

SET	A
-----	---

INDIAN SCHOOL MUSCAT
FIRST PRE BOARD EXAMINATION 2023
ARTIFICIAL INTELLIGENCE (417)

CLASS: X

Max.Marks: 50

MARKING SCHEME			
SET	QN.NO	VALUE POINTS	MARKS SPLIT UP
A	Q. 1	Answer any 4 out of the given 6 questions on Employability Skills (1x 4= 4 marks)	
	i.	d) Rahul learns from the feedback and makes his project work better.	1
	ii.	b) Computer tends to restart on its own	1
	iii.	(a) takes responsibility for her mistakes	1
	iv.	(b) Manages the business	1
	v.	c) Self-motivation	1
	vi.	(d) Time	1
	Q.2	Answer any 5 out of the given 6 questions (1x 5= 5 marks)	
	i.	b) Gender Equality	1
	ii.	c) Resolution	1
	iii.	b) Data Science	1
	iv.	$Accuracy = \frac{\text{Correct prediction}}{\text{Total cases}} * 100\%$ $Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$	1
	v.	a) Both Statements 1 and 2 are correct.	1

	vi.	If we have a value of 1 (that is 100%) for both Precision and Recall.	1
	Q.3	Answer any 5 out of the given 6 questions (1x 5= 5 marks)	
	i.	Surveys, webscraping,cameras,sensors,observations ,API (any two-1/2 mark each)	1
	ii.	(d) the problem can be identified and goals can be set.	
	iii.	c) 1	1
	iv.	Automatic Summarization:	1
	v.	b) Website recommendations	1
	vi.	(d) Regression	1
	Q.4	Answer any 5 out of the given 6 questions (1x 5= 5 marks)	
	i.	Red (R), Green (G) and Blue (B)	1
	ii.	Stemming / Lemmatisation	1
	iii.	c) AI Bias	1
	iv.	a) True Positive ,c)True Negative	1
	v.	c) Both A and R are true and R is the correct explanation of A	1
	vi.	d. Confusion Matrix	1
	Q.5	Answer any 5 out of the given 6 questions (1x 5= 5 marks)	
	i.	b. Text Classification	1
	ii.	<i>pixel value</i> describes how bright that pixel is, and/or what colour it should be.	
	iii.	(c) Recall	1
	iv.	scriptbots	1
	v.	Linguistic intelligence	1
	vi.	Stopwords are the words which occur very frequently in the corpus but do not add any value to it.	1
		SECTION B – SUBJECTIVE TYPE QUESTIONS	

		Answer any 3 out of the given 5 questions on Employability Skills (2 x 3 = 6 marks) Answer each question in 20 – 30 words.	
	Q. 6	Interests 1. Things that you like to do in your free time that make you happy. An acquired or natural capacity 2. Things you are curious about or would do even if no one asked you to do it. 3. Things you want to learn or would like to do in the future Ability 1. An acquired or natural capacity 2. Enable you to perform a particular job or task with considerable proficiency.	2
	Q. 7	(a) Use passwords to login to your computer (b) Install Anti-virus and Firewall , Anti-viruses and Firewall monitor the data coming in and out of a computer and prevent and viruses from entering. Antiirusescan also detect and clean viruses that may have entered a computer. (c) Encrypt Data , This is usually done by banks and companies in which important customer information is stored. They can encrypt their entire hard disk using encrypting feature inWindows (Bitlocker). This would force users to use a decryption password (or key) before starting the computer thus preventing unauthorisedusage. (d) Secure sites , If the site address starts with https,//and a lock symbol, then it is safe to give your credit card and bank details.	2
	Q. 8	1)Quality Education 2) Clean Water and Sanitation 3)Affordable and Clean Energy 4)Decent Work and Economic Growth 5)Reduced Inequalities (Explanation of Any two SDGs – 1 mark each	2
	Q. 9	Organiserof Society's Productive Resources Helpful in Capital Formation: Increase in Employment Opportunities: Development of New Production Techniques: Visionary Leader: Contribution of the Execution of Government Policies Higher Productivity (any 4 – 1/2 mark each)	2
	Q. 10	The use of visual components, such as pictures, graphics, and designs, to convey information and ideas to an audience is known as visual communication. (1 mark) Example of visual communication Infographic ,An infographic is a visual display of facts, figures, or other knowledge that aims to express complicated concepts simply and	2

		effectively. Graph,Chart,Diagram,Photograph. (eg. any two – 1 mark)	
		Answer any 4 out of the given 6 questions in 20- 30 words each (2 x4 = 8 marks)	
	Q. 11	<p>bag of words gives us two things:</p> <ol style="list-style-type: none"> 1. A vocabulary of words for the corpus 2. The frequency of these words (number of times it has occurred in the whole corpus).(1 mark) <p>1. Text Normalisation: Collect data and pre-process it</p> <p>2. Create Dictionary: Make a list of all the unique words occurring in the corpus. (Vocabulary)</p> <p>3. Create document vectors: For each document in the corpus, find out how many times the word from the unique list of words has occurred.</p> <p>4. Create document vectors for all the documents. (1 mark)</p>	2
	Q.12	<p>Object Detection</p> <p>Object detection is the process of finding instances of real-world objects such as faces, bicycles, and buildings in images or videos. Object detection algorithms typically use extracted features and learning algorithms to recognize instances of an object category. It is commonly used in applications such as image retrieval and automated vehicle parking systems. (1 mark)</p> <p>Instance Segmentation</p> <p>Instance Segmentation is the process of detecting instances of the objects, giving them a category and then giving each pixel a label on the basis of that. A segmentation algorithm takes an image as input and outputs a collection of regions (or segments). (1 mark)</p>	2
	Q.13	<p>The term used for the whole textual data from all the documents altogether is known as corpus. (1 mark)</p> <p>The goal of sentiment analysis is to identify sentiment among several posts or even in the same post where emotion is not always explicitly expressed. Companies use Natural Language Processing applications, such as sentiment analysis, to identify opinions and sentiment online to help them understand what customers think about their products and services(1 mark)</p>	2
	Q.14	<p>stemming is the process in which the affixes of words are removed and the words are converted to their base form. Stemming and lemmatization both are alternative processes to each other as the role of both the processes is same – removal of affixes. But the difference between both of them is that in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one. Lemmatization makes sure that lemma is a word with meaning and hence it takes a longer time to execute than stemming. (2 marks)</p>	2
	Q.15	<p>Clustering: Refers to the unsupervised learning algorithm which can cluster the unknown data according to the patterns or trends identified out of it. The patterns observed might be the ones which are known to the developer or it might even come up with some unique patterns out of it. (1 mark)</p> <p>Dimensionality Reduction: We humans are able to visualise upto 3-Dimensions only but according to a lot of theories and algorithms, there are various entities</p>	2

		which exist beyond 3-Dimensions. For example, in Natural language Processing, the words are considered to be N-Dimensional entities. Which means that we cannot visualise them as they exist beyond our visualisation ability. Hence, to make sense out of it, we need to reduce their dimensions. Here, dimensionality reduction algorithm is used.(1 mark)	
	Q.16	<p>Underfit -The model's output does not match the true function at all. Hence the model is said to be underfitting and its accuracy is lower. (1 mark)</p> <p>Overfit -model performance is trying to cover all the data samples even if they are out of alignment to the true function. This model is said to be overfitting and this too has a lower accuracy.(1 mark)</p>	2
		Answer any 3 out of the given 5 questions in 50- 80 words each (4 x3 = 12 marks)	
	Q.17	<p>A Neural Network is divided into multiple layers and each layer is further divided into several blocks called nodes.</p> <p>Each node has its own task to accomplish which is then passed to the next layer. The first layer of a Neural Network is known as the input layer. The job of an input layer is to acquire data and feed it to the Neural Network. No processing occurs at the input layer. Next to it, are the hidden layers. Hidden layers are the layers in which the whole processing occurs. Their name essentially means that these layers are hidden and are not visible to the user.</p> <p>Each node of these hidden layers has its own machine learning algorithm which it executes on the data received from the input layer. The processed output is then fed to the subsequent hidden layer of the network. There can be multiple hidden layers in a neural network system and their number depends upon the complexity of the function for which the network has been configured. Also, the number of nodes in each layer can vary accordingly. The last hidden layer passes the final processed data to the output layer which then gives it to the user as the final output. Similar to the input layer, output layer too does not process the data which it acquires. It is meant for user-interface. (2.5 marks)</p> <p>Features</p> <ol style="list-style-type: none"> 1. Neural Network systems are modelled on the human brain and nervous system. 2. They are able to automatically extract features without input from the programmer. 3. Every neural network node is essentially a machine learning algorithm. 4. It is useful for solving problems where data set is very large. <p>(any 3 features – 1.5 marks)</p>	4
	Q.18	<p>The 4W's of Problem Scoping are Who, What, Where and Why. These 4Ws helps in identifying and understanding the problem in a better and efficient manner.</p> <p>1.Who - "Who" part helps us in comprehending and categorizing who all are affected directly and indirectly with the problem and who are called the stakeholders.</p> <p>2. What - "What" part helps us in understanding and identifying the nature of the</p>	4

		<p>problem</p> <p>3. Where - "Where" does the problem arises, situation and the location.</p> <p>4. Why - "Why" is the given problem worth solving. It speaks about the benefits which the stakeholders would get from the solution and how it will benefit them as well as the society.</p> <p>(3 marks)</p> <p>The Problem Statement Template is the summary of all the key points of 4W's into one single template.</p> <p>If the same problem occurs in the future, this template helps to fix it easily.</p> <p>(1 mark)</p>	
	Q.19	<ul style="list-style-type: none"> • Kinesthetic Intelligence: It is the ability that is related to how a person uses his limbs in a skilled manner. • Existential Intelligence: It is the intelligence related to religious and spiritual awareness. • Spatial Visual Intelligence: It is the ability to perceive the visual world and the relationship of one object to another. • Intrapersonal Intelligence: It is defined as the ability to realize one's weakness, strength and his own feelings. <p>(4 marks)</p>	4
	Q.20	<p>Text Normalization- In Text Normalization, we undergo several steps to normalize the text to a lower level.</p> <p>1)Sentence Segmentation - Under sentence segmentation, the whole corpus is divided into sentences.Each sentence is taken as a different data so now the whole corpus gets reduced to sentences.</p> <p>2)Tokenisation- After segmenting the sentences, each sentence is then further divided into tokens.Tokens is a term used for any word or number or special character occurring in a sentence. Under tokenisation, every word, number and special character is considered separately and each of them is now a separate token.</p> <p>3)Removing Stop words, Special Characters and Numbers - In this step, the tokens which are not necessary are removed from the token list.</p> <p>4)Converting text to a common case -After the stop words removal, we convert the whole text into a similar case, preferably lower case. This ensures that the case-sensitivity of the machine does not consider same words as different just because of different cases.</p> <p>5)Stemming- In this step, the remaining words are reduced to their root words. In other words, stemming is the process in which the affixes of words are removed and the words are converted to their base form.</p> <p>Lemmatization -in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one.</p>	4

		With this we have normalized our text to tokens which are the simplest form of words present in the corpus. Now it is time to convert the tokens into numbers. For this, we would use the Bag of Words algorithm (4 marks)	
	Q.21	<p>Total no. of wrong predictions made by the model =FP+FN</p> <p style="text-align: right;">=25+5</p> <p style="text-align: right;">=30</p> <p>(1 mark)</p> <p>Precision = TP /(TP+FP)</p> <p style="padding-left: 40px;">=60/(60+25)</p> <p style="padding-left: 40px;">=60/85</p> <p style="padding-left: 40px;">=0.706</p> <p>(1 mark)</p> <p>Recall= TP /(TP+FN)</p> <p style="padding-left: 40px;">=60/(60+05)</p> <p style="padding-left: 40px;">=60/65</p> <p style="padding-left: 40px;">=0.923</p> <p>(1 mark)</p> <p>F1 Score = 2*Precision*Recall/ (Precision+Recall)</p> <p style="padding-left: 40px;">=2*0.706*0.923/(0.706+0.923)</p> <p style="padding-left: 40px;">=1.303/1. 629</p> <p style="padding-left: 40px;">=0.7998</p> <p>(1 mark)</p>	4

SET	B
------------	----------

**INDIAN SCHOOL MUSCAT
FIRST PRE BOARD EXAMINATION 2023
SUBJECT WITH SUBJECT CODE**

CLASS: X

Max.Marks:

MARKING SCHEME			
SET	QN.NO	VALUE POINTS	MARKS SPLIT UP
B	Q. 1		1
	vii.	(a) e-mail	1
	viii.	(c) Online predators	1
	ix.	(b) External	1
	x.	(d) patient	1
	xi.	(d) Life under Sun	1
	xii.	b. Self-motivation	1
	Q.2		
	vii.	c) Both Statements 1 and 2 are correct.	1
	viii.	c) Responsible consumption and Production	1
	ix.	we get a value of 1 (that is 100%) for both Precision and Recall.	1
	x.	$Accuracy = \frac{\text{Correct prediction}}{\text{Total cases}} * 100\%$ $Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$	1

	v.	b) resolution	1
	vi.	b)Data Science	1
	Q.3		
	vi.	(d) the problem can be identified and goals can be set	1
	vii.	Surveys, webscraping,cameras,sensors,observations ,API (any two-1/2 mark each)	1
	viii.	a) Website recommendations	1
	ix.	Automatic Summarization:	1
	x.	d) 1	1
	vi.	c) Clustering	1
	Q.4		
	vii.	AI Bias	1
	viii.	Stemming / Lemmatisation (any one)	1
	ix.	c. Confusion Matrix	1
	x.	The number of pixels in an image is sometimes called the <i>resolution</i>	1
	xi.	c)Both A and R are true and R is the correct explanation of A	1
	xii.	1)True Positive 3) True Negative	1
	Q.5		
	vii.	Recall	1
	viii.	0 to 255	1
	ix.	c. Text Classification (b) Precision	1
	x.	Script bots	1
	xi.	Deep	1
	xii.	The term used for the whole textual data from all the documents altogether is known as corpus.	1

		SECTION B – SUBJECTIVE TYPE QUESTIONS	
	Q. 6	<p>Specific, Measurable, Achievable, Realistic and Time Bound</p> <p>• Specific : A specific and clear goal answers six questions. Who is involved in the goal? What do I want to do? Where do I start? When do I start and finish? Which means do I use? Why am I doing this?</p> <p>Not a specific goal: “I would learn to speak English.”</p> <p>Specific goal: “I would learn to speak English fluently by joining coaching classes after my school everyday, and in six months I will take part in the inter-school debate competition.”</p> <p>Measureable : A measureable goal answers the questions “How much?”, “How many?” and “How do I know that I have achieved results?”</p> <p>Not measurable goal: “I want to be rich.”</p> <p>Measurable goal: “I want to have 5 times more money than what I have today in my hand at the end of this year.”</p> <p>Achievable : Breaking down big goals into smaller parts will make the goal achievable. Bigger Goal: “I want to become a teacher in my school.”</p> <p>Realistic : A realistic goal would be something that we want to achieve and can work towards.</p> <p>Example of unrealistic goal: “I will read my entire year’s syllabus in one day and get good marks.”</p> <p>Realistic goal: “I spend 3 hours every day of the year after school to revise my subjects to get good marks in the exams.”</p> <p>• Time bound : A SMART goal should have a timeframe by when the goal needs to be achieved.</p> <p>This encourages us to take actions to completely fulfill the goals.</p>	2
	Q. 7	<p>Virus- Viruses are computer programs that can damage the data and software programs or steal the information stored on a computer. Major types of viruses are Worms and Trojan Horse. (1 mark)</p> <p>Internet Scams, Sometimes you may receive very attractive offers saying you have won a huge money in a lottery and that you can claim the prize by depositing a certain amount of money. When you deposit the money using credit card or online banking, you not only lose the deposit money but your card / account information may be misused later. (1 mark)</p>	2
	Q. 8	<p>The use of body language, gestures, and other nonverbal signals to convey meaning or messages is known as nonverbal communication. Facial expressions, Eye contact, Gestures, Posture, Touch, (any two)</p>	2
	Q. 9	<p>Organiser of Society’s Productive Resources</p> <p>Helpful in Capital Formation:</p> <p>Increase in Employment Opportunities:</p>	2

		Development of New Production Techniques: Visionary Leader: Contribution of the Execution of Government Policies Higher Productivity (any 4 – 1/2 mark each)	
	Q. 10	1)Quality Education 2) Clean Water and Sanitation 3)Affordable and Clean Energy 4)Decent Work and Economic Growth 5)Reduced Inequalities (Explanation of Any two SDGs – 1 mark each)	2
	Q. 11	bag of words gives us two things: 1. A vocabulary of words for the corpus 2. The frequency of these words (number of times it has occurred in the whole corpus). 1. Text Normalisation: Collect data and pre-process it 2. Create Dictionary: Make a list of all the unique words occurring in the corpus. (Vocabulary) 3. Create document vectors: For each document in the corpus, find out how many times the word from the unique list of words has occurred. 4. Create document vectors for all the documents. (2 mark)	2
	Q.12	Classification Image Classification problem is the task of assigning an input image one label from a fixed set of categories . This is one of the core problems in CV that, despite its simplicity, has a large variety of practical applications. (1 mark) Classification + Localisation This is the task which involves both processes of identifying what object is present in the image and at the same time identifying at what location that object is present in that image. It is used only for single objects. (1 mark)	2
	Q.13	Text classification makes it possible to assign predefined categories to a document and organize it to help you find the information you need or simplify some activities. For example, an application of text categorization is spam filtering in email(1 mark) Sentiment Analysis: The goal of sentiment analysis is to identify sentiment among several posts or even in the same post where emotion is not always explicitly expressed. Companies use Natural Language Processing applications, such as sentiment analysis, to identify opinions and sentiment online to help them understand what customers think about their products and services(1 mark)	2
	Q.14	stemming is the process in which the affixes of words are removed and the words are converted to their base form. Stemming and lemmatization both are alternative processes to each other as the role of both the processes is same – removal of affixes. But the difference between both of them is that in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one. Lemmatization makes sure that lemma is a word with meaning and hence it takes a longer time to execute than stemming.	2

		<div><div>Caring</div><div>Lemmatization</div><div>Care</div></div> <div><div>Caring</div><div>Stemming</div><div>Car</div></div> <div>(2 marks)</div>											
	Q.15	<p>Classification: Where the data is classified according to the labels. For example, in the grading system, students are classified on the basis of the grades they obtain with respect to their marks in the examination. This model works on discrete dataset which means the data need not be continuous. (1 mark)</p> <p>Regression: Such models work on continuous data. For example, if you wish to predict your next salary, then you would put in the data of your previous salary, any increments, etc., and would train the model. Here, the data which has been fed to the machine is continuous.(1 mark)</p>	2										
	Q.16	<p>Underfit -The model’s output does not match the true function at all. Hence the model is said to be underfitting and its accuracy is lower. (1 mark)</p> <p>Overfit -model performance is trying to cover all the data samples even if they are out of alignment to the true function. This model is said to be overfitting and this too has a lower accuracy.(1 mark)</p>	2										
	Q.17	<table><tr><td>Problem Scoping</td><td>Understanding the problem</td></tr><tr><td>Data Acquisition</td><td>Collecting accurate and reliable data</td></tr><tr><td>Data Exploration</td><td>Arranging the data uniformly</td></tr><tr><td>Modelling</td><td>Creating Models from the data</td></tr><tr><td>Evaluation</td><td>Evaluating the project</td></tr></table> <p>The Problem Statement Template is the summary of all the key points of 4W’s into one single template. If the same problem occurs in the future, this template helps to fix it easily.</p> <p>(1 mark)</p>	Problem Scoping	Understanding the problem	Data Acquisition	Collecting accurate and reliable data	Data Exploration	Arranging the data uniformly	Modelling	Creating Models from the data	Evaluation	Evaluating the project	4
Problem Scoping	Understanding the problem												
Data Acquisition	Collecting accurate and reliable data												
Data Exploration	Arranging the data uniformly												
Modelling	Creating Models from the data												
Evaluation	Evaluating the project												
	Q.18	<p>Text Normalizationin Text Normalization, we undergo several steps to normalize the text to a lower level.</p> <p>1)Sentence Segmentation - Under sentence segmentation, the whole corpus is divided into sentences.Each sentence is taken as a different data so now the whole corpus gets reduced to sentences.</p> <p>2)Tokenisation- After segmenting the sentences, each sentence is then further divided into tokens.Tokens is a term used for any word or number or special character occurring in a sentence. Under tokenisation, every word, number and special character is considered separately and each of them is</p>	4										

		<p>now a separate token.</p> <p>3)Removing Stop words, Special Characters and Numbers - In this step, the tokens which are not necessary are removed from the token list.</p> <p>4)Converting text to a common case -After the stop words removal, we convert the whole text into a similar case, preferably lower case. This ensures that the case-sensitivity of the machine does not consider same words as different just because of different cases.</p> <p>5)Stemming- In this step, the remaining words are reduced to their root words. In other words, stemming is the process in which the affixes of words are removed and the words are converted to their base form.</p> <p>Lemmatization -in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one.</p> <p>With this we have normalized our text to tokens which are the simplest form of words present in the corpus. Now it is time to convert the tokens into numbers. For this, we would use the Bag of Words algorithm (4 marks)</p>	
	Q.19	<ul style="list-style-type: none"> • Kinesthetic Intelligence: It is the ability that is related to how a person uses his limbs in a skilled manner. • Linguistic Intelligence: It is the Language processing skills both in terms of understanding or implementation in writing or speech. • Existential Intelligence: It is the intelligence related to religious and spiritual awareness. • Spatial Visual Intelligence: It is the ability to perceive the visual world and the relationship of one object to another. <p>(4 marks)</p>	4
	Q.20	<p>A Neural Network is divided into multiple layers and each layer is further divided into several blocks called nodes.</p> <p>Each node has its own task to accomplish which is then passed to the next layer. The first layer of a Neural Network is known as the input layer. The job of an input layer is to acquire data and feed it to the Neural Network. No processing occurs at the input layer. Next to it, are the hidden layers. Hidden layers are the layers in which the whole processing occurs. Their name essentially means that these layers are hidden and are not visible to the user.</p> <p>Each node of these hidden layers has its own machine learning algorithm which it executes on the data received from the input layer. The processed output is then fed to the subsequent hidden layer of the network. There can be multiple hidden layers in a neural network system and their number depends upon the complexity of the function for which the network has been configured. Also, the number of nodes in each layer can vary accordingly. The last hidden layer passes the final processed data to the output layer which then gives it to the user as the final output. Similar to the input layer, output layer too does not process the data which it acquires. It is meant for user-interface. (2.5 marks)</p> <p>Features</p> <ol style="list-style-type: none"> 5. Neural Network systems are modelled on the human brain and nervous system. 6. They are able to automatically extract features without input from the programmer. 7. Every neural network node is essentially a machine learning 	4

		<p>algorithm.</p> <p>8. It is useful for solving problems where data set is very large.</p> <p>(any 3 features – 1.5 marks)</p>	
	Q.21	<p>Total no. of correct predictions made by the model =TP+TN</p> <p style="text-align: right;">=50+20</p> <p style="text-align: right;">=70</p> <p>(1 mark)</p> <p>Precision = TP /(TP+FP)</p> <p style="text-align: right;">=50/(50+20)</p> <p style="text-align: right;">=50/70</p> <p style="text-align: right;">=0.714</p> <p>(1 mark)</p> <p>Recall= TP /(TP+FN)</p> <p style="text-align: right;">=50/(50+10)</p> <p style="text-align: right;">=50/60</p> <p style="text-align: right;">=0.833</p> <p>(1 mark)</p> <p>F1 Score = 2*Precision*Recall/ (Precision+Recall)</p> <p style="text-align: right;">=2*0.714*0.833/(0.714+0.833)</p> <p style="text-align: right;">=1.189/1. 547</p> <p style="text-align: right;">=0.768</p> <p>(1 mark)</p>	4

**INDIAN SCHOOL MUSCAT
FIRST PRE BOARD EXAMINATION 2023
SUBJECT WITH SUBJECT CODE**

CLASS: X/ XII

Max.Marks:

				SET	C
MARKING SCHEME					
SET	QN.NO	VALUE POINTS			MARKS SPLIT UP
C	Q. 1				
	xiii.	c. Specific , Measurable, Achievable, Realistic, Time bound			1
	xiv.	(d) Trojan Horse			1
	xv.	c. Self-motivation			1
	xvi.	(c) patient			1
	xvii.	(b) Population goals			1
	xviii.	(b) External			1
	Q.2				
	xi.	d) Partnership for the Goals			1
	xii.	we get a value of 1 (that is 100%) for both Precision and Recall.			1
	xiii.	a) Both Statements 1 and 2 are correct.			1
	xiv.	c)Data Science			1
	v.	b) resolution			1

	vi.	$Accuracy = \frac{Correct\ prediction}{Total\ cases} * 100\%$ $Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$	1
	Q.3		
	xi.	(a) the problem can be identified and goals can be set	1
	xii.	Surveys, webscraping, cameras, sensors, observations ,API (any two-1/2 mark each)	1
	xiii.	b) Website recommendations	1
	xiv.	Automatic Summarization:	1
	xv.	a) 1	1
	vi.	b) Classification	1
	Q.4		
	xiii.	The number of pixels in an image is sometimes called the <i>resolution</i> Red (R), Green (G) and Blue (B)	1
	xiv.	c) Both A and R are true and R is the correct explanation of A b)	1
	xv.	a) AI Bias	1
	xvi.	b. Confusion Matrix	1
	xvii.	Stemming / Lemmatisation	1
	xviii.	b) False Positive d) False Negative	1
	Q.5		
	xiii.	c. Text Classification	1
	xiv.	Grayscale images are images which have a range of shades of gray without apparent colour.	1
	xv.	(b) Precision	1
	xvi.	Smart bots	1
	xvii.	kineasthetic intelligence	1

	xviii.	The term used for the whole textual data from all the documents altogether is known as corpus.	1
		SECTION B – SUBJECTIVE TYPE QUESTIONS	
	Q. 6	<p>Worms: These are viruses that replicate themselves and spread to all files once they attack a computer. This makes it very difficult to remove them.</p> <p>Trojan Horse: A Trojan Horse disguises itself i.e., it appears to be a useful software program but once it reaches a computer it starts behaving like a virus and destroying data.</p>	2
	Q. 7	1. Know what they want from life 2. Are focussed 3. Know what is important 4. Are dedicated to fulfill their dreams	2
	Q. 8	In the verbal communication word and language is used to transmit a message. Eg., The verbal communication can done through speaking or writing and can occur in various settings, including face-to-face conversations, phone calls, and video chats (any two)	2
	Q. 9	Organiser of Society's Productive Resources Helpful in Capital Formation: Increase in Employment Opportunities: Development of New Production Techniques: Visionary Leader: Contribution of the Execution of Government Policies Higher Productivity (any 4 – 1/2 mark each)	2
	Q. 10	1) Quality Education 2) Clean Water and Sanitation 3) Affordable and Clean Energy 4) Decent Work and Economic Growth 5) Reduced Inequalities (Explanation of Any two SDGs – 1 mark each)	2
	Q. 11	<p>Underfit -The model's output does not match the true function at all. Hence the model is said to be underfitting and its accuracy is lower. (1 mark)</p> <p>Overfit -model performance is trying to cover all the data samples even if they are out of alignment to the true function. This model is said to be overfitting and this too has a lower accuracy.(1 mark)</p>	2
	Q.12	<p>Classification</p> <p>Image Classification problem is the task of assigning an input image one label from a fixed set of categories. This is one of the core problems in CV that, despite its simplicity, has a large variety of practical applications. (1 mark)</p> <p>Classification + Localisation</p>	2

		This is the task which involves both processes of identifying what object is present in the image and at the same time identifying at what location that object is present in that image. It is used only for single objects. (1 mark)											
	Q.13	Automatic summarization is relevant not only for summarizing the meaning of documents and information, but also to understand the emotional meanings within the information, such as in collecting data from social media. Automatic summarization is especially relevant when used to provide an overview of a news item or blog post, while avoiding redundancy from multiple sources and maximizing the diversity of content obtained.(1 mark) Text classification makes it possible to assign predefined categories to a document and organize it to help you find the information you need or simplify some activities. For example, an application of text categorization is spam filtering in email(1 mark)	2										
	Q.14	stemming is the process in which the affixes of words are removed and the words are converted to their base form. Stemming and lemmatization both are alternative processes to each other as the role of both the processes is same – removal of affixes. But the difference between both of them is that in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one. Lemmatization makes sure that lemma is a word with meaning and hence it takes a longer time to execute than stemming. <div><div>Caring</div><div>Lemmatization</div><div>Care</div></div> <div><div>Caring</div><div>Stemming</div><div>Car</div></div> (2 marks)	2										
	Q.15	<table><tr><td>Problem Scoping</td><td>Understanding the problem</td></tr><tr><td>Data Acquisition</td><td>Collecting accurate and reliable data</td></tr><tr><td>Data Exploration</td><td>Arranging the data uniformly</td></tr><tr><td>Modelling</td><td>Creating Models from the data</td></tr><tr><td>Evaluation</td><td>Evaluating the project</td></tr></table>	Problem Scoping	Understanding the problem	Data Acquisition	Collecting accurate and reliable data	Data Exploration	Arranging the data uniformly	Modelling	Creating Models from the data	Evaluation	Evaluating the project	2
Problem Scoping	Understanding the problem												
Data Acquisition	Collecting accurate and reliable data												
Data Exploration	Arranging the data uniformly												
Modelling	Creating Models from the data												
Evaluation	Evaluating the project												
	Q.16	bag of words gives us two things: 1. A vocabulary of words for the corpus 2. The frequency of these words (number of times it has occurred in the whole corpus). 1. Text Normalisation: Collect data and pre-process it 2. Create Dictionary: Make a list of all the unique words occurring in the corpus. (Vocabulary) 3. Create document vectors: For each document in the corpus, find out how many times the word from the unique list of words has occurred. 4. Create document vectors for all the documents. (2 mark)	2										

Q.17	<ul style="list-style-type: none"> • Intrapersonal Intelligence: It is defined as the ability to realize one's weakness, strength and his own feelings. • Linguistic Intelligence: It is the Language processing skills both in terms of understanding or implementation in writing or speech. • Existential Intelligence: It is the intelligence related to religious and spiritual awareness. • Spatial Visual Intelligence: It is the ability to perceive the visual world and the relationship of one object to another. <p>(4 marks)</p>	4
Q.18	<p>Text Normalizationin Text Normalization, we undergo several steps to normalize the text to a lower level.</p> <p>1)Sentence Segmentation - Under sentence segmentation, the whole corpus is divided into sentences.Each sentence is taken as a different data so now the whole corpus gets reduced to sentences.</p> <p>2)Tokenisation- After segmenting the sentences, each sentence is then further divided into tokens.Tokens is a term used for any word or number or special character occurring in a sentence. Under tokenisation, every word, number and special character is considered separately and each of them is now a separate token.</p> <p>3)Removing Stop words, Special Characters and Numbers - In this step, the tokens which are not necessary are removed from the token list.</p> <p>4)Converting text to a common case -After the stop words removal, we convert the whole text into a similar case, preferably lower case. This ensures that the case-sensitivity of the machine does not consider same words as different just because of different cases.</p> <p>5)Stemming- In this step, the remaining words are reduced to their root words. In other words, stemming is the process in which the affixes of words are removed and the words are converted to their base form.</p> <p>Lemmatization -in lemmatization, the word we get after affix removal (also known as lemma) is a meaningful one.</p> <p>With this we have normalized our text to tokens which are the simplest form of words present in the corpus. Now it is time to convert the tokens into numbers. For this, we would use the Bag of Words algorithm (4 marks)</p>	4
Q.19	<p><u>Supervised Learning : 2 types</u></p> <p>Classification: Where the data is classified according to the labels. For example, in the grading system, students are classified on the basis of the grades they obtain with respect to their marks in the examination. This model works on discrete dataset which means the data need not be continuous.(1mark)</p> <p>Regression: Such models work on continuous data. For example, if you wish to predict your next salary, then you would put in the data of your previous salary, any increments, etc., and would train the model. Here, the data which has been fed to the machine is continuous.</p> <p>(1 mark)</p> <p><u>Unsupervised Learning : 2types</u></p> <p>Clustering: Refers to the unsupervised learning algorithm which can cluster the unknown data according to the patterns or trends identified out of it. The patterns observed might be the ones which are known to the developer or it might even come up with some unique patterns out of it. (1 mark)</p>	4

		<p>Dimensionality Reduction: We humans are able to visualise upto 3-Dimensions only but according to a lot of theories and algorithms, there are various entities which exist beyond 3-Dimensions. For example, in Natural language Processing, the words are considered to be N-Dimensional entities. Which means that we cannot visualise them as they exist beyond our visualisation ability. Hence, to make sense out of it, we need to reduce their dimensions. Here, dimensionality reduction algorithm is used.(1 mark)</p>	
	Q.20	<p>A Neural Network is divided into multiple layers and each layer is further divided into several blocks called nodes.</p> <p>Each node has its own task to accomplish which is then passed to the next layer.</p> <p>The first layer of a Neural Network is known as the input layer. The job of an input layer is to acquire data and feed it to the Neural Network. No processing occurs at the input layer. Next to it, are the hidden layers. Hidden layers are the layers in which the whole processing occurs. Their name essentially means that these layers are hidden and are not visible to the user.</p> <p>Each node of these hidden layers has its own machine learning algorithm which it executes on the data received from the input layer. The processed output is then fed to the subsequent hidden layer of the network. There can be multiple hidden layers in a neural network system and their number depends upon the complexity of the function for which the network has been configured. Also, the number of nodes in each layer can vary accordingly. The last hidden layer passes the final processed data to the output layer which then gives it to the user as the final output. Similar to the input layer, output layer too does not process the data which it acquires. It is meant for user-interface. (2.5 marks)</p> <p>Features</p> <ol style="list-style-type: none"> 9. Neural Network systems are modelled on the human brain and nervous system. 10. They are able to automatically extract features without input from the programmer. 11. Every neural network node is essentially a machine learning algorithm. 12. It is useful for solving problems where data set is very large. <p>(any 3 features – 1.5 marks)</p>	4
	Q.21	<p>Total no. of wrong predictions made by the model =FP+FN</p> <p style="text-align: center;">=5+10</p> <p style="text-align: center;">=15</p> <p>(1 mark)</p> <p>Precision = $TP / (TP+FP)$</p> <p style="text-align: center;">$=70/(70+5)$</p> <p style="text-align: center;">$=70/75$</p> <p style="text-align: center;">$=0.933$</p> <p>(1 mark)</p> <p>Recall= $TP / (TP+FN)$</p> <p style="text-align: center;">$=70/(70+10)$</p> <p style="text-align: center;">$=70/80$</p>	4

		<p>=0.875</p> <p>(1 mark)</p> <p>F1 Score = $2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$</p> <p>= $2 * 0.933 * 0.875 / (0.933 + 0.875)$</p> <p>= $1.632 / 1.808$</p> <p>=0.902</p> <p>(1 mark)</p>	
--	--	--	--