Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

MIXED VARIABLES CLASSIFICATION: A COMPREHENSIVE REVIEW

Kartini Kasim[1*,2], Hashibah Hamid[3], Ayu Abdul-Rahman[4]

[1] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kedah, Malaysia
Email: kartini2929@uitm.edu.my
[2] School of Quantitative Sciences, Universiti Utara Malaysia, Kedah, Malaysia
Email: ms.kartini@gmail.com
[3] Institute of Strategic Industrial Decision Modelling (ISIDM), School of Quantitative Sciences, Universiti Utara Malaysia, Kedah, Malaysia
Email: hashibah@uum.edu.my
[4] School of Quantitative Sciences, Universiti Utara Malaysia, Kedah, Malaysia
Email: ayurahman@uum.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 26.06.2025
Revised date: 16.07.2025
Accepted date: 26.08.2025
Published date: 19.09.2025

To cite this document:

Kasim, K., Hamid, H., & Abdul-Rahman, A. (2025). Mixed Variables Classification: A Comprehensive Review. *Journal of Information System and Technology Management, 10* (40), 216-229.

DOI: 10.35631/JISTM.1040015

Abstract:

Classification involving mixed categorical and continuous variables presents unique challenges due to the differences in data scales and statistical properties. These types of data are commonly encountered across various research fields. In this study, we introduce a systematic review of classification methods designed to handle mixed-type variables. A total of 52 journal articles were selected through a systematic review process based on the related topics, which were identified and screened from the Web of Science (WoS) as well as Scopus databases. The selected articles were reviewed and categorised into two main groups of classification methods: parametric and non-parametric approaches. Among parametric approaches, methods based on the location model were among the most frequently used. In contrast, non-parametric approaches, such as classification trees and machine learning methods, were often employed when distributional assumptions were not met. This review also presents the use of classification methods for mixed-type variables across diverse fields such as medical, psychology, agriculture, and biology. Therefore, this review may serve as a useful reference for researchers working with mixed-type data and emphasises the importance of ongoing methodological development to address the complexity of such data in real-world applications.

Keywords:

Classification, Mixed Variable, Systematic Review

Journal of Information System and Technology Management JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

Introduction

Classification is a process of allocating data to predetermined categories on the basis of characteristics. In many real-world scenarios, researchers are often required to classify observations that consist of both categorical as well as continuous variables, which are commonly known as mixed-type variables. The classification problem involving mixed-type data occurs in a various applications across many areas of study such as engineering (Rezvan, Hamadani, & Shalbafzadeh, 2013), healthcare (Singh & Mantri, 2024) and finance (Giri, De, Dehuri, & Cho, 2021). Over the past few decades, various statistical classification methods have been developed to cater to classification problems involving mixed variables. These methods generally fall under two broad categories: parametric approaches, which rely on specific distributional assumptions and structured model forms (e.g., the location model) and non-parametric approaches, which are more flexible and often used when such assumptions cannot be justified (e.g., classification trees or data mining-based algorithms). Each approach offers distinct advantages and limitations, following the data's nature as well as the intended application. Although there have been numerous studies introducing new methods or adapting existing techniques for mixed-type variable classification, a structured synthesis of these developments remains limited. A comprehensive review is needed to map out which classification methods are most commonly used, how they are categorised, and in what areas of application they have been implemented. The objective of this article is to provide a systematic review of classification methods for mixed-type variables and their applications across various domains. The reviewed methods are classified into non-parametric as well as parametric approaches, allowing a clearer understanding of their methodological foundations. Additionally, this review identifies the application areas in which these methods have been implemented, offering insights into how mixed-variable classification methods are used in practice. The findings serve as a foundation for researchers seeking appropriate methods for datasets containing both categorical and continuous variables.

Material and Methods

The detailed procedure for systematic literature reviews is divided into four subsections, which are identification, screening, eligibility, data abstraction, as well as analysis.

Identification

To choose a few appropriate papers to be used in this report, the steps involved in a systematic review of the reports are as follows. The beginning stage is the keywords identification as well as the search over the related, similar terms according to the thesaurus, dictionaries, and encyclopedias, along with the past research. Owing to this, once all the applicable keywords have been established, search strings have been formulated on Web of Science (WoS) and Scopus (refer to Table 1) databases. In step one of the systematic review process, the current research undertaking was able to retrieve 120 papers in both databases, and the first step concerning the systematic review was complete.

Table 1: The Search String

Database	Search string
Scopus	TITLE-ABS-KEY (("classification" OR "classify" OR "categorization") AND ("discriminant" OR "discriminative" OR "decision rule" OR "rule-based" OR "discrimination") AND ("methods" OR "techniques" OR "approaches" OR "strategies") AND ("mixed variables" OR"mixed data" OR

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

	"categorical and continuous" OR "binary AND continuous")
WoS	(("classification" OR "classify" OR "categorization") AND ("discriminant" OR "discriminative" OR "decision rule" OR "rule-based" OR "discrimination") AND ("methods" OR "techniques" OR "approaches" OR "strategies") AND ("mixed variables" OR "mixed data" OR "heterogeneous variable" OR "categorical and continuous" OR "binary AND continuous")) (All Fields) and Article (Document Types) and English (Languages)

Screening
In the first stage of screening, redundant articles are to be dropped. The second phase involved reading 105 articles guided by several inclusion and exclusion criteria, formulated by the researchers, whereas 15 articles were removed in the first phase. The first criterion was literature (research articles) since this is the primary source of beneficial information. Also, the current study does not refer to the publications that take the form of systematic reviews, meta-syntheses, reviews, meta-analyses, chapters, books, conference proceedings, as well as book series. This review study was limited to English-language studies in which 36 articles in all were disregarded based on specific criteria.

Eligibility
In the third step, referred to as eligible, 69 articles have been set up. At this stage, all the titles and most important contents of the articles were read carefully to make sure that the inclusion criteria were met and could be included in the current study with the ongoing research objectives. Therefore, 17 articles were omitted as they were not associated with mixed variable classification topics. Finally, 52 articles are accessible for review (refer to Table 2).

Table 2: The Searching Selection Criterion

Criterion	Inclusion	Exclusion
Language	English	Non-English
Literature Type	Journal	Journal
Publication	(only research articles)	(book chapter, conference proceeding, review)

Data Abstraction and Analysis
An integrative analysis was conducted in this work as a major assessment strategy and also as a synthesis of multiple research designs composed of quantitative as well as mixed strategies. The main goal is to identify relevant topics and subtopics within this research domain. The beginning phase consists of data collection for thematic development. As illustrated in Figure 1, a total of 52 publications were compiled to extract key assertions or materials related to the present topic of study. In the second stage, we focused on classifying mixed-type variables by identifying and forming significant groupings. The classification method using a discriminant rule for mixed variables is the key topic that has progressed from the method. Next, we carried on in the manner we had set up each subject, as well as the themes, notions, or ideas, henceforth. A log was held across the data analysis to assimilate any kind of views, analyses, conundrums,

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015
as well as other evidence dissimilar to data analyses. Lastly, we compared the findings to identify whether there was any inconsistency in the process of theme design.

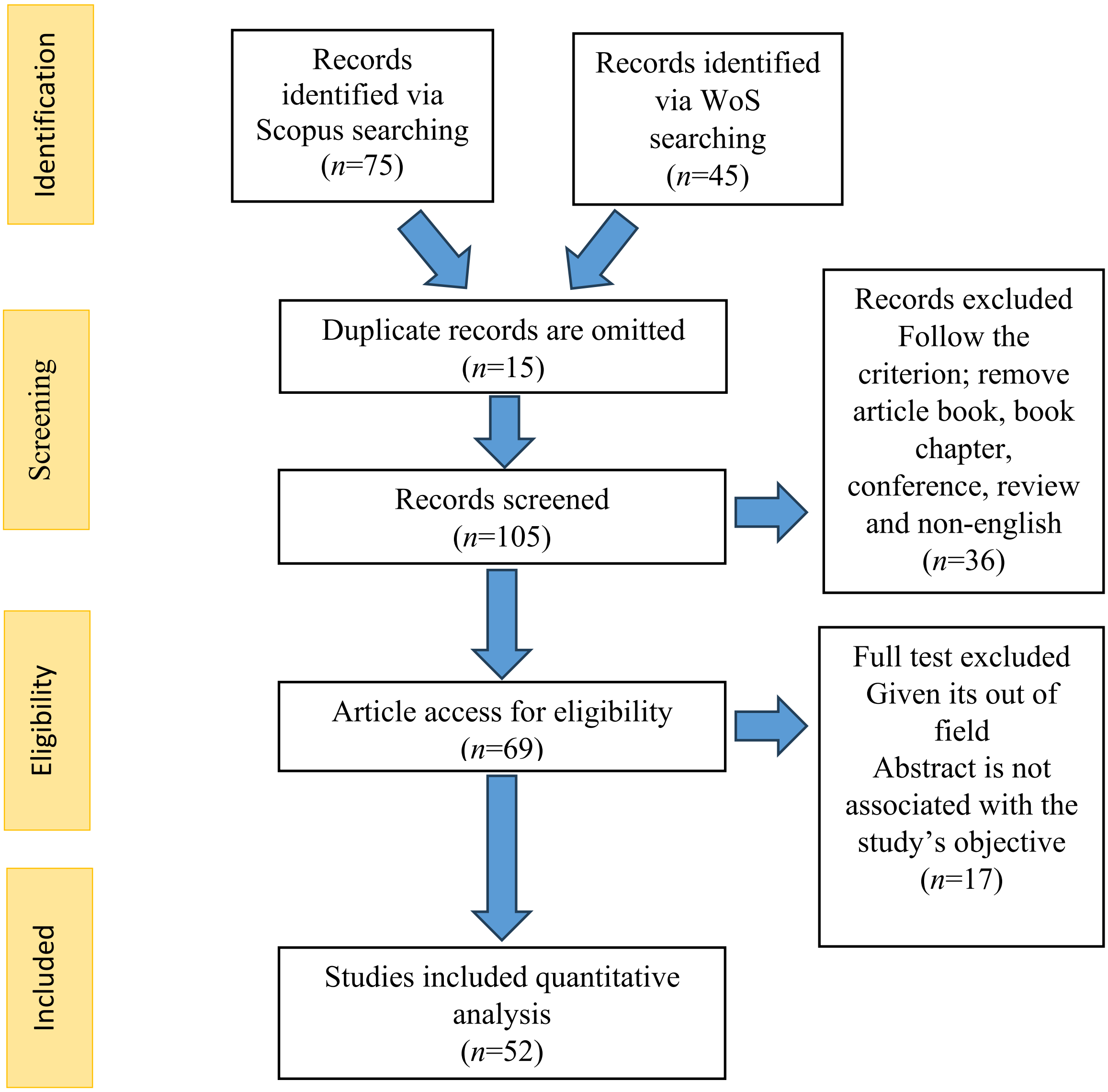

Figure 1: Flow Diagram Regarding the Proposed Searching Study Result and Finding
Source: Moher D, Liberati A, Tetzlaff J. (2009)

Findings

This review explores the classification methods employed by researchers to address mixed-variable classification problems. Through a systematic search and screening process, 53 relevant journal articles were identified and analysed. These articles were categorised under two main themes: (1) statistical approaches and (2) methodological applications. The first theme was further divided into two sub-themes; parametric approaches and non-parametric approaches. The second theme focused on the application areas in which these methods were applied, highlighting the practical usage of mixed-variable classification across multiple fields. A detailed breakdown of these findings is presented in Table 3 and Table 4.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

Table 3: The Research Article Finding Based on the Methodological Development in Mixed Variables Classification Model

No	Author	Year	Title	Approach	Classification Methods
1	Rezvan et al.	2013	Case-based reasoning for classification in the mixed data sets employing the compound distance methods	Parametric approach	LDA
2	Hamid	2010	A new approach for classifying large number of mixed variables	Parametric approach	Location model
3	Krzanowski	1976	Canonical representation of the location model for discrimination or classification	Parametric approach	Location model
4	Hamid	2018	Winsorized and smoothed estimation of the location model in mixed variables discrimination	Parametric approach	Location model
5	Christopher, Terrin, Griffith, D'agostino, & Selker	2001	Predictive performance of missing data methods for logistic regression, classification trees and neural networks	Parametric approach	Logistic regression
6	Ahsan, Mahmud, Saha, Gupta, & Siddique	2021	Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance	Parametric approach & Non-parametric approach	Machine learning (Logistic Regression, LDA, KNN, CART)
7	Krzanowski	1987	A comparison between two distance-based discriminant principles	Parametric approach	Distance-based Rule
8	Mahat, Krzanowski, & Hernandez,	1980	Variable selection in discriminant analysis based on the location model for mixed variables	Parametric approach	Location model
9	Qu et al.	2025	Comparing the predictive discrimination of machine learning models for ordinal outcomes: A case study of dehydration prediction in patients with acute diarrhea	Parametric approach	Regression, machine learning, logistic regression
10	Garnock-Jones & Langer	1980	Parahebe catarractae (scrophulariaceae): Infraspecific taxonomy	Parametric approach	Discriminant analysis
11	Christopher et al.	2001	Predictive performance of missing data methods for logistic regression, classification trees and neural networks	Parametric approach and non-parametric approach	Logistic regression. classification tree, neural network
12	Leung	1996	Error rates for classifying observations based on binary	Parametric approach	Linear discriminant function

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

No	Author	Year	Title	Approach	Classification Methods
			and continuous variables with covariates		
13	de Leon, Soo, & Williamson	2011	Classification with discrete and continuous variables via general mixed-data models	Parametric approach	Location linear discriminant function
14	Cuadras, Fortiana, & Oliva	1997	The proximity of an individual to a population with applications in discriminant analysis	Parametric approach	Linear discriminant function
15	Hamid, Mei, & Yahaya	2017	New discrimination procedure of location model for handling large categorical variables	Parametric approach	Location model
16	Mbina Mbina, Nkiet, & Eyi Obiang	2019	Variable selection in discriminant analysis for mixed continuous-binary variables and several groups	Parametric approach	Discriminant analysis
17	Wolfe et al.	2004	Using nuclear morphometry to discriminate the tumorigenic potential of cells: A comparison of statistical methods	Parametric approach	Regression, machine learning and logistic regression
18	Hamid, Mahat, & Ibrahim	2021	Adaptive Variable Extractions with LDA for Classification of Mixed Variables, and Applications to Medical Data	Parametric approach	LDA
19	Krzanowski	1975	Discrimination and classification using both binary and continuous variables	Parametric approach	Location model
20	Vlachonikolis	1990	Predictive discrimination and classification with mixed binary and continuous variables	Parametric approach	Location model
21	Peddle	1993	An empirical comparison of evidential reasoning, linear discriminant analysis, and maximum likelihood algorithms for alpine land cover classification	Parametric approach	LDA
22	Hamid, Ngu, & Alipiah,	2018	New Smoothed Location Models Integrated with PCA and Two Types of MCA for Handling Large Number of Mixed Continuous and Binary Variables	Parametric approach	Location model
23	Amiri, Khazaei, & Ganjali	2019	Mixtures of general location model with factor analyzer covariance structure for clustering mixed type data	Parametric approach	Location model

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

No	Author	Year	Title	Approach	Classification Methods
24	Leung	2002	Performance of the location linear discriminant function under across-location heteroscedasticity	Parametric approach	Linear discriminant function
25	Leung	1994	Classification of dichotomous and continuous-variables with incomplete samples	Parametric approach	Location model
26	Asparoukhov & Krzanowski	2000	Non-parametric smoothing of the location model in mixed variable discrimination	Non-parametric approach	Location model
27	Liu, Feng, & Pedrycz	2013	Extraction of fuzzy rules from fuzzy decision trees: An axiomatic fuzzy sets (AFS) approach	Non-parametric approach	Fuzzy descision tree
28	Deeva, Bubnova, & Kalyuzhnaya	2023	Advanced Approach for Distributions Parameters Learning in Bayesian Networks with Gaussian Mixture Models and Discriminative Models	Non-parametric approach	Bayesian networks
29	Mahat et al.	2007	Variable selection in discriminant analysis based on the location model for mixed variables	Non-parametric approach	Non-parametric smoothing location model
30	Akgöbek	2013	A rule induction algorithm for knowledge discovery and classification	Non-parametric approach	Data mining
31	Talwalker & Rao	1990	Modified quadratic analysis in prediction with mixed binary and continuous explanatory variables	Non-parametric approach	Modified quadratic analysis
32	Akkus, Sanisoglu, Ugurlu, & Celik,	2010	The criteria for classification tree methods in clinical researches	Non-parametric approach	Classification tree
33	Singh & Mantri,	2024	A clinical decision support system using rough set theory and machine learning for disease prediction	Non-parametric approach	Machine learning, SVM, decision tree
34	Feldman & Gross	2005	Mortgage default: Classification trees analysis	Non-parametric approach	Classification tree
35	Ahsan et al.	2021	Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance	Non-parametric approach	Machine learning
36	Asparoukhov & Krzanowski	2000	Non-parametric smoothing of the location model in mixed variable discrimination	Non-parametric approach	Non-parametric smoothing

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

No	Author	Year	Title	Approach	Classification Methods
37	Yamga et al.	2023	Interpretable clinical phenotypes among patients hospitalized with COVID-19 using cluster analysis	Non-parametric approach	Decision tree
38	Qu et al.	2025	Comparing the predictive discrimination of machine learning models for ordinal outcomes: A case study of dehydration prediction in patients with acute diarrhea	Non-parametric approach	Machine learning
39	Hamid	2019	Handling Outliers and Empty Cells Problems: Winsorized Smoothed Location Model	Non-Parametric approach	Winsorized smoothed location model

Table 4: The Research Article Application Based on the Methodological Development in Mixed Variables Classification Model

No	Author	Year	Title	Data	Area	Methods
1	Hamid et al.	2021	Adaptive Variable Extractions with LDA for Classification of Mixed Variables, and Applications to Medical Data	Cancer	Medical	LDA, MCA, PCA
2	Giri et al.	2021	Biogeography based optimization for mining rules to assess credit risk	Credit risk	Finance	Machine learning
3	Arwatchananukul et al.	2024	Acoustic response discrimination of phulae pineapple maturity and defects using factor analysis of mixed data and machine learning algorithms	Pineapple	Agriculture	Machine learning, factor analysis
4	Pota et al.	2017	Early prediction of radiotherapy-induced parotid shrinkage and toxicity based on CT	Radiotherapy	Medical	Fuzzy classification

Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

No	Author	Year	Title	Data	Area	Methods
			radiomics and fuzzy classification			
5	Shayan, Mohammad Gholi Mezerji, Shayan, & Naseri	2015	Prediction of Depression in Cancer Patients With Different Classification Criteria, Linear Discriminant Analysis versus Logistic Regression	Depression in cancer patients	Psychology	LDA, logistic regression
6	Bishnoi, Al-Ansari, Khan, Heddam, & Malik	2022	Classification of Cotton Genotypes with Mixed Continuous and Categorical Variables: Application of Machine Learning Models	Cotton (Gossypium hirsutum)	Agriculture	Machine learning
7	Stansbury, de Freitas, Wu, & Janik	2015	Can a gray seal (Halichoerus grypus) generalize call classes?	Gray seal	Biology	Generalized linear model
8	Du Pasquier, Naciri, & Jeanmonod,	2015	Morphological analysis of the Silene gigantea complex (Caryophyllaceae) across the Balkan Peninsula, south-western Turkey and Cyprus	Morphology of Silene gigantea	Botany	LDA
9	Arwatchananukul et al.	2024	Acoustic response discrimination of phulae pineapple maturity and defects using factor analysis of mixed data and machine learning algorithms	Fruit maturity (pineapple)	Agriculture	Machine learning

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

No	Author	Year	Title	Data	Area	Methods
10	Baykalov, Bodner, Ostonen, & Rewald	2025	Spectral clustering analysis: discrimination of grass-herb roots and live-dead roots in VISNIR and SWIR regions	Roots of Alopecurus pratensis (meadow foxtail), Urtica dioica (nettle), as well as the rhizosphere	Biology	PCA, K-means clustering, partial least squares discriminant analysis, generalised linear model, distributed random forest
11	Garnockjones & Langer	1980	Parahebe-catarractae (scrophulariaceae) - infraspecific taxonomy	Morphological character of P. catarractae (plant)	Botany	PCA
12	Ahsan et al.	2021	Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance	Heart disease	Medical	Logistic Regression, LDA, KNN, CART, Naive Bayes, XGBoost, SVM, Classifier Gradient, Random Forest Boost, AdaBoost, Extra Tree Classifier

Discussion

In this systematic review, a total of 120 journal articles were screened, of which 52 were identified as directly relevant to the classification of mixed-type variables. This review provided a comprehensive overview of the statistical methods and modelling strategies that have been developed or applied to handle datasets containing both categorical and continuous variables. Among the parametric approaches, the classification method based on the location model was found to be one of the most widely used for simultaneous classification of mixed-type variables. Its popularity is likely due to both its strong theoretical foundation and its ability to simultaneously handle mixed-type variables (Amiri et al., 2019; Hamid, Mei, & Yahaya, 2017; Krzanowski, 1976) making it suitable for various real-world applications. In contrast, classification trees and machine learning were identified as one of the most commonly employed methods under the non-parametric approach, valued for their simplicity, flexibility, and ability to handle mixed variables without strict distributional assumptions (Akkus et al., 2010; Feldman & Gross, 2005; Qu et al., 2025). This review also revealed that studies involving mixed-type variables have been conducted across various research domains. Several classification methods have been

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015
applied to address mixed-variable classification problems in fields such as medicine, psychology, agriculture, and plant biology.

Conclusion

In summary, this review has achieved its main objective of providing a systematic review of classification methods for mixed-type data and their applications across various research domains. These methods can be broadly grouped into parametric and non-parametric approaches, with the location model, classification trees, and machine learning among the most frequently applied. The review also provides useful guidance for practitioners across research domains where mixed-type data are frequently encountered and methodological choices directly affect decision-making. In line with recent advances in mixed-type classification, future research may explore promising directions for handling data complexity more effectively (Azevedo, Rocha, & Pereira, 2024; Imam, Musilek, & Reformat, 2024). Overall, future studies could benefit from developing adaptable models that are better suited to the challenges posed by real-world datasets involving mixed-type data.

Acknowledgement

The authors would like to sincerely thank the reviewers for their insightful opinions and constructive comments, which have significantly enhanced this manuscript. This research obtained no particular grant from any funding agency in the commercial, public, or not-for-profit organizations.

References

Ahsan, M. M., Mahmud, M. A. P., Saha, P. K., Gupta, K. D., & Siddique, Z. (2021). Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. *Technologies*, *9*(3).

Akkus, Z., Sanisoglu, S. Y., Ugurlu, M., & Celik, M. Y. (2010). The criteria for classification tree methods in clinical researches. *Journal of International Dental and Medical Research*, *3*(2), 88–92.

Amiri, L., Khazaei, M., & Ganjali, M. (2019). Mixtures of general location model with factor analyzer covariance structure for clustering mixed type data. *Journal of Applied Statistics*, *46*(11), 2075–2100.

Arwatchananukul, S., Chaiwong, S., Aunsri, N., Kittiwachana, S., Luengwilai, K., Trongsatitkul, T., Saengrayap, R. (2024). Acoustic response discrimination of phulae pineapple maturity and defects using factor analysis of mixed data and machine learning algorithms. *Smart Agricultural Technology*, *9*.

Asparoukhov, O., & Krzanowski, W. J. (2000). Non-parametric smoothing of the location model in mixed variable discrimination. *Statistics and Computing*, *10*(4), 289–297.

Azevedo, B. F., Rocha, A. M. A. C., & Pereira, A. I. (2024). Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Machine Learning*, *113*(7).

Baykalov, P., Bodner, G., Ostonen, I., & Rewald, B. (2025). Spectral clustering analysis: discrimination of grass-herb roots and live-dead roots in VISNIR and SWIR regions. *Plant and Soil*.

Bishnoi, S., Al-Ansari, N., Khan, M., Heddam, S., & Malik, A. (2022). Classification of Cotton Genotypes with Mixed Continuous and Categorical Variables: Application of Machine Learning Models. *Sustainability (Switzerland)*, *14*(20).

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015

Chin, J. L. (2011). Women and Leadership: Transforming Visions and Current Contexts. *Forum on Public Policy: A Journal of the Oxford Round Table*, (2), 1–12.

Christopher, H. S., Terrin, N., Griffith, J. L., D'agostino, R. B., & Selker, H. P. (2001). Predictive performance of missing data methods for logistic regression, classification trees and neural networks. *Journal of Statistical Computation and Simulation*, *71*(2), 115–140.

Cuadras, C. M., Fortiana, J., & Oliva, F. (1997). The proximity of an individual to a population with applications in discriminant analysis. *Journal of Classification*, *14*(1), 117–136.

de Leon, A. R., Soo, A., & Williamson, T. (2011). Classification with discrete and continuous variables via general mixed-data models. *Journal of Applied Statistics*, *38*(5), 1021–1032.

Deeva, I., Bubnova, A., & Kalyuzhnaya, A. V. (2023). Advanced Approach for Distributions Parameters Learning in Bayesian Networks with Gaussian Mixture Models and Discriminative Models. *Mathematics*, *11*(2).

Du Pasquier, P.-E., Naciri, Y., & Jeanmonod, D. (2015). Morphological analysis of the Silene gigantea complex (Caryophyllaceae) across the Balkan Peninsula, south-western Turkey and Cyprus. *Plant Systematics and Evolution*, *301*(8), 2025–2042. Feldman, D., & Gross, S. (2005). Mortgage default: Classification trees analysis. *Journal of Real Estate Finance and Economics*, *30*(4), 369–396.

Garnock-Jones, P. J., & Langer, H. J. (1980). Parahebe catarractae (scrophulariaceae): Infraspecific taxonomy. *New Zealand Journal of Botany*, *18*(2), 285–298.

Garnockjones, P. J., & Langer, H. J. (1980). Parahebe-catarractae (scrophulariaceae) - infraspecific taxonomy. *New Zealand Journal of Botany*, *18*(2), 285–298.

Giri, P. K., De, S. S., Dehuri, S., & Cho, S. B. (2021). Biogeography based optimization for mining rules to assess credit risk. *Intelligent Systems in Accounting Finance & Management*, *28*(1), 35–51.

Hamid, H. (2010). A new approach for classifying large number of mixed variables. *World Academy of Science, Engineering and Technology*, *46*, 156–161.

Hamid, H. (2018). Winsorized and smoothed estimation of the location model in mixed variables discrimination. *Applied Mathematics and Information Sciences*, *12*(1), 133–138.

Hamid, H. (2019). Alternative Methodology of Location Model for Handling Outliers and Empty Cells Problems: Winsorized Smoothed Location Model. *Journal of Mechanics of Continua and Mathematical Sciences*, 90–108.

Hamid, H., Mahat, N. I., & Ibrahim, S. (2021). Adaptive Variable Extractions with LDA for Classification of Mixed Variables, and Applications to Medical Data. *Journal of Information and Communication Technology-Malaysia*, *20*(3), 305–327.

Hamid, H., Mei, L. M., & Yahaya, S. S. S. (2017). New discrimination procedure of location model for handling large categorical variables. *Sains Malaysiana*, *46*(6), 1001–1010.

Hamid, H., Ngu, P. A. H., & Alipiah, F. M. (2018). New Smoothed Location Models Integrated with PCA and Two Types of MCA for Handling Large Number of Mixed Continuous and Binary Variables. *Pertanika Journal of Science and Technology*, *26*(1), 247–260.

Imam, F., Musilek, P., & Reformat, M. Z. (2024). Parametric and Nonparametric Machine Learning Techniques for Increasing Power System Reliability: A Review. *Information (Switzerland)*, Vol. 15.

Krzanowski, W. J. (1975). Discrimination and classification using both binary and continuous variables. *Journal of the American Statistical Association*, *70*(352), 782–790.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015
Krzanowski, W. J. (1976). Canonical representation of the location model for discrimination or classification. *Journal of the American Statistical Association*, *71*(356), 845–848.

Krzanowski, W. J. (1987). A comparison between two distance-based discriminant principles. *Journal of Classification*, *4*(1), 73–84.

Leung, C. Y. (1994). Classification of dichotomous and continuous-variables with incomplete samples. *Communications in Statistics-Theory Aad Methods*, *23*(6), 1581–1592.

Leung, C. Y. (2002). Performance of the location linear discriminant function under across-location heteroscedasticity. *Communications in Statistics-Theory and Methods*, *31*(6), 1031–1044.

Leung, C.-Y. (1996). Error rates for classifying observations based on binary and continuous variables with covariates. *Computational Statistics and Data Analysis*, *21*(6), 627–645.

Liu, X., Feng, X., & Pedrycz, W. (2013). Extraction of fuzzy rules from fuzzy decision trees: An axiomatic fuzzy sets (AFS) approach. *Data and Knowledge Engineering*, *84*, 1–25.

Mahat, N. I., Krzanowski, W. J., & Hernandez, A. (2007). Variable selection in discriminant analysis based on the location model for mixed variables. *Advances in Data Analysis and Classification*, *1*(2), 105–122.

Mbina Mbina, A., Nkiet, G. M., & Eyi Obiang, F. (2019). Variable selection in discriminant analysis for mixed continuous-binary variables and several groups. *Advances in Data Analysis and Classification*, *13*(3), 773–795.

Moher D, Liberati A, Tetzlaff J, A. D. (2009). PRISMA 2009 Flow Diagram. *The PRISMA Statement*, Vol. 6, p. 1000097.

Peddle, D. R. (1993). An empirical comparison of evidential reasoning, linear discriminant analysis, and maximum likelihood algorithms for alpine land cover classification. *Canadian Journal of Remote Sensing*, *19*(1), 31–44.

Pota, M., Scalco, E., Sanguineti, G., Farneti, A., Cattaneo, G. M., Rizzo, G., & Esposito, M. (2017). Early prediction of radiotherapy-induced parotid shrinkage and toxicity based on CT radiomics and fuzzy classification. *Artificial Intelligence in Medicine*, *81*, 41–53.

Qu, K. X., Gainey, M., Kanekar, S. S., Nasrim, S., Nelson, E. J., Garbern, S. C., Schmid, C. H. (2025). Comparing the predictive discrimination of machine learning models for ordinal outcomes: A case study of dehydration prediction in patients with acute diarrhea. *Plos Digital Health*, *4*(5).

Rezvan, M. T., Hamadani, A. Z., & Shalbafzadeh, A. (2013). Case-based reasoning for classification in the mixed data sets employing the compound distance methods. *Engineering Applications of Artificial Intelligence*, *26*(9), 2001–2009.

Shayan, Z., Mohammad Gholi Mezerji, N., Shayan, L., & Naseri, P. (2015). Prediction of depression in cancer patients with different classification criteria, Linear discriminant analysis versus logistic regression. *Global Journal of Health Science*, *8*(7), 41–46.

Singh, K. N., & Mantri, J. K. (2024). A clinical decision support system using rough set theory and machine learning for disease prediction. *Intelligent Medicine*, *4*(3), 200–208.

Stansbury, A. L., de Freitas, M., Wu, G.-M., & Janik, V. M. (2015). Can a gray seal (Halichoerus grypus) generalize call classes? *Journal of Comparative Psychology*, *129*(4), 412–420.

Talwalker, S., & Rao, B. R. (1990). Modified quadratic analysis in prediction with mixed binary and continuous explanatory variables. *Journal of Statistical Planning and Inference*, *26*(1), 47–57.

Vlachonikolis, I. G. (1990). Predictive discrimination and classification with mixed binary and continuous variables. *Biometrika*, *77*(3), 657–662.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 10 Issue 40 (September 2025) PP. 216-229
DOI: 10.35631/JISTM.1040015
Wolfe, P., Murphy, J., McGinley, J., Zhu, Z., Jiang, W., Gottschall, E. B., & Thompson, H. J. (2004). Using nuclear morphometry to discriminate the tumorigenic potential of cells: A comparison of statistical methods. *Cancer Epidemiology Biomarkers and Prevention*, *13*(6), 976–988.

Yamga, E., Mullie, L., Durand, M., Cadrin-Chenevert, A., Tang, A., Montagnon, E., Chassé, M. (2023). Interpretable clinical phenotypes among patients hospitalized with COVID-19 using cluster analysis. *Frontiers in Digital Health*, *5*.