ⇒ *not real time*
 - Graph API
 - Twitter streaming API** → *live data*
 - PRAW
- Which of the following statement(s) is/are true about Twitter and its API?
 - S1: Twitter keys can be created from developer.twitter.com
 - S2: Twitter API has no rate limit. We can collect any number of tweets using Twitter API.
 - S3: Streaming API can collect all the tweets related to certain events.
- Consider the following examples. Match the categories: Fundamentalists, pragmatists and unconcerned with each example.
 - Example 1: a business leader who makes decisions based on what will be most profitable for the company rather than adhering to a particular business philosophy.
 - Example 2: A person who strictly adheres to the literal interpretation of religious texts and refuses to consider other perspectives or modern interpretations.
 - Example 3: an individual is unconcerned about the environment and takes no steps to reduce their carbon footprint or conserve resources.

- ✓ **Example 1: Pragmatist, Example 2: Fundamentalists, Example 3: unconcerned**
- Example 1: pragmatist, Example 2: unconcerned, Example3:unconcerned
 - Example 1: unconcerned, Example 2: pragmatists, Example3:unconcerned
 - Example 1: Fundamentalist, Example 2:unconcerned, Example3:unconcerned

- Which of the following statement(s) is/are true about fundamentalists?
 - They generally avoid being identified, avoid being in pictures
 - They generally use technologies to avoid tracking them
 - They generally don't care about their personal information
 - They generally care about their personal information but don't care about professional ones.
- The basic method for adding data to MongoDB is Insert. To insert a single document, use the collection's insertOne method.
 - Insert, insertOne**
 - Insert, insertMany
 - Add, addOne
 - None of the above

- Which of the following python command(s) is/are correct w.r.t to connection with MongoDB ?

- `from pymongo import MongoClient`
- `client = MongoClient(port=27017)`
- `import pymongo` → all the func in the library.
`client = pymongo.MongoClient("localhost:27017")`
- `from pymongo import MongoClient`
`client = MongoClient('localhost:27017')`
- `from pymongo import MongoClient`
`db = client.business`

default IP = localhost
 (port = 27017)
 "IP: port"
 ("localhost : 27017")
 'IP: port'
 ('localhost : 27017')

Given below is the sample data in JSON format. Here x randomly named restaurants with corresponding ratings will be created in the MongoDB database called “restaurants”. These restaurants are created in a single collection called “reviews”. Choose the correct command to find the total number of restaurants rated with a 1.

```
restaurants = {u'rating': 1,
  u'_id': ObjectId('58e66688hdtg67jf195'),
  u'name': u'pizza wings',
  u'cuisine': u'Italian'}
```

- db.reviews.find({'rating': 1}).count()
- db.reviews.find_one({'rating': 1})
- db.reviews.find_all({'rating': 1})
- db.reviews.find({'rating': 1}).count

- 1) Which of the following statement(s) is/are true about Twitter and its API?

S1: Twitter keys can be created from developer.twitter.com

S2: Twitter API has no rate limit. We can collect any number of tweets using Twitter API.

S3: Streaming API can collect all the tweets related to certain events.

- Only S1 is true
- S2 and S3 are true
- All S1, S2, S3 are true
- S1 and S3 are true

No, the answer is incorrect.

Score: 0

Accepted Answers:

Only S1 is true

2) Consider the following statements. Which of the following options is false?

- MySQL supports dynamic queries and is schema-free, whereas Mongo DB is schema-dependent.
- MongoDB is a document-based data model.
- MongoDB is a better choice than MySQL when your data has a lot of relations.
- MongoDB is faster than MySQL.

Yes, the answer is correct.

Score: 1

Accepted Answers:

MySQL supports dynamic queries and is schema-free, whereas Mongo DB is schema-dependent.

3) Consider the following statements. Which option(s) is false in understanding privacy perceptions? 1 point

- If you receive a friend request on your most frequently used OSN, you are more likely to add a person of the opposite gender as a friend.
- Many people feel about the personal information that even if they have specified their privacy settings, their data is not secure from a privacy breach.
- If you receive a friend request on your most frequently used OSN, you are less likely to add a person of the opposite gender as a friend.
- If you receive a friend request on your most frequently used OSN, you are more likely to add anyone as a friend.

No, the answer is incorrect.

Score: 0

Accepted Answers:

If you receive a friend request on your most frequently used OSN, you are less likely to add a person of the opposite gender as a friend.

If you receive a friend request on your most frequently used OSN, you are more likely to add anyone as a friend.

5) Consider the following examples. Match the categories: Fundamentalists, pragmatists and unconcerned with each example. 1 point

Example 1: a business leader who makes decisions based on what will be most profitable for the company rather than adhering to a particular business philosophy.

Example 2: A person who strictly adheres to the literal interpretation of religious texts and refuses to consider other perspectives or modern interpretations.

Example 3: an individual is unconcerned about the environment and takes no steps to reduce their carbon footprint or conserve resources.

- Example 1: Pragmatist, Example 2: Fundamentalists, Example 3: unconcerned
- Example 1: pragmatist, Example 2: unconcerned, Example3:unconcerned
- Example 1: unconcerned, Example 2: pragmatists, Example3:unconcerned
- Example 1: Fundamentalist, Example 2:unconcerned, Example3:unconcerned

Yes, the answer is correct.

7) Which of the following is true?

- TweetCred is a chrome extension that helps to find tweet credibility.
- Web of Trust score is used to denote a website's reputation.
- Both a) and b)
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Both a) and b)

1) In the context of Facebook, what does the term 'FBI' stand for?

- Federal Bureau of Investigation
- Facebook Inspector
- Facebook Bureau of Investigation
- Facebook Branch of Investment

No, the answer is incorrect.

Score: 0

Accepted Answers:

Facebook Inspector

Accepted Answers:

Facebook Inspector

2) Which of the following is Twitter's equivalent of FBI?

- Twitter API
- Twitter Search
- Tweetcred
- Web of Trust

No, the answer is incorrect.

Score: 0

Accepted Answers:

Tweetcred

3) Which of the following is true regarding Web of Trust?

- It provides the trustworthiness of a given website
- It provides the confidence of its rating
- It can classify whether a website is malicious or not
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

4) What is Westin's definition of a "Fundamentalist" in the context of privacy concerns?

- Someone who is concerned about privacy issues depending on the situation, based on how fair the tradeoff is.
- Someone who doesn't care about privacy issues at all.
- Someone who is very concerned about privacy, and is unlikely to give away private information at all.
- None of the above.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Someone who is very concerned about privacy, and is unlikely to give away private information at all.

5) What is Westin's definition of a "Pragmatist" in the context of privacy concerns?

- Someone who is concerned about privacy issues depending on the situation, based on how fair the tradeoff is.
- Someone who doesn't care about privacy issues at all.
- Someone who is very concerned about privacy, and is unlikely to give away private information at all.
- None of the above.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Someone who is concerned about privacy issues depending on the situation, based on how fair the tradeoff is.

6) What is Westin's definition of "Unconcerned" in the context of privacy concerns?

- Someone who is concerned about privacy issues depending on the situation, based on how fair the tradeoff is.
- Someone who doesn't care about privacy issues at all.
- Someone who is very concerned about privacy, and is unlikely to give away private information at all.
- None of the above.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Someone who doesn't care about privacy issues at all.

7) What is the population distribution of Westin's 3 categories?

- Fundamentalists-25%, Pragmatists-60%, Unconcerned-15%
- Fundamentalists-5%, Pragmatists-80%, Unconcerned-15%
- Fundamentalists-50%, Pragmatists-40%, Unconcerned-10%
- Fundamentalists-25%, Pragmatists-20%, Unconcerned-55%

No, the answer is incorrect.

Score: 0

Accepted Answers:

Fundamentalists-25%, Pragmatists-60%, Unconcerned-15%

9) In the survey conducted to understand privacy perceptions in India, which of the following was the most common choice in response to the question "If you receive a friend request on your most frequently used OSN, which of the following people will you add as friends?" 1 point

- Person of the opposite gender
- People from my hometown
- A person with a nice profile picture
- Strangers (people you do not know)

No, the answer is incorrect.

Score: 0

Accepted Answers:

Person of the opposite gender

8) In the survey conducted to understand privacy perceptions in India, which of the following was the most common choice in response to the question "What do you feel about the privacy of your personal information on your OSN?" 1 point

- Even though I have specified my privacy settings I am concerned about the privacy of my data.
- It is a concern, but I still share my personal information.
- It is not a concern at all.
- Since I have specified my privacy settings, my data is secure from a privacy breach.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Since I have specified my privacy settings, my data is secure from a privacy breach.

10) In the survey conducted to understand privacy perceptions in India, which of the following was the least common choice in response to the question "If you receive a friend request on your most frequently used OSN, which of the following people will you add as friends?" **1 point**

- Person of the opposite gender
- People from my hometown
- A person with a nice profile picture
- Strangers (people you do not know)

No, the answer is incorrect.

Score: 0

Accepted Answers:

Strangers (people you do not know)

1) What are the different forms of Privacy?

1 point

- Information
- Communication
- Territorial
- Bodily

Yes, the answer is correct.

Score: 1

Accepted Answers:

Information

Communication

Territorial

Bodily

2) "Please Rob Me" highlighted issues specifically stemming from:

1 point

- Public Self-Disclosure of Information on Social Media by users
- Data leaks from social media websites due to hacking attempts
- Sharing passwords with your friends on Social Media
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Public Self-Disclosure of Information on Social Media by users

3) Using an unidentified source of information to extract an individual's sensitive information is called :

1 point

- Re-identification
- Personally Identifiable Information
- Self-Disclosure
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Re-identification

4) Considering each attribute in isolation, which of the following are examples of Personally Identifiable Information (PII): **1 point**

- Race
- Age
- Aadhaar Number
- College Student ID

No, the answer is incorrect.

Score: 0

Accepted Answers:

Aadhaar Number

College Student ID

5) What is the purpose of k-anonymity as discussed in the lecture? **1 point**

- Protect PII's of re-identified individuals
- Anonymise data without losing utility
- Make datasets larger for analysis
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Protect PII's of re-identified individuals

Anonymise data without losing utility

6) Latanya Sweeney's experiments aimed to : **1 point**

- Re-identify individuals from Match.com using their Facebook profiles
- Identification of individuals walking on a campus with data from Facebook
- Predict social security number from publicly available data
- Study behaviors of individuals aged 14-18 on Instagram

Yes, the answer is correct.

Score: 1

Accepted Answers:

Re-identify individuals from Match.com using their Facebook profiles

Identification of individuals walking on a campus with data from Facebook

Predict social security number from publicly available data

7) What are the advantages of NumPy Arrays over Lists in Python?

1 point

- NumPy arrays are faster and more compact than Python lists.
- NumPy uses much less memory to store data and it provides a mechanism of specifying the data types.
- This allows the code to be optimized even further.
- None of the above.

Partially Correct.

Score: 0.66

Accepted Answers:

NumPy arrays are faster and more compact than Python lists.

NumPy uses much less memory to store data and it provides a mechanism of specifying the data types.

This allows the code to be optimized even further.

8) Write Python code to convert the list [1,2,3,4] to a NumPy array.

1 point

- import numpy as np
arr = np.array([1,2,3,4])
- import numpy as np
arr = numpy.array([1,2,3,4])
- import numpy as np
arr = np.array(1,2,3,4)
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

import numpy as np

arr = np.array([1,2,3,4])

9) What is the output of the following code:

1 point

```
import numpy as np
a = np.array([range(i,i+3) for i in [1,2,3]])
print(a)
```

- [[1 2 3]
[2 3 4]
[3 4 5]]
- [[1 2 3]
[1 2 3]
[1 2 3]]]
- [[1 2 3]
[4 5 6]
[7 8 9]]]
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

[[1 2 3]

[2 3 4]

[3 4 5]]

10) Which of the following codes can be used to make a NumPy array of 10 random numbers? **1 point**

- import numpy as np
import random
- I = []
for i in range(10):
 I.append(random.random())
- zeros_array = np.array(I)
print(zeros_array)
- import numpy as np
random_array = np.random.rand(10)
- import numpy as np
random_array = np.random.random(10)
- None of the above

Partially Correct.

Score: 0.67

Accepted Answers:

*import numpy as np
import random*

*I = []
for i in range(10):
 I.append(random.random())*

*zeros_array = np.array(I)
print(zeros_array)*

*import numpy as np
random_array = np.random.rand(10)*

*import numpy as np
random_array = np.random.random(10)*

11) Which of the following NumPy functions can be used to generate an NxN diagonal matrix ? **1 point**

- np.eye(5)
- np.diagonal(np.eye(5))
- np.diagonal(5)
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

np.eye(5)

12) Which of the following Python code can be used to read a CSV file using Pandas **1 point**

- import pandas as pd
data = pd.read_excel('file.csv')
- import pandas as pd
data = pandas.readcsv('file.csv')
- import pandas as pd
data = pd.read_csv('file.csv')
- None of the above.

Yes, the answer is correct.

Score: 1

Accepted Answers:

```
import pandas as pd  
data = pd.read_csv('file.csv')
```

13) Consider the below structure of a dataframe.

1 point

Vehicle Number	Offence Date	Offence At	Offence Detail	amount	offence_day_name	offence_Month	offence_hour
DL3CCF 08XX	2015-11-02 13:10:00	DASHRAT HPURI M/S (JPC)	REFUSAL TO GO BY TSR/TAXI	100	Monday	11	13

Which of the following Python snippets can be used to plot a histogram of the amount column ?

Assume df is defined as the above dataframe with 100 rows.

- df['amount'].plot.hist()
- df[df.columns[4]].plot.hist()
-
- import matplotlib.pyplot as plt
plt.hist(df['amount'])
plt.xlabel('Amount')
plt.ylabel('Frequency')
plt.title('Histogram of Amount')
plt.show()
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

```
df['amount'].plot.hist()  
df[df.columns[4]].plot.hist()
```



14) What is the output of the following code ?

1 point

```
pd.Series(['A', 'B', 'C'], index=[1, 2, 3])
```

- 1 A
- 2 B
- 3 C
- A 1
- B 2
- C 3
- A B
- C 1
- 2 3
- None of the above

Yes, the answer is correct.

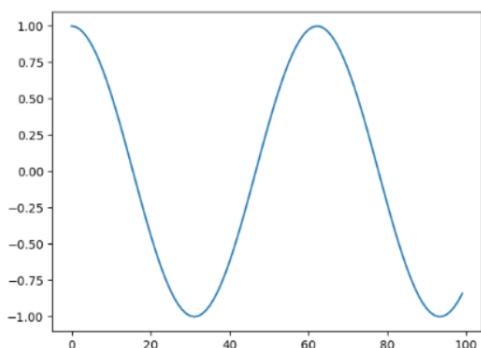
Score: 1

Accepted Answers:

```
1 A  
2 B  
3 C
```

15) Select the correct option for plotting the below function using matplotlib

1 point



```
import numpy as np  
  
x = np.linspace(0, 10, 100)  
  
//Write your code here
```

- plt.plot(np.cos(x))
plt.show()
- plt.scatter(x, np.sin(x))
plt.show()
- plt.scatter(np.cos(x), x)
plt.show()
- None of the above

1. What is the critical challenge that Latanya Sweeney's research on k-anonymity addresses?

-
- a) make datasets larger for analysis
 - b) eliminate data from datasets
 - c) **anonymize data without losing utility**
 - d) ignore research utility for better privacy

Suppose you're working on a project that involves analyzing sales data for a chain of stores. You are provided with a list of daily sales figures for a particular product over some time. The task is to calculate various statistics such as the total sales, average sales, and maximum sales for that product. Answer the following questions [2-4]:

2) Which data structure can significantly enhance the efficiency and simplicity of your calculations?

-
- a) Python numpy list
 - b) **Python numpy array**
 - c) Python variables
 - d) None of these

3. Fill in the correct function calculate_total_sales(sales_list) which calculates the total sales from a list of daily sales figures using NumPy in the below code:

```
import numpy as np
def calculate_total_sales(sales_list):
    <<<write your code here>>>
    return total_sales
daily_sales = [1200, 1500, 1300, 1400, 1800, 1600]
total_sales = calculate_total_sales(daily_sales)
print("Total sales:", total_sales)
```

a) sales_array = np.array(sales_list) total_sales = np.sum(sales_array)	b) sales_array = np.list(sales_list) total_sales = np.sum(sales_array)
c) sales_array = np.array(sales_list) total_sales = np.sum(np.concatenate(sales_array,sales_list))	d) None of the above

4. Fill in the correct function calculate_maximum_sales(sales_list) which calculates the maximum sales from a list of daily sales figures using NumPy in the below code:

```
import numpy as np
def calculate_maximum_sales(sales_list):
    <<<write your code here>>>
    return max_sales
daily_sales = [1200, 1500, 1300, 1400, 1800, 1600]
maximum_sales = calculate_maximum_sales(daily_sales)
print("Maximum sales:", maximum_sales)
```

a) sales_array = np.array(sales_list) max_sales = np.max(sales_array)	b) sales_array = np.list(sales_list) max_sales = np.max(sales_array)
c) sales_array = np.array(sales_list) max_sales = np.max(np.concatenate(sales_array,sales_list))	d) None of the above

Imagine you have a NumPy array representing your monthly expenses for the past year. Each array element corresponds to the total expenses (in dollars) for a specific month. The task is to analyze and extract specific information.

```
import numpy as np
expenses_array = np.array([1200, 1500, 1300, 1400, 1800, 1600, 1400, 1250,
1350, 1500, 1700, 1900])
```

Answer the following questions from [5-7]:

5. How would you extract expenses for the second half of the year (July to December)?

- a) second_half_expenses = expenses_array[6:]
- b) second_half_expenses = expenses_array[5:]
- c) second_half_expenses = expenses_array[12]
- d) second_half_expenses = expenses_array[7:]

6. Identify the months where expenses exceeded \$1600.

- a) `high_expense_months = np.where(expenses_array < 1600)[0]`
- b) `high_expense_months = np.where(expenses_array > 1600)[0]`
- c) `high_expense_months = np.where(expenses_array > 1600))`
- d) `high_expense_months = np.where(expenses_array > 1600)[0] + 1`

7. Calculate the average expenses for the first quarter of year.

- a) `first_quarter_average = np.mean(expenses_array[:4])`
- b) `first_quarter_average = np.mean(expenses_array[:2])`
- c) `first_quarter_average = np.mean(expenses_array[:3])`
- d) `first_quarter_average = np.mean(expenses_array[:6])`

a.

	Category	Total Revenue	Average Price
0	clothing	105	35.00
1	electronics	1630	543.33
2	home	270	90.00

b.

	Category	Total Revenue	Average Price
0	clothing	1630	35.00
1	electronics	105	543.33
2	home	270	90.00

c.

	Category	Total Revenue	Average Price
0	clothing	105	35.00
1	electronics	270	90.00
2	home	1630	543.33

d) None of the above

8. Choose the correct option for “result”
import pandas as pd
data = {
'Product': ['A', 'B', 'A', 'B', 'C', 'A', 'C', 'B',
'C', 'A'],
'Category': ['Electronics', 'Clothing',
'Electronics', 'Clothing', 'Home',
'Electronics', 'Home', 'Clothing', 'Home',
'Electronics'],
'Price': [500, 40, 600, 35, 100, 550, 80, 30,
90, 480]
}
sales_df = pd.DataFrame(data)
grouped = sales_df.groupby('Category')
result = grouped.agg({'
'Price': ['sum', 'mean']}
).reset_index()
Rename the columns
result.columns = ['Category', 'Total
Revenue', 'Average Price']
print(result)

9. What is the primary goal of k-anonymity in data privacy?

- a) Ensure that every attribute in the dataset is unique.
- b) Prevent unauthorized access to the dataset.
- c) Minimize the amount of data collected.
- d) Making it difficult to link specific individuals to their records.

10. What will be the output of the below code:

```
import numpy as np
import matplotlib.pyplot as plt
time = np.linspace(0, 2 * np.pi, 1000)
amplitude = 1.0
frequency = 1.0
sine_wave = amplitude * np.sin(frequency * time)
plt.figure(figsize=(8, 6))
plt.plot(time, sine_wave, label='Sine Wave', color='blue')
plt.title('Sine Wave')
plt.xlabel('Time')
plt.ylabel('Amplitude')
plt.legend()
plt.grid(True)
plt.show()
```

- a) A plot consisting of sine wave
- b) A plot consisting of cosine wave
- c) Logical error in code
- d) None of the above

Recently, there was a privacy policy change by Whatsapp; how did Whatsapp go about the situation?

- o They notified the users with the revised policy and gave them a timeline to accept the policy. ✓
- o They only tell about the policy change to newly registered users. ✗
- o They hide this information from all users. ✗
- o They notify the users with the revised policy and give them a timeline, where either they can accept or reject it and continue using the product.

Which of the following statement(s) is/are true about k-anonymity?

- o The information for each individual in the released anonymized dataset cannot be distinguished from at-least $k-1$ individuals whose information also appears in the anonymized dataset.
- o Any record in a k -anonymized data set has a maximum probability $1/k$ of being re-identified
- o K -anonymity provides privacy even if the sensitive attributes in the dataset lack diversity

Assume all the required libraries (if any) are already imported? What would be the output of the following code (python2.7+)?

```
online= lambda a : a + 15
print(online(10)) = ①
socialmedia = lambda b, c : b * c
print(socialmedia(12, 4)) = ②
```

lambda \Rightarrow online (a) { return $a + 15$ }

① online (10) \Rightarrow a = 10
↳ returns $10 + 15 = 25$

② social media (12, 4)
↳ returns $12 \times 4 = 48$

Given below the following, Identify the correct option(s) regarding python list and python numpy array.

- 1) [2, 'dog', 3, 4.0, 5] = diff. data type → list. ↗ np array, list
- 2) [0., 1.23452321, 1.98675643, 3.45675423, 7.12354378] = same data type
- o 1) can be a python list whereas 2) can be a python numpy array = only same data type
 - o 1) can be a python numpy array whereas 2) can be a python list
 - o 1) can be a python list whereas 2) can be a python list → same or diff. "
 - o 1) can be a python numpy array whereas 2) can be a python numpy array

What is the output of the following code?

[[[3 4 5]

[9 14 19]]

import numpy as np

[[69 78 87]

import numpy.matlib

[99 112 125]]]

media = np.arange(12).reshape(2,2,3)

social = np.arange(8).reshape(2,2,2)

print(np.matmul(social, media))

What would be the 'output' of the below code

```
import numpy as np  
privacy= np.array([ 4, 5, 6 ] ) ✓  
security=np.array([ 1, 2, 3 ] ) ✓  
output = np.hstack((security, privacy))
```

- o [4 5 6 1 2 3] [1,2,3 , 4,5,6]
- ✓ [1 2 3 4 5 6]
- o [1 3 2 4 6 5]
- o [1 4 5 6 2 3]

- 2) How can you start the jupyter notebook server without opening a browser?

- jupyter notebook --no-browser
- jupyter notebook --browser
- jupyter notebook --not-browser
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

jupyter notebook --no-browser

5) What is the problem statement of LATANYA SWEENEY's research work, as discussed in the lecture?

1 point

- Given person-specific field-structured data, produce a release of the data with scientific guarantees that the individuals who are the subjects of the data cannot be reidentified while the data remain practically useful.
- Identifying important attributes for the re-identification of a person, given any dataset, only ZIP, Birth Date, and Sex are the important attributes.
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Given person-specific field-structured data, produce a release of the data with scientific guarantees that the individuals who are the subjects of the data cannot be reidentified while the data remain practically useful.

6) Which of the following statement(s) is/are true about k-anonymity?

1 point

- The information for each individual in the released anonymized dataset cannot be distinguished from at-least $k-1$ individuals whose information also appears in the anonymized dataset.
- Any record in a k-anonymized data set has a maximum probability $1/k$ of being re-identified
- K-anonymity provides privacy even if the sensitive attributes in the dataset lack diversity

Yes, the answer is correct.

Score: 1

Accepted Answers:

The information for each individual in the released anonymized dataset cannot be distinguished from at-least $k-1$ individuals whose information also appears in the anonymized dataset.

Any record in a k-anonymized data set has a maximum probability $1/k$ of being re-identified

8) What are the correct guidelines to raise awareness about oversharing on social media?

- Develop platforms like "Pleaserobme.com" to create awareness
- Change the privacy of your posts and profiles periodically
- Do not make birthdate, address, children's names related information public.
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the above

16) Photo/video taken by CCTV cameras is of more concern than someone's smartphone according to the participants, in the context of Mobile phones?

- True
- False

No, the answer is incorrect.

Score: 0

Accepted Answers:

False

1) A guy tweets - "Chillin' at Goa :D" along with a selfie from his public twitter account. He puts his geolocation on the tweet as well. How would you classify the person as per Westin's category?

- Pragmatists
- Unconcerned
- Fundamentalists
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Unconcerned

2) Which of the following constitute acts of self disclosure of information on social networks?

- Uploading selfies
- Posting travel details
- Sharing work and home location
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

3) Which of the following is an identified source of information?

- Flickr
- CCTV
- Government website
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Government website

12) Tapping of phone calls leads to what kind of privacy breach?

- Information
- Communication
- Territorial
- Bodily

No, the answer is incorrect.

Score: 0

Accepted Answers:

Communication

13) Which of following is/are raising concern for privacy issues?

- Increasing public self-disclosures through online social networks
- Improving accuracy in Face recognition
- Re-identification techniques getting better
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

14) What information did Latanya Sweeney use to map medical list with voter list?

- Zip
- Birth date
- Sex
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

15) Pittpatt has been acquired by:

- Yahoo
- Google
- Microsoft
- Amazon

No, the answer is incorrect.

Score: 0

Accepted Answers:

Google

16) Mechanical turk is a kind of:

- Crowdsourcing platform
- Machine learning platform
- Data visualization platform
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Crowdsourcing platform

17) Which of the following is/are example of re identification?

- Mapping tinder user with the facebook profile
- Mapping linkedin user with the shaadi.com profile
- Finding twitter profile from aadhar details
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

1) What are the outcomes of publicizing police interactions on X (formerly Twitter) ? **1 point**

- Appreciation of the Police by the Public
- Increased Trust in the Police
- Quick Resolutions of Public Grievances
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Appreciation of the Police by the Public

Increased Trust in the Police

Quick Resolutions of Public Grievances

2) As discussed in the lectures, which of the following campaigns on social media to **1 point** create trust in the police backfired ?

- #myNYPD
- #myLAPD
- #myDelhiPolice
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

#myNYPD

3) As of January 1, 2024 : Which of the following verification check marks are applied **1 point** to Law Enforcement Accounts on X (formerly Twitter) ?

-  Blue Check Mark
-  Grey Check Mark
-  Gold Check Mark
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:



Grey Check Mark

4) Which of the following attributes is the most effective in differentiating a real Police **1 point** account from a fake Police account on X (formerly Twitter) as discussed in the lecture ?

- User Handle
- Profile Picture
- Verification Status of the Account
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Verification Status of the Account

5) Analyzing data without any preconceptions is called _____

1 point

- Inductive Thematic Analysis
- Lexical Analysis
- Natural Language Processing
- None of the above

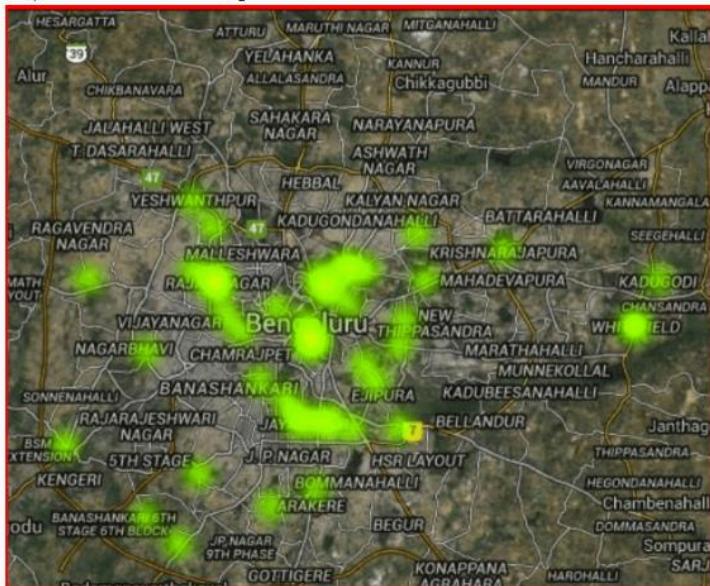
Yes, the answer is correct.

Score: 1

Accepted Answers:

Inductive Thematic Analysis

6) Which of the following is/are true about the visualization shown below ?



- The visualization is called a Heatmap
- The data used to plot the above visualization is called Temporal Data
- This type of visualization falls under the category of Geospatial Analysis
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

The visualization is called a Heatmap

This type of visualization falls under the category of Geospatial Analysis

7)

Saturday evening.

"Time – between 5.30 pm and 6pm. Location: The circle between Freedom Park and the route that goes into Cubbon Park, towards Century Club. Not a single police posted here. I was waiting for an auto at the circle and these two guys rode by asking if they could drop me. . . . Please ensure there are police put here for safety . . ."

What type of actionable information is highlighted in (bold) in the above example?

- Temporal data
- Spatial data
- Linguistic data
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Temporal data

8) Consider the following tweet : "There is a huge pothole on the main road, Kindly **1 point** repair it Address:- Opposite Kachiguda police station, Hyderabad". This is an example of _____

- Actionable Information
- Unactionable Information

Yes, the answer is correct.

Score: 1

Accepted Answers:

Actionable Information

9) As discussed in the lectures, communication of the police with the citizens on **1 point** Social Media is _____

- Formal
- Informal

Yes, the answer is correct.

Score: 1

Accepted Answers:

Formal

10) What is meant by an n-gram ?

1 point

- A sequence of n alternative characters in a text or speech corpus
- A method for optimizing databases using caching
- A continuous sequence of n items extracted from a text or speech corpus
- A method for optimizing big data queries using indexing

Yes, the answer is correct.

Score: 1

Accepted Answers:

A continuous sequence of n items extracted from a text or speech corpus

11) All students taking the PSOSM course are divided into “groups” of students from each state of India. This can be best achieved using _____. **1 point**

- n-gram
- k-means clustering
- LIWC
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

k-means clustering

12) As discussed in the lecture, which of the following measures of behaviors can be studied while analyzing communications between the police and citizens on social media. **1 point**

- Topic Characteristics
- Engagement Measures
- Emotional Measures
- Social and Cognitive Measures

Yes, the answer is correct.

Score: 1

Accepted Answers:

Topic Characteristics

Engagement Measures

Emotional Measures

Social and Cognitive Measures

13) LIWC can categorize an input corpus into the following categories: **1 point**

- Positive Emotions
- Negative Emotions
- Anxiety
- Sadness

Yes, the answer is correct.

Score: 1

Accepted Answers:

Positive Emotions

Negative Emotions

Anxiety

Sadness

14) Select the correct option(s) below while analyzing a post on social media:

1 point

- Valence refers to the negative or positive sentiment of the post
- Valence refers to the intensity of the sentiment expressed in the post
- Arousal refers to the negative or positive sentiment of the post
- Arousal refers to the intensity of the sentiment expressed in the post

Yes, the answer is correct.

Score: 1

Accepted Answers:

Valence refers to the negative or positive sentiment of the post

Arousal refers to the intensity of the sentiment expressed in the post

15) Given below are two types of communication as discussed in the lecture.

My Vehicle KA-02-HW-3183 white color Honda Dio was stolen from Kadamba Hotel (Near Modi Hospital), RajajiNagar on Friday (25th July) evening between 6:30-7:45PM. Please help in tracing my vehicle.

A

Dear BCP, though I stay at JP Nagar, but being part of KSFC Layout RWA (Banaswadi Police station) , I got to know that there are frequent problem at KSFC Layout near BBMP Hall . . . 8-10 youths were creating a public nuisance > (shouting, boozing, manhandling, etc.) yesterday i.e. on 27/7/14, Sunday night 9 pm. We need your support to help avoid a molestation scenario at KSFC Layout. Thank you

B

Select the correct labels for A and B :

- A = Direct Communication ; B = Indirect Communication
- A = Indirect Communication ; B = Indirect Communication
- A = Direct Communication ; B = Direct Communication
- A = Indirect Communication ; B = Direct Communication

Yes, the answer is correct.

Score: 1

Accepted Answers:

A = Direct Communication ; B = Indirect Communication

1. How do Online Social Networks improve public Communication between police and residents?

- a) provide actionable information and collective action
- b) provide mutual accountability
- c) understand fear and victimization effects
- d) help to contact concerned authorities

2. Which of the following are examples of actionable information?

- a) OSN post regarding traffic issues on the road
- b) OSN post regarding open potholes on the road
- c) OSN post sharing about Azaadi ka Amrit Mahotsav celebrations in the locality
- d) None of the above

4. Given below are two types of communication as discussed in the lecture.

- a) Formal and Formal
- b) Informal and Formal
- c) Informal and Informal
- d) Formal and Informal

"Dear Sir, Request to take action on Railway Station parking contractors they are not issuing parking slips . . . today @Yashwantpur Railway Station Tumkur Roadside Entrance Parking."

Kudos to the Banaswadi Traffic Police Team. My Salute and sincere thanks to the Banaswadi Traffic Inspector XXX. . . . An ANGEL in disguise

5. How would you identify emotional and psychological states based on the written text?

- a) LIWC
- b) K-means clustering
- c) N-gram analysis
- d) All the above

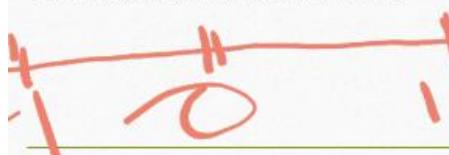
6. Consider a sentence "ISRO launched Chandrayaan and conducted experiments on thermophysical properties and environment". What is the thematic content and subject matter being discussed?

- a) The linguistic diversity of the sentence.
- b) The emotional tone is conveyed in the sentence.
- c) The grammatical structure of the sentence.
- d) The activities and research conducted by ISRO.

7. What does an “n” represent in the n-gram analysis of a sentence?

- a) The number of emotions in the analysis.
- b) The number of words in each n-gram.
- c) The number of characters in each n-gram.
- d) The number of different languages in the sentence.

8. A sentiment analysis study analyses a text passage for valence and arousal values. On a scale of -1 to 1, the valence and arousal score is 0.75 and 0.85. What do these values suggest about the emotional characteristics of the text?



Valence
→ -0
Arousal Score

- a) Negative sentiment, low emotional intensity
- b) Positive sentiment, high emotional intensity
- c) Negative sentiment, high emotional intensity
- d) Positive sentiment, low emotional intensity

9. A social media analyst is studying the engagement characteristics (likes, comments, and shares) of three different posts on a platform. The engagement counts for each post over a week are as follows:

Post 1: 120 likes, 30 comments, 15 shares

Post 2: 80 likes, 45 comments, 20 shares

Post 3: 100 likes, 25 comments, 10 shares

Calculate the mean engagement (average likes, comments, and shares) for each post and then find the standard deviation of engagement for the entire set of posts.

Based on the standard deviation, which post shows the highest variability in engagement characteristics?

- a) Post 1
- b) Post 2
- c) Post 3
- d) All posts show similar variability

Here's how to analyze the engagement data and answer the question about the post with the highest variability:

1. Calculate mean engagement for each post:

Post 1: $(120 \text{ likes} + 30 \text{ comments} + 15 \text{ shares}) / 3 = 55$ likes/comments/shares per day

Post 2: $(80 \text{ likes} + 45 \text{ comments} + 20 \text{ shares}) / 3 = 48.33$ likes/comments/shares per day

Post 3: $(100 \text{ likes} + 25 \text{ comments} + 10 \text{ shares}) / 3 = 45$ likes/comments/shares per day

2. Calculate the standard deviation of engagement for all posts:

Combine the engagement values for all posts into a single dataset:

120, 30, 15, 80, 45, 20, 100, 25, 10]

Then, calculate the standard deviation of this dataset:

Standard deviation = 38.12 (approximately)



3. Analyze variability based on standard deviation:

Now, compare the mean engagement of each post to the standard deviation:

Post 1: $55 - 38.12 = 16.88$

Post 2: $48.33 - 38.12 = 10.21$

Post 3: $45 - 38.12 = 6.88$

10. What kind of post and sentiment is depicted in the below example? "Absolutely disgusted by the state of our local park! It's been months since we reported the broken playground equipment, and nothing has been done. Our kids deserve better!"

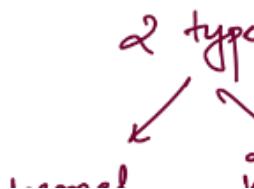
- a) Citizen-initiated and positive
- b) Public authority-initiated and negative
- c) Citizen-initiated and negative
- d) Public authority-initiated and positive

Choose the correct option for a research question: "How do citizens and police engage in social media discussion threads?"

- Topical characteristics
- Engagement characteristics
- Emotional Exchanges
- Cognitive and Social Orientation

As discussed in the lecture, what is the communication style of posts done by citizens and police on social media?

- **Informal, formal by citizen and formal by police**
- Informal, formal by citizen and Informal, formal by police
- formal by citizen and formal by police
- Informal by citizen and informal by police



Which features (from options) can be used to measure behaviour on social media (shown in the figure below)?

Engagement features

- Linguistic features
- Social features
- Cognitive features

No. of police and citizen who comment in posts
Distinct citizens who comment in posts
Average no. of likes and comments

2) Select the pair(s) of social platforms that are adopted by different police departments to engage with people?

- Whisper, Yak-Yak
- Twitter, Youtube
- Yak-Yak, Tik-tok
- Facebook, Twitter

Yes, the answer is correct.
Score: 1

Accepted Answers:
Twitter, Youtube
Facebook, Twitter

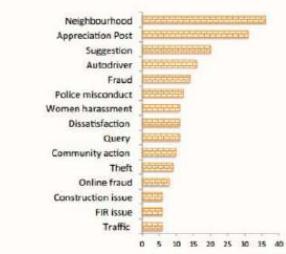
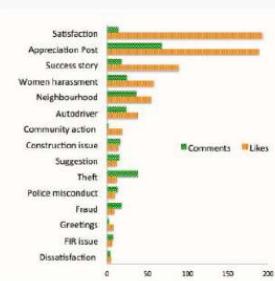
5) Which of the following pieces of information about a user is/are returned by the Twitter API?

- The user's username
- The user's contact to sign up with Twitter
- The number of followers of user
- The Geo-location of user

Yes, the answer is correct.
Score: 1

Accepted Answers:
The user's username
The number of followers of user
The Geo-location of user

6) Given below the graphs: Here X-axis shows the total number of posts of a particular type. Y-axis shows the post type.



Which is the most important issue that is reflected in the graph?

- Theft
- Traffic
- Neighborhood
- Dissatisfaction

Yes, the answer is correct.
Score: 1

Assignment 5

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

1) Which is the most popular Police Department in India?

- Bangalore
- Delhi
- Chennai
- Hyderabad

No, the answer is incorrect.

Score: 0

Accepted Answers:

Bangalore

2) Which is the official and verified account of the Bangalore Police?

- @bangalorepolice
- @bangalore_police
- @blrctypolice
- @blrcty_police

No, the answer is incorrect.

Score: 0

Accepted Answers:

@blrctypolice

3) Which of the following can be seen as use for social media by the police in India?

- Inform about certain decisions (ex. Diversions in certain routes)
- Take complaints
- Take feedback
- Promoting events

No, the answer is incorrect.

Score: 0

Accepted Answers:

Inform about certain decisions (ex. Diversions in certain routes)

Take complaints

Take feedback

4) Which of these platforms have been adopted by the police departments to engage with the public?

- Whisper
- Youtube
- Facebook
- Twitter

No, the answer is incorrect.

Score: 0

Accepted Answers:

Youtube

Facebook

Twitter

7) According to the graph given above, complete the following statement: "Autodriver posts...."

- Are more frequent than suggestion posts, and have more engagement than suggestion posts
- Are less frequent than suggestion posts, and have less engagement than suggestion posts
- Are more frequent than suggestion posts, and have less engagement than suggestion posts
- Are less frequent than suggestion posts, and have more engagement than suggestion posts

No, the answer is incorrect.

Score: 0

Accepted Answers:

Are less frequent than suggestion posts, and have more engagement than suggestion posts

9) What are the various types of posts that come from police accounts?

- Thank you
- Acknowledgment
- Follow Up
- Reply

No, the answer is incorrect.

Score: 0

Accepted Answers:

Acknowledgment

Follow Up

Reply

10) What can be said about accountability of posts to/from police accounts?

- The police should be completely accountable
- The Citizens should be completely accountable
- There should be a good balance of accountability between police and citizens
- The social network must be accountable

No, the answer is incorrect.

Score: 0

Accepted Answers:

There should be a good balance of accountability between police and citizens

11) "@delhipolice Help! My friend is in danger!"- Is this post considered actionable by the Delhi police?

- Yes
- No
- Maybe

No, the answer is incorrect.

Score: 0

Accepted Answers:

No

12) In the context of policing, actionable information is

- Specific, consistent and credible information that can be used to make a decision
- Information that informs the police of an action
- Information that will cause the police to take action against the poster
- Information posted to the Facebook page of a police department

No, the answer is incorrect.

Score: 0

Accepted Answers:

Specific, consistent and credible information that can be used to make a decision

13) Which of the following information about a user is returned by the Twitter API?

- The user's username on Twitter
- The telephone number used by the user to sign up with Twitter
- The number of followers the user has on Twitter
- The location used by the user to sign up with Twitter

No, the answer is incorrect.

Score: 0

Accepted Answers:

The user's username on Twitter

The number of followers the user has on Twitter

The location used by the user to sign up with Twitter

17) Police posts are mostly on which of the following:

- Advisories**
- Celebrity Promotions**
- Making police data public**
- Status of different cases investigated**

No, the answer is incorrect.

Score: 0

Accepted Answers:

Advisories

Status of different cases investigated

18) Citizens initiated threads show which of the following characteristics:

- Higher negative sentiment**
- Low arousal**
- High arousal**
- Lower negative sentiment**

No, the answer is incorrect.

Score: 0

Accepted Answers:

Higher negative sentiment

High arousal

Consider the following table:

Total Posts	w/ ≥ 1 Comment	P&C	C
 85,408	46,845	5,519 $P_{P\&C}$	41,326 P_C
 47,474	24,984	17,196 $C_{P\&C}$	7,788 C_C

$P_{P\&C}$ -> Police initiated discussion threads where both police (P) and citizens (C) left comments and were part of discussion

P_C -> Police initiated discussion threads where only citizens left comments

$C_{P\&C}$ -> Citizen initiated discussion threads where police and citizens both left comments

C_P -> Citizen initiated discussion threads where only citizens left comments

19) What percentage of total posts made by Citizens and Police, in total received interactions from both of them?

1 point

- 8.04
- 17.09
- 54.33
- 43.02

No, the answer is incorrect.

Score: 0

Accepted Answers:

17.09

20) "Police-made posts are more popular than citizen-made posts?" Which of the following arguments support the above statement?

- Information that the police gives is more useful
- Police is more important than the citizens
- Police accounts have a far more varied reach than individual citizen accounts
- Twitter gives more priority to all Police accounts

No, the answer is incorrect.

Score: 0

Accepted Answers:

Information that the police gives is more useful

Police accounts have a far more varied reach than individual citizen accounts

21) What percentage of police initiated posts, have at least one comment?

- 22%
- 55%
- 95%
- 15%

No, the answer is incorrect.

Score: 0

Accepted Answers:

55%

1) You receive an email from "Facebook Technical Support" from an unverified email, **1 point** not related to Facebook.com, asking you to reset your password via a suspicious form. This is called _____.

- Phishing
- Clickbaiting
- Compromised Account
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Phishing

2)



1 point

While browsing a social media website, you come across the post shown above. According to the lecture, the post falls under :

- Work From Home Scam
- Hashtag Hijacking
- Compromised Account
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Work From Home Scam

3) A group of hackers decide to target the Managing Director of a Company with the **1 point** intention of stealing his/her credit card details by sending suspicious emails. As discussed in the lecture, this is called _____.

- Clickbaiting
- Whaling
- Account Compromise
- None of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Whaling

4)

 **HDFC Bank Cares** 
@HDFCBank_Cares

Follow

1 point

Have a question? We're here to help you between 7AM-11PM. Please don't share any personal or banking information publicly, instead you can DM us those details.



HDFC Bank Cares

@hdfcbank_rb

Follow



HDFC Bank Cares

@hdfcbank_rbl

Follow

Customer service Twitter handle support cares



HDFC Bank Cares

@HDFCBank_26

Follow

Customer Service Twitter Handle Support Cares

The above image is an example of _____.

- Hashtag Hijacking
- Compromised Account
- Fake Customer Service Accounts
- None of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Fake Customer Service Accounts



5) As discussed in the lecture, which of the following helps in improving the Pagerank **1 point** of a website (say, abc.com) on search engines ?

- Reputed websites creating reciprocal links with abc.com
- abc.com creating unidirectional links to many websites
- Number of links on abc.com's homepage
- None of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Reputed websites creating reciprocal links with abc.com

6) Spamming the index of search engines is called

1 point

- Spamdexing
- Spamexing
- Spamming
- None of the above

Partially Correct.

Score: 0.5

Accepted Answers:

Spamdexing

Spamexing

7) Abhishek and Suresh are college batchmates. Abhishek sees Suresh's profile on **1 point** Instagram and follows him. Suresh after seeing this notification, feels compelled to follow back Abhishek. This phenomenon is called _____.

- Spamming
- Reciprocity
- Link farming
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Reciprocity

8) Consider three accounts on Twitter:

1 point

Anushka with 500 followers and 250 following

Ramesh with 1,500 followers and 1,000 following

Himanshu with 500 followers and 200 following

Consider, all of them tweeted about #BLRrains expressing their happiness experiencing rain.

According to the lecture, consider the indegree of these accounts and select which of the following accounts are likely to rank higher while searching for #BLRrains on Twitter.

- Anushka
- Ramesh
- Himanshu

Yes, the answer is correct.

Score: 1

Accepted Answers:

Ramesh



9) Which of the following is/are the most widely used formats to represent node-edge **1 point** graphs on computers ?

- Adjacency Matrix
- CSV
- GraphML
- None of the above

Partially Correct.

Score: 0.67

Accepted Answers:

Adjacency Matrix

CSV

GraphML

10)

$$\begin{matrix} & \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{1} & \left[\begin{matrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{matrix} \right] \\ \mathbf{2} & \\ \mathbf{3} & \end{matrix}$$

1 point

Figure 1

The above adjacency matrix (Figure 1) is an example of _____

- Directed Graph
- Undirected Graph

Yes, the answer is correct.

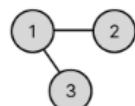
Score: 1

Accepted Answers:

Directed Graph

11)

Undirected Graph



1 point

A	1	2	3
1	0	1	0
2	0	0	1
3	0	1	0

B	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

C	1	2	3
1	0	1	1
2	0	0	1
3	0	1	0

D	1	2	3
1	0	1	1
2	1	0	0
3	1	0	0

Figure 2

Consider the above image (Figure 2), and select the correct adjacency matrix representation of the above shown undirected graph.

- A
- B
- C
- D

Yes, the answer is correct.

Score: 1

Accepted Answers:

D

12) An NxN adjacency matrix implies there are _____ nodes in the graph.

1 point

- N/2
- N-1
- 2N-1
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

None of the above

13) Observe the graph below (Figure 3).

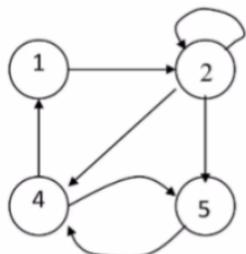


Figure 3

The indegree of Node 2 is _____.

- 1
- 2
- 3
- 4

Yes, the answer is correct.

Score: 1

Accepted Answers:

2

14)

1 point

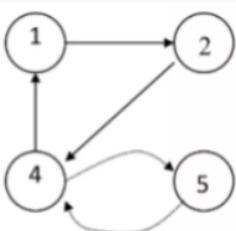


Figure 4

Consider each node to be an account on Instagram in the graph above (Figure 4). Which of the above nodes is/are the most influential ?

- Node 1
- Node 2
- Node 4
- Node 5

Yes, the answer is correct.

Score: 1

Accepted Answers:

Node 4

15) Which of the following is/are graph centrality measures?

1 point

- Node degree
- Betweenness centrality
- Closeness centrality
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Node degree

Betweenness centrality

Closeness centrality

1. Consider the following scenario on a social network

- A is friends with B and C.
- B is friends with A and D.
- C is friends with A and D.
- D is friends with B and C.

Calculate the betweenness centrality for A, B, C, and D. Identify the individual having the highest betweenness centrality.

-
- a) A
 - b) B
 - c) C
 - d) All of them have the same betweenness centrality

2. What does indegree centrality measure in the social media and network analysis context?

-
- a) The level of activity a user has on the platform.
 - b) The number of outgoing connections a user has with other users.
 - c) A user's popularity based on their followers or incoming connections.
 - d) The diversity of content shared by a user.

3. In a social network, which individual is most likely to be considered the "go-to person" or the one to bridge communication gaps between different groups?

-
- a) Someone with the most followers
 - b) Someone who follows the most people
 - c) Someone who is mentioned most often in conversations between different groups
 - d) Someone who comments on many different posts across the network

4. What is link farming?

- a) Agricultural practice that involves cultivating crops in a vertical arrangement.
- b) Rapidly growing links on a website through unethical or manipulative means.
- c) A strategy for developing social connections by attending networking events.
- d) All the above

5. Which of the following practices is considered unethical and attempts to manipulate search engine results through deceptive means?

- a) White hat SEO
- b) **Black hat SEO**
- c) Organic search
- d) Paid search advertising

6. Calculate the Engagement Rate based on the following engagement metrics for a user on online social media.

Number of retweets: 50

Number of likes: 150

Number of comments: 100

Number of followers: 500

Number of mentions: 30

- a) 60
- b) 50
- c) 20
- d) 90

The Engagement Rate is typically calculated as a percentage and represents the level of interaction relative to the total number of followers. The formula for calculating the Engagement Rate is:

$$\text{Engagement Rate} = \left(\frac{\text{Likes} + \text{Comments} + \text{Shares}}{\text{Followers}} \right) \times 100$$

Given the engagement metrics you provided:

Likes = 150, Comments = 100, Shares (Retweets) = 50, Followers = 500

$$\text{Engagement Rate} = \left(\frac{150+100+50}{500} \right) \times 100$$

$$\text{Engagement Rate} = \left(\frac{300}{500} \right) \times 100$$

$$\text{Engagement Rate} = 0.6 \times 100 = 60\%$$

7. Which of the following is considered as “social etiquette” on online social media such as Twitter?

- a) Following back
- b) Comment on the post of an influential person
- c) Giving likes to the post of an influential person
- d) Saying “Thank you” to any appreciation by an influential person

8. A psychological principle where individuals feel compelled to return favors and interactions they receive is called

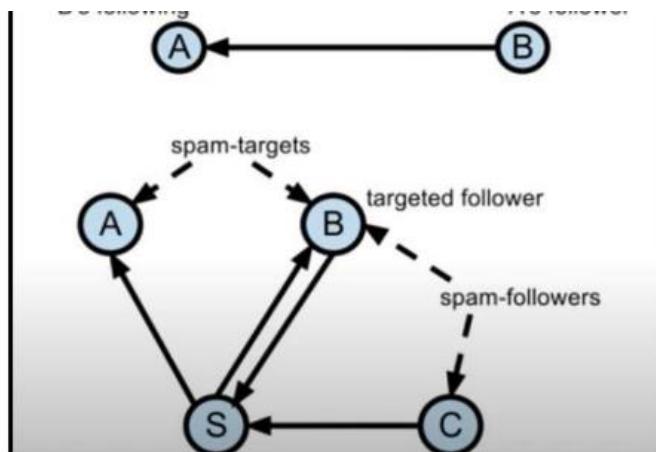
- a) Spaming
- b) Reciprocity
- c) Link farming
- d) All of the above

9. Which of the following is/are graph centrality measures?

- a) Eigen centrality
- b) Node degree
- c) Betweenness centrality
- d) Closeness centrality

10. Which of the following are spam targets?

- a) A
- b) B
- c) C
- d) S



The Facebook team asked you to find the most popular person in this friendship network, which centrality measure is the most appropriate?

- Betweenness
- Degree**
- Closeness
- None of the above

What does network diameter signify?

- The hierarchical size of a network
- The linear size of a network**
- Variation in the node membership
- Number of communities in the network

3) What does betweenness centrality depict about the node?

1 point

- One who is quickly approachable
- Who is close to everyone
- One who is most influential
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

One who is quickly approachable

3) An Adjacency matrix and XML file format are preferable for representing

- Graph data
- Tabular data
- Both a and b
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Graph data

5) [True / False] Spamming can get you suspended on Twitter?

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

6) What is a legit way to identify the popularity of a web page?

- Link farming
- PageRank

Yes, the answer is correct.

Score: 1

Accepted Answers:

7) Select the true statement(s):

- Impersonation account has access to the user id and password of the target person.
- Klout score is influenced by twitter indegree
- Act of tricking someone into handing over her login credentials to exploit personal information
- Impersonation in the form of Fan accounts is a crime.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Klout score is influenced by twitter indegree

Act of tricking someone into handing over her login credentials to exploit personal information

9) Which of the following statements are true for Link Farmers?

- Top Link farmers have high indegree as compared to Spammers
- Top Link farmers have lower indegree as compared to Spammers
- Node indegree of Link Farmers are the same as any Random samples
- Node indegree of Link Farmers are higher than any Random samples

Yes, the answer is correct.

Score: 1

Accepted Answers:

Top Link farmers have high indegree as compared to Spammers

Node indegree of Link Farmers are higher than any Random samples

3) Match the following networks with their main measure of social influence:

1. Facebook	A. Views
2. Amazon	B. Reviews
3. LinkedIn	C. Likes
4. YouTube	D. Endorsements

- 1-C, 2-B, 3-D, 4-A
- 1-C, 2-A, 3-D, 4-B
- 1-C, 2-B, 3-A, 4-D
- 1-B, 2-D, 3-C, 4-A

No, the answer is incorrect.

Score: 0

Accepted Answers:

1-C, 2-B, 3-D, 4-A

6) Which of the following statements about link farming is TRUE?

- Link farming involves hunting for relevant links across the web by companies
- Link farming involves creating new URLs for the same page for deceptive purposes
- Link farming involves spamming the index of a search engine
- Link farming is used only on Twitter

No, the answer is incorrect.

Score: 0

Accepted Answers:

Link farming involves spamming the index of a search engine

7) Which of the following statements about pagerank is FALSE?

- It is used to evaluate the importance and authority of a page/post.
- It depends on the quantity and quality of the links/pages linking to the page/post
- The higher the pagerank, the older the page/post
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

The higher the pagerank, the older the page/post

8) What scale does Klout score use?

- 1 - 100
- 1 - 10
- 0 - 100
- 0 - 10

No, the answer is incorrect.

Score: 0

Accepted Answers:

1 - 100

9) Which of the following is not a form of phishing?

- Whaling
- Spear Phishing
- Clone Phishing
- Sharking

No, the answer is incorrect.

Score: 0

Accepted Answers:

Sharking

10) Which of the following can be the vector of attack in clickbaiting?

- News links that lead to irrelevant articles
- Thumbnails of videos which lead to fake sites
- Advertisements of miraculous products
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

11) What is the difference between Impersonation attacks and Compromised accounts? 1 point

- Impersonation occurs on social media and accounts are compromised on email
- Individual accounts are impersonated, whereas organization accounts are compromised.
- Impersonation can occur to anyone, but accounts of only public figures are compromised.
- Impersonation involves creating a fake account of a person, whereas a compromised account was created by the legitimate owner but hijacked later

No, the answer is incorrect.

Score: 0

Accepted Answers:

Impersonation involves creating a fake account of a person, whereas a compromised account was created by the legitimate owner but hijacked later.

13) On Twitter if a spam account S, follows A and B, and is followed by A and C, which of the following are spam-targets? 1 point

- A only
- A and C
- A and B
- A, B and C

No, the answer is incorrect.

Score: 0

Accepted Answers:

A and B

14) What constitutes as hashtag hijacking or hashjacking? 1 point

- A hashtag that allows people to post about any accounts being hijacked.
- A hashtag that allows hijackers to mark that they have successfully gained control of the account
- A phenomenon where companies promote their products using existing popular hashtags that are not relevant to the product
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

A phenomenon where companies promote their products using existing popular hashtags that are not relevant to the product

1) According to the study discussed in the lecture, which of the following users were 1 point most likely to reciprocate links from spammers ?

- Users with less than 100 followers
- Users with less than 1,000 followers
- Users with more than 10,000 followers
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Users with more than 10,000 followers

2) According to the study discussed in the lecture, which of the following users usually **1 point** engaged themselves in Link Farming ?

- People talking about Business, Internet Marketing, Money etc.
- People promoting their own business or content or trends in a domain.
- Equally distributed among Random Samples of users all across Twitter.
- None of the above

No, the answer is incorrect.

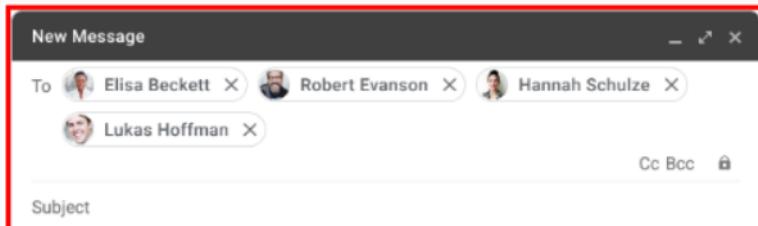
Score: 0

Accepted Answers:

People talking about Business, Internet Marketing, Money etc.

People promoting their own business or content or trends in a domain.

3) The above image from Gmail is an example of the ideas proposed by Facemail **1 point** from MIT.



- True
- False

No, the answer is incorrect.

Score: 0

Accepted Answers:

True

Observe the image below and answer the following questions [4-5] :



4) Which of the following nudge has been shown in the picture above ? **1 point**

- Sentiment Nudge
- Timer Nudge
- Picture Nudge
- None of the above

5) Select the utility of the above type of nudge : **1 point**

- Making sure that you are sharing content with the intended audience/people.
- Giving you the opportunity to cancel your post during a certain timeframe.
- Making sure your posts are not perceived as negative by other users.
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Making sure that you are sharing content with the intended audience/people.

6) As discussed in the lecture, which of the following is/are not a social media nudge? **1 point**

- Picture Nudge
- Timer Nudge
- Voice Nudge
- Video Nudge

No, the answer is incorrect.

Score: 0

Accepted Answers:

Voice Nudge

Video Nudge

7) Which of the following statements is/are correct about nudges ? **1 point**

- Interventions help users make better decisions about posting on social media.
- The primary aim of nudges is to reduce user generated content on social media.
- Nudges help individuals avoid regrettable online disclosures.
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Interventions help users make better decisions about posting on social media.

Nudges help individuals avoid regrettable online disclosures.

8) Hemant wanted to steal data from a rival business by going to their office and transferring the required files by connecting a USB Flash Drive to the target computer. This type of attack is called _____.

- Physical Attack
- Semantic Attack
- Syntactic Attack
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Physical Attack

9) Denial of Service Attacks and Buffer Overflow Attacks fall under : **1 point**

- Physical Attacks
- Semantic Attacks
- Syntactic Attacks
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Syntactic Attacks

10) Shivani, in order to steal login credentials from Anisha, sends an email to her with **1 point** the subject "Check your NPTEL Certification Exam Results Here !" which directs Anisha towards a login form with a website domain not associated with NPTEL. This is a type of _____ attack.

- Syntactic
- Semantic
- Physical
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Semantic

11) Himesh, a college student, gets a malicious SMS from an unidentified source **1 point** saying "Your Bank Account has been suspended : Click here to re-activate your account". This example most closely resembles :

- Whaling
- Vishing
- Smsishing
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Smsishing

12) Consider the following code snippet :

1 point

```
import nltk  
nltk.download("all") #for downloading nltk data packages, ignore if done  
nltk.tokenize.word_tokenize("Welcome to PSOSM.")
```

What is the output from line 3 (word_tokenize function) above ?

- ['W', 'e', 'l', 'c', 'o', 'm', 'e', '', 't', 'o', '', 'P', 'S', 'O', 'S', 'M', '.']
- ['Welcome', 'to', 'PSOSM', '.']
- ['Welcome', 'to', 'PSOSM']
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

['Welcome', 'to', 'PSOSM', '.']

13) Consider the following list of words generated using nltk

1 point

[our, are, yourself, yourselves, apple, orange, have, doesn't, aren't, out, most, where]

Which of the following are not stop words ?

- apple
- orange
- yourself
- most

No, the answer is incorrect.

Score: 0

Accepted Answers:

apple

orange

14) Consider the below code snippet, consider the variable token holds 1000+ words extracted from a text corpus.

1 point

```
p = str.maketrans("", "", string.punctuation)  
tokens = [i.translate(p) for i in tokens]
```

What is the purpose of the translate function as shown above ?

- To map uppercase characters to lowercase characters in the tokens list
- To map all punctuations to a null character in the tokens list
- To map all null characters to punctuations in the tokens list
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

To map all punctuations to a null character in the tokens list

15) Which of the following code snippets can be used to count the number of words in **0 points** a list named tokens with 1000+ elements(words).

L = {}
for i in tokens:
 if i not in L:
 L[i] = 1
 else:
 L[i] += 1
print()
 from collections import Counter
print(Counter(tokens))
 Counter(tokens).count()
 None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

L = {}

```
for i in tokens:  
    if i not in L:  
        L[i] = 1  
    else:  
        L[i] += 1  
print()  
from collections import Counter  
print(Counter(tokens))
```

1. Consider the following Python code:

```
punctuations = "!"|"-|{|}|;|'|<|>|/|?|@|#|$%^&*|~|\""  
my_str = "Hello everyone!!!, Are you enjoying psosm course --- and learning new  
skills."  
no_punct = ""  
for char in my_str:  
    if char not in punctuations:  
        no_punct += char  
print(no_punct)
```

What is the primary purpose of this code?

- a) To count the total number of punctuation marks in the string. X
- b) To convert all letters in the string to uppercase. X
- c) To create a new string with all punctuation marks removed. ✓
- d) To reverse the order of characters in the original string. X

2. Consider the following code:

```
from collections import Counter  
x = [2, 1, 3, 4, 5, 2]  
token = ['social', 'media', 'nptel', 'nltk', 'python', 'tutorial']  
pairs= Counter(dict(zip(token, x))).most_common(2)
```

Which of the following is the correct option?

- a) This code is used to retrieve an ascending sorted list having two tuples.
- b) This code is used to retrieve a descending sorted list having two tuples. ✓
- c) This code retrieves two tuples in a sorted or unsorted manner.
- d) This code returns a logical error.

3. Which of the following are not stopwords in the following list?

stopwords = ["i", "me", "my", "myself", "we", "our", "ours", "ourselves", "you", "your", "yours", "yourself", "yourselves", "he", "him", "his", "himself", "she", "her", "hers", "herself", "it", "its", "itself", "they", "them", "their", "theirs", "themselves", "what", "which", "who", "whom", "this", "that", "these", "those", "am", "is", "are", "was", "were", "be", "been", "being", "have", "has", "had", "having", "do", "does", "did", "doing", "apple", "an", "the", "and", "but", "if", "or", "because", "as", "until", "while", "of", "at", "by", "for", "with", "about", "against", "between", "into", "through", "during", "before", "after", "above", "below", "to", "from", "up", "down", "in", "out", "on", "off", "over", "under", "again", "further", "then", "once", "here", "there", "when", "where", "why", "how", "all", "any", "both", "each", "few", "more", "most", "other", "some", "such", "no", "nor", "not", "only", "own", "same", "so", "than", "too", "very", "s", "t", "can", "will", "just", "don", "should", "now"]

- a) All the above are stopwords. X
- b) Further, again, apple X
- c) While, apple X
- d) apple X

4. Which elements in the below sample mail make it potentially phishing email, as discussed in the lecture?

- a) Urgency in subject
- b) Regrettable information in the body
- c) Potential phishing link along with body
- d) All of the above

From: sri-phd-students-indr-bounces@mailman.sri.cs.cmu.edu on behalf of eBay Inc [mailto:fun804194205199@eBay.com]
To: sri-people@cs.cmu.edu
Cc:
Subject: eBay: urgent security notice [Sun, 05 Feb 2006 18:54:02 -0400]



Dear eBay Member,

We regret to inform you that your eBay account could be suspended if you don't re-update your account information.

To resolve this problem please visit link below and re-enter your account information:

https://signin.ebay.com/ws/eBaySAPI.dll?SignIn&sid=verify&co_partnerid=2&steid=0

If your problems could not be resolved your account will be suspended for a period of 24 hours, after this period your account will be terminated.

For the User Agreement, Section 9, we may immediately issue a warning, temporarily suspend, indefinitely suspend or terminate your membership and refuse to provide our services to you if we believe that your actions may cause financial loss or legal liability for you, our users or us. We may also take these actions if we are unable to verify or authenticate any information you provide to us.

Due to the suspension of this account, please be advised you are prohibited from using eBay in any way. This includes the registering of a new account. Please note that this suspension does not relieve you of your agreed-upon obligation to pay any fees you may owe to eBay.

Regards,
SafeHarbor Department eBay, Inc.
The eBay Team

5. Which of the following is correct statement as discussed in the lecture?

- a) Younger participants are more vulnerable to phishing attacks. ✓
- b) Older participants are more vulnerable to phishing attacks. X
- c) Younger and older participants are equally vulnerable to phishing attacks. X
- d) No one is vulnerable to phishing attacks. X

6. In a scientific experiment, which group refers to manipulating the independent variable?

-
- a) Observational group
 - b) Comparison group
 - c) Control group
 - d) **Experimental group**

7. Which of the following is/are the advantages of having privacy policies?

-
- a) To help users make informed decisions with the information that is presented to them
 - b) To help individuals avoid regrettable online disclosures
 - c) **Both a and b are correct**
 - d) None of the above

9. As discussed in the lecture, which of the following is/are not a social media nudge?

-
- a) Timer Nudge
 - b) Picture Nudge
 - c) Sentiment Nudge
 - d) **Voice Nudge**

10. Which of the following has the property of having high indegree and high outdegree on online social media?

-
- a) **Top Link farmers**
 - b) Random social network users
 - c) Top Spammers
 - d) All of the above

11. What will this code print?

```
def word_count(text):
    words = text.split()
    return len(words)
sentence = "The quick brown fox jumps over the lazy dog."
number_of_words = word_count(sentence)
print(f"The sentence contains {number_of_words} words.")
```

- a) The sentence contains 9 words.
- b) The sentence contains 7 words.
- c) The sentence contains 8 words.
- d) The sentence contains 5 words.

12. What will this code print?

```
fruits = ["apple", "banana", "cherry", "mango", "mango"]
unique_fruits = set(fruits)
print(len(unique_fruits))
```

- a) 4
- b) 5
- c) 2
- d) 3

6
a = (

13. What will this code print?

```
def double_letters(word):
    for i in range(len(word) - 1):
        if word[i] == word[i + 1]:
            return True
    return False
result = double_letters("mississippi")
print(result)
```

- a) True
- b) False
- c) "mississippi"
- d) An error message

How are top link farmers different from spammers and random social network users?

- **High indegree and high outdegree**
- Low indegree and high outdegree
- Low indegree and low outdegree
- High indegree and low outdegree

What topics are dominant in link farmers?

- **Markets, business**

- Life, friends
- Love, person
- Music, world

If a URL is part of a link farm, what is the value of - the number of external URLs mentioned vs the number of mentions received

- **Almost 1**

- Much higher than 1
- Far lesser than 1
- Can not say

Nudges while using any application offers a great user experience:

- True

 **False**

Semantic Attacks takes advantage of the:

- **context as perceived by humans.**
- technology used by software.
- time of day.
- very weak wall of a server room.

Who among younger and older participants were more vulnerable to phishing attack as discussed in the lecture?

- **Younger**
- Older

Select all methods by which we can avoid phishing:

- **By educating the users**
- **By using browser extensions**
- By disclosing all private information on Facebook
- By opening all links present in an email

1) How are top link farmers different from spammers and random social network users?

- High indegree and high outdegree
- Low indegree and high outdegree
- Low indegree and low outdegree
- High indegree and low outdegree

Yes, the answer is correct.

Score: 1

Accepted Answers:

High indegree and high outdegree

4) Select the communication corresponding to the most number of successful phishing attacks:

- Male to Male
- Female to Male
- Male to Female
- Female to Female

Yes, the answer is correct.

Score: 1

Accepted Answers:

Male to Female

1) What is reciprocity?

- The probability of someone following an account
- The probability of an account following someone
- The probability of someone following me back, given that I follow that person
- Both A and B

No, the answer is incorrect.

Score: 0

Accepted Answers:

The probability of someone following me back, given that I follow that person

2) In the context of spamming, which of the following are true?

- Users with low indegree don't reciprocate to links from spammers
- Responsiveness increases with the number of followers
- Users with large indegree tend to follow back more
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

3) According to Pagerank, who is the top link farmer?

- Britney Spears
- Barack Obama
- UK Prime Minister
- JetBlue Airways

No, the answer is incorrect.

Score: 0

Accepted Answers:

Barack Obama

4) Which of the following groups is likely to have the highest indegree on Twitter?

- Spammers
- A randomly selected sample of users
- Top 100k link farmers
- All of the above have the same in degree on average

No, the answer is incorrect.

Score: 0

Accepted Answers:

Top 100k link farmers

5) Which of the following groups is likely to have the highest outdegree on Twitter?

- Spammers
- A randomly selected sample of users
- Top 100k link farmers
- All of the above have the same in degree on average

No, the answer is incorrect.

Score: 0

Accepted Answers:

Top 100k link farmers

6) Approximately, what is the most common indegree/outdegree ratio amongst the top link farmers?

- 1
- 0.5
- 1.5
- 0.75

No, the answer is incorrect.

Score: 0

Accepted Answers:

1

7) What is Facemail?

- A clone of Facebook
- A program that would show a person's face when an email is addressed to him/her.
- An email client for Android
- A program used to detect faces

No, the answer is incorrect.

Score: 0

Accepted Answers:

A program that would show a person's face when an email is addressed to him/her.

9) What does IRB stand for in the context of getting the required permissions before conducting a study?

- Information Review Board
- Institutional Review Board
- Internal Review Board
- Industrial Review Board

No, the answer is incorrect.

Score: 0

Accepted Answers:

Institutional Review Board

10) Which of the following "nudges" were included for the experiment on Facebook?

- Timer nudge
- Sentiment nudge
- Picture nudge
- Sensitivity nudge

No, the answer is incorrect.

Score: 0

Accepted Answers:

Timer nudge

Sentiment nudge

Picture nudge

11) Which of the following is/are true regarding the semantic barrier?

- If the barrier is large, you are more likely to fall for semantic attacks
- If the barrier is large, you are less likely to fall for semantic attacks
- It is the difference between what you think a system is doing, and what a system thinks you're doing.
- It is a firewall to protect yourself from semantic attacks.

No, the answer is incorrect.

Score: 0

Accepted Answers:

If the barrier is large, you are more likely to fall for semantic attacks

It is the difference between what you think a system is doing, and what a system thinks you're doing.

12) Which of the following is/are types of phishing attacks?

- Whaling
- Vishing
- SMSishing
- Spear phishing

No, the answer is incorrect.

Score: 0

Accepted Answers:

Whaling

Vishing

SMSishing

Spear phishing

13) In the phishing experiment talked about in the lecture, which of the following had the highest percentage of successful phishes?

- From male to female
- From female to male
- From male to male
- From female to female

No, the answer is incorrect.

Score: 0

Accepted Answers:

From male to female

1) Which of the following is/are location-based services?

1

- Foursquare
- Ola
- Yelp
- Uber

Yes, the answer is correct.

Score: 1

Accepted Answers:

Foursquare

Ola

Yelp

Uber

2) Mayorships on Foursquare are given to:

- The first person to check in at a location
- The person who leaves the most tips at a location
- The person who checks in most frequently at a location in a given period
- People who write the most reviews on a location

Yes, the answer is correct.

Score: 1

Accepted Answers:

The person who checks in most frequently at a location in a given period

3) Which of the following is publicly available data on Foursquare?

- Mayorships
- Tips
- Dones
- Checkins

No, the answer is incorrect.

Score: 0

Accepted Answers:

Mayorships

Tips

Dones

4) Look at the figure (Figure 1) given below for the Cumulative Distribution of the Number of Mayorships, Tips and Dones per User: 1 point

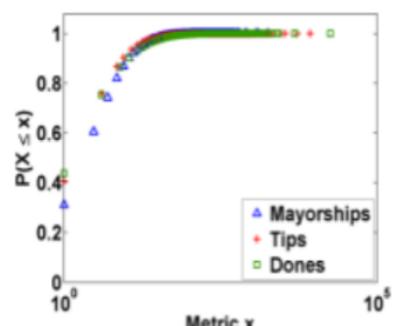


Figure 1

Majority of the users on Foursquare do not contribute to the Mayorships, Tips or Dones. This statement is :

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

5) In the paper as discussed in the lecture, a high correlation between the number of mayorships and the number of tips across cities, with a Spearman's correlation coefficient ρ equal to 0.78 indicates that :

- There is a high monotone increasing relation between the number of mayorships and the number of tips.
- There is a high monotone decreasing relation between the number of mayorships and the number of tips.
- There is no relation regarding monotonicity between the number of mayorships and the number of tips.
- None of the above.

Yes, the answer is correct.

Score: 1

Accepted Answers:

There is a high monotone increasing relation between the number of mayorships and the number of tips.

6) Look at the figure (Figure 2) given below for the Cumulative Distribution of Displacements Between Consecutive Tips/Dones Posted per User:

1 point

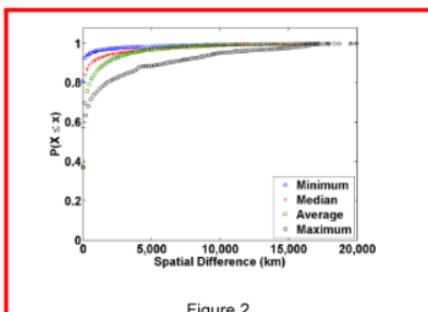


Figure 2

About _____ % of the users have spatial differences between their tips/dones at more than or equal to 6,000km on average.

- 10
- 20
- 30
- 40

No, the answer is incorrect.

Score: 0

Accepted Answers:

10

7) Look at the (Figure 3) for the Cumulative Distribution of Time Interval Between Consecutive Tips/Dones Posted per User :

1 point

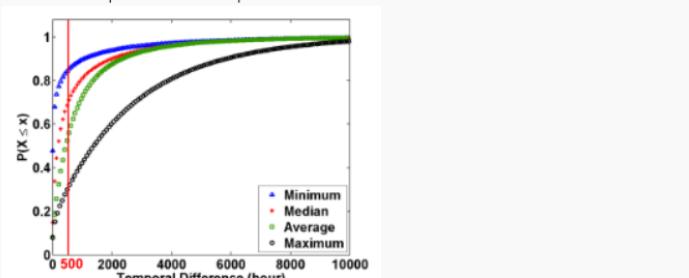


Figure 3

50% of the users post tips/dones in intervals of less than 500 hours on average. This statement is :

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

8) What is the difference between Pearson and Spearman rank correlation? **1 point**

Hint: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

- Pearson correlation assesses linear relationships, while Spearman correlation evaluates monotonic relationships.
- Spearman correlation is always between -1 and 1, while Pearson correlation can have values outside this range.
- Spearman correlation is calculated using a formula involving the sum of squared differences, whereas Pearson correlation is based on the product of z-scores.
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Pearson correlation assesses linear relationships, while Spearman correlation evaluates monotonic relationships.

9) Which plot is shown in the below figure (Figure 4) ?

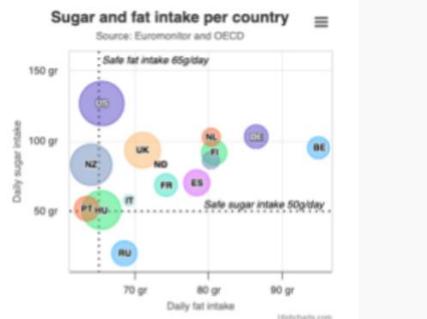


Figure 4

- Scatter Plot
- Bar Chart
- Bubble Plot
- Histogram

Yes, the answer is correct.

Score: 1

Accepted Answers:

Bubble Plot

10) Which of the following charts can be generated using highcharts ? **1 point**

- Bar Chart
- Line Chart
- Scatter Plot
- Bubble Chart

Yes, the answer is correct.

Score: 1

Accepted Answers:

Bar Chart

Line Chart

Scatter Plot

Bubble Chart

11) Which of the following is the correct **chart type** in the options dictionary to print a horizontal bar chart? **1 point**

- bar
- column
- horizontal_bar
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

bar

12) Refer to the options dictionary for a line chart as shown in Figure 5 and complete the code with the correct option for enabling the legend **1 point**

```
10 options = {  
11     "chart": {  
12         "type" : "line"  
13     },  
14     "title": {  
15         "text": "Line Chart"  
16     },  
17     #  
18     # WRITE YOUR CODE HERE  
19     #-----  
20     "xAxis": {  
21         "categories": ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"],  
22         "title": {  
23             "text": "Months"  
24         }  
25     },  
26     "yAxis": {  
27         "title": {  
28             "text": "Number of Followers"  
29         }  
30     },  
31 }
```

Figure 5

- "legend": {"enabled": True},
- "legend": True,
- "legend": {"legend": True},
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

"legend": {"enabled": True},



13) Ravi wanted to plot his number of followers on Instagram for 2023 in a line chart using **0 point** Highcharts. When he ran the code given below, it gave an unexpected output as shown below in (Figure 6)



Figure 6

Study the code below (Figure 7) and determine the line due to which he might be getting the unexpected output (assume all necessary imports have already been made) :

```
1 chart = Highchart()
2
3 options = {
4     "chart": {"type": "line"},
5     "title": {"text": "Line Chart"},
6     "xAxis": {
7         "categories": ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"],
8         "title": {"text": "Months"}
9     },
10    "yAxis": {
11        "title": {"text": "Number of Followers"}
12    }
13 }
14
15 data = [80, 70, 60, 50, 40, 30, 20, 10, 20, 30, 40, 50]
16
17 chart.set_dict_options(options)
18
19 chart.add_data_set(data, "line", "Data 1")
20
21 chart.save_file('./line')
```

Figure 7

- Line 11
- Line 12
- Line 22
- Line 26

No, the answer is incorrect.

Score: 0

Accepted Answers:

Line 26

1. Which of the following is/are location-based services?

- a) Foursquare
- b) Yelp
- c) Gowalla
- d) Uber

3. You have two datasets and want to measure the strength and direction of the relationship between them. One set of data consists of the ranks or ordinal values of the observations, while the other set consists of continuous numerical data. Which correlation coefficient is most appropriate for this scenario?

- a) Spearman's rank correlation coefficient
- b) Pearson's correlation coefficient
- c) Either Spearman or Pearson correlation can be used interchangeably.
- d) None of the above because correlation coefficients are not applicable in this scenario.

4. Which of the following data fields in a social media platform are open text fields, whose validity is not enforced by the system and may carry noise or invalid information?

- a) User's date of birth
- b) Profile picture
- c) **Bio/about me section**
- d) Number of followers

Noise → D
IR

5. Imagine you're developing a social media platform with location-based features. You want to allow users to share interesting places they've visited. Which of the following methods would provide the most accurate and reliable information about these places for other users?

- a) Open text field for users to describe the place, including location details.
- b) Multiple choice selection from a pre-defined list of popular landmarks/attractions.
- c) Automatic location detection based on the user's current GPS coordinates.
- d) **Combination of options a) and b), allowing users to choose from a list and further elaborate with open text.**

6. What is the key difference between a scatter plot and a bar chart?

- a) Scatter plots display categorical data, while bar charts are used for numerical data.
- b) **Scatter plots show the relationship between two numerical variables, while bar charts display the frequency or distribution of categorical data.**
- c) Scatter plots are only used for displaying data with a single variable, while bar charts can visualize data with multiple variables.
- d) All the above

7. Imagine you're a data scientist analyzing tweets related to current events. You want to identify the most influential users in this conversation. Which of the following metrics would be the most accurate indicator of user influence in this specific context?

-
- a) Total number of followers
 - b) **Average number of retweets per tweet**
 - c) Number of hashtags used per tweet
 - d) Sentiment analysis score of their tweets

8. Imagine you're analyzing data on the number of books read by students in a university class. The x-axis shows the number of books read, and the y-axis shows the number of students who read that many books. Which type of plot would be most appropriate to visualize this data?

Both data →
numerical

- a) Scatter plot
- b) **Bar chart**
- c) Bubble plot
- d) Histogram

9. What is the primary characteristic of Location-Based Social Networks (LBSNs)?

-
- a) They allow users to share text-based posts with their friends and followers.
 - b) They provide a platform for online gaming and virtual reality experiences.
 - c) **They enable users to connect with others and share their physical locations and activities.**
 - d) They focus exclusively on professional networking and job searching.

10. Imagine you're analyzing data from a location-based social network (LBSN) focused on food recommendations. Users can share tips about restaurants, cafes, and other food establishments. The platform also tracks user check-ins at these locations. You're interested in understanding how users discover new food places through the platform. Which of the following insights would be most valuable for this purpose?

- a) The average number of tips each user reads before checking into a new restaurant.
- b) **The percentage of users who discover new restaurants through check-ins from friends.**
- c) The correlation between the number of followers a user has and the diversity of cuisines they try.
- d) **The proportion of users who mention "friends' recommendations" in their tips about new restaurants.**

Gephi allows the user to configure the following parameters:

- o **filter nodes and edges** ✓
- o **Add edges manually** ✓
- o **calculate statistics of graph** ✓
- o **Change node size and colour** ✓

"Nidhi Razdan received an offer to teach journalism at Harvard University. After a few months, she realized that it was not genuine and meant to access her personal information".

- o Which of the following statements is/are true?
 - It is a phishing attack.**
 - It is a semantic attack.**

What are the potential reasons users change their usernames?

- o Suit a trending event
- o Gain / lose anonymity
- o Avoid boredom
- o **All of the above**

What is correct, when talking about the username change frequency?

- o It is strongly and positively correlated to the indegree of the users
- It is weakly and positively correlated to the indegree of the users**
- o It is strongly and negatively correlated to the indegree of the users
- o It is weakly and negatively correlated to the indegree of the users

“ Link farming will not be effective in increasing your social capital and influence.”

- True
- False

Why are usernames important for linking users across social media platforms?

- Usernames do not change
- Restrictions on character and length
- User can have many usernames
- Publicly available attribute.

- In the context of text processing, Match the following preprocessing steps with the reasons for carrying that preprocessing? Mark the most likely option.

i. Removal of punctuations	1. Filtering out words like RTs, QT
ii. Lower casing	2. do not add any semantic meaning to the text
iii. Removal of stop words	3. To suppress the frequent words like "a", "the"
iv. Removing words with small lengths	4. To get reliable counts of a word

i - 2, ii - 4, iii - 3, iv - 1

- Which among the following options can be used to create a DataFrame in Pandas?

- A pandas series
- An ndarray
- A python dict
- All of the above

- 6) Which of the following option(s) is/are the correct syntax for reading a .csv file in python?

import pandas as pd
data=pd.read_csv("filename.csv",header = None, encoding='utf-8')

import pandas as pd
data=pd.read_csv(filename.csv)

import pandas as pd
file="filename.csv"
data=pd.read_csv(file,header = None, encoding='utf-8')

import pandas as pd
data=pd.read_csv(filename.csv, header = None, encoding='utf-8')

Yes, the answer is correct.

Score: 1

Accepted Answers:

*import pandas as pd
data=pd.read_csv("filename.csv",header = None, encoding='utf-8')*

*import pandas as pd
file="filename.csv"
data=pd.read_csv(file,header = None, encoding='utf-8')*

9) Which of the following is an example of a 3-D Plot?

- Scatter chart
- Bar chart
- Bubble chart
- Histogram

Yes, the answer is correct.

Score: 1

Accepted Answers:

Bubble chart

1) Which of the following social networks uses location information for related services?

- Facebook
- Foursquare
- Tinder
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

3) According to the lecture, what is the majority opinion about mobile operators using local language being privacy invasive?

- Strongly agree
- Strongly disagree
- Agree
- Neutral

No, the answer is incorrect.

Score: 0

Accepted Answers:

Agree

17) How do you install highcharts on ubuntu?

- sudo pip install python-highcharts
- sudo apt-get install highcharts
- sudo pip install highcharts
- All of the above are valid

No, the answer is incorrect.

Score: 0

Accepted Answers:

sudo pip install python-highcharts

18) What does line 'chart = Highchart()' do in creating highcharts?

- Creates a highchart with the name as chart
- Creates a container where options and data for highcharts can be added
- Doesn't do anything, statement is not required
- Creates highchart dictionary with the name as chart

No, the answer is incorrect.

Score: 0

Accepted Answers:

Creates a container where options and data for highcharts can be added

19) Which of the following statement is true?

- Tooltip tells the value of a particular chart point when we hover over it.
- Legends depicts what each colour code corresponding to a series represents.
- Scatter charts are used for 2-D plots
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

All of the above

20) Bubble chart is used for:

- 1-D plotting
- 2-D plotting
- 3-D plotting
- None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

3-D plotting

1) Which of the following refers to profile linking on social media as discussed in the **1 point** lectures ?

- The ability to connect profiles of a user on different social networks.
- Posting on multiple social media from a single platform.
- A common verification check mark across two or more social networks.
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

The ability to connect profiles of a user on different social networks.

2) Which of the following OSNs contain quality and descriptive information for the purposes of profile linking, as discussed in the lecture ? **1 point**

- 4Chan
- Reddit
- Whisper
- LinkedIn

Yes, the answer is correct.

Score: 1

Accepted Answers:

LinkedIn

3)

1 point

The above X(formerly Twitter) profile helps in profile linking through self-identification.

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True



Registration: t1 @nitinsgr
Observation: t2 @nitinsgr
Observation: t3 @nitinsgr

@nitinsgr
@explorer_nitin
@logicalIndian

Attribute Evolution

Select the correct observations for the above image.

- The above image is a representation of username changes across Facebook and Twitter over time
- As time passes, similarity between usernames across Facebook and Twitter decreases.
- As time passes, similarity between usernames across Facebook and Twitter increases.
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

The above image is a representation of username changes across Facebook and Twitter over time

As time passes, similarity between usernames across Facebook and Twitter decreases.

5) Which of the following feature(s) can be used for profile linking on social media. **1 point**
Say, between Facebook and X(formerly Twitter) for a boy called Arjun ?

- Profile Picture
- Username
- Similarity between people followed by the Arjun on both platforms
- Similarity between people who follow Arjun on both platforms

Partially Correct.
Score: 0.75

Accepted Answers:

Profile Picture

Username

Similarity between people followed by the Arjun on both platforms

Similarity between people who follow Arjun on both platforms

6) Why is studying username change behavior is/are preferable for social media researchers? **1 point**

- This is one of the unique attributes of the user
- It is usually restricted in character length, allowing for smaller datasets
- It is the publicly available attribute
- None of the above

Partially Correct.

Score: 0.67

Accepted Answers:

This is one of the unique attributes of the user

It is usually restricted in character length, allowing for smaller datasets

It is the publicly available attribute

7) Which of the following is/are the reason(s) behind using anonymous social networks ? **1 point**

- Increasing awareness of privacy amongst social media users
- Preventing use of personal information available publicly for surveillance purposes
- Avoiding consequences of controversial political opinions posted on social media
- None of the above

Partially Correct.

Score: 0.66

Accepted Answers:

Increasing awareness of privacy amongst social media users

Preventing use of personal information available publicly for surveillance purposes

Avoiding consequences of controversial political opinions posted on social media

8) Which of the following is/are anonymous social networks? **1 point**

- Whisper
- Yik Yak
- Blind
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Whisper

Yik Yak

Blind

9) "Whispers" on Whisper are equivalent to _____ on X (formerly Twitter)

1 point

- Replies
- Likes
- Tweets
- None of the above

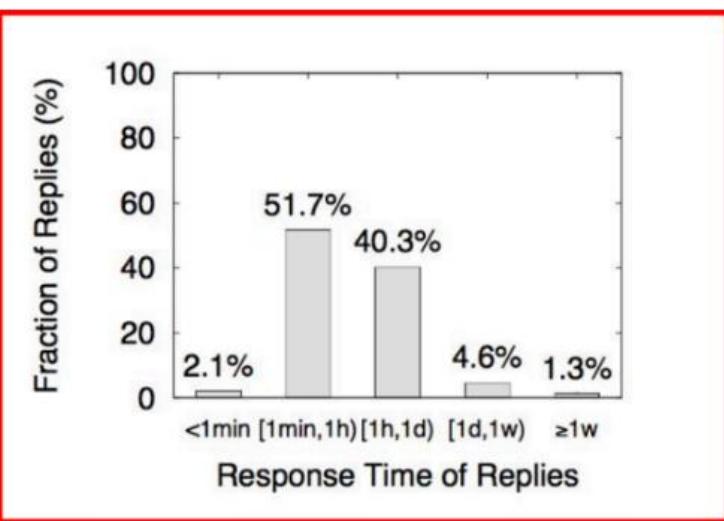
Yes, the answer is correct.

Score: 1

Accepted Answers:

Tweets

10) Observe the graph below regarding replies on Whisper and answer the question **1 point** below.



"If a whisper does not get attention shortly after posting, it is unlikely to get attention later." This statement is :

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

Observe the table given below and answer the following questions [11-13]

Graph	# of Nodes	# of Edges	Avg. Degree	Clustering Coef.	Avg. Path Length	Assortativity Coef.
Whisper	690K	6,531K	9.47	0.033	4.28	-0.011
Facebook	707K	1,260K	1.78	0.059	10.13	0.116
Twitter	4,317K	16,972K	3.93	0.048	5.52	-0.025

11) For the same number of Nodes/Users on Whisper, Facebook and Twitter, users on **1 point** Whisper create more connections/edges than on Facebook/Twitter.

- True
 False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

12) Users on Whisper tend to create more close-knit groups than on Facebook/Twitter. **1 point**

- True
 False

Yes, the answer is correct.

Score: 1

Accepted Answers:

False

13) Assortativity Coefficient being closer to zero implies the graph is more random than **1 point** "small world"

- True
 False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

14) The measure of how close a graph is to complete, with 0 indicating an incomplete **1 point** and 1 indicating a complete graph is called _____.

- Average Degree
- Clustering Coefficient
- Graph Density
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

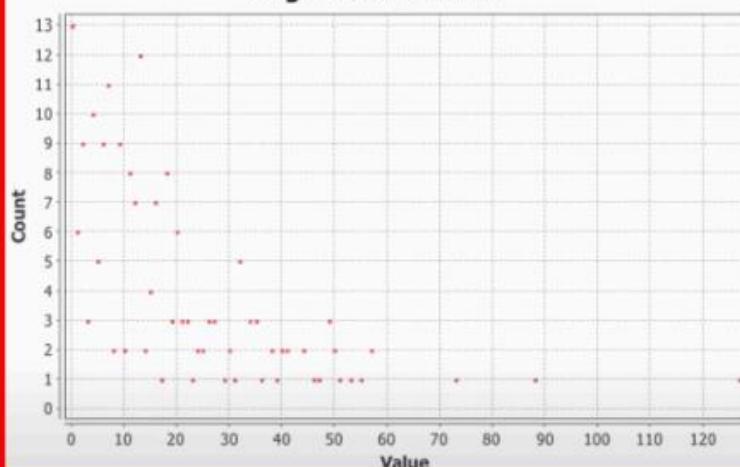
Graph Density

Graph Density

15)

Degree Distribution

1 point



Consider the above Degree Distribution graph generated in Gephi. Which of the following statements is/are correct about this ?

- The average degree of the graph is greater than 70
- The average degree of the graph is lesser than 50
- The average degree of the graph is 125
- The average degree of the graph is less than 10

Yes, the answer is correct.

Score: 1

Accepted Answers:

The average degree of the graph is lesser than 50

1. What are the different approaches to profile linking?

- a) Study and analyze common attributes (place of work, profile picture, name, etc.) across different social media platforms ✓
- b) Check important attributes such as the same emails, locations, and contact numbers (if available) used across multiple platforms ✓
- c) Only looking at the user posting behavior on social media platforms will help in profile linking. ✗
- d) It is not possible to find the same user profiles on different social media platforms. ✗

2. Which of the following statements is/are true regarding username change on social media?

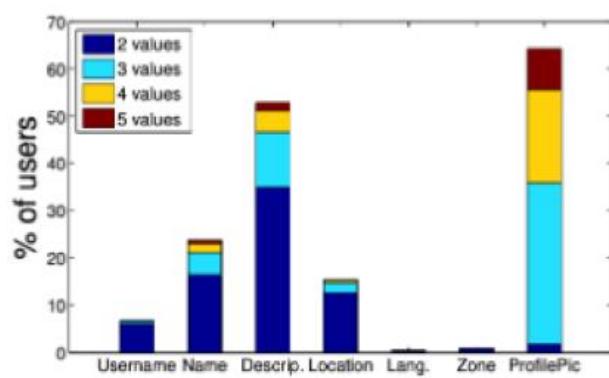
- S1: Changes to any profile attribute other than username do not lead to unreachability to the user profile. ✓
- S2: The potential benign reasons for username change include space gain, suiting a trending event, gaining/losing anonymity, adjusting to real-life events, and avoiding boredom. ✓
- S3: The malicious intentions for username change may include obscured username promotion and username squatting. ✓
- S4: Search by the user's old username results in a failed attempt to reach the user's profile, potentially making others falsely believe that the user account has been deactivated. ✓

- a) Only S1 and S4 are correct ✗
- b) Only S1, S2, and S3 are correct ✗
- c) S1, S2, and S4 are correct ✗
- d) All S1, S2, S3, S4 are correct. ✗

3. Below is the graph of username change.

Which of the following is correct inference from the graph (as discussed in the lecture)?

- a) Around 35 per cent, of the users change their profile picture at least 3 times. ✓
- b) Around 60-70 percent of the users change their usernames atleast once. ✗
- c) Around 35 percent of the users change their profile picture almost 3 times. ✗
- d) Around 50-55 per cent of users have changed their descriptions thrice. ✗



4. Why do whisper users have a low clustering coefficient over the network?

- a) Whisper users are likely to interact with complete strangers who are highly unlikely to interact with each other.
- b) Users interact with a large sample of other users
- c) New users make 20% of the contribution in the whisper content.
- d) All the above

5. What is the primary purpose of Gephi software?

- a) Video Editing
- b) 3D Modeling
- c) Network Visualization and Analysis
- d) Music Composition

6. Why studying username change behavior is/are useful for social media researchers?

- a) This is one of the unique attributes of the user
- b) It is usually homogeneous
- c) The number of characters and length of the username is restricted
- d) It is the publicly available attribute

7. In a directed graph, which of the following statements about a vertex's indegree and outdegree is true?

- a) The indegree is always equal to the outdegree.
- b) The indegree refers to the number of edges leaving the vertex, and the outdegree refers to the number of edges entering the vertex.
- c) The indegree refers to the number of edges entering the vertex, and the outdegree refers to the number of edges leaving the vertex.
- d) The concept of indegree and outdegree only applies to undirected graphs.

8. Which of the following is/are anonymous social networks?

- a) Secret
- b) Wickr
- c) Yak yak
- d) Blind



9. Consider two social media platforms: Twitter and Whisper. Which platform is more likely to exhibit a scale-free network structure with a power law degree distribution?

- a) Twitter, with its large number of celebrity and influencer accounts that attract many followers.
- b) **Whisper, where users predominantly interact with anonymous strangers in temporary connections.**

10. Consider the below situation

Imagine you're a student in a university, and you decide to survey to determine the number of friends each of your classmates has on social media. You ask everyone to report the number of friends they have on one popular social media platform (say Facebook). When you collect the data and calculate the average number of friends reported by your classmates, 300 friends. However, when you individually examine the number of friends reported by each of your classmates, you notice that many of them have fewer than 300 friends, perhaps ranging from 50 to 200 friends.

Why did you observe such a phenomenon?

- a) Most of the individuals have reported it wrong in the report.
- b) **Few individuals have exceptionally high friend counts (outliers) that skew the average upward.**
- c) There is an error in individual calculations. It is impossible to have 300 friends on average if many individuals have fewer than 300 friends.
- d) None of the above

1) Why is username a good candidate to link profiles?

- Username is public
- Username is heterogeneous
- Username is homogeneous
- Username cannot be changed

Yes, the answer is correct.

Score: 1

Accepted Answers:

Username is public

Username is homogeneous

2) Clustering coefficient is the measure of:

- Tendency of nodes to form clusters
- If a cluster is possible or not
- Edge similarity
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Tendency of nodes to form clusters

3) Which network among the following has the maximum clustering coefficient, as discussed in the lecture?

- Instagram
- Facebook
- Twitter
- Gab

4) 4chan is a / an:

- Anonymous social network
- Online social network
- Location based social network
- Ephemeral social network

Yes, the answer is correct.

Score: 0

Accepted Answers:

Anonymous social network

Online social network

5) On anonymous networks we can see who liked a post?

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

False

6) Whisper provides: [select all applicable]

- Archive of a user history
- Links to personal information given user ID
- Persistent Social links
- Replies

Yes, the answer is correct.

Score: 1

Accepted Answers:

Replies

7) How to link the profile of a user on Facebook to 4chan?

- By searching for user name and doing a semantic matching
- Comparing friends of a user on both the networks
- Both a and b
- Impossible

Yes, the answer is correct.

Score: 1

Accepted Answers:

Impossible

8) We can use the visualization software Gephi to get Modularity of a given graph?

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

9) Which of the following layout options is NOT available in Gephi ?

- Contraction
- Expansion
- Random
- Classic

Yes, the answer is correct.

Score: 1

Accepted Answers:

Classic

10) Why is visualizing graphs necessary?

- It gives quantitative measures that are otherwise impossible to find
- Human deducible decision making
- To calculate the number of clusters
- None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Human deducible decision making

10) Refer to the figure below (Figure 5) about a partner account promoting another in a **1 point** tweet and answer the following question:

Figure 5

Which of the following phenomena is shown in the above figure as discussed in the paper ?

- Username Squatting
- Obscured Username Promotion
- Adjust to Events
- Space Gain

Yes, the answer is correct.
Score: 1
Accepted Answers:
Obscured Username Promotion

1. Which social network was primarily designed for sharing location-based information and recommendations?

-
- a) **Foursquare**
 - b) Google+
 - c) Twitter
 - d) All of the above

2. If you were developing a mobile app that combines location-based recommendations with real-time social updates, which of these platforms would you draw inspiration from the most?

-
- a) **Foursquare and Twitter**
 - b) Google+ and Foursquare
 - c) Combine elements from all three platforms.
 - d) Choose a different platform entirely.

3. Which of the following best defines the term "Valid Ambiguous Geographical Information (VAGI)?"

- a) Geographical information that is factually incorrect or demonstrably false.
- b) **Geographical information that is open to multiple interpretations or lacks clarity.**
- c) Geographical information that is intentionally deceptive or misleading.
- d) Geographical information that is outdated or no longer relevant.

4. What is the primary function of a Valid Artificial Geographical Interface (VAGI)?

- a) To provide comprehensive access to all available geographical data in a readily usable format.
- b) **To assist users in navigating complex physical environments using AI-powered tools.**
- c) To automatically verify the accuracy and validity of all online geographical information.
- d) To translate spoken descriptions of locations into precise geographical coordinates.

5. You want to compare engagement for two social media campaigns. Campaign A has 100,000 views and 500 comments, while Campaign B has 50,000 views and 1,000 comments. Which campaign had higher engagement per view?

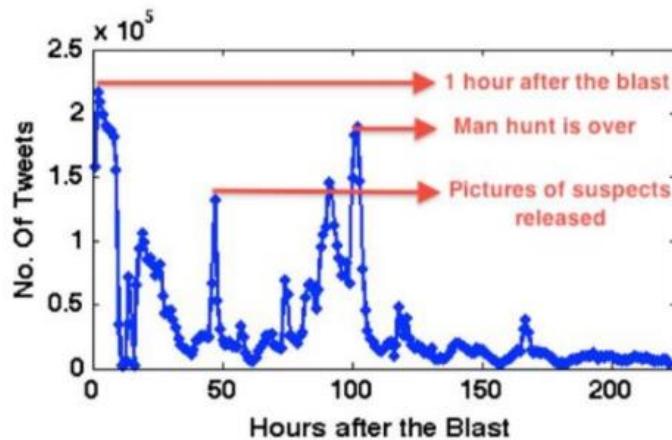
- a) Campaign A
- b) **Campaign B**
- c) They have the same engagement per view (5 comments/view).
- d) Insufficient data to determine.

6. Analyzing follower growth on a social media account shows a consistent increase of 1,000 new followers every month. Which statistic would best represent this growth rate?

- a) Average followers per day
- b) Percentage increase from the previous month
- c) Annual projected follower count
- d) None of the above

7. What are the major contributions of the figure. What does the following figure depict?

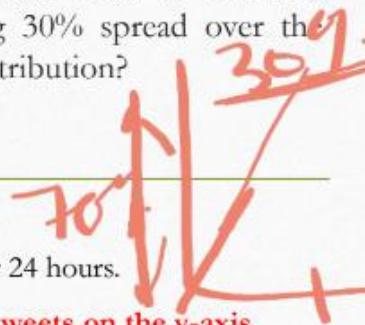
- a) The figure shows the temporal distribution of tweets after the Boston blast.
- b) The figure shows the spatial distribution of tweets after the Boston blast.
- c) Both above.
- d) None of the above.



8. You track click-through rates (CTR) for two different links shared on a social media post. Link A has a CTR of 5%, while Link B has a CTR of 0.5%. Which statement is most accurate?

- a) **Link A is 10 times more effective than Link B.**
- b) Link A receives 5 times as many clicks as Link B.
- c) Both links have equally ineffective CTRs.
- d) Click-through rates alone cannot indicate link effectiveness.

9. Analyzing retweet patterns for a popular tweet shows 70% of retweets occurred within the first 24 hours, with the remaining 30% spread over the following week. Which graph would best visualize this distribution?



- a) Bar chart with equal spacing for each day.
- b) Pie chart showing proportions of retweets within and after 24 hours.
- c) **Line graph with time on the x-axis and cumulative retweets on the y-axis.**
- d) Stacked bar chart with one section for first 24 hours and another for the rest of the week.

10. You discover a statistically significant correlation between the number of hashtags used in a social media post and its engagement. What does this finding suggest?

- a) Using more hashtags guarantees higher engagement.
- b) Hashtags directly cause increased engagement.
- c) **There might be an underlying factor influencing both hashtag usage and engagement.**
- d) The correlation cannot be interpreted without understanding the cause-and-effect relationship.

What are the different type of attacks that could happen:

- i. Pragmatic
- ii. Syntactic
- iii. Physical

Physical 2
will
rela

- Only i and iii
- Only iii
- Only ii and iii**
- All of the above

What is the term used for Generating usernames and monetizing them, when necessary?

- Username Sale
- Username hopping
- Username squatting**
- All of the above

What is meant by the Pareto distribution of username change?

- 20% users frequently change usernames in short intervals and 80% change rarely after long intervals**
- 80% users frequently change usernames in short intervals and 20% change rarely after long intervals
- 20% users rarely change usernames in short intervals and 80% change frequently after long intervals
- 80% users rarely change usernames in short intervals and 20% change frequently after long intervals

For anonymous social networks which of the following statements hold true?

- i. Clustering coefficient is different
- ii. User interact with strangers

- Only i
- Only ii
- Both i and ii**
- None of the above

me tie → lesser cluster

clustering w.
each node.

Which of the following statements are false about spammers:

- i. Users of less in-degree do not reciprocate to spammers → True
- ii. Users with small in-degree tend to follow back more. → False
- iii. Top link farmers have very high in-degree compared to spammers ✓

- Only ii**
- Only i
- Only iii
- None of the above

“For linking user profiles across different social media, having profile information from past will not help in improving the efficiency or accuracy” Is this statement true or False?

- True
- False**

Mark all the true statements with respect to Fake information on social media?

- Truthful information comes onto social media faster than fake information
- ↳ **Various user and user's network features can be used to classify fake information.**
- **Community of users posting fake content are usually closed.**
- Techniques used to classify fake content on twitter would work equally well on facebook post, without any need for tweaking.

↳ purpose •

Which of the following is/are Location-based social networks?

- Gab
- **Foursquare**
- Koo
- Whisper

- In the context of e-crime match the following:

1. Clickbaiting	i. creating fake account with photos and details sourced from social media
2. Phishing	ii. using hashtag to post unrelated content to sell product
3. Hashtag Hijacking	iii. getting you to click on links which are not legitimate.
4. Impersonation	iv. Act of tricking someone into handing login credentials

- For matplotlib, which of the following parameters can be configured?

- ✓ i. Minor axis label
 - ✓ ii. Intervals on both axis
 - ✓ iii. Size of figure
- iv. The number of subplots in a single figure

- Only i and ii
- Only i
- Only i, ii, iv
- ✓ All of the above

2) What is the main difficulty while collecting user-location based data through a social network, say Twitter?

- API doesn't give us the location of the post
- Of all the tweet that we collect, very few of them have geo-tagged information available
- The locations given are very vague
- Twitter only gives us, country-level information

No, the answer is incorrect.

Score: 0

Accepted Answers:

Of all the tweet that we collect, very few of them have geo-tagged information available

3) Out of the following 4, identify the 'valid' location-based features?

- Near Agra
- Near Taj Mahal
- Latitude and Longitude
- 'At my friend's home'

No, the answer is incorrect.

Score: 0

Accepted Answers:

Latitude and Longitude

4) Out of the following 4, identify the 'ambiguous' location-based features?

- Near Agra
- Near Taj Mahal
- Latitude and Longitude
- 'At my friend's home'

No, the answer is incorrect.

Score: 0

Accepted Answers:

Near Agra

Near Taj Mahal

Near Taj Mahal

5) Out of the following 4, identify the 'non-geographic information' location-based feature?

- Near Agra
- Near Taj Mahal
- Latitude and Longitude
- 'At my friend's home'

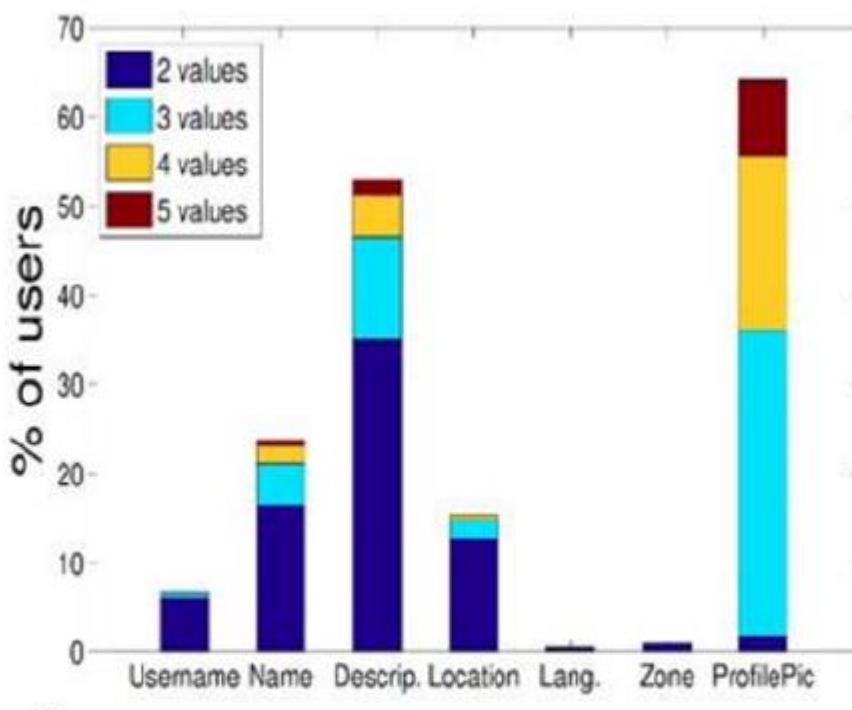
No, the answer is incorrect.

Score: 0

Accepted Answers:

'At my friend's home'

7) What can be inferred from the following graph?



- Around 50% User change their time zones very frequently
- Users change their usernames more than they change their descriptions
- Around 65% of the users change at least one profile attribute
- Users change their profile pictures multiple times.

No, the answer is incorrect.

Score: 0

Accepted Answers:

Around 65% of the users change at least one profile attribute

Users change their profile pictures multiple times.

9) Which features would you use to analyse Suspended profiles?

- Spatial identification
- Retweet / reply network
- Temporal analysis
- Impact /vs/ growth

No, the answer is incorrect.

Score: 0

Accepted Answers:

Retweet / reply network

Temporal analysis

10)

What can be inferred from the following graph?

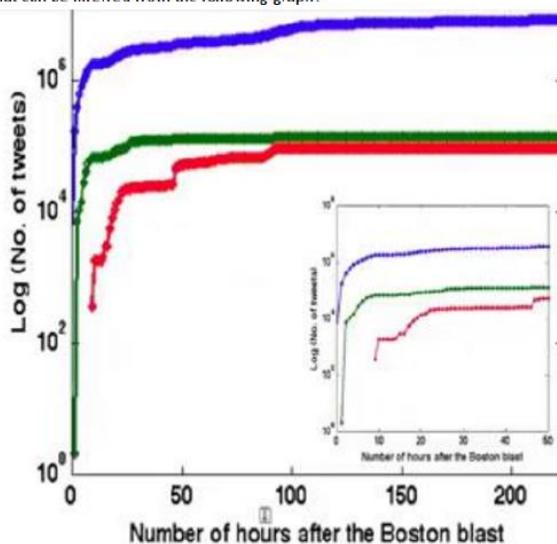


Fig: Log distributions of number of total face, true information and NA category tweets for the Boston Blast

- True information starts spreading much later than the rumours
- Rumors often surpass the actually true content
- Rumors starts spreading much later than the true informations
- True information often surpasses the rumors

No, the answer is incorrect.

Score: 0

Accepted Answers:

*True information starts spreading much later than the rumours
Rumors often surpass the actually true content*

1) Which of the following can be used for writing Python Programs in an Ubuntu Terminal ?

- vim
- neovim
- vi
- All of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the Above

2) What is used to install and manage packages for Python?

- pip
- npm
- apt get
- Anaconda

Yes, the answer is correct.

Score: 1

Accepted Answers:

pip

3) Why is a "Graph" API called so?

- It allows users to draw and create graphs and charts
- It represents social connections and relationships as a graph data structure
- It primarily deals with geographical mapping and location-based services
- It is used exclusively for graph theory and mathematical computations

Yes, the answer is correct.

Score: 1

Accepted Answers:

It represents social connections and relationships as a graph data structure

4) What characteristics did the researchers find among top link farmers?

- They had low in-degree compared to random samples
- They were unpopular and inactive users
- They had high in-degree compared to random samples
- None of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

They had high in-degree compared to random samples

5) Which of the following are/were virtual worlds in Social Media ?

- Club Penguin
- Habbo
- Twitter
- All of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

Club Penguin

Habbo

6) Which V of social media represents "Amount of content generated across social media platforms." 1

- Value
- Velocity
- Volume
- Veracity

Yes, the answer is correct.

Score: 1

Accepted Answers:

Volume

7) Which of the following is/are e-crimes on online social media ?

- Phishing
- Clickbaiting
- Work From Home Scam
- All of the Above

Yes, the answer is correct.

Score: 1

Accepted Answers:

All of the Above

8) Which of the following is/are the goals of a phishing attack ?

- Stealing personal information
- Stealing your financial information
- Stealing your identity
- Attacking your social networking contacts

Yes, the answer is correct.

Score: 1

Accepted Answers:

Stealing personal information

Stealing your financial information

Stealing your identity

Attacking your social networking contacts

9) Hitesh has 400 followers on Twitter, 500 followers on Facebook and 1400 followers **1 point** on Instagram. The act of finding the actual number of unique followers across all his social media profiles is called de-duplicating. This statement is:

- True
- False

Yes, the answer is correct.

Score: 1

Accepted Answers:

True

10) A node in a graph has an in-degree of 2 and an out-degree of 4. The degree of this **1 point** node is:

- 2
- 4
- 6
- 8

Yes, the answer is correct.

Score: 1

Accepted Answers:

6

11) Which of the following is the correct command for initializing a git repository in a newly created empty folder ? **1 point**

- git initialize
- git create
- git initiate
- git init

Yes, the answer is correct.

Score: 1

Accepted Answers:

git init

12) Latanya Sweeney's research highlighted how two seemingly unrelated pieces of information can be linked to reidentify an individual. Which term best describes this concept? **1 point**

- Data Anonymization
- Data Aggregation
- Reidentification
- All of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Reidentification

13) Which of the following OSNs have an API ?

- X (formerly Twitter)
- Facebook
- YouTube
- None of the above

Partially Correct.

Score: 0.66

Accepted Answers:

X (formerly Twitter)

Facebook

YouTube

14) Which of the following are not stopwords in the following list?

1 point

stopwords = ["i", "me", "my", "myself", "we", "our", "ours", "ourselves", "you", "your", "yours", "yourself", "yourselves", "he", "him", "his", "himself", "she", "her", "hers", "herself", "it", "its", "had", "having", "do", "does", "did", "doing", "mango", "a", "an", "the", "and", "but", "if", "or", "because", "as", "until", "while", "of", "at", "by", "for", "with", "about", "against", "between", "into", "now"]

- because, against, mango, yourself
- mango, now
- mango
- All the words in the list are stopwords

No, the answer is incorrect.

Score: 0

Accepted Answers:

mango

15) A sentiment analysis study analyzes a text passage for valence and arousal **1 point**
values. On a scale of -1 to 1, the valence and arousal score is 0.80 and 0.90. What do these values suggest about the emotional characteristics of the text?

- Negative sentiment, low emotional intensity
- Positive sentiment, High emotional intensity
- Negative sentiment, high emotional intensity
- Positive sentiment, low emotional intensity

Yes, the answer is correct.

Score: 1

Accepted Answers:

Positive sentiment, High emotional intensity

1. Given below the statement. The statement is “An API is a set of rules and protocols that allows different software applications to communicate with each other, while a programming language is a formal language used to write instructions for a computer to perform specific tasks.”

-
- a) True
 - b) False
 - c) Can't say

2. Which of the following features is supposed to be a part of tweet content on Twitter?

- a) Presence of pronouns
- b) No. of followers
- c) No. of retweets
- d) Mention of self words (I; my; mine)

3. Which of the following features is most likely to be manipulated by bots on social media platforms?

- a) Number of comments on a post.
- b) Profile picture and username.
- c) **Frequency of posting and engagement.**
- d) Sentiments expressed in the post content.

4. Which of the following is username creation behavior?

- a) **Static behavior patterns**
- b) Occasional reuse patterns
- c) **Temporal behavior patterns**
- d) Frequent reuse patterns

6. Why is "Graph" API called like that?

- a) it allows users to draw and create graphs and charts.
- b) **it represents social connections and relationships as a graph data structure.**
- c) it primarily deals with geographical mapping and location-based services.
- d) it is used exclusively for graph theory and mathematical computations.

8. What is the primary purpose of an API key on Twitter?

- a) To limit the number of characters in a tweet.
- b) To verify a user's identity on the platform.
- c) **To authenticate and access Twitter's APIs for development purposes.**
- d) To encrypt and secure direct messages on Twitter.

9. What is the primary purpose of link farming in the context of SEO (Search Engine Optimization)?

- a) To create high-quality, relevant backlinks to improve a website's search engine ranking.
- b) To increase organic traffic by optimizing on-page content and meta tags.
- c) **To manipulate search engine algorithms by generating many low-quality backlinks.**
- d) To enhance the user experience by improving website navigation and design.

1. You analyze your Facebook friends' data with Python and discover that most are within your age range and location. What privacy concern does this highlight?

-
- a) Filter bubble creation and potential for biased information exposure
 - b) Lack of diversity in your social network connections
 - c) Facebook's data collection practices and potential for unintended inferences
 - d) **All the above**

2. You're building a Twitter bot to automatically retweet posts with a specific hashtag. What ethical considerations should you prioritize?

-
- a) Transparency about the bot's purpose and functionality
 - b) Avoiding manipulation of trending topics or hashtags
 - c) Disclosing the bot's identity to avoid misleading users
 - d) **All the above**

3. You receive an SMS claiming your bank account is compromised and urging you to click a link to verify your details. What are the red flags suggesting a potential phishing scam?

- a) Urgency, pressure to act quickly, and unfamiliar language
- b) Grammatical errors, suspicious domain names in the link, and unsolicited communication
- c) Requesting personal information directly through the link, lack of official branding, and typos
- d) **All the above**

4. You analyze YouTube video comments with NLTK to identify hate speech. What limitations should you be aware of in this approach?

-
- a) Context and sarcasm can be misunderstood by machines
 - b) Cultural nuances and slang might be misinterpreted
 - c) Bias in training data can lead to unfair or inaccurate detections
 - d) **All the above**

5. You discover a website claiming to sell verified Instagram accounts. What potential issues arise from purchasing such accounts?

-
- a) Violation of Instagram's terms of service and account deactivation
 - b) Risk of purchasing stolen accounts linked to illegal activities
 - c) Lack of control over account security and potential for password compromise
 - d) **All the above**

6. You analyze Facebook ad targeting options and notice specific demographic filters based on income level. What privacy concern does this highlight?

-
- a) Potential for discriminatory advertising based on sensitive user data
 - b) Lack of transparency in Facebook's ad targeting algorithms
 - c) Risks of data misuse by advertisers for profiling and manipulation
 - d) **All the above**

7. You receive an email claiming your Netflix subscription is about to expire, urging you to update your payment details. What are the signs suggesting a potential phishing attack?

- a) Unfamiliar sender address, generic greeting, and unusual formatting
- b) Threats of service interruption, excessive use of emojis, and pressure to act immediately
- c) Requesting login details directly through the email, typos and grammatical errors, and lack of personalization
- d) **All the above**

8. You analyze Twitter trending topics and notice unusual patterns of coordinated activity. What might this indicate?

- a) Organic trending based on genuine public interest
- b) Potential for bot manipulation and artificial amplification of hashtags
- c) Coordinated disinformation campaigns aimed at influencing public opinion
- d) **All the above**

9. You develop a browser extension that tracks users' online activity across websites. What ethical considerations must you prioritize?

- a) Transparency about data collection practices and user consent
- b) Limiting data collection to specific purposes and avoiding unauthorized sharing
- c) Implementing strong security measures to protect user data from breaches and misuse
- d) **All the above**

10. You discover a data leak exposing millions of Instagram user email addresses and phone numbers. What are the potential consequences of this breach?

-
- a) Increased risk of phishing attacks and spam targeting exposed users
 - b) Potential for identity theft and unauthorized access to accounts
 - c) Reputational damage for Instagram and loss of user trust
 - d) **All the above**