Innovations in Intelligent Systems and Advanced Engineering
Vol. 1, No. 2, 2025 (August)

Optimized Information Integration in Data Mining Using Ensemble Classification

C Priya
Department of Electrical and Electronics Engineering, Sri Sairam Engineering College, Chennai, Tamil Nadu, India
priya.eee@sairam.edu.in

K. Kumuthapriya
Department of Electronics and Communication Engineering, Tagore Engineering College,
Chennai, Tamil Nadu, India
kumuthapriya875@gmail.com

R. Sankar
Department of Electrical and Electronics Engineering, Chennai Institute of Technology, Chennai, Tamil Nadu, India
sankar3673@yahoo.com

Anantha Raman Rathinam
Department of Computer Science and Engineering, Malla Reddy College of Engineering,
Secunderabad, Hyderabad, Telanagana, India
granantha.raman@gmail.com

M. Venkatesan
Department of Artificial Intelligence and Data Science, Panimalar Engineering College. Chennai, Tamil Nadu, India
venkatesan5488@gmail.com

K. Jeevitha
Department of Electronics and Communication Engineering, R.M.K. Engineering College, Chennai, Tamil Nadu, India
kja.ece@rmkec.ac.in

Mohd Miskeen Ali
Department of Computer Science and Engineering, Sphoorthy Engineering College, Hyderabad, Telangana, India
info.miskeen@gmail.com

Dennis Surendar
Mortgage Banking, Fannie Mae, Granite Park VII, 5600 Granite Pkwy, Plano, Texas, USA 75024
dennis.surendar@gmail.com

Abstract: Effectively integrating categorisation rules is crucial in data mining to improve predicted consistency and robustness. Traditional methods, including ensemble techniques and weighted rule aggregation, frequently do not maintain the structural integrity of classifier parameters. This study introduces an optimised information integration framework utilising maximum entropy classifiers, wherein classifier fusion is accomplished via the preservation of probabilistic parameters. The method combines non-parametric wave functions such as Dirichlet and Wishart for handling continuous distributions and uses regression analysis statistics across regulated input dimensions for generated classification. The wave parameters are systematically categorised first-order or higher-order populations, facilitating scalable implementation. Fusion is achieved by the multiplication of hyper-distributions, resulting in streamlined assignment formulae that preserve probabilistic attributes. The proposed strategy guarantees the preservation of critical hyper-distributions during integration, allowing their effective application in future organised training phases. This maximum entropy-based fusion methodology improves classifier interoperability and provides a reliable method for optimised information integration in complex data mining contexts.

Keywords: Data mining, fusion, classification, rule-based mining, maximum entropy

I. INTRODUCTION

In most machine learning platforms, the process of extracting insights, such as classification regulations, from sample information is organised into sequentially parts. Typical instances include intelligent sensor nodes, robotic teams, and virtual agents that adapt dynamically to their environments. At a pivotal juncture, the aggregated information acquired from several classifiers necessitates fusion prior to its successful application to novel input data. An illustrative example of this methodology is seen in computer network intrusion detection systems, as emphasised in [1]. In these situations, the transmission of raw data is restricted by limited communication capacity, and dependence on a centralised unit creates risk, since a singular failure might incapacitate the entire system. For this reason, most data extraction tasks are decomposed into subtasks to address constraints such as memory usage or processing time. The locally derived knowledge is then reinforced or aggregated at a later stage, while minimizing communication overhead. Figure 1 shows the fundamental node structure that is utilised in this procedure.

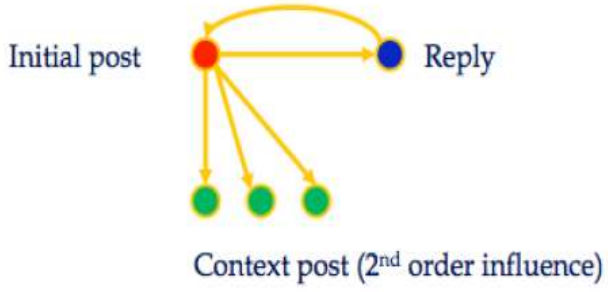

Figure 1: Impact of Initial, Reply, and Contextual Posts in Discussion Threads

The discussion centers on classification problems, highlighting the role of probabilistic generative classifiers in addressing them. These classifiers produce outputs interpretable as conditional probabilities of classes given specific input samples. Generative models aim to capture the underlying mechanisms responsible for generating observed data, typically formulated through Bayes' theorem. In this framework, the input space is represented by a random multivariate vector a (e.g., $a \in T^C$), while c denotes the class variable ($e.g., x_i \dots X$). The distribution $q(a)$ is often expressed using a Density Network (DN), whereas c is assumed to follow a mixing distribution [2]. In contrast, discriminative classifiers, such as support vector machines, do not attempt to model data generation. Instead, they focus directly on maximizing classification accuracy for unseen samples by learning decision boundaries that best separate classes.

Generative and discriminative classifiers each offer distinct strengths and limitations, and in many practical scenarios, they are applied in combination [4]. The posterior probability $q(x|a)$ plays a central role in such integrations, particularly within ensemble methods. When multiple classifiers are combined, posterior probabilities enable rejection strategies by withholding decisions if no class probability surpasses a set threshold or when unfamiliar conditions are encountered in complex environments. Generative classifiers, however, face challenges such as overfitting due to the often-large number of parameters. Furthermore, mismatches between assumed and actual data distributions can lead to significant declines in performance, making their reliability highly dependent on the application context [5]. In the context of fusing or merging generative probabilistic classifiers, three principal strategies can be distinguished:

1. At the output level: Classifiers are combined based on their posterior probabilities. This approach is straightforward for probabilistic classifiers, since outputs are already expressed as probabilities.

2. At the model level: Classifiers are integrated by combining model components, often through convergence and re-normalization of probability distributions.

3. At the parameter level: Modules or rules are fused by averaging parameters when they are adequately comparable.

This paper focuses on the third approach, proposing a new method for fusing variable parameters. Although parameter averaging may seem simple, challenges arise when dealing with multivariate distributions, such as Gaussian distributions with covariance matrices. In these cases, covariance matrices must be decomposed into components representing scaling and rotation. The uncertainty around each parameter's real value is represented by a hyper-distribution. This includes mixture weights, means, and covariance matrices. A more sophisticated method is made possible by second-order distributions. The normal-gamma distribution, seen in Figure 1, is a second-order distribution that maps to the normal distribution's "mean" and "variance" parameters. A second-order distribution is required for every part of a Gaussian classifier [6].

II. RELATED WORKS

The area of knowledge fusion is the one that is being discussed here. The following levels or categories can experience fusion:

- ✓ To provide a more comprehensive set of results, data or information produced from data might be combined.
- ✓ Distributed models provide the ability to merge models or parts derived from experimental data.
- ✓ These models can be coupled to gain insights or produce findings across geographical and temporal dimensions, such as in temporal and spatial data mining.

This study concentrates on the second category. The concept of probabilistic information fusion is particularly relevant [7]. Several variants exist in the literature, but the most notable approaches are probabilistic concurrent estimation techniques or the fusion of independent probability distributions, such as in the independent probability pool method [8]. Applications of this technique have been explored in fields such as visualization, automation, and fault detection. It is fundamentally distinct from first-order approaches, as it involves second-order distributions [9].

Combining classifier at classification stage is as simple as combining their labels or feeding their outputs into pre-trained models of decision-making [10]. Work in the second category often relies heavily on the representation of information. Studies have focused on the merging of constraints, which is one way that knowledge is commonly presented. Two primary research directions emerge from this perspective [11]. However, graphical fused frameworks provide another way to depict information such as Bayesian networks, topic graphs, or related structures. Research into the fusion of the "foundations" of systems is relatively limited. For example, in neural networks, studies have explored radial basis function networks and probabilistic fusion methods that rely on hidden-layer neurons and conjugate priors. Such research provides a more detailed framework than consensus-based neural network approaches, particularly in terms of formula derivation and implementation for classifier fusion [12].

The Iowa Gambling Task has been used in related research to examine impulsiveness, loss sensitivity and reward sensitivity, making decisions, with the BIS/BAS scale applied to assess the effects of impulsivity [13]. In computer vision, texture features extracted from Local Binary Patterns (LBP) and its completed modelling have been applied for classification models [14]. SVM classifiers have been used to determine optimal hyperplanes that separate feature spaces with maximum margin. Complex texture features have also been applied to fundus image classification [15].

III. PROPOSED METHODOLOGY

Multiple components show the proposed classifier's understanding of one data process "generation". If classifier rules are used, every cell divided collection fulfils practice. This paper explains how two or more

classification models and their information may become a classifier using combination and mixed measures. Training hyper distributions are used in the fusion method. During fusion, these hyper-distributions were preserved with certain benefits in comparison to linear combining the proposed classifier parameters. For uncertain fused classifier, add the high energy variance. Define and multiply the hyper-distributions of two classifiers (e.g., process models) to fuse them. As seen in the classifier, components from the first and second classifiers are combined if they do not exist in each other. This means this element sets are combined and mixing coefficients are adjusted. Figure 2 shows the fusion.

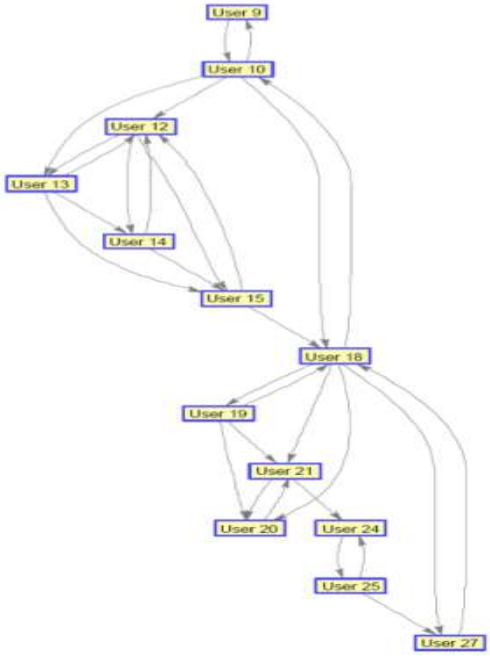

Figure 2: Fusion Process

The general classifier is the one that is derived from all the mixture data. The algorithm explains the proposed technique in a very casual way. In this part, when the necessary equations have been derived, the fusion formulas will be described. The connected nature of each operator makes it possible to quickly merge multiple classifiers into a single total classifier via repeatedly merging pairs of classifiers. It is possible to combine the two classifiers into one. Here, it is assumed, among other things, that the input spaces and categorised and continuous dimensions of the classifiers to be integrated are equal. Figure 3 depicts the proposed paradigm. Some similar components may also be present in two simulated classification methods that are otherwise identical. The ability to mix any two identical components is now an absolute must. Joint probability parameters are available in the sets of data for training and their respective representations when two classifiers self-train using separate datasets. It was presumed that all brackets are analogous, given the use of identical previous data in all scenarios and the ambiguity surrounding the measurement criteria, for instance. It is not a stringent restriction, either.

This encompasses two succeeding populations. In conclusion, the succeeding background possesses the same structural form as the two merged posters. Important for a pair of causes: first, it makes it easier to quickly set back criteria, and second, the integrated rear makes it possible to recover a designation. This is important. We now have the basic idea of how to combine two brow ridges. It is important to identify which part of the initial categorisation depends on which component since every aspect has a posterior probability. See whether the two things in the second category are comparable by doing a comparison test. An element is responsible for making

this happen. When the degree of similarity between two components is greater than an end-user identification, the components are fused. When deciding whether to fuse two components, the similarity test must be symmetrical. It is also possible to use other methods. One such use case is to ignore the covariance matrices and instead focus on the separation between the two standard distribution centres. Although it must be decided in accordance with the order, this subject is yet unattended.

IV. RESULTS AND DISCUSSIONS

The hyper-distributions that are generated during training are utilised by the fusing approach. Compared to just combining the algorithm's parameters linearly, unique hyper-distributions remain intact during fusion and offer clear benefits. When parameters of fused classifier are indeterminate, such as variance in high-dimensional data, these can also be included into the process. Table 1 presents the accuracy analysis values of the proposed system. Typically, two classifiers may be integrated by delineating and multiplying their corresponding hyper-distributions, which reflect the fundamental generating mechanisms. If certain elements of one classifier are absent in the other, they are combined and preserved in the resultant fused classifier. As seen in Figure 3, this results in the formation of a composite set of components, with the mixing coefficients modified correspondingly.

Table 1: Accuracy Analysis

Training Accuracy (1st Classifier)	Training Accuracy (2nd Classifier)
82.56	81.59
78.63	72.58
85.56	79.35
86.23	77.63
90.23	74.53

Figure 3: Classification Outcome

The classifier obtained from all mixture parts is referred to as a universal classification. The algorithm outlined above provides a general summary of the proposed methodology. This section delineates the related fusion formulae after the derivation of the requisite equations. Table 2 presents the examination of execution time dependent on parameters. Due to the associative nature of the proposed operators, it is possible to efficiently merge pairs of classifiers to create a single unified classifier. Thus, two classifiers can be sequentially integrated into a singular consolidated classifier.

Table 2: Parameter Based Execution Time Analysis

Number of Parameters	Execution Time (ms)
2	214.681
4	220.048
8	225.549
12	231.187

This section operates under a fundamental assumption: all classifiers to be integrated must utilise the same reference framework, with equal locations, scales, and uniform dimensions. If the optimisation model yields identical schemes, multiple-member agreements may arise. The examination of execution time related to this assumption is depicted in Figure 4. This stipulation guarantees the efficient fusion of all pairings of analogous components. The situation where two CMMs are trained independently using different portions of the training data, their resultant models may include shared joint probability parameters. In these instances, it is presumed that all components are comparable either due to being trained with identical prior distributions or because uncertainty over the estimated parameters must be explicitly reflected.

Figure 4: Execution Time Analysis

This conclusion is not straightforward; it includes two consecutive distributions. In conclusion, the new posterior will maintain the identical structural structure as the two combined posteriors. This is important for two reasons. Initially, it facilitates the more effective establishment of the criteria for the posterior. Second, the classifier output can be directly inferred from the fused posterior, which is of practical importance. Thus far, the fundamental concept of fusing two posterior distributions has been outlined. Since each component has its own posterior distribution, it is necessary to determine which component of the first classifier corresponds to, or depends on, a component of the second classifier. The response time analysis values for this process are presented in Table 3.

Table 3: Response Time Analysis

Datasets	Response Time (ms)
Dataset-A	90.235
Dataset-B	91.265
Dataset-C	91.478

Similarity Assessment - The objective is to ascertain if two elements of the secondary classifier exhibit adequate similarity. This is accomplished by evaluating separate components, which are merged only when their resemblance surpasses a certain threshold set by the user. To find out if two CMM parts should be combined, the similarity test must be balanced, rather than depending on a unilateral criterion such as "H." Various approaches may be employed for this objective, and their respective reaction times across datasets are depicted in Figure 5. In certain situations, similarity may be assessed predominantly by the distance between the centres of two normal

distributions, disregarding their covariance matrices. This method is context-dependent and must be selected based on the application's unique requirements.

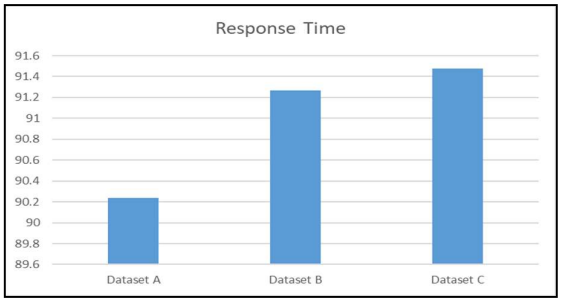

Figure 5: Response Time for Different Datasets

V. CONCLUSION

A unique approach is presented for integrating two maximum entropy classifications using a fusion process that operates directly at the hyper-distribution level. The methodology yields a CMM trained by the Bayesian variational inference technique, guaranteeing consistency in probabilistic integration. The extension beyond two classifiers is accomplished seamlessly, as the knowledge-fusion process may be applied repeatedly without modifying the core architecture. Fusion is regulated by uniform decision factors, such as a predetermined threshold, which stay constant during transformations and facilitate methodical integration. Although the notion seems technically straightforward, its practical use for extensive applications necessitates more investigation. Modifications to the fusion formulae are feasible when classifiers are predetermined, enabling adaptability across various settings. The concept is particularly pertinent for distributed data mining, where partitioned datasets require efficient management under stringent communication limitations. Additionally, it facilitates situations where information is processed locally at the source of creation. A possible use is collaborative learning, in which locally generated rules are shared to improve bigger distributed systems, thus augmenting scalability and knowledge transfer.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1]. M. Zhang, H. Song, S. Lv, Y. Li, X. Yu, and J. Bao, "Research on the Multi-sensors Information Fusion Technique Based on the Neural Networks and its Application," Int. Workshop on Knowledge Discovery and Data Mining, pp. 93–96, 2009.

[2]. B. Verma, and A. Rahman, "Cluster Oriented Ensemble Classifier: Impact of Multi-cluster Characterisation on Ensemble Classifier Learning," IEEE Tr. on Knowledge and Data Engineering, vol. 24, no. 4, pp. 605–618, 2011.

[3]. X. Ceamanos, B. Waske, J. A. Benediktsson, J. Chanussot, M. Fau- vel, and J. R. Sveinsson, "A Classifier Ensemble Based on Fusion of Support Vector Machines for Classifying Hyperspectral Data," International Journal of Image and Data Fusion, vol. 1, no. 4, pp. 293–307, 2010.

[4]. P. Gray, A. Preece, N. Fiddian, W. Gray, T. Bench-Capon, M. Shave, N. Azarmi, I. Wiegand, M. Ashwell, M. Beer, Z. Cui, B. Diaz, S. Embury, K. Hui, A.C. Jones, D.M. Jones, G.J.L. Kemp, E.W. Lawson, K. Lunn, P. Marti, J. Shao, P.R.S. Visser, "KRAFT: Knowledge Fusion from Distributed Databases and

Knowledge Bases," in Proceedings of the 8th International Workshop on Database and Expert Systems Applications, pp. 682–691, 1997.

[5]. K. Ying Hui, and P. Gray, "Constraint and Data Fusion in a Distributed Information System," in Advances in Databases, ser. Lecture Notes in Computer Science, S. Embury, N. Fiddian,W.Gray, and A. Jones, Eds. Springer Berlin, Heidelberg, vol. 1405, pp. 181–182, 1998.

[6]. K. Y. Hui, "Knowledge Fusion and Constraint Solving in a Distributed Environment," Ph.D. dissertation, Department of Computing Science, University of Aberdeen, 2000.

[7]. G. Pavlin, P. De Oude, M. Maris, J. Nunnink, and T. Hood, "A Multi-Agent Systems Approach to Distributed Bayesian Information Fusion," Information Fusion, vol. 11, no. 3, pp. 267–282, 2010

[8]. E. Santos Jr, J. Wilkinson, and E. Santos, "Bayesian Knowledge Fusion," in Proc. of the 22nd Int. FLAIRS Conf., 2009, pp. 559–564, 2009.

[9]. Y. Wang, B. Wu, and J. Hu, "A Semantic Knowledge Fusion Method Based on Topic Maps," in Workshop on Intelligent Information Technology Application, pp. 74–76, 2007.

[10]. H. Lu and B. Feng, "An Intelligent Topic Map-Based Approach to Detecting and Resolving Conflicts for Multi-Resource Knowledge Fusion," Information Technology Journal, vol. 8, no. 8, pp. 1242–1248, 2009.

[11]. A. Smirnov, M. Pashkin, N. Chilov, and T. Levashova, "KSNET- Approach to Knowledge Fusion from Distributed Sources," Computing and Informatics, vol. 22, no. 2, pp. 105–142, 2003.

[12]. O. Buchtala, and B. Sick, "Techniques for the Fusion of Symbolic Rules in Distributed Organic Systems," in IEEE Mountain Workshop on Adaptive and Learning Systems, pp. 85-90, 2006.

[13]. G. Prakash, and A. Khan, "Investigate the Role of Impulsivity in Decision Making During Gambling Task: Case Study," International Journal of MC Square Scientific Research, vol. 4, no. 1, 2013

[14]. S. Murugan, T.R. Ganesh Babu, and C. Srinivasan, "Underwater Object Recognition Using KNN Classifier," International Journal of MC Square Scientific Research vol. 9, no. 3, pp. 48-52, 2017.

[15]. A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," International Journal of Biomedical Imaging, vol. 2013, no. 1, pp. 1-9, 2013.