NEURAL-NETWORK LIKE LOGICAL-COMBINATORIAL STRUCTURE OF DATA AND THE POSSIBILITIES OF ITS APPLICATION FOR CONSTRUCTING CONCEPT LATTICES

Xenia Naidenova, Sergey Curbatov, and Vladimir Parkhomenko

Annotation. A generalization of algorithm is proposed for implementing the well-known effective inductive method of constructing sets of cardinality (q+1) ((q+1)-sets) from their subsets of cardinality q ((q)-sets). A new neural network-like combinatorial data-knowledge structure supporting this algorithm is advanced. This structure can drastically increase the efficiency of inferring functional and implicative dependencies as like as association rules from a given dataset. Various modes of functioning the network for constructing concept lattices are considered. Some algorithms for constructing concept lattice, inferring good maximally redundant and irredundant classification tests are given with using a generalization process based on Galois's connections and a direct and backward wave of network activity propagation. A method of initial approximation of the network is given for mining classification tests.

The level-wise method of (q+1)-sets' construction is also used for association rule mining. The same principle underlies the algorithm Titanic for generating key patterns and the algorithm TANE for discovering functional dependencies. In all enumerated problems, the same algorithm deals with different sets of elements (items (values of attributes), itemsets, attributes, object descriptions, indices of itemsets) and checks the different properties of generated subsets. These properties can be, for example: "to be a frequent (large) itemset", "to be a key pattern", "to be a test for a given class of examples", "to be a good test for a given class of examples", and some others. If a constructed subset does not possess a required property, then it is deleted from consideration. This deletion reduces drastically the number of subsets to be built at all greater levels. Generally, this algorithm solves the task of inferring all maximal subsets of a set S (i.e., such subsets that cannot be extended) possessing a given PROPERTY. The set S can be interpreted depending on the context of a considered problem. The neural-network like logical-combinatorial structure can be also used in the all enumerated problems.

The possibilities of application of proposed network in text mining are also considered, for example, for extracting associative dependencies between words, extracting topic of text and contexts of topic.