CS607 - Final Term Papers Fall 2012

2 mark Questions:

1. What is hypothesis space?

Answer:-

The hypothesis space used by a machine learning system is the set of all hypotheses that might possibly be returned by it. It is typically dened by a hypothesis language, possibly in conjunction with a language bias.

2. What is machine learning?

Answer:- (Page 160)

Generally speaking, the goal of machine learning is to build computer systems that can learn from their experience and adapt to their environments. Obviously, learning is an important aspect or component of intelligence.

3. What is a strip?

Answer:- (Page 197)

STRIP is one of the founding languages developed particularly for planning. Let us understand planning to a better level by seeing what a planning language can represent.

4. What is differing between height and depth of a tree?

Answer:-

Height and depth of a tree is equal but height and depth of a node is not equal because the height is calculated by traversing from leaf to the given node depth is calculated from traversal from root to the given node.

4 questions of 3 marks

2. What do you know about Mamdani's fuzzy inference method?

Answer:- (Page 153)

Mamdani's fuzzy inference method is the most commonly seen fuzzy methodology. Mamdani's method as among the first control systems built using fuzzy set theory. It was proposed in 1975 by Ebrahim Mamdani as an attempt to control a steam engine and boiler combination by synthesizing a set of linguistic control rules obtained from experienced human operators. Mamdani's effort was based on Lotfi Zadeh's 1973 paper on fuzzy algorithms for complex systems and decision processes.

اللہ کاخوف سب سے بڑی دانائی ہے

3. What is the basic key of candidate elimination algorithm?
Answer:- (Page 173)

The key idea in Candidate-Elimination algorithm is to output a description of the set of all hypotheses consistent with the training examples. This subset of all hypotheses is actually the version space with respect to the hypothesis space H and the training examples D, because it contains all possible versions of the target concept.

4. Predicate action?
Answer:- (Page 198)

Action is a predicate used to change states. It has three components namely, the predicate itself, the pre-condition, and post-condition predicates. For example, the action to buy something item can be represented as,
Action:
buy(X)

Pre-conditions:
at(Place) _ sells(Place, X)

Post-conditions/Effect:
have(X)

4 questions of 5 marks.
   1. **Differ between supervised and unsupervised learning? 5 marks**
 Answer:-

Unsupervised Learning
• The model is not provided with the correct results during the training.
• Can be used to cluster the input data in classes on the basis of their sta?s?cal proper?es only.
• Cluster significance and labeling.
• The labeling can be carried out even if the labels are only available for a small number of objects representative of the desired classes.

Supervised Learning
• Training data includes both the input and the desired results.
• For some examples the correct results (targets) are known and are given in input to the model during the learning process.
• The construction of a proper training, validation and test set (Bok) is crucial.

• These methods are usually fast and accurate.
• Have to be able to generalize: give the correct results when new data are given in input without knowing a priori the target.

دنیامیں سب سے مشکل کام اپنی اصلاح اور سب سے آسان کام دوسروں پر نکتہ چینی کرنا ہے

Muhammad Moaaz Siddiq – MCS(4th)
Moaaz.pk@gmail.com
Campus:- Institute of E-Learning & Moderen Studies
**(IEMS) Samundari**

2. What are the linear separable line and non linear separable line?5
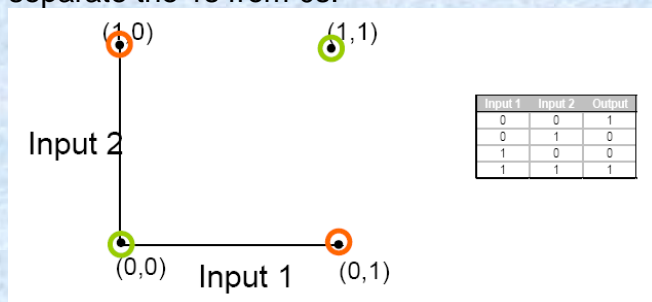
Answer:- (Page 184)

There is a whole class of problems which are termed as linearly separable. This name is given to them, because if we were to represent them in the input space, we could classify them using a straight line. The simplest examples are the logical AND or OR. We have drawn them in their input spaces, as this is a simple 2D problem. The upper sloping line in the diagram shows the decision boundary for AND gate, above which, the output is 1, below is 0. The lower sloping line decides for the OR gate similarly.

Input 1	Input 2	AND
0	0	0
0	1	0
1	0	0
1	1	1

Input 1	Input 2	OR
0	0	0
0	1	1
1	0	1
1	1	1
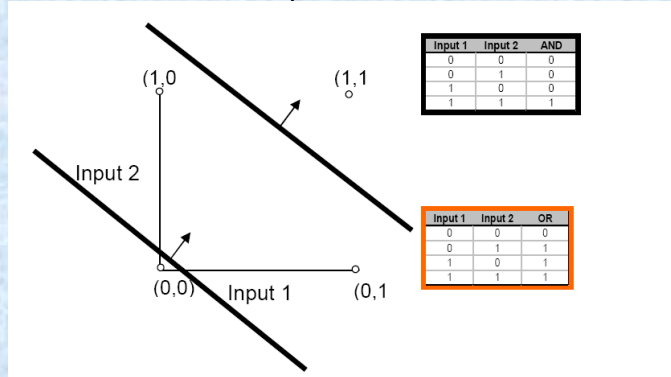
A single perceptron simply draws a line, which is a hyper plane when the data is more then 2 dimensional. Sometimes there are complex problems (as is the case in real life). The data for these problems cannot be separated into their respective classes by using a single straight line. These problems are not linearly separable.

Linearly Separable

Linearly Non Separable /
Non linear decision region
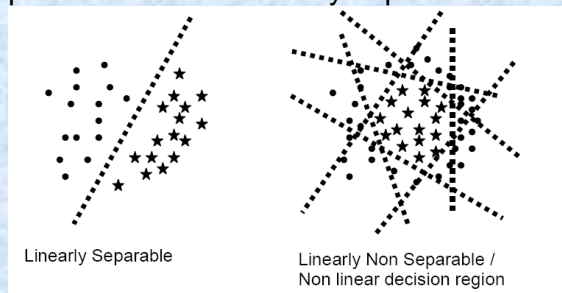
Another example of linearly non-separable problems is the XOR gate (exclusive OR). This shows how such a small data of just 4 rows, can make it impossible to draw one line decision boundary, which can separate the 1s from 0s.

Input 1	Input 2	Output
0	0	1
0	1	0
1	0	0
1	1	1

Muhammad Moaaz Siddiq – MCS(4th)
Moaaz.pk@gmail.com
Campus:- Institute of E-Learning & Moderen Studies
(IEMS) Samundari

3. Generic Algorithms?

Answer:- (Page 77)

The genetic algorithm technology comes from the concept of human evolution. Genetic Algorithms is a search method in which multiple search paths are followed in parallel. At each step, current states of different pairs of these paths are combined to form new paths. This way the search paths don't remain independent, instead they share information with each other and thus try to improve the overall performance of the complete search space.

4. Write different knowledge Acquisition techniques.

Answer:- (Page 184)

• Knowledge elicitation by interview
• Brainstorming session with one or more experts. Try to introduce some structure to this session by defining the problem at hand, prompting for ideas and looking for converging lines of thought.
• Electronic brainstorming
• On-site observation
• Documented organizational expertise, e.g. troubleshooting manuals

CS607 - Final Term Papers Fall 2012

2 marks.

Out of different general components of and AI cycle which two components are best closed coupled.

Answer:- (Page 89)

Knowledge representation (KR) and reasoning are closely coupled components; each is intrinsically tied to the other.

How can you elaborate the condition predicate in STRIPS.

Answer:- (Page 197)

Condition predicates are the predicates that define states. For example, a predicate that specifies that we are at location 'X' is given as.
at(X)

Define deductive learning.

Answer:- (Page 162)

Deductive learning works on existing facts and knowledge and deduces new knowledge from the old.
This is best illustrated by giving an example. For example, assume:
A = B
B = C
Then we can deduce with much confidence that:
C = A

3 Marks

Difference between fuzzy logic and Crisp value.

Answer:-

The input variables in a fuzzy control system are in general mapped by sets of membership functions similar to this, known as "fuzzy sets". The process of converting a crisp input value to a fuzzy value is called "fuzzification".

Who does neural network resemble the human brain.

Answer:- (Page 187)

It resembles the brain in two respects:
• Knowledge is acquired by the network through a learning process (called training)
• Interneuron connection strengths known as synaptic weights are used to store the knowledge

How can you elaborate the importance and need of hypothesis space.

Answer:- Rep

Write down fuzzy statement in everyday life.

Answer:-

Fuzziness can be found in many areas of daily life, such as in engineering, medicine, meteorology, manufacturing and others. it is particularly frequent in all areas in which human judgment, evaluation and decision are important. These are the areas of decision making, reasoning, and learning and so on.

5 Marks.

Write down the task for which connectionist approach is well suited.

Answer:- (Page 181)

Tasks for which connectionist approach is well suited include:
• Classification
    • Fruits – Apple or orange

• Pattern Recognition
    • Finger print, Face recognition

• Prediction
    • Stock market analysis, weather forecast

How can you elaborate the conflict resolution strategies?

Answer:- (Page 125)

To overcome the conflict problem stated above, we may choose to use on of the following conflict resolution strategies:
• Fire first rule in sequence (rule ordering in list). Using this strategy all the rules in the list are ordered (the ordering imposes prioritization). When more than one rule matches, we simply fire the first in the sequence
• Assign rule priorities (rule ordering by importance). Using this approach we assign explicit priorities to rules to allow conflict resolution.

• More specific rules (more premises) are preferred over general rules. This strategy is based on the observation that a rule with more premises, in a sense, more evidence or votes from its premises, therefore it should be fired in preference to a rule that has less premises.
• Prefer rules whose premises were added more recently to WM (timestamping). This allows prioritizing recently added facts over older facts.
• Parallel Strategy (view-points). Using this strategy, we do not actually resolve the conflict by selecting one rule to fire. Instead, we branch out our execution into a tree, with each branch operation in parallel on multiple threads of reasoning. This allows us to maintain multiple view-points on the argument concurrently

Why Genetic algorithm is used in the biological theory of Human evolution? Also write its functions.
Answer:- Rep

CS607 - Final Term Papers Fall 2012

What is Mean Squared error? 2 marks
Answer:- (Page 189)
The most common technique for measuring the total error in each iteration of the neural network (epoch) is Mean Squared Error (MSE).

What is sate in STRIP? Give Example? 2 marks
Answer:- (Page 197)
State is a conjunction of predicates represented in well-known form, for example, a state where we are at the hotel and do not have either cash or radio is represented as,

```
at(hotel) ∧ ¬have(cash) ∧ ¬have(radio)
```

What is the basic key of candidate elimination algorithm? 3 marks
Answer:- Rep

Differ between supervised and unsupervised learning? 5 marks
Answer:- Rep

Five parts of fuzzy inference process? 5 marks
Answer:- (Page 154)
• Fuzzification of the input variables
• Application of fuzzy operator in the antecedent (premises)
• Implication from antecedent to consequent
• Aggregation of consequents across the rules
• Defuzzification of output

بری صحبت سے تنہائی بہتر ہے اور تنہائی سے نیک صحبت بہتر ہے

Find S, candidate elimination and decision tree representation, from these algorithms which algorithm is best? 5 marks

Answer:- (Page 172)

FIND-S finds the maximally specific hypothesis possible within the version space given a set of training data. How can we use the general to specific ordering of hypothesis space to organize the search for a hypothesis consistent with the observed training examples? One way is to begin with the most specific possible hypothesis in H, then generalize the hypothesis each time it fails to cover an observed positive training example. (We say that a hypothesis "covers" a positive example if it correctly classifies the example as positive.)

The key idea in Candidate-Elimination algorithm is to output a description of the set of all hypotheses consistent with the training examples. This subset of all hypotheses is actually the version space with respect to the hypothesis space H and the training examples D, because it contains all possible versions of the target concept.

The Candidate-Elimination algorithm represents the version space by storing only its most general members (denoted by G) and its most specific members (denoted by S). Given only these two sets S and G, it is possible to enumerate all members of the version space as needed by generating the hypotheses that lie between these two sets in general-to-specific partial ordering over hypotheses.

Decision trees give us disjunctions of conjunctions, that is, they have the form:
(A AND B) OR (C AND D)
Where A, B, C and D are the attributes for the problem. This tree gives a positive output if either A AND B attributes are present in the instance; OR C AND D attributes are present. Through decision trees, this is how we reach the final hypothesis. This is a hypothetical tree. In real problems, every tree has to have a root node. There are various algorithms like ID3 and C4.5 to find decision trees for learning problems.

9 .what is usage of math and biology in AI? 2marks

Answer:- (Page 9)

From over 400 years of mathematics we have formal theories of logic, probability, decision-making and computation.
From biology we have information about the network structure of a human brain and all the theories on functionalities of different human organs.

10. What is validation and training? 2 Marks

Answer:- (Page 160)

Training: a training set of examples of correct behavior is analyzed and some representation of the newly learnt knowledge is stored. This is often some form of rules.

Validation: the rules are checked and, if necessary, additional training is given. Sometimes additional test data are used, but instead of using a human to validate the rules, some other automatic knowledge based component may be used. The role of tester is often called the critic.

ایماندار کو غصہ دیر سے آتا ہے اور جلدی دور ہو جاتا ہے

12 .steps of ESDLC? 5 Marks.
Answer:- (Page 129)
• Feasibility study
• Rapid prototyping
• Alpha system (in-house verification)
• Beta system (tested by users)
• Maintenance and evolution

13. Difference between supervised and unsupervised learning? 5 Marks.
Answer:- Rep

14 .Linearly separable and linearly non separable difference with examples? 5 Marks
Answer:- Rep

15 .steps of back ward chaining? 5 Marks.
Answer:- (Page 126)
1. Start with the goal.
2. Goal may be in WM initially, so check and you are done if found!
3. If not, then search for goal in the THEN part of the rules (match conclusions, rather than premises). This type of rule is called goal rule.
4. Check to see if the goal rule's premises are listed in the working memory.
5. Premises not listed become sub-goals to prove.
6. Process continues in a recursive fashion until a premise is found that is not supported by a rule, i.e. a premise is called a primitive, if it cannot be concluded by any rule
7. When a primitive is found, ask user for information about it. Back track and use this information to prove sub-goals and subsequently the goal.

زندگی میں کامیابی کا یہی راز ہے کہ پریشانیوں سے پریشان مت بنو

Muhammad Moaaz Siddiq – MCS(4th)
Moaaz.pk@gmail.com
Campus:- Institute of E-Learning & Moderen Studies (IEMS) Samundari

CS607 - Final Term Papers Spring 2012

1- Advantages of neural network
Answer:- (Page 187)
• Excellent for pattern recognition
• Excellent classifiers
• Handles noisy data well
• Good for generalization

2- what is the difference between knowledge analysis and evaluation function
Answer:- (Page 79 & 131)
An evaluation function is the criteria that check various individuals/ solutions for being better than others in the population. Notice that mutation can be as simple as just flipping a bit at random or any number of bits.
The goal of knowledge analysis is to analyze and structure the knowledge gained during the knowledge acquisition phase.

3- Stages of life cycle of ESDLC
Answer:- Rep

4- steps of knowledge elicitation

5- FIS applications
Answer:- (Page 153)
Fuzzy inference system (FIS) is the process of formulating the mapping from a given input to an output using fuzzy logic. This mapping then provides a basis from which decisions can be made, or patterns discerned
Fuzzy inference systems have been successfully applied in fields such as automatic control, data classification, decision analysis, expert systems, and computer vision. Because of its multidisciplinary nature, fuzzy inference systems are associated with a number of names, such as fuzzy-rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and simply (and ambiguously !!) fuzzy systems. Since the terms used to describe the various parts of the fuzzy inference process are far from standard, we will try to be as clear as possible about the different terms introduced in this section.

6- Difference between supervised and unsupervised learning
Answer:- Rep

8- CNF conversion
Answer:- (Page 107)
Resolution requires all sentences to be converted into a special form called conjunctive normal form (CNF). A statement in conjunctive normal form (CNF) consists of ANDs of Ors.

9

CS607 - Final Term Papers Spring 2012

What is the prime role of expert system designer?   2 marks

Answer:- (Page 117)

One of the prime roles of the expert system designer is to act as a knowledge engineer. As a knowledge engineer, the designer must overcome the knowledge acquisition bottleneck and find an effective way to get information from the expert and encode it in the knowledge base, using one of the knowledge representation techniques we discussed in KRR.

What is inductive learning?   2 marks

Answer:- (Page 162)

Inductive learning takes examples and generalizes rather than starting with existing knowledge. For example, having seen many cats, all of which have tails, one might conclude that all cats have tails.

Among Entropy and information gain which is the best to measure effectiveness? 2 marks

Answer:- (Page 177)

information gain

Why clustering algorithm are categorized as unsupervised? 2 marks

Answer:- (Page 190)

• A cluster represents some specific underlying patterns in the data

• Useful for finding patterns in large data sets

Why we use Candidate Elimination Algorithm among many other methods? 3 marks

Answer:- Rep

Among Mutation, crossover and combination of both which is the best strategy to get our desired result sets? 3 marks

Answer:- (Page 82)

The best technique in general turns out to be a combination of both, i.e., crossover with mutation.

دنیا کی سب سے بڑی فتح نفس پر قابو رکھنا ہے

Which term is best suited for a function that is used specially to drive out the degree of truth in fuzzy sets? Write its two characteristics. 3 marks

Answer:- (Page 149)

The degree of truth that we have been talking about, is specifically driven out by a function called the membership function.

Some characteristics of the membership functions are:
• It is represented by the Greek symbol 1
• Truth values range between 0.0 and 1.0
    o Where 0.0 normally represents absolute falseness

    o And 1.0 represent absolute truth

What is Clustering? 3 marks

Answer:- (Page 205)

Clustering is a form of unsupervised learning, in which the training data is available but without the classification information or class labels. The task of clustering is to identify and group similar individual data elements based on some measure of similarity.

Five parts of fuzzy inference process?   5 marks

Answer:- Rep

Does Alpha Beta pruning evaluate each and every node in the tree? 5 marks

Answer:- (Page 64)

All nodes/situations have not yet been previously evaluated for their static evaluation score. Only two leaf nodes have been evaluated so far.
It is a procedure that reduces both the tree branches that must be generated and the number of evaluations. This procedure is called Alpha Beta pruning which "prunes" the tree branches thus reducing the number of static evaluations.

In a situation we have four neurons in a hidden layer then how many lines should we draw in order to classify the data in a linearly non separable problem?   5 marks

Answer:- (Page 64)

Multiple layer perceptrons achieve this task by the introduction of one or more hidden layers. Each neuron in the hidden layer is responsible for a different line. Together they form a classification for the given problem.

جھوٹ انسان اور ایمان دونوں کا دشمن ہے
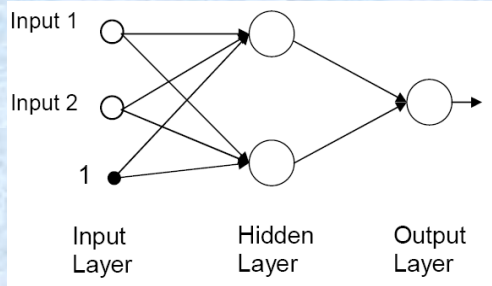
Each neuron in the hidden layer forms a different decision line. Together all the lines can construct any arbitrary non-linear decision boundaries. These multilayer perceptrons are the most basic artificial neural networks.

Genetic Algorithm is inspired by the structure and/or functional aspects of the biological neural networks and it consist of an interconnected group of artificial neurons. Do you agree or not? 5 marks

Answer:- Rep

CS607 - Final Term Papers Spring 2012

define induction learning 2 marks

Answer:- Rep

Define strips   2marks

Answer:- Rep

Define vision space   2 marks

Answer:- (Page 170)

Version space is a set of all the hypotheses that are consistent with all the training examples. When we are given a set of training examples D, it is possible that there might be more than one hypotheses from the hypothesis space that are consistent with all the training examples.

What is machine learning   3 marks

Answer:- Rep

عقل مند کہتا ہے میں کچھ نہیں جانتا جبکہ بے وقوف کہتا ہے کہ میں سب کچھ جانتا ہوں

How expert system work   3 marks

Answer:- (Page 114)

Expert systems may be used in a host of application areas including diagnosis, interpretation, prescription, design, planning, control, instruction, prediction and simulation.

How to lead machine learning in developing the expert system 5 marks

Answer:- (Page 163)

Many AI applications are built with rich domain knowledge and hence do not make use of machine learning. To build such expert systems, it is critical to capture knowledge from experts. However, the fundamental problem remains unresolved, in the sense that things that are normally implicit inside the expert's head must be made explicit. This is not always easy as the experts may find it hard to say what rules they use to assess a situation but they can always tell you what factors they take into account. This is where machine learning mechanism could help. A machine learning program can take descriptions of situations couched in terms of these factors and then infer rules that match expert's behavior.

Write disadvantages of Artificial Neural Networks. 5marks

Answer:- (Page 187)

• The power of ANNs lie in their parallel architecture
  – Unfortunately, most machines we have are serial (Von Neumann architecture)
• Lack of defined rules to build a neural network for a specific problem
  – Too many variables, for instance, the learning algorithm, number of neurons per layer, number of layers, data representation etc
• Knowledge is implicit

• Data dependency

CS607 - Final Term Papers Spring 2012

BIKE is heavy is this statement contain uncertain facts?
Answer:- (Page 94)
It is fuzzy fact.
Fuzzy facts are ambiguous in nature, e.g. the book is heavy/light. Here it is unclear what heavy means because it is a subjective description. Fuzzy representation is used for such facts. While defining fuzzy facts, we use certainty factor values to specify value of "truth". We will look at fuzzy representation in more detail later.

<div dir="rtl">خود کو تمھیں سے بڑھ کر کوئی اچھا مشورہ نہیں دے سکتا</div>

which measurement issued as tool error for the stopping criteria in neural network for epochs
Answer:- (Page 190)
• Done through MSE. We define a low threshold usually 0.01, which if reached stops the training data.
• Another stopping criterion is the number of epochs, which defines how many maximum times the data can be presented to the network for learning.

in softcomputing the most combinations are of
  I) genetic -fuzzy system
  II) n eural-fuzzy system
  III) genetic algorithm -neural network

 the question was identify and describe which combination is best suited   5 marks
Answer:- (Page 205)
Softcomputing is naturally applied in machine learning applications. For example one usage of genetic-fuzzy system is of 'searching' for an acceptable fuzzy system that conforms to the training data. In which, fuzzy sets and rules combined, are encoded as individuals, and GA iterations refine the individuals i.e. fuzzy system, on the basis of their fitness evaluations. The fitness function is usually MSE of the individual fuzzy system on the training data. Very similar applications have been developed in the other popular neuro-fuzzy systems, in which neural networks are used to find the best fuzzy system for the given data through means of classical ANN learning algorithms.

CS607 - Final Term Papers Spring 2012

Compare between the supervised and un-supervised learning methodology.(Marks 5)
Answer:- Rep

Write the general stages of expert system development life cycle or ESDLC. (Marks 5)
Answer:- Rep

"Boolean logic is a subset of fuzzy logic." Do you agree with the statement or not? Give reason to support your answer.(Marks 3)
Answer:- (Page 147)
Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth -- truth values between "completely true" and "completely false".
For example, There are two persons. Person A is standing on the left of person B. Person A is definitely shorter than person B. But if boolean gauge has only two readings, 1 and 0, then a person can be either all or short. Let's say if the cut off point is at 5 feet 10 inches then all the people having a height greater than this limit are taller and the rest are short.

Write down the name of category with which these algorithms belong to and also describe that category briefly. These algorithms are listed below:
1) **Self-organizing maps (SOM)**
2) **k-means**
3) **Linear vector quantization (Marks 3)**

Answer:- (Page 205)

These algorithms are belongs to clustering.

Clustering is a form of unsupervised learning, in which the training data is available but without the classification information or class labels. The task of clustering is to identify and group similar individual data elements based on some measure of similarity. So basically using clustering algorithms, classification information can be 'produced' from a training data which has no classification data at the first place. Naturally, there is no supervision of classification in clustering algorithms for their learning/clustering, and hence they fall under the category of unsupervised learning.

***Information gain is a measure of the impurity in a collection of training examples*. Do you agree with the statement or not? If not, then write a correct statement. (Marks 3).**

Answer:- (Page 177)

Disagree.

Entropy as a measure of the impurity in a collection of training examples.

In order to define information gain precisely, we begin by defining a measure commonly used in statistics and information theory, called entropy, which characterizes the purity/impurity of an arbitrary collection of examples.

What are the three phases of machine learning?(Marks 3)

Answer:- (Page 160)

1. Training
2. Validation
3. Application

Which term is best suited for a situation in which we perform unsupervised learning on training data without having classification information or class labels? (Marks 2)

Answer:- (Page 205)

Clustering is a form of unsupervised learning, in which the training data is available but without the classification information or class labels

Out of different characteristics that an intelligent machine will show which two characteristics do you think are most important as compare to other characteristics for making that machine more artificially intelligent?(Marks 2)

Answer:- (Page 14)

To create intelligent machines we first need to understand how the real brain functions

جو شخص ناکامیوں سے ڈر کر بھاگتا ہے کامیابی اُس سے ڈر کر بھاگتی ہے

Differentiate briefly between training and validation in machine learning?(Marks 2)

Answer:- Rep

Draw the truth table (for two operands A and B) for Fuzzy operator AND. (Marks 2)

Answer:- (Page 150)

A	B	A and B
0	0	0
0	1	0
1	0	0
1	1	1

AND

بد صورت چہرہ بد صورت دماغ سے بہتر ہے

جو لوگوں کے سامنے فخر کرتا ہے وہ لوگوں کی نظروں سے گر جاتا ہے

عقل مند اپنے عیب خود دیکھتا ہے اور بیوقوفوں کے عیب دنیا دیکھتی ہے