INFORMATION AND KNOWLEDGE ENGINEERING

Abstract

Knowledge acts as a kind of glue to bind learning and information together. We grasp a subject better when we are comfortable with it. Key concepts and core ideas in knowledge engineering, intellectual capital, and organizational learning include knowledge, information, and data. The different dimensions that the terms "information and knowledge engineering" entail are outlined in this article. The reader will get a solid foundation in a wide range of technologies, together with: expert systems, neural networks, genetic algorithms, case-based reasoning systems, data mining, and intelligent agents. This chapter will also be a great source of information for anyone interested in or working with advanced information systems and knowledge engineering advancements.

Keywords: Information, Knowledge Engineering, Expert System, Neural Network, Data Mining, Genetic Algorithms ,Intelligent Agents ,etc

Authors

R. Hemalatha
Assistant Professor
Department of Computer Science
PSG College of Arts & Science
Coimbatore, India
latharaji123@gmail.com

Dr. R. Amutha
Assistant Professor
Department of Information Technology
PSG College of Arts & Science
Coimbatore, India
amutharajan@gmail.com

Copyright © 2022 Authors

I. INTRODUCTION

Information and knowledge engineering examines how knowledge is represented, acquired, reasoned, decided, and applied. Replicating the choices and behaviors of a human expert in a particular discipline through knowledge engineering. This article is intended to aid readers in understanding a wide range of knowledge-based strategies for planning and decision support as well as KBS and their architecture. Activities and self-assessments are included throughout the text to provide readers the chance to reflect on and solidify their comprehension of the concepts as they go along. This article can be utilized in a variety of ways to develop systems in one's own area of expertise as well as to grasp the principles of knowledge-based systems. Here, we described the various knowledge-based systems, such as expert systems, neural nets, case-based reasoning, genetic algorithms, intelligent agents, and data mining.

II. CONTENT

1. Knowledge-Based Systems

- Data, Information and Knowledge
- Difference Between Information and Knowledge
- Importance of Knowledge Engineering
- Knowledge Engineer Skills
- Types of Knowledge
- Process of Knowledge Engineering
- Benefits of Knowledge Engineering
- Types of Knowledge-Based Systems
  - Expert Systems
  - Neural Networks
  - Case Based Reasoning
  - Genetic Algorithms
  - Intelligent Agents
  - Data Mining

2. Data/Information/Knowledge

- **Data:** Data must be processed in order to be meaningful because it is a raw, unorganized fact. Data are unique symbols and characters in their most basic form, and their significance is only recognized when they are coupled to a context. It could be viewed as a collection of records and information used for research or analysis. Data are separate pieces of information. Variables are used to represent data during analytical operations. A human or a machine constantly analyzes data to determine its meaning. Data is thus meaningless. Data consists of numbers, statements, and characters.

  An example of data 17091985 - The only way to start is with these numbers. However, the number series can be encrypted and used to represent information if the information is connected to a context (date of birth: 17.09.1985).

Copyright © 2022 Authors

- **Information:** Information is a collection of data that has been processed to meet the specified needs in a meaningful way. It is altered, structured, or delivered in a specific manner to be meaningful and helpful. Information gives the data some meaning and increases the accuracy of the information. It makes undesirableness certain and reduces uncertainty. As a result, when the data is turned into information, it never contains any irrelevant information. It contains data that is relevant, appropriate, and meaningful. Furthermore, it entails the analysis of raw data, from which knowledge eventually emerged. Take a look at the data examples provided:
  - 3,6,9,15
  - lion, tiger, elephant, rhinoceros,
  - 45.3,72,55.5,62.3,70.0

  The data don't become information until we give it a context or meaning. Everything becomes clear when you are informed that,

  - 3,6,9,15 are the first four answers in the 3 x table.
  - Lion, Tiger, Elephant, Rhinoceros are examples of wild animals.
  - 45.3,72,55.5,62.3 and70.0 are the weights of boy students that are 15 years old.

- **Knowledge:** The term "knowledge" refers to the information that has been collected on a particular fact or individual. The combination of knowledge, experience, and perception benefits the person or the business. Making informed decisions and solving problems are made possible by knowledge. Knowledge influences people's thinking and how they behave. Machines are also capable of making decisions that are mostly based on newly acquired knowledge. Processing information is absolutely necessary for learning.

  Let's assume the following if you want to learn more about this information:
  - 3,6,9,15 are the first four answers in the 3 x table (due to the fact that 3 x table begins at 3 and is going up in threes and the 5 x table have to begin at 5 and go up in fives).
  - A cat isn't considered a wild animal because it's not on the list and it lives as home pet.
  - The biggest boy student is 72 kg .

  The following level will show the variations between those Knowledge Based System components.

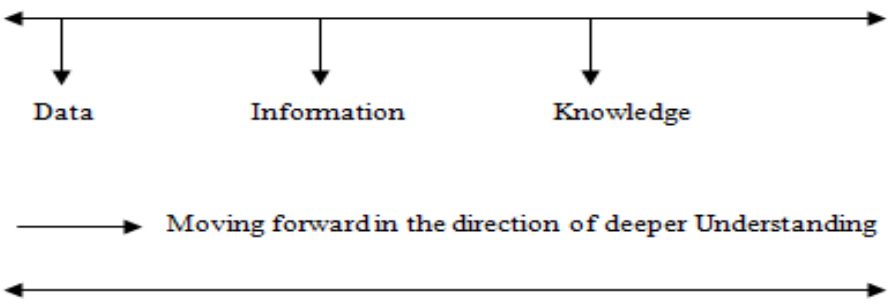

Figure 1: Knowledge level

Copyright © 2022 Authors

1. **Difference between information and knowledge:** The differences of information and knowledge is represented in the following table 1

Table 1: Difference between Information and Knowledge

	Information	**Knowledge**
1	Refined data is information.	The information is valuable knowledge.
2	Data and context	Information, experience, and intuition
3	The result of it is comprehension.	The result of it is understanding.
4	Simple to transfer	To transfer you require learning
5	Makes representation better	Increases consciousness
6	Not all information need be knowledge.	All knowledge is information.
7	Reproduction of information is possible.	Reproduction of knowledge is impossible.
8	Making predictions is impossible with just information.	Making predictions is possible with necessary knowledge.
9	A flow of meaningful messages	Resulting commitments and convictions from these messages
10	Information is a refined form of data that helps in understanding.	Knowledge is the relevant data that aids in drawing conclusions.
11	A statement that transforms the user's perception	It includes memories, principles, perceptions, and background information.

2. **Importance of knowledge engineering:** Knowledge engineering is important for several reasons. Some of the them are,

   - **Speeds decision-making:** These systems' AI capabilities quicken the processing of information and decision-making. A knowledge engineering system may recognize a task and use its knowledge to move towards a logical conclusion.
   - **Handles massive data sets:** As businesses deal with more and more amounts of data, they need a mechanism to properly process it and make decisions, which knowledge engineering provides.
   - **Develops expert systems:** Knowledge engineering results in the expert systems which is used in many sectors, such as engineering, finance, and medical. For instance, financial expert systems provide advice on funding choices, portfolio management, and risk evaluation.
   - **Creates decision-support systems:** Knowledge Engineering provides individuals with the data and instruments they need to make better decisions. Many industries use decision-support systems, including healthcare, manufacturing, and retail.

3. **Knowledge engineer skills:** A knowledge engineer is a professional who works in the field of sophisticated logic programming for computer systems that mimic human decision-making and complex cognitive tasks. One or more pieces of the "knowledge" that is eventually integrated into the technology are provided by a knowledge engineer.

Copyright © 2022 Authors

The following are some of the skills of a knowledge engineer:

- Excellent knowledge of programming languages, such as Python, Java, etc.
- Knowledge of relational database concepts and database development tools.
- A fundamental understanding of software delivery, including release management.
- Working experience of technologies for standards-based knowledge engineering, such as SQL, RDF, SPARQL, OWL, and SHACL.
- Information Structure (IS) and Business Intelligence (BI) knowledge (BI).
- Knowledge of semantic norms and expertise in creating semantic models.
- Understanding of the system development lifecycle and Natural Language Processing (NLP).
- Programming knowledge of Clojure, ClojureScript, or another Lisp language.
- Knowledge with RDBMS, such as Postgres.
- Modeling of business processes and task algorithms.
- The capacity to comprehend and communicate technical information.
- Excellent attention to details and natural interest about technology.
- The capacity to multitask under time constraints.

4. Types of Knowledge

- **Posteriori knowledge:** Posteriori knowledge is information that we gain directly from our own personal experiences. For instance, some types of knowledge are gained via logical or abstract thinking rather than from direct observation. Posteriori is a Latin word that means "that which follows."

- **A prior knowledge:** The opposite of posteriori knowledge is a priori knowledge. Information and facts are true despite the absence of personal experience. A priori knowledge is when someone can determine something to be true just by understanding the meaning of the statement's words or by thinking about what the statement means. Since the phrase "a priori" has Latin roots, it is italicised. A priori is a Latin phrase that translates to "from the former" or "from the one before."

  For instance, it is true that "All dogs are animals," and one can tell this by looking up the definition of "dog," which very certainly mentions this fact. If a person is aware of what the word "dog"

- **Dispersed or distributed knowledge:** Knowledge that is spread out means that no one individual can fully comprehend it. The information is dispersed among numerous individuals. We need to gather a team of subject-matter experts to contribute their knowledge in order to accomplish our aims if we want to combine a great deal of knowledge. For instance: A banker can be aware of the tighter funding conditions in the company's sector. The company's customer service has substantially decreased, as a passerby might note.

- **Domain or expert knowledge:** Deep understanding of a specific field or subject is known as domain knowledge. We could also refer to it as specialist knowledge. Although they may only have a general understanding of everything else, someone

Copyright © 2022 Authors

with domain expertise is extremely knowledgeable within their field. A medical specialist in general medicine may have different talents than one with domain knowledge in surgery. Numerous software developers pursue additional domain knowledge in one particular field. They can work more effectively for an employer on a full-time basis or as independent contractors if they have this additional subject knowledge. It takes domain expertise in the relevant business for software developers to build websites and applications for the health care, finance, marketing, and other industries.

- **Empirical knowledge:** Qualitative or quantitative observations, experiments, or measurements are the sources of this knowledge. Because empirical knowledge can only be acquired through direct sensory experience, it differs from A Posteriori knowledge. A philosophical, introspective, dream, or other a posteriori experience is not allowed. Empirical is a translation of the Greek word empeiria. The Greek word "Empeiria" is translated as experience. Empirical data can be used to demonstrate the fallacy of a claim or to confirm its veracity.

- **Encoded knowledge:** Knowledge that has been encoded is knowledge that has been written down in symbols and signs, such as those found in books, papers, manuals, notes, and rules of practise. By doing this, the information is made accessible to those who can decode it in the future. This information can be used to create an organization's incorporated and predictable pattern of output and behaviour.

- **Tacit Knowledge:** Tacit knowledge is a knowledge that you possess but cannot express is referred to as tacit knowledge. The phrase "tacit knowledge" refers to the skills, information, and talents that people gain through experience but find difficult to communicate orally or in any other way. Typically, someone with tacit knowledge has had it for so long that they are unable to recall when or why they first acquired it. They only acknowledge that what they already know to be true. Learning a new language, invention, aesthetic perception, leadership, and sales are examples of tactics.

- **Explicit knowledge:** The opposite of tacit knowledge is explicit knowledge. It is information that is simple to explain to outsiders. It is information that we might rapidly verbalise and convey to others. Explicit knowledge is information and data that is easy to articulate, record, and exchange. Consider that you have developed a new WordPress application and provided thorough instructions

- **Metaknowledge:** Metaknowledge is information about information. Any information we have about knowledge, such as its characteristics, classification, loss, and acquisition, is referred to as metaknowledge. The phrase is used to define elements that characterise knowledge, including tags, models, and taxonomies. Metaknowledge also encompasses a number of academic disciplines, such as epistemology (the philosophy of knowledge) and bibliography (the study of books). accordingly. In order to create a transactive memory, a group must understand how to apply this metaknowledge.

Copyright © 2022 Authors

- **Descriptive knowledge:** Descriptive knowledge is, by definition, the knowledge of specific facts as they may be expressed by our theories, conceptions, principles, schemas, and ideas. It is not necessary to have a lot of hands-on experience in the field to acquire descriptive knowledge; it can be obtained by memory. A person's comprehension of the world and how it functions is comprised of their descriptive knowledge.

- **Known unknowns:** When there are "known unknowns," we are aware that there is information we do not yet have or fully comprehend. Known unknowns are typically understandable by us because we are aware of what we need to learn and have the ability to look up information. The data analyst, researcher, or web searcher looks for information they desire but don't have: They devise a strategy to gain access to the knowledge they are unaware of. When someone uses a search engine, they frequently ask for information that they know is out there, like how to make bread or write a will, or where they can see a certain movie in their community.

- **L. Unknown Unknowns:** Facts about which we are neither aware of nor aware that we are ignorant are known as unknown unknowns. This knowledge may not be understandable to us and may even exceed our wildest dreams.

  The idea of unknowable unknowns is something we have never even contemplated.

5. **Process of knowledge engineering:** Knowledge engineering has different phases of processing. Various phases of Knowledge engineering represented in fig 2.

- **Task identification:** The work that has to be completed is described in this stage of the knowledge engineering process. A specific problem in a domain or a combination of many problems could be chosen. To enable future procedure, this task must be achievable and the subject matter expert must have a clear understanding of what it involves.

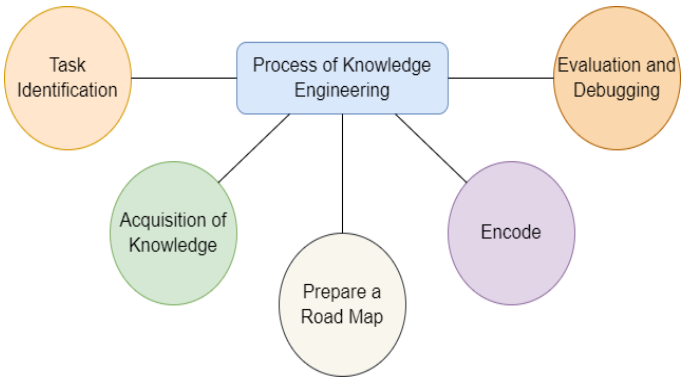

Figure 2: Process of Knowledge Engineering

- **Acquisition of knowledge:** The next step is to extract, structure, and organize knowledge from a single source, typically human specialists, after the problem has been accurately described. Knowledge is obtained from resources such as human

Copyright © 2022 Authors

experts, books, documents, sensors, or computer files. Standard records are utilized for some problems and should be gathered. For instance, a problem involving a heat exchanger requires the standard steam table data of what the value of enthalpy will be at x temperature and y pressure.

- **Prepare a road map:** Once the goal and knowledge base are established, the next step is to create the roadmap by splitting the goal into feasible segments utilizing surveys and the pertinent information base. Here, the subject-matter expert considers how he may decide and what factors might be taken into account at the relevant phases. There may be multiple ways to handle a given issue, and each must be taken into account.
- **Encode:** Now that we have this knowledge, we need to translate it into computer language. In this instance, the information is encoded through the use of various functions, and in a few instances, an algorithm is used to develop a model for a certain task. These models have the same ability to make decisions based on available parameters as an expert, but they must be trained and tested on a large enough set of records and data.

- **Evaluation and debugging:** Every phase of the creation of an expert system requires evaluation and debugging of the model before it is presented to workflow. After each tiny assignment has been analyzed, they are all put together to form a single expert system. This system is once more evaluated using problems that are similar to them, and so any problems are debugged.

6. **Benefits of knowledge engineering:** People no longer need to wait around for a human specialist to help them with their problems in today's environment. They require immediate answers to their questions, which is where knowledge engineering based expert systems, may shine. The following are some benefits of knowledge engineering represented in fig3.

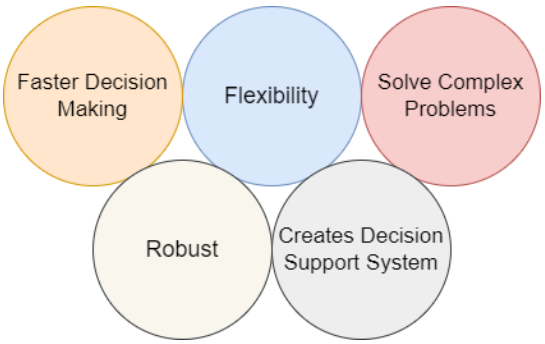

Figure 3: Benefits of Knowledge Engineering

- **Flexibility:** It enables the development of better, flexible, smarter, more effective expert systems.
- **Faster decision making:** It makes it possible to draw knowledge from several areas together to deal with complicated issues and make decisions.
- **Solve complex problems:** A knowledge-based system is a system that supports human decision-making, learning, and action by using artificial intelligence techniques to solve complex problems.

Copyright © 2022 Authors

- **Robust:** Expert systems are models that are produced through knowledge engineering methods have a tendency to be more robust.
- **Creates decision-support system:** Expert systems may analyze questions and provide solutions when they are created using knowledge engineering and Natural Language Programming (NLP), which is similar to how a chatbot functions.

7. **Types of knowledge-based systems:** Knowledge Based System (KBS) manages complicated problems by using reasoning and a knowledge base. The term is inclusive and applies to a variety of system kinds. All knowledge-based systems share two features: an attempt to formally represent knowledge, and a reasoning mechanism that enables it to infer new knowledge. Knowledge-based systems come in a variety of forms, including:

- Expert Systems
- Neural Networks
- Case Based Reasoning
- Genetic Algorithms
- Intelligent Agents
- Data Mining

- **Expert systems:** One of the most common types of knowledge-based systems is the expert system. These systems are helpful for complex analyses, calculations, and predictions because they replicate the decision-making processes of human experts. They offer distinctive explanations for the problems they might be solving in addition to the answers they offer. The knowledge base and the inference engine are two components of an expert system. Data and rules are represented in the knowledge base. To infer new data, the inference engine applies the rules to the previously known data. Working of expert systems represented in fig4. Additionally, inference engines may provide explanation and bug-fixing features .

  There are numerous instances of expert systems. Some of them are given below,

  - **MYCIN:** One of the early backward chaining-based expert systems. It can identify numerous microbes that can result in severe infections and can also recommend medications based on a person's weight.
  - **DENDRAL** was an artificial intelligence-based expert system for chemical analysis. It used spectrographic information to predict a substance's molecular structure.
  - **R1/XCON:** It should select a specific software program to create the computer system the user wants.
  - **PXDES:** Based on the patient's information, it should be simple to determine the kind and severity of lung cancer.
  - **CaDet:** This clinical support system helps patients with early cancer detection.

  The following are the expert system benefits

  - More output and productivity;
  - Better quality
  - Less downtime obtaining limited expertise

Copyright © 2022 Authors

- Adaptability and dependability
- Comprehensive knowledge
- Ease of training
- Improving problem-solving skills

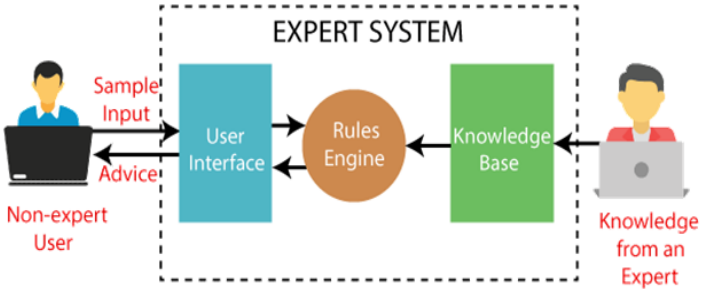

Figure 4: Working of Expert System

- **Neural networks:** An artificial neural network is a computational model that borrows concepts from biology and joins various processing elements and sections. The structure of an artificial neuron is generated by inputs with attached weights, an input function that computes the aggregated input signal to a neuron from all of its inputs, and an activation (signal) function that computes the activation level of a neuron as a function of its aggregated input signal and (possibly) of its previous state. An output signal equal to the activation value is emitted by the neuron's output (the axon). They are also known as connectionist models due to the fundamental role connections play in neural networks represented in fig5..

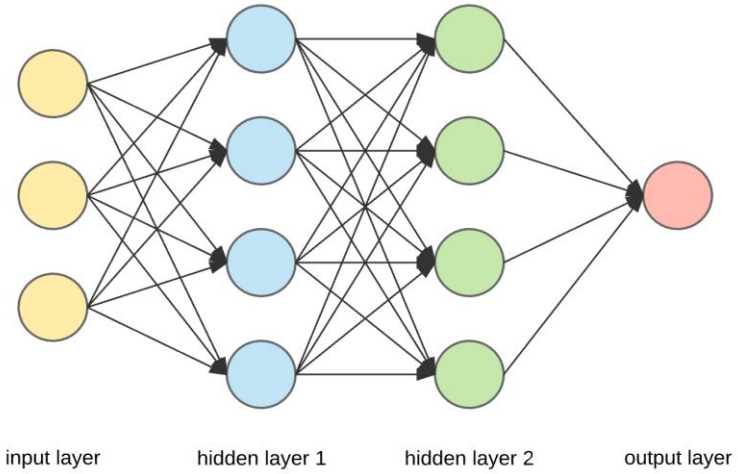

Figure 5: Layered Structure of Neural Networks

The important characteristics of a neural network are,

Learning: A network can be trained using input-output pairs of records (supervised training) or only input records from "no knowledge" at the beginning (unsupervised training). Learning may also require repetition. Learning occurs when the link weights change in a way that teaches the network to produce preferred outputs for known

Copyright © 2022 Authors

inputs. In general, neural networks carry out supervised learning tasks, gaining knowledge from data sets where the correct response has already been predetermined. The networks then gain knowledge by tweaking themselves to determine the correct response on their own, improving the precision of their forecasts.

Generalization: We must be mindful of something known as the generalisation of the neural network whenever we create our own neural networks. This refers specifically to how efficiently our model is able to put the information it has acquired elsewhere after learning from the data given.In the process of training a neural network, some data will be used for training and some data will be saved for measuring the neural network's performance. We can say that a neural network has properly generalised if it performs well on data that it hasn't been trained on.

Massive potential parallelism: It is only a method of distributing the massive and memory-intensive data over several processors, allowing us to train and analyze the information more quickly. Model Parallelism: There are instances when neural network models grow too large to be saved on a single processor. While computation, several neurons "fire" continuously.

Robustness: Even if a few neurons "go wrong," the system as a whole might still function effectively.

These important characteristics make neural networks advantageous for knowledge engineering. Expert systems can be built using neural networks. They might reflect the "hidden" knowledge of an expert system because they could be educated using a set of examples (records).

- **Case based reasoning:** Case-based reasoning (CBR) is an approach that draws on experience to solve new problems by applying previously effective solutions to problems that are similar. Memory, learning, creating plans, and problem-solving are all addressed by CBR, which provides a foundation for a new technology of intelligent computer systems that could address problems and adapt to changing circumstances. The concept behind CBR's "intelligent" reuse of knowledge from previously solved issues or cases is that the more similar problems are, the more similar the solutions will be.

  The case-based reasoning procedure often entails:
  - Retrieve: Drawing from memory an instance that is most similar to the current problem.
  - Reuse: Reusing an answer involved taking a suggestion based on experience and changing it to fit the demands of the new situation.
  - Revision: Analyzing how the solution should be used in the new situation.
  - Retain: Retaining this novel technique to problem-solving in the memory system.

Copyright © 2022 Authors

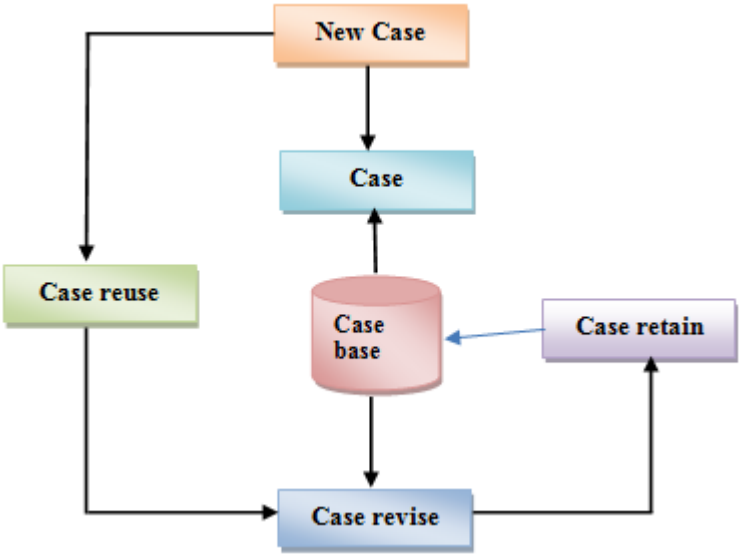

Figure 6: Case based Reasoning

- **Genetic algorithms:** The principles of genetics and natural selection are the foundation of the Genetic Algorithm (GA), a search-based optimization technique. It is frequently used to find the most reliable or close to ideal solutions to complex problems that, in any other circumstance, may take a lifetime to solve. For analyzing vast and intricate data sets, genetic algorithms are fantastic. Since they are excellent at resolving both limited and unconstrained optimization problems, they are regarded as being capable of determining viable solutions to complicated problems. Data centers, electronic circuit design, code-breaking, picture processing, and artificial creativity are just a few of the real-world uses of it.

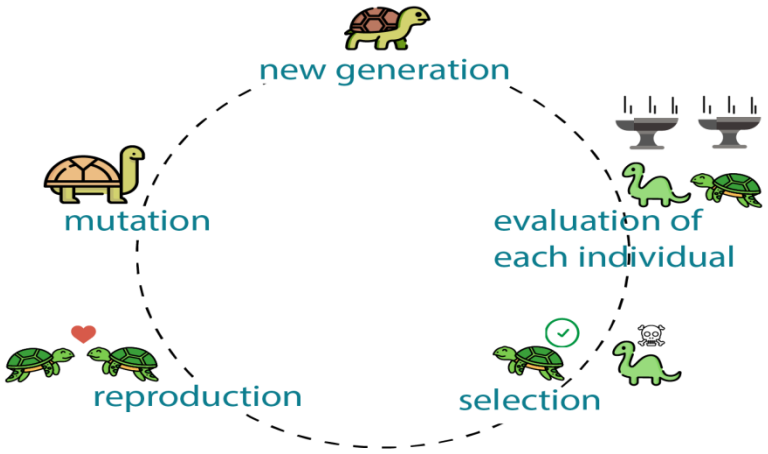

Figure 7: Workflow of Genetic Algorithm

Genetic Algorithm (GA) is a heuristic search algorithm used to tackle search and optimization issues. A subset of evolutionary algorithms, which are utilized in computation, includes this algorithm. These algorithms are more intelligent than

Copyright © 2022 Authors

random search algorithms because they direct the search to the best performing region of the solution space using historical information.

A genetic algorithm has the following benefits:

- It has excellent parallel capabilities.
- It can optimize a variety of problems, including discrete functions, multi-objective problems, and continuous functions.
- It offers solutions that get better over time.
- It requires no derivative information.

- **Intelligent agents:** An Intelligent Agent (IA) is an entity that decides and enables the use of artificial intelligence. It can also be described as a software program entity that, after sensing its environment, performs operations around users or applications. Actuators are used to start actions in that environment. This agent can perform particular, predictable, and repetitive tasks for users or applications because of a limited degree of autonomy. It is also referred to as "intelligent" because of its ability to learn while carrying out tasks. Multiple examples of real-world use of intelligent agents include,

  - **Search, retrieval, and navigation of information:** Intelligent agents improve record access and navigation. This is done by using search engines to locate information. The internet has a lot of data items, so it could take users a while to find a certain data object. This task is quickly completed on behalf of users by intelligent agents.

  - **Repetitive work tasks:** To cut down oncosts, several companies have computerized a few administrative tasks. Customer service and sales are two functional domains that have been digitized. The productivity of workplaces has also been increased by using intelligent agents.

  - **Autonomous driving:** Intelligent agents let self-driving automobiles operate more efficiently. Many sensors are used in autonomous driving to collect data from the environment. These include radar, GPS, and cameras. The environment in this application could be pedestrians, various vehicles, roadways, or street signs. Actions are started using a variety of actuators. For instance, the car's brakes are used to stop it.

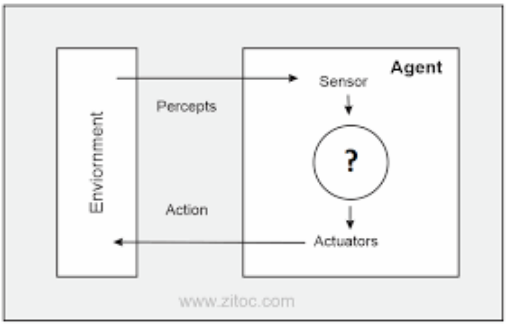

Figure 8

Copyright © 2022 Authors

- **Data mining:** We collect and keep data, and knowledge is what enables us to make informed decisions. Data mining is the process of drawing knowledge from data. The process of collecting data from massive collections of records in order to find patterns, trends, and useful information that can allow the organization to make a decision based on data is known as data mining. Finding knowledge is data mining's ultimate goal. In databases, data mining is known as knowledge discovery (KDD). Data cleansing, Data integration, Data selection, Data transformation, Data mining, Pattern assessment, and Knowledge presentation make the knowledge discovery process.

  The following types of data can be mined for information:

  - **Relational database:** A relational database is a collection of many data sets that are formally organised through tables, records, and columns. Data from a relational database can be accessed in a variety of ways without the need to comprehend the database tables.

  - **Data warehouses:** A data warehouse is a piece of technology that compiles data from several organisational sources to provide useful business insights. The vast amount of data is derived from a variety of sources, including marketing and finance.

  - **Data repositories:** A location for data storage is typically referred to as a data repository. However, a lot of IT professionals use the phrase more specifically to refer to a specific kind of arrangement inside an IT organisation.

  - **Object-relational database:** An object-relational model combines an object-oriented database paradigm and a relational database model. It supports objects, inheritance, classes, etc. Reducing the distance between relational databases and the object-oriented model techniques commonly used in many programming languages, such as C++, Java, C#, and others, is one of the main objectives of the object-relational data model.

  - **Transactional database:** A database management system (DBMS) that has the ability to undo a database transaction if it isn't carried out correctly is referred to as a transactional database.

Knowledge Discovery in Databases

Knowledge Discovery in Databases, refers to the general process of discovering knowledge in data and places emphasis on the "high-level" application of specific data mining techniques. Researchers in the fields of artificial intelligence, databases, statistics, machine learning, pattern recognition, knowledge acquisition for expert systems, and data visualization may be interested in it.

Extraction of knowledge from data in the setting of sizable databases is the overarching objective of the KDD process. The essential steps of KDD (Knowledge Discovery in Databases) are represented in fig:

Copyright © 2022 Authors

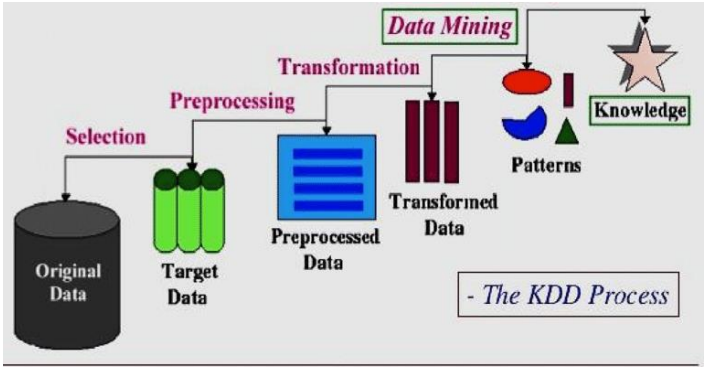

Figure 9: Steps in KDD Process

He essential steps of KDD (Knowledge Discovery in Databases) are:

1. **Understanding original data set:** The ability to understand the challenges we will confront and have the knowledge necessary to provide workable and practical answers is not limited to mathematics and statistics. It's critical to comprehend the features, boundaries, and requirements of the data or information being studied as well as the intended outcomes.

2. **Data selection:** The available data must be gathered to carry out the study and integrate them into one that can help to meet the objectives of the analysis from the set of data obtained and the objectives to be reached already specified. This information is easily obtainable from the same source or is distributed as well.

3. **Pre-processing:** By performing actions that ensure the data's utility, the reliability of the information is assessed at this point. For this, data cleaning is carried out (treatment of lost data or removing outliers). This entails removing variables or attributes that have missing data or information that isn't relevant for this kind of activity, like text or photos.

4. **Data transformation:** At this point, transformations that either entail dimensionality reduction (cutting the number of variables in the data collection) or transformations like changing the values that are integers to categorical help to improve the quality of the data (discretization).

5. **Select the appropriate data mining task:** Depending on the goals that have been established for the process, the appropriate data mining process — whether classification, regression, or grouping — can be selected in this step.

6. **Evaluation:** The next step is to assess the patterns produced and the performance attained after the algorithms have been applied to the data set in order to make sure that it achieves the objectives stated in the first phases. Cross-Validation is a technique that performs data partitioning, splitting the data into training (which will be used to develop the model) and test, to carry out this evaluation.

Copyright © 2022 Authors

7. **Interpretation:** The final step is to simply apply the newly acquired information to the context and start addressing its issues if all the previous processes were carried out properly and the evaluation's findings were satisfactory. If the findings are not satisfactory in any other case, it is required to go back and make adjustments in the earlier stages, starting with the selection of the data and ending with the evaluation stage.

REFERENCES

[1] Gupta, B., Iyer, L.S., & Aronson, J.E. (2000). “Knowledge Management: Practices And Challenges “. Industrial Management and Data Systems, Vol.100, Nos. 1 & 2.
[2] Rajendra Akerkar, Priti Sajja ,"Knowledge-Based Systems", August 2009,published by Jones & Bartlett Learning
[3] D. C. Nazário, M. A. R. Dantas and J. L. Todesco, "Knowledge Engineering: Survey of Methodologies Techniques and Tools", *IEEE Latin America Transactions*, vol. 12, December 2014.
[4] Martin James, "Information Engineering: Introduction" in, New Jersey:Prentice Hall: Engelwood Cliffs, 1990.
[5] Despres, Ch., Chauvel, D. (1999). “Knowledge management(s)”, Journal of Knowledge Management, MCB University Press, Vol.3, No.2: 110-120
[6] Gupta, B., Iyer, L.S., & Aronson, J.E.(2000). “Knowledge Management: Practices And Challenges “. Industrial Management and Data Systems, Vol.100, Nos. 1 & 2.
[7] Speel, P., Schreiber, A. Th., van Joolingen, W., and Beijer, G. (2001), "Conceptual Models for Knowledge-Based Systems", in Encyclopedia of Computer Science and Technology, Marcel Dekker Inc, New York
[8] Studer, R., Benjamins, R.V., and Fensel, D(1998), Knowledge Engineering: Principles and Methods. Data & Knowledge Engineering, 25: p. 161-197.
[9] V. R. Benjamins: Problem Solving Methods for Diagnosis and Their Role in Knowledge Acquisition. International Journal of Expert Systems: Research and Application 8 (2), 1995, pp. 93-120.
[10] D. S. Darai, S. Singh, and S. Biswas, “Knowledge engineering: An overview,” International Journal of Computer Science and Information Technologies, vol. 1, no. 4, pp. 230-234, 2010.
[11] L. F. Lai, “A knowledge engineering approach to knowledge management,” Information Sciences: An International Journal, vol.177, no. 19, pp. 4072-4094, October 2007

Copyright © 2022 Authors