Lecture 1. History, subject, goals of data mining and decision making

The purpose of the lecture: an introduction to a brief history, subject and value of intellectual analysis as a methodology, scientific field, technological discipline and principle of thinking.

Lecture plan:
Introduction
1 History, subject, goals of data mining and decision making
Conclusion

Keywords: systems analysis, subject area, functor, formal language, meaning, word, mathematics, synergetics, informatics, cognitology, branch, systems thinking, definition, theoretical method, information, game theory, author, division, algorithmic aspect, organization, dialectics, deduction, linearization, prototyping, algorithmization, substance, energy, self-organization, man, space, time, transport network, activity, queue, systemic, identification, objective, thinking, systemic, potential, subject analyst, analyst, efficiency criterion, principle deductive consistency, the principle of integrated consideration, the principle of coordinating resources and goals, the principle of conflict-free, knowledge, activity, intelligence, component, reverse, analysis.

Contents of the lecture:

Introduction

We can talk about the onset of the stage of a scientific, system-interdisciplinary approach to the problems of science, education, technology and technology, a stage that focuses not only on material-energy, but also on system-interdisciplinary aspects, construction and research of the system-information picture of the world, about the onset of the stage of systemic paradigms.

Systems analysis, whose foundations are quite ancient, is still a relatively young science (comparable in age, for example, to cybernetics). Although it is actively developing, its defining concepts and terms are not sufficiently formalized (if it is possible at all). Systems analysis is applied in any subject area, including both private and general research methods and procedures.

This science, like any other, aims to study new connections and relationships of objects and phenomena. But, nevertheless, the main problem of our science is the study of connections and relationships in such a way that the objects under study would become more controllable, studied, and the mechanism of interaction of these objects "discovered" as a result of the study - more applicable to other objects and phenomena. The tasks and principles of the systems approach do not depend on the nature of objects and phenomena.

1 History, subject, goals of data mining and decision making

When laying out the foundations of analysis, synthesis and modeling of systems, two main approaches are possible: formal and conceptual-meaningful. The formal approach uses a formal mathematical apparatus of various levels of rigor and generality (from simple relations to operators, functors, categories, algebras). Conceptual and content-based approach - concentrates on the basic concepts, ideas, approach, concepts, opportunities, on the basic methodological principles, uses a "semi-formal" introduction to the essence of the ideas and concepts under consideration. Many ideas and principles of systems analysis, although more accurate, are strict in the formal language of presentation, nevertheless, retain their strength, relevance, and the possibility of effective use in a meaningful language. It should be noted that often one good, understandable example is more important for understanding these principles than rigorous mathematical definitions. In addition, the uncertainty factor in systems analysis limits the applicability of rigorous mathematical formulations and conclusions. Below we will adhere mainly to the conceptual approach, applying where it is deemed necessary, formal definitions and provisions, although we clearly understand that a high degree of formalization is required to present the foundations of a science claiming to be a methodological one, up to the creation of axioms. With this approach, we want to expand the circle of readers to whom this course of lectures will be available and useful. Despite the meaningful formulations and algorithmic procedures of some of the given basic provisions and facts, they are based on a fairly formal foundation.

The word "system" (organism, structure, union, whole, made up of parts) originated in ancient Greece about 2000 years ago. Ancient scientists (Aristotle, Democritus, Plato and others) considered complex bodies, processes and myths of the universe as composed of various systems (for example, atoms, metaphors). The development of astronomy (Copernicus, Galileo, Newton and others) made it possible to move to the heliocentric system of the world, to categories such as "thing and properties", "whole and part", "substance and attributes", "similarity and difference", etc. Further development of the system analysis is influenced by various philosophical views, theories about the structure of knowledge and the possibility of prediction (Bacon, Hegel, Lambert, Kant, Fichte and others). As a result of this development, systems analysis takes the position of methodological science. Naturalists of the XIX-XX centuries. (Bogdanov, Bertalanffy, Wiener, Ashby, Zwicky and others) not only actualized the role of model thinking and models in natural science, but also formed the basic system-forming principles, the principles of the consistency of scientific knowledge, "combined" the theory of open systems, philosophical principles and the achievements of natural science. Systems theory and systems analysis have received modern development under the influence of the achievements of both classical fields of science (mathematics, physics, chemistry, biology, history, etc.) and non-classical fields (synergetics, computer science, cognitology, theories of nonlinear dynamics and dynamic chaos, catastrophes, neuro-mathematics, neuroinformatics, etc.). It is necessary to emphasize the influence of technology (since ancient times) and technology (modernity) on the development of systems analysis, in particular, on its applied branch - systems engineering, on the methodology for designing complex technical systems. This influence is mutual: the development of technology and technology enriches system analysis with new methods, models, environments.

The era of the origin of the foundations of systems analysis was characterized by the consideration of most often systems of physical or philosophical (epistemological) origin. At the same time (Aristotle's) postulate: "The importance of the whole is above the importance of its components" was later replaced by a new postulate (Galileo): "The whole is explained by the properties of its components."

The greatest contribution to the origin and development of systems analysis, systems thinking was made by such scientists as R. Descartes, F. Bacon, I. Kant, I. Newton, F. Engels, A.I. Berg, A.A. Bogdanov, N. Wiener, L. Bertalanffy, Ch. Darwin, I. Prigogine, E. Ashby, A.A. Lyapunov, N.N. Moiseev and others. The ideas of systems analysis were also developed by A. Averyanov, R. Akoff, V. Afanasyev, R. Abdeev, I. Blauberg, N. Belov, L. Brillouin, N. Buslenko, V. Volkova, D. Gvishiani, V. Geodakyan, K. Gein, J. van Guig, A. Denisov, E. Dubrovsky, V. Zavadsky, Y. Klimontovich, D. Kolesnikov, E. Kveid, V. Kuzmin, O. Lange, E. Lutsenko, V. Lektorsky, V. Lefebvre, Y. Liebig, A. Malinovsky, M. Mesarovich, V. Mogilevsky, K. Negoitse, N. Ovchinnikov, S. Optner, J. Paterson, F. Peregudov, D. Pospelov, A. Rapoport, L. Rastrigin, S. Rodin, L. Rosenbluth, V. Sadovsky, V. Segal, V. Simankov, B. Sovetov, V. Solodovnikov, F. Tarasenko, K. Timiryazev, A. Uemov, Y. Chernyak, G. Haken, J. Haldane, G. Shuster, A. Shileiko, G. Shchedrovitsky, E. Yudin, S. Yakovlev, S. Young and many others.

Subject area - a branch of science that studies the subject aspects of system processes and system aspects of subject processes and phenomena. This definition can be considered a systemic definition of the subject area.

System analysis - a set of concepts, methods, procedures and technologies for the study, description, implementation of phenomena and processes of various nature and nature, interdisciplinary problems; it is a set of general laws, methods, techniques for studying such systems.

Systems analysis is a methodology for studying complex, often not well-defined problems of theory and practice.

Strictly speaking, there are three branches of science studying systems:

1. *systemology (systems theory) which studies theoretical aspects and uses theoretical methods (information theory, probability theory, game theory, etc.);*
2. *systems analysis (methodology, theory and practice of systems research), which explores methodological and often practical aspects and uses practical methods (mathematical statistics, operations research, programming, etc.);*
3. *systems engineering, systems engineering (practice and technology of design and research of systems).*

The author is responsible for the term systems technology. This division is rather arbitrary.

All these branches have in common a systematic approach, a systemic principle of research - the consideration of the studied set not as a simple sum of components (linearly interacting objects), but as a set of nonlinear and multilevel interacting objects.

Any subject area can also be defined as systemic.

Example. Informatics is a science that studies information-logical and algorithmic aspects of system processes, system aspects of information processes. This definition can be considered a systemic definition of computer science.

Syst

++

Any subject area can also be defined as systemic.

Example. Informatics is a science that studies information-logical and algorithmic aspects of system processes, system aspects of information processes. This definition can be considered a systemic definition of computer science.

Systems analysis is closely related to synergetics. Synergetics is an interdisciplinary science that studies general ideas, methods and patterns of organization (changes in structure, its space-time complication) of various objects and processes, invariants (unchanging essences) of these processes. "Synergistic" in translation means "joint, acting in concert." This is the theory of the emergence of new qualitative properties, structures at the macroscopic level.

Systems analysis is closely related to philosophy. Philosophy provides general methods of meaningful analysis, and system analysis provides general methods of formal, interdisciplinary analysis of subject areas, identification and description, and the study of their system invariants. It is possible to give a philosophical definition of systems analysis: systems analysis is an applied dialectic.

System analysis provides the following system methods and procedures for use in various sciences and systems:

1. *abstraction and concretization;*
2. *analysis and synthesis, induction and deduction;*
3. *formalization and specification;*
4. *composition and decomposition;*
5. *linearization and selection of non-linear components;*
6. *structuring and restructuring;*
7. *prototyping;*
8. *reengineering;*
9. *algorithmicization;*
10. *modeling and experiment;*
11. *programmed control and regulation;*
12. *recognition and identification;*
13. *clustering and classification;*
14. *peer review and testing;*
15. *15.verification*

16. *and other methods and procedures.*

There are the following main types of resources in nature and in society.

1. *Substance is the most well-studied resource, which is mainly represented by D.I. Mendeleev is quite complete and is replenished not so often. Substance acts as a reflection of the constancy of matter in nature, as a measure of the homogeneity of matter.*
2. *Energy is an incompletely studied type of resource, for example, we do not own a controlled thermonuclear reaction. Energy acts as a reflection of the variability of matter, transitions from one type to another, as a measure of the irreversibility of matter.*
3. *Information is a poorly understood type of resource. Information acts as a reflection of the order, structuredness of matter, as a measure of order, self-organization of matter (and society). Now this concept we designate some messages; below we devote a more detailed discussion to this concept.*
4. *Man - acts as a bearer of the highest level of intelligence and is in the economic, social, humanitarian sense the most important and unique resource of society, is considered as a measure of reason, intelligence and purposeful action, a measure of social beginning, the highest form of reflection of matter (consciousness).*
5. *Organization (or organization) acts as a form of resources in society, a group that determines its structure, including the institutions of human society, its superstructure, is used as a measure of the ordering of resources. The organization of the system is associated with the presence of some causal relationships in this system. The organization of a system can have various forms, for example, biological, informational, ecological, economic, social, temporal, spatial, and it is determined by cause-and-effect relationships in matter and society.*
6. *Space - a measure of the extent of matter (events), its (his) distribution in the environment.*
7. *Time is a measure of reversibility (irreversibility) of matter, events. Time is inextricably linked with changes in reality.*

We can talk about various fields in which a person is "placed" - material, energy, informational, social, about their spatial, resource (matter, energy, information) and temporal characteristics.

Example. Let's consider a simple task - go to school in the morning. This problem, often solved by a student, has all aspects:

1. *material, physical aspect - the student needs to move some mass, for example, textbooks and notebooks to the required distance;*
2. *energy aspect - the student needs to have and spend a specific amount of energy for movement;*
3. *informational aspect - information about the route of movement and the location of the university is needed and it needs to be processed along the way;*
4. *human aspect - movement, in particular, movement by bus is impossible without a person, for example, without a bus driver;*

5. *organizational aspect - suitable transport networks and routes, stops, etc .;*
6. *spatial aspect - moving a certain distance;*
7. *time aspect - this movement will take time (during which the corresponding irreversible changes in the environment, in relationships, in connections) will take place.*

All types of resources are closely related and intertwined. Moreover, they are impossible without each other, the actualization of one of them leads to the actualization of the other.

Example. When wood is burned in a stove, heat energy is released, heat energy is used for cooking, food is used to obtain biological energy of the body, biological energy is used to obtain information (for example, to solve a problem), to move in time and space. A person even during sleep spends his biological energy to maintain information processes in the body; moreover, sleep is a product of such processes.

The social organization and activity of people improves information resources, processes in society, the latter, in turn, improve production relations.

If classical natural science explains the world on the basis of movement, interconversions of matter and energy, then now the real world, objective reality can be explained only taking into account the accompanying systemic, and especially systemic-informational and synergetic processes.

A special type of thinking is systemic, inherent in the analyst, who wants not only to understand the essence of the process, phenomenon, but also to control it. Sometimes it is identified with analytical thinking, but this identification is not complete. An analytical mind can be, and a systems approach is a methodology based on systems theory.

Subject-oriented (subject-oriented) thinking is a method (principle) with the help of which it is possible to purposefully (as a rule, for the purpose of study) identify and actualize, cognize cause-effect relationships and patterns in a number of particular and general events and phenomena. Often this is a technique and technology for studying systems.

Systemic (system-oriented) thinking is a method (principle), with the help of which it is possible to purposefully (as a rule, for the purpose of management) identify and actualize, learn causal relationships and patterns in a number of general and universal events and phenomena. This is often a systems research methodology.

In systemic thinking, a set of events, phenomena (which may consist of various constituent elements) is actualized, investigated as a whole, as one event organized according to general rules, a phenomenon whose behavior can be predicted, predicted (as a rule) without clarifying not only the behavior of the constituent elements, but also the quality and quantity of themselves. Until it is clear how the system functions or develops as a whole, no knowledge of its parts will give a complete picture of this development.

Example. In accordance with the principle of systems thinking, society consists of people (and, of course, of public institutions). Each person is also a system (physiological, for example). A person, in turn, has systems inherent in him as an organism, for example, the circulatory system. When people interact with other people, new systems are formed - family, ethnic group, etc. This interaction can occur at the level of social institutions, individuals (for example, social interactions) and even individual circulatory systems (for example, with direct blood transfusion).

In accordance with the principle of a systems approach, each system influences the other system. The entire surrounding world is interacting systems. The goal of systems analysis is to find out these interactions, their potential and "direct them to the service of man."

A subject analyst (subject-oriented or simply an analyst) is a person, a professional who studies, describes a certain subject area, a problem in accordance with the principles and methods, technologies of this area. This does not mean "narrow" consideration of this problem, although this is often the case.

System (system-oriented) analyst - a person, a high-level professional (expert), who studies, describes systems in accordance with the principles of the systems approach, analysis, i.e. studying the problem comprehensively. He has a special mindset based on multi-knowledge, a fairly wide outlook and experience, a high level of intuition for foresight, and the ability to make reasonable resource-based decisions. Its main task is to help the subject analyst to make the correct (consistent with other systems, not "worsening" them) solution when solving subject problems, identifying and studying the criteria for the effectiveness of their solution.

Necessary attributes of systems analysis as scientific knowledge:

1. *the presence of a subject area - systems and systemic procedures;*
2. *identification, systematization, description of general properties and attributes of systems;*
3. *identification and description of patterns and invariants in these systems;*
4. *actualization of patterns for the study of systems, their behavior and relationships with the environment;*
5. *accumulation, storage, updating of knowledge about systems (communicative function).*

System analysis is based on a number of general principles, including:

1. *the principle of deductive sequence - a sequential consideration of the system in stages: from the environment and connections with the whole to the connections of the parts of the whole (see the stages of system analysis in more detail below);*
2. *the principle of integrated consideration - each system should be integral as a whole, even when considering only individual subsystems of the system;*
3. *the principle of coordinating resources and purposes of consideration, updating the system;*
4. *the principle of non-conflict - the absence of conflicts between parts of the whole, leading to a conflict of goals of the whole and the part.*

Everything in the world is systemic: practice and practical actions, knowledge and the process of cognition, the environment and connections with it (in it). System analysis as a methodology of scientific knowledge structures all this, allowing to explore and reveal invariants (especially hidden) of objects, phenomena and processes of various nature, considering their general and different, complex and simple, whole and parts.

Any human intellectual activity must be inherently a systemic activity, involving the use of a set of interrelated systemic procedures on the way from setting a task, goals, planning resources to finding and using solutions.

Example. Any economic decision should be based on the fundamental principles of systems analysis, economics, informatics, management and take into account human behavior in the socio-economic environment, i.e. should be based on rational, socially and economically sound norms of behavior in this environment.

Failure to use systematic analysis does not allow knowledge (laid down by traditional education) to turn into skills and skills for their application, into skills for conducting systemic activities (building and implementing purposeful, structured, resource-based constructive procedures for solving problems). A systemically thinking and acting person, as a rule, predicts and considers the results of his activities, measures his desires (goals) and his capabilities (resources), takes into account the interests of the environment, develops intelligence, develops the correct worldview and correct behavior in human groups.

The world around us is infinite in space and time; a person exists for a finite time, having finite resources (material, energy, informational, human, organizational, spatial and temporal) in order to achieve the goal.

The contradictions between the unlimited human desire to know the world and the limited (resources, uncertainty) ability to do this, between the infinity of nature and the finiteness of human resources, have many important consequences, including for the very process of human cognition of the surrounding world. One of these features of cognition, which allows you to gradually, step by step resolve these contradictions: the use of an analytical and synthetic way of thinking, i.e. dividing the whole into parts and presenting the complex in the form of a collection of simpler components, and vice versa, connecting simple ones and building, thus, complex. This also applies to individual thinking, and to social consciousness, and to all knowledge of people, and to the process of cognition itself.

Example. The analytic nature of human knowledge is manifested in the existence of various sciences, and in the differentiation of sciences, and in a deeper study of increasingly narrow questions, each of which is interesting, and important, and necessary in itself. At the same time, the reverse process of knowledge synthesis is just as necessary. This is how "borderline" sciences arise - bionics, biochemistry, synergetics and others. However, this is only one form of synthesis. Another, higher form of synthetic knowledge is realized in the sciences about the most general properties of nature. Philosophy reveals and describes the general properties of all forms of matter; mathematics studies some, but also general relations. The synthetic sciences include systems analysis, informatics, cybernetics, etc., combining formal, technical, humanitarian and other knowledge.

Conclusion

So, the dissection of thinking into analysis, synthesis and interconnection of these parts is an obvious sign of the systematic nature of knowledge.

The process of cognition structures the systems, the world around us. Everything that is not cognized at a given moment of time creates "chaos in the system", which, being inexplicable within the framework of the theory under consideration, makes one look for new structures, new information, new forms of representation and description of knowledge, leads to the emergence of new branches of knowledge; this chaos also provides an incentive for the development of skills and abilities of the researcher.

A systematic approach to the study of problems, system analysis is a consequence of the scientific and technological revolution, as well as the need to solve its problems using the same approaches, methods, technologies. Such problems arise in economics, computer science, biology, politics, etc.

Control questions

See the manual on the organization of students' independent work.