NEUTROSOPHIC FUZZY SETS ANALYSIS IN DATA MINING

Egamberdiev N.A., Toshtemirov Z.N., Berdiyev M.R.

Abstract: in this article, Neutrosoft fuzzy set analysis is considered in the intellectual analysis of data. For this, statistical analysis, machine learning methods, Data Mining, Fuzzy Sets, Neutrosophic Fuzzy sets and Membership Functions are widely used. In addition, currently the most widely used methods of intellectual analysis are mentioned.

Keywords: data mining, fuzzy sets, neutrosophic fuzzy sets, membership functions, machine learning.

In the realm of data mining, extracting valuable insights from vast amounts of data has become a pivotal challenge. Traditional data mining techniques often struggle with uncertainty, ambiguity, and imprecision inherent in real-world datasets. To address these complexities, researchers have turned to advanced methodologies that can handle imprecise and uncertain information. One such promising approach is the application of Neutrosophic Fuzzy Sets (NFS) analysis in data mining. Combining the power of fuzzy logic and neutrosophic reasoning, this methodology provides a robust framework to deal with imprecise and uncertain data, offering a new dimension for knowledge discovery [1-4].

Understanding Neutrosophic Fuzzy Sets:

Neutrosophic Fuzzy Sets (NFS) originated from the fusion of fuzzy logic and neutrosophy, introduced by Florentin Smarandache. Fuzzy logic allows for the representation of imprecision, vagueness, and uncertainty, while neutrosophy handles the concept of indeterminacy, including contradictory or unknown information. By integrating these two theories, NFS offers a more comprehensive representation of real-world uncertainty [5-8].

NFS Analysis in Data Mining:

Handling Uncertainty: In data mining, uncertainty arises due to missing data, noisy measurements, or incomplete information. NFS analysis provides an effective way to model and manage this uncertainty by capturing the imprecision and vagueness associated with data attributes. It allows for flexible membership functions that can assign degrees of membership to different classes, even in the presence of overlapping or contradictory information [9-12].

NFS Membership Function:

The NFS membership function assigns a degree of membership to each element of a universe of discourse. It consists of three components: truth (t), indeterminacy (i), and falsity (f). The membership value ranges from 0 to 1.

$\mu(x) = (t, i, f)$, where $0 \le t, i, f \le 1$.

Fuzzy Rule-Based Classification: NFS analysis enhances traditional fuzzy rule-based classification by incorporating neutrosophic reasoning. By considering the three components of truth, indeterminacy, and falsity, NFS extends the classification capabilities. This facilitates the handling of uncertain and ambiguous situations where precise categorization is challenging. It enables more accurate and nuanced decision-making, resulting in improved classification accuracy [13-15].

NFS analysis extends traditional fuzzy rule-based classification to incorporate neutrosophic reasoning. The classification rules are defined using NFS membership functions. The classification process involves combining the rules and their associated membership values to make decisions.

If x_1 is a_1 and x_2 is a_2 then y is b, where a_1 and a_2 are NFS membership functions for input variables x_1 and x_2, b is the NFS membership function for the output variable b.

Neutrosophic Clustering: Traditional clustering algorithms often struggle with the uncertainty inherent in data. Neutrosophic clustering, an application of NFS analysis, addresses this limitation. It allows for the identification of meaningful clusters in datasets with imprecise boundaries, where data points can simultaneously belong to multiple clusters with varying degrees of membership. Neutrosophic clustering provides a more flexible and adaptive approach, accommodating overlapping and uncertain cluster assignments.

Feature Selection: Selecting relevant features is crucial for effective data mining. However, in real-world datasets, attributes may contain uncertain or irrelevant information. NFS analysis enables the integration of fuzzy logic and neutrosophy in feature selection, allowing for the evaluation of attribute relevance considering imprecise and uncertain data. This helps in identifying informative features and reducing the impact of noisy or irrelevant attributes on the mining process.

Neutrosophic clustering is an application of NFS analysis that handles uncertainty and imprecise boundaries in clustering algorithms. It assigns degrees of membership to data points for multiple clusters, considering overlapping memberships.

Let $x=\{x_1,x_2,...,x_n\}$ be the dataset, and $c=\{c_1,c_2,...,c_k\}$ be the set of clusters.

$\mu(c_i,x_j)=(t,i,f)$, where t,i and f represent the degrees of membership of data point x_j to cluster c_i.

Feature Selection with NFS:

NFS analysis can also be applied to feature selection, where attribute relevance is evaluated considering imprecise and uncertain data. Feature selection algorithms incorporate NFS membership functions to assess the significance of attributes in the mining process.

Let $x=\{x_1,x_2,...,x_n\}$ be the dataset, and $a=\{a_1,a_2,...,a_m\}$ be the set of attributes.

$\mu(a_j)=(t,i,f)$, where t,i and f represent the degrees of relevance of attribute a_j.

Benefits and Challenges:

The application of Neutrosophic Fuzzy Sets analysis in data mining offers several benefits. It enables the representation and management of uncertainty, leading to more accurate and robust knowledge discovery. NFS analysis also enhances decision-making capabilities by considering imprecise, vague, and indeterminate information. It provides a flexible framework that accommodates complex and uncertain real-world datasets.

However, there are also challenges associated with the adoption of NFS analysis in data mining. The complexity of neutrosophic reasoning and the computational requirements for large-scale datasets remain areas of active research. Further developments in algorithms and tools are necessary to fully exploit the potential of NFS analysis.

Neutrosophic Fuzzy Sets analysis presents a promising avenue for data mining in handling uncertainty, imprecision, and ambiguity. By combining the power of fuzzy logic and neutrosophy, this methodology offers a comprehensive approach to deal with real-world datasets.

References:

[1] R.A.Aliev, B.Fazlollahi, R.R.Aliev, "Soft Computing and its applications in business and economics", Springer, Berlin, Heidelberg, New York, 2004.

[2] Dilnoz Muhamediyeva, Nodir Egamberdiyev. An application of Gauss neutrosophic numbers in medical diagnosis // International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities http://www.icisct2021.org/ ICISCT 2021, November 3-5, 2021.

[3] Egamberdiyev N.A., FUZZY REGRESSION ALGORITHM FOR CLASSIFICATION OF WEAKLY FORMED PROCESSES // SCIENCE AND PRACTICE: IMPLEMENTATION TO MODERN SOCIETY MANCHESTER, GREAT BRITAIN, 26-28.12.2020.

[4] D.Mukhamediyeva, N.Egamberdiev, ALGORITHM OF CLASSIFICATION OF MEDICAL OBJECTS ON THE BASIS OF NEUTROSOPHIC NUMBERS, Proceedings of the 4th International Scientific and Practical Conference SCIENCE, EDUCATION, INNOVATION: TOPICAL ISSUES AND MODERN ASPECTS TALLINN, ESTONIA, 4-5.10.2021, pp. 374-380.

[5] Mukhamedieva D.T., Egamberdiev N.A., Zokirov J.Sh., Mathematical support for solving the classification problem using neural network algorithms // Turkish Journal of Computer and Mathematics Education. Vol.12 No.10 (2021).

[6] D.T.Mukhamedieva and N.A.Egamberdiev, APPROACHES TO SOLVING OPTIMIZATION TASKS BASED ON NATURAL CALCULATION ALGORITHMS, Scientific-technical journal, 3(2) 2020, pp. 58-67.

[7] N.A.Egamberdiev, O.T.Xolmuminov, X.R.Ochilov, ANALYSIS OF CLASSICAL MODELS OF CLASSIFICATION OF SLOWLY FORMED PROCESSES, International Scientific-Online Conference: SOLUTION OF SOCIAL PROBLEMS IN MANAGEMENT AND ECONOMY", Spain, October 7, 2022, pp.12-16.

[8] N.A.Egamberdiev, O.T.Xolmuminov, X.R.Ochilov, CHOOSING AN EFFICIENT ALGORITHM FOR SOLVING THE CLASSIFICATION PROBLEM, International Scientific Online Conference: THEORETICAL ASPECTS IN THE FORMATION OF PEDAGOGICAL SCIENCES, October 10, 2022, pp.154-158.

[9] D.Muhamedieva, N.Egamberdiev, O.Xolmuminov, APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR CREDIT RISK ASSESSMENT, «Science and innovation» xalqaro ilmiy jurnali. 2022, 6-soni. 388-395 betlar.

[10] D.Muhamediyeva, N.Egamberdiyev, An application of Gauss neutrosophic numbers in medical diagnosis, International Conference on Information Science and Communications Technologies ICISCT 2021, Tashkent, Uzbekistan, 2021, pp. 1-4.

[11] D.Muhamediyeva, N.Egamberdiyev, A.Bozorov, FORECASTING RISK OF NON-REDUCTION OF HARVEST, Proceedings of the 2nd International Scientific and Practical Conference, SCIENTIFIC COMMUNITY: INTERDISCIPLINARY RESEARCH, Hamburg, Germany, 26-28.01.2021, pp.694-698.

[12] Ф.Нуралиев, О.Нарзуллоев, Н.Эгамбердиев, С.Тастанова, V Международная научно-практическая конференция, RECENT SCIENTIFIC INVESTIGATION, Осло, Норвегия, 26-28 апреля, 2022, с. 447-451.

[13] Мухамедиева Д.Т., Эгамбердиев Н.А., Подходы к решению задач оптимизации на основе алгоритмов природных вычислений, Фарғона политехника институти илмий-техника журнали, 2020, Том 24, №2, с. 75-84.

[14] N.A.Egamberdiyev, M.M.Kamilov, A.Sh.Hamroyev, O‘quv tanlamalar uchun k-o‘lchovli fiksirlangan tayanch to‘plamlar tizimini aniqlash algoritmini ishlab chiqish, Muhammad al-xorazmiy avlodlari, 1(7) 2019, 45-48 betlar.

[15] Xo‘jayev O.Q., N.A.Egamberdiyev, Sh.N.Saidrasulov, Sinflarga ajratish masalasini yechishda samarali usulni tanlash algoritmi, Axborotkommunikasiyalar: Tarmoqlar, Texnologiyalar, Yyechimlar. 1(49)2019. Har choraklik ilmiy-texnik jurnal, 39-43 betlar.