NAME – Shaurya Teotia

BATCH – 6th June

DOMAIN – Data Science with ML/AI

ASSESSMENT

Q1. Mention some of the advantages of python?

Ans. Python is an easy to understand and learn programming language and has a lot of libraries which we can import to perform endless tasks like machine learning, data science and implement AI models. It is also free to use and is open source.

Q2. What are local variables and global variables in Python?

Ans. Local variables are variables that are declared within a particular scope, e.g. a function and are accessible only there. Global variables on the other hand are accessible throughout the program as they are declared outside any local scope.

e.g.

```

a=5

def func:

b=7

return

```

Here a is a global variable and can be accessed anywhere, including in the function func, whereas b is a local variable and is local to the function func, i.e. it can be accessed only within func.

Q3. What is Lambda Functions in Python?

Ans. Lambda functions are functions without names that can be written in a single line.

e.g.

```

c=lambda a,b : a+b

c(6,3)

```

The output will be 9 as this lambda function adds the given two arguments.

Q4. What is a Negative Index in Python?

Ans. Negative index is index from the last element of the object we are indexing instead of the first. So, for e.g. if we have a list a = [10,22,13,44,65], then a[-1] will give us 65 and a[-2] will give us 44.

Q5. What is the difference between tuples and lists?

Ans. Tuples are similar to lists but are unchangable, i.e. we cannot add/remove elements and cannot change the values of elements inside the tuples.

On the other hand, lists can append new elements, remove existing ones and change values of existing elements too.

Also, tuples are denoted by () and lists by [].

Q6. What is a dynamically typed language?

A dynamically typed language is one where the data type of a variable is determined automatically when the program is run, rather than the programmer explicitly providing the data type of the variable they are declaring.

Python is an example. e.g. if in Python we write a = "abcd", then the data type of a is automatically determined as string, rather than us having to state that.

Q7. What are the data types available in python? Brief about it?

int: integers, i.e. ...,-4,-3,-2,-1,0,1,2,3,4,...

str: strings (sequence of characters), enclosed by single or double quotes

bool: True or False logical variables

list: collection of elements of various data types, and can be changed, denoted by []

tuple: collection of elements of various data types, that cannot be changed, denoted by ()

dict: short for dictionary, collection of key-value pairs, denoted by {} with key-value pairs represented as key:value

set: unordered, unchangeable collection of unique elements, denoted by {}, new elements can be added/removed

Q8. Why we have to use functions in Python? Explain with an example?

Ans. We need to use functions for tasks that we have to perform again and again, so that we do not need to retype the same code every time but can call the corresponding function instead.

e.g. the following function combines the two lists given to it and returns the resulting list.

def combine\_lists (a,b):

c=list()

for i in a:

c.append(i)

for i in b:

c.append(i)

return c

a=[1,2,3]

b=['a','b','c']

c=combine\_lists(a,b)

c

Q9. What is the difference between function and generators? Give an example?

Ans. Generators are like functions but instead of return they use yield, which appends elements to the generator object which is then returned. We can use for loop to iterate through the returned generator object.

They are advantageous over functions when the input is very large, as functions will take a lot of time and memory but generators will display the output in the console only, consuming much lesser time and memory. Functions will compute the entire thing before returning, but generators will go one element at a time.

e.g.

```

def gen(n):

for i in range(n):

yield i

```

This generator will generate a sequence of the first n whole numbers

def gen(n):

for i in range(n):

yield i

a = gen(5)

for i in a:

print (i)

Q10. Explain brief about conditional statements with an example?

Ans. Conditional statements are statements that check whether or not the given condition is true, and based on whether it is true or false, it runs the corresponding block of code.

e.g. in the following code the conditional statement checks the condition a < 5. As a is 10, so the condition is false, and it goes to the 'else' part and computes that. If it was true it would have computed the first block instead.

a = 10

if (a<5):

print('This prints if the statement is true')

else:

print('This prints if the statement is false')

Q11. Why we have to use exceptional handling in python? What are the keyword used to handle the error , give one example?

Ans. We use exception handling to handle cases where an abnormal event occurs, i.e. an exception occurs, that can crash the program. This could be an input that should not be entered, like making the divisor in a division 0, or a size for a list so large that it isn't possible to make that list due to memory limitations. To avoid crashing the program, we use exception handling to handle these exceptions and ask the user to retry, or do not compute that line, etc.

The keywords used are try to check for exceptions and except to handle the exception, and raise can also be used to raise a specific error.

e.g. the following code is for division in which the code handles the case where the division doesn't happen successfully, like in the case the user enters 0 as the divisor, and asks them to retry.

while True:

try:

a=int(input('Dividend: '))

b=int(input('Divisor: '))

print(a,'/2',b,'=',a/b)

break

except:

print('Error. Try again')

Q12. Define class and objects with an example?

Ans. Classes are user defined data types that can have their own members and functions, which can be used to keep related functions together and data together for a specific purpose. e.g. we can make a student class which can have student details as members, e.g. name and email, along with functions like register for a course, etc.

Objects are instances of classes. Thus, if we define the student class described above and then make a variable of student type, then that variable is an object of student type.

e.g.

class student:

def \_\_init\_\_(self,name,email):

self.name=name

self.email=email

s = student('Priyam','abc@xyz.com')

print(s.name,s.email)

print('Q13. Write a program to swap the two numbers in python?')

a=5

b=10

print('Before: a = ',a,", b = ",b)

(a,b)=(b,a)

print('After: a = ',a,'b = ',b)

print('Q14. Write a palindrome program in python')

def palindrome (a):

l=len(a)

for i in range(l):

if a[i]!=a[l-i-1]:

return False

return True

print(palindrome('abcabca'))

print(palindrome('abcba'))

Q15. What is the use of numpy array over list?

Ans. 1. Numpy takes lesser memory than lists.

2. Numpy is more convenient. For e.g., we can directly add two numpy arrays while we have to use list comprehension in the case of lists

3. Numpy is faster than lists

Q16. Explain in brief about reshape in numpy with an example?

Ans. Reshape changes the dimensions of the numpy array into the given dimensions. For example, the following code changes the given array from a 4x4 array to a 8x2 array.

import numpy as np

a = np.array([[1,2,3,4],[5,6,7,8],[9,10,11,12],[13,14,15,16]])

print(a,'\n')

a=np.reshape(a,(8,2))

print(a)

Q17. What is the use of pandas explain in brief?

Ans. Pandas is a powerful python library that can be imported and then used to analyse, manipulate and clean data.

We can use pandas to import data in the form of data frames and then clean it to prepare it for further processing, e.g. machine learning. We can analyse the data by filtering it, printing summaries and stats of the data, and make modifications to it in order to clean it.

Q18. What do you mean by supervised machine learning and unsupervised machine learning? Explain in brief?

Ans. Supervised machine learning is the type of machine learning that uses labeled data, i.e. both independent and dependent variables, to create a model that can then predict the label of newly input data.

Examples include regression (linear, multilinear, logistic etc.), decision tree, random forest, KNN etc.

Unsupervised learning on the other hand uses data that is unlabeled, i.e. contains only independent variables and not dependent variables to create a model that figures out clusters or similar data points on its own.

Examples include k-means clustering, etc.

Q19. Differenciate between Decision tree and Random forest?

Ans. A decision tree is the basic unit of a random forest.

A decision tree is a structure in the form of one root nodes, several decision nodes and leaf nodes that contain the final output. We start from the root node, follow the branch that satisifies our condition and continue till we reach a leaf node, that contains the output.

A random forest on the other hand is a collection of several decision trees.

Q20. What do you mean by forward propagation and backward propagation in ANN?

Ans. Forward propagation in an ANN is when we move forward, i.e. left to right, i.e. towards the direction of the output layer.

Backward propagation is when we move in the opposite direction, towards the input layer, i.e. the output of a hidden layer or the output layer itself is used as the input in a previous layer. This is done to update weights and reduce loss.

Q21. Explain the working flow of ANN and CNN?

Ans. ANN consists of input layer, hidden layers, activation functions, bias, weights and output layer.

We start from the input layer, that adds all the inputs multiplied by their weights, adds the bias and uses the activation function to produce an output which goes on to the next layer, and so on, until we reach the final output layer. During this process we also update weights to reduce loss.

CNN consists of an input image, that goes through optional padding. Then several rounds (as required) of convolution, and pooling happen along with filters, to produce a fully connected layer. Then the output produced is flattened. This flattened image is used as input to the fully connected layer to produce a final output which could be a classification of the original image.

Q22. What is the use of activation function and optimizers?

Ans. Activation functions are used to transform inputs to outputs. They introduce non linearity to the network.

Optimizers are used to minimize loss.

Q23. What is the difference between R square and adjusted Rsquare?

Ans. R square is simply the sum of squares of difference between actual and predicted values.

Adjusted R square on the other hand is the R square value adjusted for the number of independent variables used to calculate the predictions. It is equal to 1 - [(1-Rsquare)(n-1)]/(n-k-1)], where n=no. of data points, and k=no. of independent variables. It helps us know if adding independent variables is benefitting our model or not.

Q24. What do you mean by LSTM in RNN?

Ans. LSTM is short for long short term memory in RNN (recurrent neural networks).

It is an improved version of RNN that uses feedback to remember long term dependencies unlike usual RNN that remembers only short term dependencies.

Q25. What is the use of image processing?

Ans. Image processing is used to analyse images, build models for image classifications, manipulate images and modify their pixels, etc. It can be used in image classification, or editing of images to for example remove the background, etc.

Q26. What are the steps involved in NLP?

Ans. The steps involved in natural language processing are:

1. Lexical analysis

2. Syntactical analysis

3. Semantic analysis

4. Discourse integration

5. Pragmatic analysis

Q27. What do you mean by bias and variance?

Ans. Bias is the model error that happens due to under fitting, or not viewing the data enough times to draw correct patterns from it.

Variance on the other hand happens with over fitting, when the data is viewed too many times and results in the model being very accurate with the training data but failing with new data.

Q28. What are the steps involved in machine learning to create model?

Ans. To create a model, we:

1. Import the dataset using which we have to create the model

2. Perform EDA (exploratory data analysis), in which we understand the dataset

3. Clean the dataset to handle NULL values, check the data types of columns, etc.

4. Pre-processing in which we perform label encoding, binary encoding, scaling etc. and split our dataset into training and testing data

5. Model creation using linear regression, logistic regression, KNN, etc. as required

6. Evaluation and validation of model using metric analysis which includes metrics like accuracy, confusion matrix, etc.