BSc.(Information Technology)
(Semester VI)
2019-20

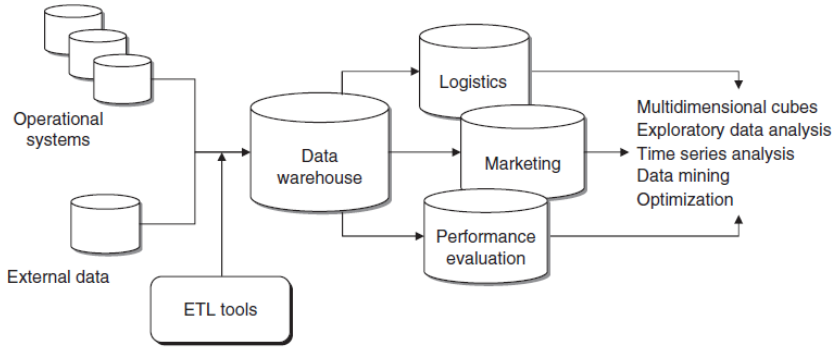
Business Intelligence
(USIT 603 Core)
University Paper Solution

By
Hrishikesh Tendulkar

Question 1

Q1a. Describe the architecture of business intelligence.

Ans: Business intelligence may be defined as a set of mathematical models and analysis methodologies that exploit the available data to generate information and knowledge useful for complex decision-making processes.



Data sources: In a first stage, it is necessary to gather and integrate the data stored in the various primary and secondary sources, which are heterogeneous in origin and type. The sources consist for the most part of data belonging to operational systems, but may also include unstructured documents, such as emails and data received from external providers.

Data warehouses and data marts: Using extraction and transformation tools known as extract, transform, load (ETL), the data originating from the different sources are stored in databases intended to support business intelligence analyses.

Data exploration: The tools for performing a passive business intelligence analysis, which consist of query and reporting systems, as well as statistical methods. These are referred to as passive methodologies because decision makers are requested to generate prior hypotheses or define data extraction criteria, and then use the analysis tools to find answers and confirm their original insight.

Data mining: Purpose is the extraction of information and knowledge from data. Their purpose is instead to expand the decision makers' knowledge. These include mathematical models for pattern recognition, machine learning and data mining techniques

Optimization: By moving up one level in the pyramid we find optimization models that allow us to determine the best solution out of a set of alternative actions, which is usually fairly extensive and sometimes even infinite.

Decisions: Finally, the top of the pyramid corresponds to the choice and the actual adoption of a specific decision and in some way represents the natural conclusion of the decision-making process. Even when business intelligence methodologies are available and successfully adopted, the choice of a decision pertains to the decision makers, who may also take advantage of informal and unstructured information available to adapt and modify the recommendations and the conclusions achieved through the use of mathematical models.

Q1b. Explain Data Information and Knowledge

Ans: **Data:** Generally, data represent a structured codification of single primary entities, as well as of transactions involving two or more primary entities. For example, for a retailer data refer to primary entities such as customers, points of sale and items, while sales receipts represent

the commercial transactions.

Information: Information is the outcome of extraction and processing activities carried out on data, and it appears meaningful for those who receive it in a specific domain. For example, to the sales manager of a retail company, the proportion of sales receipts in the amount of over € 100 per week, or the number of customers holding a loyalty card who have reduced by more than 50% the monthly amount spent in the last three months, represent meaningful pieces of information that can be extracted from raw stored data.

Knowledge: We can think of knowledge as consisting of information put to work into a specific domain, enhanced by the experience and competence of decision makers in tackling and solving complex problems. For a retail company, a sales analysis may detect that a group of customers, living in an area where a competitor has recently opened a new point of sale, have reduced their usual amount of business.

Q 1c. Define system. Explain how system can be characterized

Ans: System is made up of a set of components that are in some way connected to each other so as to provide a single collective result and a common purpose. Every system is characterized by boundaries that separate its internal components from the external environment.

A system is said to be open if its boundaries can be crossed in both directions by flows of materials and information.

When such flows are lacking, the system is said to be closed.

In general terms, any given system receives specific input flows, carries out an internal transformation process and generates observable output flows

Q1d. What is business intelligence? Why effective and timely decision is important

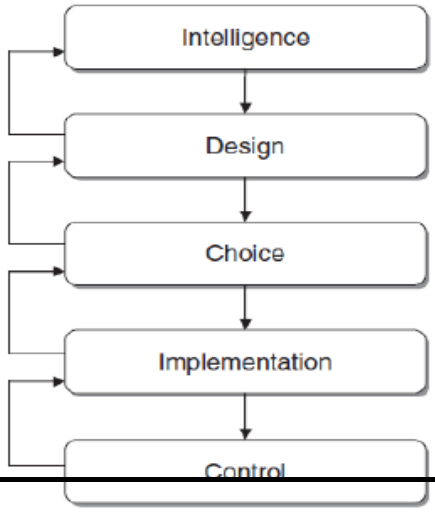
Ans: Business intelligence may be defined as a set of mathematical models and analysis methodologies that exploit the available data to generate information and knowledge useful for complex decision-making processes.

Effective decisions: The application of rigorous analytical methods allows decision makers to rely on information and knowledge which are more dependable. As a result, they are able to make better decisions and devise action plans that allow their objectives to be reached in a more effective way. Indeed, turning to formal analytical methods forces decision makers to explicitly describe both the criteria for evaluating alternative choices and the mechanisms regulating the problem under investigation. Furthermore, the ensuing in-depth examination and thought lead to a deeper awareness and comprehension of the underlying logic of the decision-making process.

Timely decisions: Enterprises operate in economic environments characterized by growing levels of competition and high dynamism. As a consequence, the ability to rapidly react to the actions of competitors and to new market conditions is a critical factor in the success or even the survival of a company.

Q1e. Explain the phases of decision making process system.

Ans:



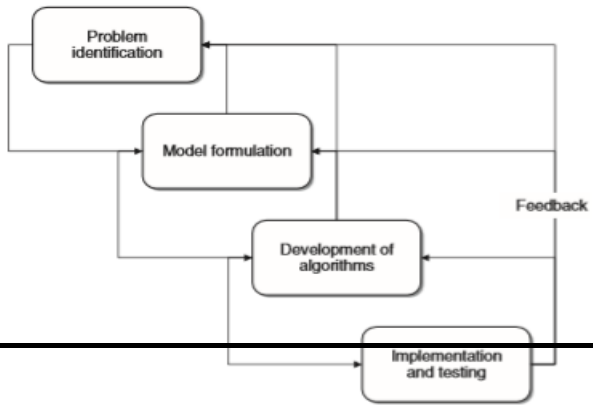
Intelligence: In the intelligence phase the task of the decision maker is to identify, circumscribe and explicitly define the problem that emerges in the system under study. The analysis of the context and all the available information may allow decision makers to quickly grasp the signals and symptoms pointing to a corrective action to improve the system performance.
Design: In the design phase actions aimed at solving the identified problem should be developed and planned. At this level, the experience and creativity of the decision makers play a critical role, as they are asked to devise viable solutions that ultimately allow the intended purpose to be achieved.
Choice: Once the alternative actions have been identified, it is necessary to evaluate them on the basis of the performance criteria deemed significant. Mathematical models and the corresponding solution methods usually play a valuable role during the choice phase.
Implementation: When the best alternative has been selected by the decision maker, it is transformed into actions by means of an implementation plan. This involves assigning responsibilities and roles to all those involved into the action plan.
Control: Once the action has been implemented, it is finally necessary to verify and check that the original expectations have been satisfied and the effects of the action match the original intentions. In particular, the differences between the values of the performance indicators identified in the choice phase and the values actually observed at the end of the implementation plan should be measured.

Q1f. Explain the major potential advantage derived from adoption of a DSS
Ans: Major potential advantages deriving from the adoption of a DSS:

- An increase in the number of alternatives or options considered
- An increase in the number of effective decisions devised
- A greater awareness and a deeper understanding of the domain analyzed and the problems investigated
- The possibility of executing scenario and what-if analyses by varying the hypotheses and parameters of the mathematical models
- An improved ability to react promptly to unexpected events and unforeseen situations;
- A value-added exploitation of the available data
- an improved communication and coordination among the individuals and the organizational departments
- More effective development of teamwork
- A greater reliability of the control mechanisms, due to the increased intelligibility of the decision process.

Question 2

Q2a. Explain the primary phases of model.
Ans: It is possible to break down the development of a mathematical model for decision making into four primary phases, shown in the following figure. The figure also includes a feedback mechanism



which takes into account the possibility of changes and revisions of the model.

Problem identification - First of all, the problem at hand must be correctly identified. The observed critical symptoms must be analysed and interpreted in order to formulate hypotheses for investigation.

Model formulation – Once the problem to be analyzed has been properly identified, effort should be directed toward defining an appropriate mathematical model to represent the system. A number of factors affect and influence the choice of model, such as the time horizon, the decision variables, the evaluation criteria, the numerical parameters and the mathematical relationships.

1. **Time horizon** – Usually a model includes a temporal dimension.
2. **Evaluation criteria**: Appropriate measurable performance indicators should be defined in order to establish a criterion for the evaluation and comparison of the alternative decisions.

Development of algorithms

Once a mathematical model has been defined, one will naturally wish to proceed with its solution to assess decisions and to select the best alternative.

Implementation and test

When a model is fully developed, then it is finally implemented, tested and utilized in the application domain. It is also necessary that the correctness of the data and the numerical parameters entered in the model be preliminarily assessed. These data usually come from a data warehouse or a data mart previously set up.

Q2b. What is predictive and optimization model.

Ans: **P*redictive* models** play a primary role in business intelligence systems, since they are logically placed upstream with respect to other mathematical models and, more generally, to the whole decision-making process. Predictions allow input information to be fed into different decision-making processes, arising in strategy, research and development, administration and control, marketing, production and logistics. Basically, all departmental functions of an enterprise make some use of predictive information to develop decision making, even though they pursue different objectives.

Optimization models many decision-making processes faced by companies or complex organizations can be cast according to the following framework: given the problem at hand, the decision maker defines a set of *feasible* decisions and establishes a criterion for the evaluation and comparison of alternative choices, such as monetary costs or payoffs. At this point, the decision maker must identify the *optimal* decision according to the evaluation criterion defined, that is, the choice corresponding to the minimum cost or to the maximum payoff.

Optimization models represent a fairly substantial class of optimization problems that are derived when the objective of the decision-making process is a function of the decision variables, and the criteria describing feasible decisions can be expressed by a set of mathematical equalities and inequalities in the decision variables.

Q2c. Explain some areas where data mining is used

Ans: The term data mining indicates the process of exploration and analysis of a dataset, usually of large size, in order to find regular patterns, to extract relevant knowledge and to obtain meaningful recurring rules. Data mining plays an ever-growing role in both theoretical studies and applications. Data mining methodologies can be applied to a variety of domains, from marketing and manufacturing process control to the study of risk factors in medical diagnosis, from the evaluation of the effectiveness of new drugs to fraud detection.

Relational marketing – Data mining applications in the field of relational marketing, have significantly contributed to the increase in the popularity of these methodologies. Some relevant applications within relational marketing are:

- Identification of customer segments that are most likely to respond to targeted marketing campaigns, such as cross-selling and up-selling
- Identification of target customer segments for retention campaigns
- Prediction of the rate of positive responses to marketing campaigns
- Interpretation and understanding of the buying behavior of the customers
- Analysis of the products jointly purchased by customers, known as market basket analysis.

Fraud detection – Fraud detection is another relevant field of application of data mining. Fraud may affect different industries such as telephony, insurance (false claims) and banking (illegal use of credit cards and bank checks; illegal monetary transactions).

Risk evaluation – The purpose of risk analysis is to estimate the risk connected with future decisions, which often assume a dichotomous form. For example, using the past observations available, a bank may develop a predictive model to establish if it is appropriate to grant a monetary loan or a home loan, based on the characteristics of the applicant.

Text mining – Data mining can be applied to different kinds of texts, which represent unstructured data, in order to classify articles, books, documents, emails and web pages. Examples are web search engines or the automatic classification of press releases for storing purposes. Other text mining applications include the generation of filters for email messages and newsgroups.

Image recognition – The treatment and classification of digital images, both static and dynamic, is an exciting subject for both its theoretical interest and the great number of applications it offers. It is useful to recognize written characters, compare and identify human faces, apply correction filters to photographic equipment and detect suspicious behaviors through surveillance video cameras.

Web mining – Web mining applications, are intended for the analysis of so-called clickstreams – the sequences of pages visited and the choices made by a web surfer. They may prove useful for the analysis of e-commerce sites, in offering flexible and customized pages to surfers, in caching the most popular pages or in evaluating the effectiveness of an e-learning training course.

Medical diagnosis – Learning models are an invaluable tool within the medical field for the early detection of diseases using clinical test results. Image analysis for diagnostic purpose is another field of investigation that is currently burgeoning.

Q2d. Write short notes on analysis methodology of data mining.

Ans: Data mining activities can be subdivided into a few major categories, based on the tasks

and the objectives of the analysis.

Supervised learning – In a supervised (or direct) learning analysis, a target attribute either represents the class to which each record belongs, For example on loyalty in the mobile phone industry, a measurable quantity, such as the total value of calls that will be placed by a customer in a future period. As a second example of the supervised perspective, consider an investment management company wishing to predict the balance sheet of its customers based on their demographic characteristics and past investment transactions. Supervised learning processes are therefore oriented toward prediction and interpretation with respect to a target attribute.

Unsupervised learning – Unsupervised (or indirect) learning analyses are not guided by a target attribute. Therefore, data mining tasks in this case are aimed at discovering recurring patterns and affinities in the dataset. As an example, consider an investment management company wishing to identify clusters of customers who exhibit homogeneous investment behaviour, based on data on past transactions. In most unsupervised learning analyses, one is interested in identifying clusters of records that are similar within each cluster and different from members of other clusters.

Q2e. What is meant by data validation? Explain different kinds of data validation

Ans: The quality of input data may prove unsatisfactory due to incompleteness, noise and inconsistency.

Incompleteness – Some records may contain missing values corresponding to one or more attributes, and there may be a variety of reasons for this. It may be that some data were not recorded at the source in a systematic way, or that they were not available when the transactions associated with a record took place. In other instances, data may be missing because of malfunctioning recording devices. It is also possible that some data were deliberately removed during previous stages of the gathering process because they were deemed incorrect. Incompleteness may also derive from a failure to transfer data from the operational databases to a data mart used for a specific business intelligence analysis.

Noise – Data may contain erroneous or anomalous values, which are usually referred to as outliers. Other possible causes of noise are to be sought in malfunctioning devices for data measurement, recording and transmission. The presence of data expressed in heterogeneous measurement units, which therefore require conversion, may in turn cause anomalies and inaccuracies.

Inconsistency – Sometimes data contain discrepancies due to changes in the coding system used for their representation, and therefore may appear inconsistent. For example, the coding of the products manufactured by a company may be subject to a revision taking effect on a given date, without the data recorded in previous periods being subject to the necessary transformations in order to adapt them to the revised encoding scheme.

The purpose of data validation techniques is to identify and implement corrective actions in case of incomplete and inconsistent data or data affected by noise.

Q2f. Write short notes on data transformation

Ans: In most data mining analyses it is appropriate to apply a few transformations to the dataset in order to improve the accuracy of the learning models subsequently developed.

Standardization - Most learning models benefit from a preventive standardization of the data, also called normalization. The most popular standardization techniques include the decimal scaling method, the min-max method and the z-index method.

Decimal Scaling – Decimal scaling is based on the transformation

$$x'_{ij} = \frac{x_{ij}}{10^h},$$

where *h* is a given parameter which determines the scaling intensity. In practice, decimal scaling corresponds to shifting the decimal point by *h* positions toward the left. In general, *h* is fixed at a value that gives transformed values in the range [−1, 1].

Min-Max. Min-max standardization is achieved through the transformation

$$x'_{ij} = \frac{x_{ij} - x_{\min,j}}{x_{\max,j} - x_{\min,j}}(x'_{\max,j} - x'_{\min,j}) + x'_{\min,j}, \qquad (6.3)$$

Where

$$x_{\min,j} = \min_i x_{ij}, \quad x_{\max,j} = \max_i x_{ij}, \qquad (6.4)$$

are the minimum and maximum values of the attribute $\mathbf{a}_j$ before transformation, while $x'_{\min,j}$ $x'_{\max,j}$ and are the minimum and maximum values that we wish to obtain after transformation. In general, the extreme values of the range are defined so that

$$x'_{\min,j} = -1 \text{ and } x'_{\max,j} = 1 \text{ or } x'_{\min,j} = 0 \text{ and } x'_{\max,j} = 1.$$

z-index – z-index based standardization uses the transformation

$$x'_{ij} = \frac{x_{ij} - \bar{\mu}_j}{\bar{\sigma}_j},$$

Where $\bar{\mu}_j$ and $\bar{\sigma}_j$ are respectively the sample mean and sample standard deviation

of the attribute $\mathbf{a}_j$. If the distribution of values of the attribute $\mathbf{a}_j$ is roughly normal, the *z*-index based transformation generates values that are almost certainly within the range *(*−3, 3*)*.

Question 3

Q3a. Explain Phases and Taxonomy of classification model

Ans: The development of a classification model consists therefore of **three main phases.**

Training phase: During the training phase, the classification algorithm is applied to the examples belonging to a subset T of the dataset D, called the training set , in order to derive classification rules that allow the corresponding target class y to be attached to each observation x.

Test phase: In the test phase, the rules generated during the training phase are used to classify the observations of D not included in the training set, for which the target class value is already known. To assess the accuracy of the classification model, the actual target class of each instance in the test set V = D − T is then compared with the class predicted by the classifier.

Prediction phase: The prediction phase represents the actual use of the classification model to assign the target class to new observations that will be recorded in the future. A prediction is obtained by applying the rules generated during the training phase to the explanatory variables that describe the new instance.

There are four main categories of classification models:

1. Heuristic models
2. Separation models
3. Regression models
4. Probabilistic models

Heuristic models: Heuristic methods make use of classification procedures based on simple and intuitive algorithms. This category includes nearest neighbor methods, based on the concept of distance between observations, and classification trees, which use divide-and-conquer schemes to derive groups of observations that are as homogeneous as possible with respect to the target class.

Separation models: Separation models divide the attribute space R^n into H disjoint regions $\{S_1, S_2, \ldots, S_H\}$, separating the observations based on the target class. The observations x_i in region S_h are assigned to class $y_i = v_h$. In general, it is hard to determine a collection of simple regions that exactly subdivide the observations of the dataset based on the value of the target class. Therefore, a loss function is defined to take the misclassified points into account, and an optimization problem is solved in order to derive a subdivision into regions that minimizes the total loss.

Regression models: Regression models, are used for the prediction of continuous target variables, make an explicit assumption concerning the functional form of the conditional probabilities $P_{y|x}(y|x)$, which correspond to the assignment of the target class by the supervisor

Probabilistic models: In probabilistic models, a hypothesis is formulated regarding the functional form of the conditional probabilities $P_{x|y}(x|y)$ of the observations given the target class, known as class-conditional probabilities. Subsequently, based on an estimate of the prior probabilities Py(y) and using Bayes' theorem, the posterior probabilities $P_{y|x}(y|x)$ of the target class assigned by the supervisor can be calculated.

Q3b. Write short notes on evaluation of classification model.

Ans: Classification methods can be evaluated based on several criteria, as follows:

Accuracy - Evaluating the accuracy of a classification model is crucial for two main reasons. First, the accuracy of a model is an indicator of its ability to predict the target class for future observations. Based on their accuracy values, it is also possible to compare different models in order to select the classifier associated with the best performance. Let T be the training set and V the test set, and t and v be the number of observations in each subset, respectively. The relations D = T ∪ V and m = t + v obviously hold. The most natural indicator of the accuracy of a classification model is the proportion of observations of the test set V correctly classified by the model. If yi denotes the class of the generic observation xi ∈ V and f (xi) the class predicted through the function f ∈ F identified by the learning algorithm A = AF, the following loss function can be defined:

$$L(y_i, f(\mathbf{x}_i)) = \begin{cases} 0, & \text{if } y_i = f(\mathbf{x}_i), \\ 1, & \text{if } y_i \neq f(\mathbf{x}_i). \end{cases}$$

The accuracy of model A can be evaluated as:

$$\text{acc}_A(\mathcal{V}) = \text{acc}_{A_\mathcal{F}}(\mathcal{V}) = 1 - \frac{1}{v}\sum_{i=1}^{v} L(y_i, f(\mathbf{x}_i))$$

In some cases, it is preferable to use an alternative performance indicator given by the proportion of errors made by the classification algorithm:

errA(V) = errAF(V) = 1 − accAF(V) = ∑ (( ))

Speed - Some methods require shorter computation times than others and can handle larger problems. However, classification methods characterized by longer computation times may be applied to a small-size training set obtained from a large number of observations by means of random sampling schemes. It is not uncommon to obtain more accurate classification rules in this way.

Robustness - A classification method is robust if the classification rules generated, as well as the corresponding accuracy, do not vary significantly as the choice of the training set and the test set varies, and if it is able to handle missing data and outliers.

Scalability - The scalability of a classifier refers to its ability to learn from large datasets, and it is inevitably related to its computation speed. Therefore, the remarks made in connection with sampling techniques for data reduction, which often result in rules having better generalization capability, also apply in this case.

Interpretability - If the aim of a classification analysis is to interpret as well as predict, then the rules generated should be simple and easily understood by knowledge workers and experts in the application domain.

Q3c. Assume your own training dataset and predict the class label of an unknown sampling using Naïve Bayesian classification

Ans:

area	numin	timein	numout	Pothers	Pmob	Pland	numsms	numserv	numcall	diropt	churner
2	1	1	2	1	4	1	3	2	2	0	1
1	1	3	3	2	4	1	4	2	3	0	0
3	2	1	2	2	4	1	3	2	1	0	0
1	2	3	2	3	4	1	1	2	1	0	0
2	3	4	4	4	1	1	3	2	1	0	0
3	3	4	1	4	2	1	4	3	1	0	0
3	3	3	4	4	3	1	4	3	1	1	0
1	1	1	1	1	3	2	1	1	1	0	1
2	2	2	2	1	3	2	2	3	1	1	1
4	2	1	3	2	3	2	1	2	1	1	1
3	1	1	2	2	2	2	2	2	1	0	1
4	3	4	4	2	1	2	2	4	1	1	1
2	1	1	3	2	3	2	2	4	1	1	0
4	2	1	2	3	2	2	2	2	1	0	0
3	3	4	4	3	2	2	2	4	1	1	0
1	1	2	1	4	2	2	2	2	1	0	0
4	1	1	2	4	2	2	4	2	1	0	1
1	1	1	1	1	1	3	1	1	1	0	0
3	1	1	1	1	1	3	1	1	1	0	1
2	3	4	3	1	1	3	1	1	1	0	1
1	3	3	3	1	2	3	4	2	1	0	0
4	2	2	2	2	2	3	1	1	1	0	1
3	3	2	1	4	1	3	1	1	1	0	0

Consider the data given above. The relative frequencies of the sample attribute values given the target class are as follows

area

P (area = 1 | churner = 0) = 5/13
P (area = 2 | churner = 0) = 2/13
P (area = 3 | churner = 0) = 5/13
P (area = 4 | churner = 0) = 1/13

P (area = 1 | churner = 1) = 1/10
P (area = 2 | churner = 1) = 3/10
P (area = 3 | churner = 1) = 2/10
P (area = 4 | churner = 1) = 4/10

numin

P (numin = 1 | churner = 0) = 4/13
P (numin = 2 | churner = 0) = 3/13
P (numin = 3 | churner = 0) = 6/13

P (numin = 1 | churner = 1) = 5/10
P (numin = 2 | churner = 1) = 3/10
P (numin = 3 | churner = 1) = 2/10

timein

P (timein = 1 | churner = 0) = 4/13 P (timein = 1 | churner = 1) = 6/10
P (timein = 2 | churner = 0) = 2/13 P (timein = 2 | churner = 1) = 2/10
P (timein = 3 | churner = 0) = 4/13 P (timein = 3| churner = 1) = 0
P (timein = 4 | churner = 0) = 3/13 P (timein = 4 | churner = 1) = 2/10

numout

P (numout = 1 | churner = 0) = 4/13 P (numout = 1 | churner = 1) = 2/10
P (numout = 2 | churner = 0) = 3/13 P (numout = 2 | churner = 1) = 5/10
P (numout = 3 | churner = 0) = 3/13 P (numout = 3 | churner = 1) = 2/10
P (numout = 4 | churner = 0) = 3/13 P (numout = 4 | churner = 1) = 1/10

Pothers

P (Pothers = 1 | churner = 0) = 2/13 P (Pothers = 1 | churner = 1) = 5/10
P (Pothers = 2 | churner = 0) = 3/13 P (Pothers = 2 | churner = 1) = 4/10
P (Pothers = 3 | churner = 0) = 3/13 P (Pothers = 3 | churner = 1) = 0
P (Pothers = 4 | churner = 0) = 5/13 P (Pothers = 4 | churner = 1) = 1/10

Pmob

P (Pmob = 1 | churner = 0) = 3/13 P (Pmob = 1 | churner = 1) = 3/10
P (Pmob = 2 | churner = 0) = 5/13 P (Pmob = 2 | churner = 1) = 3/10
P (Pmob = 3 | churner = 0) = 2/13 P (Pmob = 3 | churner = 1) = 3/10
P (Pmob = 4 | churner = 0) = 3/13 P (Pmob = 4 | churner = 1) = 1/10

Pland

P (Pland = 1 | churner = 0) = 6/13 P (Pland = 1 | churner = 1) = 1/10
P (Pland = 2 | churner = 0) = 4/13 P (Pland = 2 | churner = 1) = 6/10
P (Pland = 3 | churner = 0) = 3/13 P (Pland = 3 | churner = 1) = 3/10

numsms

P (numsms = 1 | churner = 0) = 3/13 P (numsms = 1 | churner = 1) = 5/10
P (numsms = 2 | churner = 0) = 4/13 P (numsms = 2 | churner = 1) = 3/10
P (numsms = 3 | churner = 0) = 2/13 P (numsms = 3 | churner = 1) = 1/10
P (numsms = 4 | churner = 0) = 4/13 P (numsms = 4 | churner = 1) = 1/10

numserv

P (numserv = 1 | churner = 0) = 2/13 P (numserv = 1 | churner = 1) = 4/10
P (numserv = 2 | churner = 0) = 7/13 P (numserv = 2 | churner = 1) = 4/10
P (numserv = 3 | churner = 0) = 2/13 P (numserv = 3 | churner = 1) = 1/10
P (numserv = 4 | churner = 0) = 2/13 P (numserv = 4 | churner = 1) = 1/10

numcall

P (numcall = 1 | churner = 0) = 12/13 P (numcall = 1 | churner = 1) = 9/10
P (numcall = 2 | churner = 0) = 0 P (numcall = 2 | churner = 1) = 1/10
P (numcall = 3| churner = 0) = 1/13 P (numcall = 3| churner = 1) = 0

diropt

P (diropt = 0 | churner = 0) = 10/13 P (diropt = 0 | churner = 1) = 7/10
P (diropt = 1 | churner = 0) = 3/13 P (diropt = 1 | churner = 1) = 3/10

Once the conditional probabilities of each attribute given the target class have been estimated, suppose that we wish to predict the target class of a new observation, represented by the vector x = (1, 1, 1, 2, 1, 4, 2, 1, 2, 1, 0). With this aim in mind, we compute the posterior probabilities P(x|0) and P(x|1):

P(x|0) = 5/13 * 4/13 * 4/13 * 3/13 * 2/13 * 3/13 * 4/13 * 3/13 * 7/13 * 12/13 * 10/13
$=0.81* 10^{-5}$

P (x|1) = 1/10 * 5/10 * 6/10 * 5/10 * 5/10 * 1/10 * 6/10 * 5/10 * 4/10 * 9/10 * 7/10
$= 5.67 * 10^{-5}$

Since the relative frequencies of the two classes are given by

P (churner = 0) = 13/23 = 0.56, P (churner = 1) = 10/23 = = 0.44,
We have
P (churner = 0|x) = P (x|0) P (churner = 0) = 0.81 * 10^{-5} * 0.56 = 0.46 * 10^{-5}
P (churner = 1|x) = P (x|1) P (churner = 1) = 5.67 * 10^{-5} * 0.44 = 2.495 * 10^{-5}
The new example x is then labeled with the class value {1}, since this is associated with the maximum a posteriori probability.

Q3d. Differentiate between the following clustering methodologies:
i) Partitioning method ii) Hierarchical method

Ans Partition methods develop a subdivision of the given dataset into a predetermined number K of non-empty subsets**.** Hierarchical methods carry out multiple subdivisions into subsets based on a tree structure and characterized by different homogeneity thresholds. Partition methods require the number of clusters to be predetermined. Hierarchical methods do not require the number of clusters to be predetermined.

Partitioning method is faster than clustering whereas Hierarchical method is slower than partitioning method

Partition methods are heuristic in nature as they are based on greedy methods where as Hierarchical methods are algorithmic in nature

Partition methods start with an initial assignment of the m variable observations to the K clusters and then iteratively apply a reallocation technique whose purpose is to place some observations in a different cluster so that the overall quality of the subdivision is improved. In order to evaluate the distance between two clusters most hierarchical methods resort to minimum distance or maximum distance or mean distance or ward distance

Partition methods has two algorithms K-means and K-medoids where as Hierarchical methods have Agglomerative and Divisive methods

Q3e. Explain evaluation of clustering model

Ans: To evaluate a clustering method it is first necessary to verify that the clusters generated correspond to an actual regular pattern in the data. It is therefore appropriate to apply other clustering algorithms and to compare the results obtained by different methods. In this way it is also possible to evaluate if the number of identified clusters is robust with respect to the different techniques applied. At a subsequent phase it is recommended to calculate some performance indicators. Let *C* = {*C*1,*C*2, . . . , *CK*} be the set of *K* clusters generated. An indicator of homogeneity of the observations within each cluster *Ch* is given by the *cohesion*, defined as

$$\mathrm{coh}(C_h) = \sum_{\substack{\mathbf{x}_i \in C_h \\ \mathbf{x}_k \in C_h}} \mathrm{dist}(\mathbf{x}_i, \mathbf{x}_k).$$

The overall cohesion of the partition *C* can therefore be defined as

$$\mathrm{coh}(\mathcal{C}) = \sum_{C_h \in \mathcal{C}} \mathrm{coh}(C_h).$$

One clustering is preferable over another, in terms of homogeneity within each cluster, if it has a smaller overall cohesion. An indicator of inhomogeneity between a pair of clusters is given by the *separation*, defined as

$$\mathrm{sep}(C_h, C_f) = \sum_{\substack{\mathbf{x}_i \in C_h \\ \mathbf{x}_k \in C_f}} \mathrm{dist}(\mathbf{x}_i, \mathbf{x}_k).$$

Again the overall separation of the partition *C* can be defined as

$$\mathrm{sep}(\mathcal{C}) = \sum_{\substack{C_h \in \mathcal{C} \\ C_f \in \mathcal{C}}} \mathrm{sep}(C_h, C_f).$$

Q3f. Explain k-means method

Ans: The *K-means* algorithm receives as input a dataset *D*, a number *K* of clusters to be generated and a function dist(**x***i* , **x***k)* that expresses the inhomogeneity between each pair of observations, or equivalently the matrix **D** of distances between observations. Given a cluster *Ch*, *h* = 1, 2, . . . *K*, the *centroid* of the cluster is defined as the point **z***h* having coordinates equal to the mean value of each attribute for the observations belonging to that cluster, that is,

$$z_{hj} = \frac{\sum_{\mathbf{x}_i \in C_h} x_{ij}}{\text{card}\{C_h\}}.$$

***K*-means algorithm**

1. During the initialization phase, *K* observations are arbitrarily chosen in *D* as the centroids of the clusters.
2. Each observation is iteratively assigned to the cluster whose centroid is the most similar to the observation, in the sense that it minimizes the distance from the record.
3. If no observation is assigned to a different cluster with respect to the previous iteration, the algorithm stops.
4. For each cluster, the new centroid is computed as the mean of the values of the observations belonging to the cluster, and then the algorithm returns to step 2.

The algorithm described in the above procedure starts by arbitrarily selecting K observations that constitute the initial centroids. At each subsequent iteration each record is assigned to the closer

As stated earlier, if the attributes are numerical or ordinal categorical it is possible to transform them into a standardized representation in *n*-dimensional Euclidean space. In this case, one may also express the function of the overall heterogeneity between each observation and the point **w*h*** representing the cluster *Ch* to which it is assigned, by means of the minimization of the squared error

Question 4

Q4a.What is marketing decision process? Explain relational marketing in details.

Ans: Marketing decision process are characterized by a high level of complexity due to simultaneous presence of multiple objectives and countless alternative actions resulting from the combination of the major choice options available to decision makers. The importance of mathematical models for marketing has been further strengthened by the availability of massive databases of sales transactions that provide accurate information on how customers make use of services or purchase products. In order to fully understand the reasons why enterprises develop relational marketing initiatives, consider the following three examples: an insurance company that wishes to select the most promising market segment to target for a new type of policy; a mobile phone provider that wishes to identify those customers with the highest probability of churning, that is, of discontinuing their service and taking out a new contract with a competitor, in order to develop targeted retention initiatives; a bank issuing credit cards that needs to identify a group of customers to whom a new savings management service should be offered. These situations share some common features: a company owning a massive database, which describes the purchasing behavior of its customers, and the way they make use of services wishes to extract from these data useful and accurate knowledge to develop targeted and effective marketing campaigns. The aim of a *relational marketing* strategy is to initiate, strengthen, intensify and

preserve over time the relationships between a company and its stakeholders, represented primarily by its customers, and involves the analysis, planning, execution and evaluation of the activities carried out to pursue these objectives.

Q4b. Write a short note on sales force management

Ans: Most companies have a sales network and therefore rely on a substantial number of people employed in sales activities, who play a critical role in the profitability of the enterprise and in the implementation of a relational marketing strategy. The term *salesforce* is generally taken to mean the whole set of people and roles that are involved, with different tasks and responsibilities, in the sales process. A preliminary taxonomy of salesforces is based on the type of activity carried out, as indicated below.

Residential - Residential sales activities take place at one or more sites managed by a company supplying some products or services, where customers go to make their purchases. This category includes sales at retail outlets as well as wholesale trading centres and *cash-and-carry* shops.

Mobile - In mobile sales, agents of the supplying company go to the customers' homes or offices to promote their products and services and collect orders. Sales in this category occur mostly within B2B relationships, even though they can also be found in B2C contexts.

Telephone - Telephone sales are carried out through a series of contacts by telephone with prospective customers. They can be subdivided into a few main categories:

- Designing the sales network;
- Planning the agents' activities;
- Contact management;
- Sales opportunity management;
- Customer management;
- Activity management;
- Order management;
- Area and territory management;
- Support for the configuration of products and services;
- Knowledge management with regard to products and services.

Designing the sales network and planning the agents' activities involve decision-making tasks that may take advantage of the use of optimization models.

Q4c. Explain market basket analysis

Ans: The purpose of *market basket analysis* is to gain insight from the purchases made by customers in order to extract useful knowledge to plan marketing actions. It is mostly used to analyze purchases in the retail industry and in e-commerce activities, and is generally amenable to unsupervised learning problems. It may also be applied in other domains to analyze the purchases made using credit cards, the complementary services activated by mobile or fixed telephone customers, the policies or the checking accounts acquired by a same household. The data used for this purpose mostly refer to purchase transactions, and can be associated with the time dimension if the purchaser can be tracked through a loyalty card or the issue of an invoice. Each transaction consists of a list of purchased items. This list is called a *basket*, just like the baskets available at retail points of sale. If transactions cannot be connected to one another, say because the purchaser is unknown, one may then apply association rules, The rules extracted in this way can then be used to support different decision-making processes, such as assigning the location of the items on the shelves, determining the layout of a point of sale, identifying which items should be included in promotional flyers, advertisements or coupons distributed to customers. If customers are

individually identified and traced, besides the above techniques it is also possible to develop further analyses that take into account the time dimension of the purchases

Q4d. List revenue management system. Explain any one in detail.

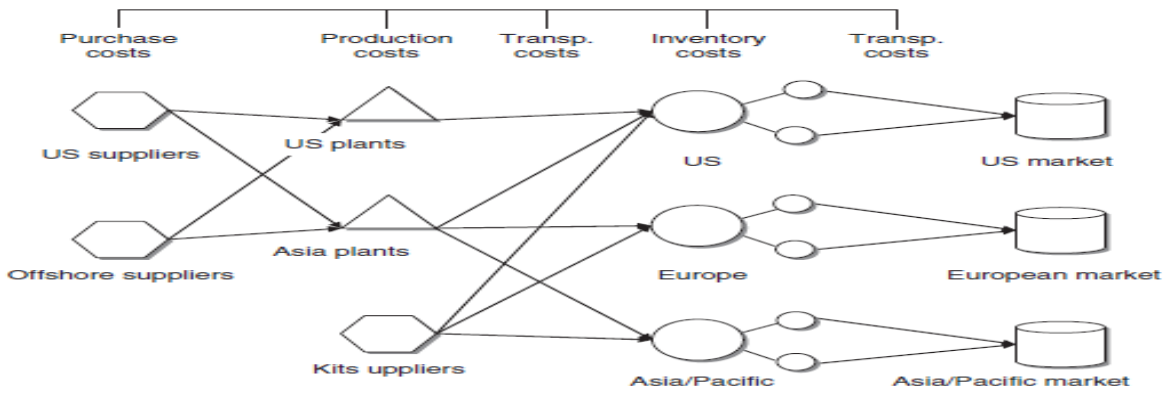
Ans: *Revenue management* is a managerial policy whose purpose is to maximize profits through an optimal balance between demand and supply. It is mainly intended for marketing as well as logistic activities and has found growing interest in the service industry, particularly in the air transportation, tourism and hotel sectors. The purpose of revenue management is to maximize profits, aligning the offer of products and services to the expected demand, using both the major levers of the marketing mix (e.g. prices, promotions, assortment) and the levers of logistics (e.g. efficiency and timeliness). . Revenue management affects some highly complex decision-making processes like

- market segmentation, by product, distribution channel, consumer type and geographic area, performed using data mining models
- prediction of future demand, using time series and regression models;
- identification of the optimal assortment, i.e. the mix of products to be allocated to each point of sale;
- definition of the market response function, obtained by identifying models and rules that explain the demand based on company actions, the initiatives of competitors and other exogenous contextual events;
- management of activities aimed at determining the price of each product (*pricing*) as well as the timing and the amount of markdowns;
- planning, management and monitoring of sales promotions, and assessment of their effectiveness;
- sales analysis and control, and use of the information gathered to evaluate market trends;
- material procurement and stock management policies, such as control policy, frequency of issued orders, reorder quantities;
- integrated management of the different sales and distribution channels

The adoption of revenue management methods and tools requires a few prerequisite conditions to be satisfied within a company, since without them the expected results are unlikely to be achieved. As with any innovation project, it is the people and the organization that constitute a key success factor rather than the use of specific software tools. In this case too, the culture and the structure of the processes within an organization must be prepared to adopt powerful tools that may turn out to be unsafe and disrupting if improperly used. It is therefore necessary to develop within the enterprise an information culture, particularly among those knowledge workers who operate in the marketing and logistics departments, more directly involved with the application of revenue management strategies. This means that all marketing data must be systematically gathered, controlled, normalized, integrated and stored in a data mart. To segment the market and to create micro-segments, business intelligence methods and analytical models should be used. It is therefore advisable for an enterprise turning to revenue management to have already developed relational marketing initiatives or at least to be able to carry out data mining analyses.

Q4e. What is supply chain optimization?

Ans: A *supply chain* may be defined as a network of connected and interdependent organizational units that operate in a coordinated way to manage, control and improve the flow of materials and information originating from the suppliers and reaching the end customers, after going through the procurement, processing and distribution subsystems of a company, as shown in figure below



The aim of the integrated planning and operations of the supply chain is to combine and evaluate from a systemic perspective the decisions made and the actions undertaken within the various sub-processes that compose the logistic system of a company. Many manufacturing companies, such as those operating in the consumer goods industry, have concentrated their efforts on the integrated operations of the supply chain, even to the point of incorporating parts of the logistic chain that are outside the company, both upstream and downstream. The major purpose of an integrated logistic process is to minimize a function expressing the total cost, which comprises processing costs, transportation costs for procurement and distribution, inventory costs and equipment costs. The need to optimize the logistic chain, and therefore to have models and computerized tools for medium-term planning and for capacity analysis, is particularly critical in the face of the high complexity of current logistic systems, which operate in a dynamic and truly competitive environment.

Optimization models represent a powerful and versatile conceptual paradigm for analysing and solving problems arising within integrated supply chain planning, and for developing the necessary software. Due to the complex interactions occurring between the different components of a logistic production system, other methods and tools intended to support the planning activity seem today inadequate, such as electronic spreadsheets, simulation systems and planning modules at infinite capacity included in enterprise resource planning software.

Conversely, optimization models enable the development of realistic mathematical representations of a logistic production system, able to describe with reasonable accuracy the complex relationships among critical components of the logistic system, such as capacity, resources, plans, inventory, batch sizes, lead times and logistic flows, taking into account the various costs. Moreover, the evolution of information technologies and the latest developments in optimization algorithms mean that decision support systems based on optimization models for logistics planning can be efficiently developed

Q4f. Explain CCR model in detail.

Ans: Using data envelopment analysis, the choice of the optimal system of weights for a generic DMU*j* involves solving a mathematical optimization model whose decision variables are represented by the weights *ur* , *r* ∈ *K*, and *vi* , *i* ∈ *H*, associated with each output and input. The best-known of which is probably the Charnes–Cooper–Rhodes (CCR) model. The CCR model formulated for DMU*j* takes the form

$$\max \quad \vartheta = \frac{\sum_{r\in\mathcal{K}} u_r y_{rj}}{\sum_{i\in\mathcal{H}} v_i x_{ij}},$$

$$\text{s.to} \quad \frac{\sum_{r\in\mathcal{K}} u_r y_{rj}}{\sum_{i\in\mathcal{H}} v_i x_{ij}} \le 1, \qquad j \in \mathcal{N},$$

$$u_r, v_i \ge 0, \qquad r \in \mathcal{K}, i \in \mathcal{H}.$$

The objective function involves the maximization of the efficiency measure for DMU*j*..Constraints require that the efficiency values of all the units, calculated by means of the weights system for the unit being examined, be lower than one. Finally, conditions guarantee that the weights associated with the inputs and the outputs are non-negative. In place of these conditions, sometimes the constraints $ur, vi \geq \delta$, $r \in K$, $i \in H$ may be applied, where $\delta > 0$, preventing the unit from assigning a null weight to an input or output. Model can be linearized by requiring the weighted sum of the inputs to take a constant value, for example 1. This condition leads to an alternative optimization problem, the *input-oriented* CCR model, where the objective function consists of the maximization of the weighted sum of the outputs

$$\begin{aligned} \max \quad & \vartheta = \sum_{r \in \mathcal{K}} u_r y_{rj}, \\ \text{s.to} \quad & \sum_{i \in \mathcal{H}} v_i x_{ij} = 1, \\ & \sum_{r \in \mathcal{K}} u_r y_{rj} - \sum_{i \in \mathcal{H}} v_i x_{ij} \leq 0, \qquad j \in \mathcal{N}, \\ & u_r, v_i \geq 0, \qquad r \in \mathcal{K}, i \in \mathcal{H}. \end{aligned}$$

Let ϑ^* be the optimum value of the objective function corresponding to the optimal solution $(\mathbf{v}^*,\mathbf{u}^*)$ of problem. DMU*j* is said to be *efficient* if $\vartheta^* = 1$ and if there exists at least one optimal solution $(\mathbf{v}^*,\mathbf{u}^*)$ such that $\mathbf{v}^* > \mathbf{0}$ and $\mathbf{u}^* > \mathbf{0}$. By solving a similar optimization model for each of the *n* units being compared, one obtains *n* systems of weights. The flexibility enjoyed by the units in choosing the weights represents an undisputed advantage, in that if a unit turns out to be inefficient based on the most favourable system of weights, its inefficiency cannot be traced back to an inappropriate evaluation process.

Question 5

Q5a. What is meant by knowledge management system?

Ans: Knowledge management is a process that helps organizations identify, select, organize and transfer important information and expertise that are the part of the organization's memory and that reside within the organization in an unstructured manner. Knowledge Management is the systematic and active management of ideas, information and knowledge residing in an organization's employees. The structuring of knowledge enables effective and efficient problem solving, dynamic learning, strategic planning and decision making. KM initiatives focus on identifying knowledge explicating it in such a way that it can be shared in a formal manner and leveraging its value through reuse. The information technologies that make KM available throughout an organization are referred as KM systems. Knowledge management systems refer to the use of modern IT to systemize, enhance and expedite intra and interim KM. KM systems are intended to help an organization cope with turnover, rapid change and downsizing by making the expertise of the organization's human capital widely accessible

Q5b. Explain knowledge management activities

Ans: Knowledge management activities include creation of knowledge, sharing of knowledge and seeking and the use of knowledge

Knowledge Creation - Knowledge creation is the generation of new insights, ideas, or routines described knowledge creation as an interplay between tacit and explicit knowledge and as a growing spiral as knowledge moves among the individual, group, and organizational levels. The four modes of knowledge creation are socialization,

externalization, internalization, and combination. The socialization mode refers to the conversion of tacit knowledge to new tacit knowledge through social interactions and shared experience among organization members (e.g., mentoring). The combination mode refers to the creation of new explicit knowledge by merging, categorizing, reclassifying, and synthesizing existing explicit knowledge (e.g., statistical analyses of market data). The other two modes involve interactions and conversion between tacit and explicit knowledge.

Knowledge Sharing - Knowledge sharing is the willful explication of one person's ideas, insights, solutions, experiences to another individual either via an intermediary, such as a computer-based system, or directly. However, in many organizations, information and knowledge are not considered organizational resources to be shared but individual competitive weapons to be kept private. Organizational members may share personal knowledge with fear they perceive that they are of less value if their knowledge is part of the organizational public domain. Research in organizational learning and knowledge management suggests that some facilitating conditions include trust, interest, and shared language fostering access to knowledgeable members and a culture marked by autonomy, redundancy, requisite variety, intention, and fluctuation.

Knowledge Seeking – Knowledge seeking, also referred to as knowledge sourcing is the search for and use of internal organizational knowledge. Lack of time or lack of reward may hinder the sharing of knowledge, and the same is true of knowledge seeking. Individuals may sometimes prefer to not reuse knowledge if they feel that their own performance review is based on the originality or creativity of their ideas. Such was the case for marketing employees in a global consumer goods organization.

Q5c. Explain the role of people in knowledge management

Ans: KMS is an enterprise-wide effort, where many people are involved. They include

The Chief Knowledge Officer - Most firms developing KMS have created a knowledge management officer chief knowledge officer (CKO)—at the senior level. The objectives of the CKO's role are to maximize the firm's knowledge assets, design and implement KM strategies, effectively exchange knowledge assets internally and externally, and promote system use. The CKO is responsible for developing processes that facilitate knowledge transfer. The CKO is responsible for defining the area of knowledge within the firm that will be the focal point, based on the firm's mission and objectives. The CKO is responsible for standardizing the enterprise-wide vocabulary and controlling the knowledge directory. This is critical in areas that must share knowledge across departments, to ensure uniformity. The CKO must get a handle on the company's repositories of research, resources, and expertise, including where they are stored and who manages and accesses them. Then the CKO must encourage pollination among disparate workgroups with complementary resources. The CKO is responsible for creating an infrastructure and cultural environment for knowledge sharing. He or she must assign or identify the knowledge champions within the business units. The CKO's job is to manage the content the champions' groups produce, continually add to the knowledge base, and encourage colleagues to do the same. Successful CKOs should have the full and enthusiastic support of their managers and of top management. Ultimately, the CKO is responsible for the entire knowledge management project while it is under development and then for management of the system and the knowledge after it is deployed. A CKO needs a range of skills to make knowledge management initiatives succeed. These attributes are indispensable, according to CKOs and consultants.

- Interpersonal communication skills to convince employees to adopt cultural changes
- Leadership skills to convey the knowledge management vision and passion for it

- Business acumen to relate knowledge management efforts to efficiency and profitability
- Strategic thinking skills to relate knowledge management efforts to larger goals
- Collaboration skills to work with various departments and persuade them to work together
- The ability to institute effective educational programs
- An understanding of IT and its role in advancing knowledge management

The CEO - The CEO is responsible for championing a knowledge management effort. He or she must ensure that a competent and capable CKO is found and that the CKO can obtain all the resources (including access to people with knowledge sources) needed to make the project a success. The CEO must also gain organization-wide support for contributions to and use of the KMS. The CEO must also prepare the organization for the cultural changes that are about to occur. Support is the critical responsibility of the CEO. The CEO is the primary change agent of the organization. The officers generally must make available to the CKO the resources needed to get the job done.

The CFO - The chief financial officer (CFO) must ensure that the financial resources are available.

The COO - The chief operating officer (COO) must ensure that people begin to embed knowledge management practices into their daily work processes. There is a special relationship between the CKO and chief information officer (CIO).

The CIO - The CIO is responsible for the IT vision of the organization and for the IT architecture, including databases and other potential Knowledge sources. The CIO must cooperate With the CKO in making these resources available. KMS are expensive propositions, and it is wise to use existing systems if they are available and capable.

Managers - Managers must also support the knowledge management effort and provide access to sources of knowledge. In many KMS, managers are an integral part of the communities of practice.

Communities of Practice - A community of practice (COP) is a group of people in an organization with a common professional interest. Ideally all the KMS users should each be in at least one COP

Q5d. Compare and contrast between Artificial intelligence versus Natural intelligence

Ans: Artificial intelligence has several advantages over Natural intelligence

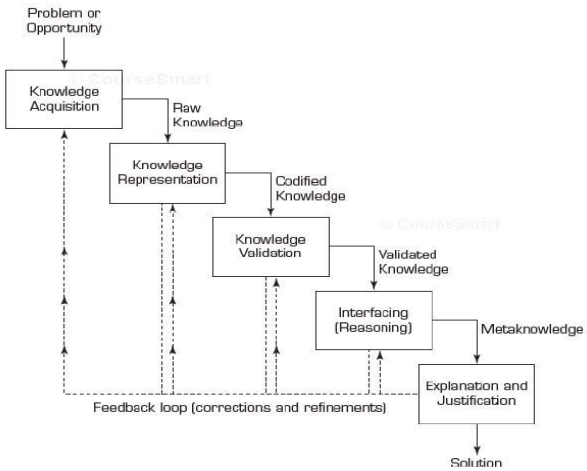
- AI is more permanent. Natural intelligence is perishable from a commercial standpoint, in that workers can change their place of employment or forget information. However, AI is permanent as long as the computer systems and programs remain unchanged.
- AI offers ease of duplication and dissemination. Transferring a body of knowledge from one person to another usually requires a lengthy process of apprenticeship; even so, expertise can seldom be duplicated completely. However, when knowledge is embedded in a computer system, it can easily be transferred from that computer to any other computer on the Internet or on an intranet.
- AI can be less expensive than natural intelligence. There are many circumstances in which buying computer services costs less than having corresponding human power carry out the same tasks. This is especially true when knowledge is disseminated over the Web.
- AI, being a computer technology, is consistent and thorough. Natural intelligence is erratic because people are erratic, they do not always perform consistently.
- AI can be documented. Decisions made by a computer can be easily documented by tracing the activities of the system. Natural intelligence is difficult to document. For example, a person may reach a conclusion but at some later date may be unable to re-

create the reasoning process that led to that conclusion or to even recall the assumptions that were part of the decision.

- AI can execute certain tasks much faster than a human can.
- AI can perform certain tasks better than many or even most people.
- Natural intelligence does have some advantages over AI, such as the following:
- Natural intelligence is truly creative, whereas AI is uninspired. The ability to acquire knowledge is inherent in human beings, but with AI knowledge must be built into a carefully constructed system constrained by a large number of assumptions.
- Natural intelligence enables people to benefit from and use sensory experience directly in a synergistic way, whereas most AI systems must work with numeric and/or symbolic inputs in a sequential manner with predetermined representational forms.

Q5e. Write a short note on knowledge engineering.

Ans: The collection of intensive activities encompassing the acquisition of knowledge from human experts and conversion of this knowledge into a repository are called knowledge engineering. Knowledge engineering requires cooperation and close communication between the human experts and the knowledge engineer to successfully codify and explicitly represent the rules used to solve a problem within a specific domain. The figure below shows the process of knowledge engineering



Following are the five major activities in knowledge engineering:

Knowledge Acquisition – It involves the acquisition of knowledge from human experts, books, documents, sensors or computer files. The knowledge may be specific to the problem domain or to the problem solving procedures.

Knowledge Representation – Acquired knowledge is organized so that it will be ready for use in activity called knowledge representation. This activity involves preparation of a knowledge map and encoding of knowledge in the knowledge base

Knowledge Validation – It involves validating and verifying the knowledge until its quality is acceptable

Inferencing – This activity involves the design of software to enable the computer to make inferences based on the stored knowledge and specifics of the problem. The system can then provide advice to no expert users

Explanation and Justification – This step involves the design and programming of an explanation capability

Q5f. What are the applications of expert system?

Ans: **Interpretation Systems** – Systems that infer situation descriptions from observations. This category includes surveillance, speech understanding, image analysis, signal interpretation and many kinds of intelligence analysis. An interpretation system explains observed data by assigning them symbolic meanings that describe the situation

Prediction Systems – These systems include weather forecasting, demographic predictions, economic forecasting, traffic predictions, crop estimates, and military, marketing and financial forecasting

Diagnostic Systems – These systems include medical, electronic, mechanical and software diagnoses. Diagnostic systems typically relate observed behavioral irregularities to underlying causes

Design Systems – These systems develop configurations of objects that satisfy the constraints of the design problem. Such problems include circuit layout, building design and plant layout. Design systems construct descriptions of objects in various relationships with one another and verify that these configurations conform to stated constraints.

Planning Systems – These systems specialize in planning problems such as automatic programming. They also deal with short and long term planning areas such as project management, routing, communications, product development, military applications and financial planning

Monitoring Systems – These systems compare observations of system behavior with standards that seem crucial for successful goal attainment. These crucial features correspond to potential flaws in the plan.

Debugging Systems – These systems rely on planning, design and prediction capabilities for creating specifications or recommendations to correct a diagnosed problem

Repair Systems – These systems develop and execute plans to administer a remedy for certain diagnosed problem. Such systems incorporate debugging, planning and execution capabilities.

Instruction Systems – Systems that incorporate diagnosis and debugging subsystems that specifically address students need. These systems begin by constructing a hypothetical description of the student that interprets his behavior. They then diagnosis his weakness and identify appropriate remedies to overcome the deficiencies. Finally they plan a tutorial interaction intended to deliver remedial knowledge to the student

Control Systems – Systems that adaptively govern the overall behavior of the system. To do this a control system must repeatedly interpret the current situation predict the future, diagnose the cause of anticipated problem, formulate a remedial plan and monitor its execution to ensure success.