

Iterative fuzzy-prototype conceptualization of multifactorial constrained prior data spaces for intelligent decision systems*

Serge Dolgikh^{1,†}, Oksana Mulesa^{2,3,*} and Volodymyr Sabadosh^{3,†}

¹ National Aviation University, Lubomyra Huzara 1, Kyiv, Ukraine

² University of Presov, Presov, Slovakia

³ Uzhhorod National University, Universytetska St 14, Uzhhorod, Ukraine

Abstract

In the field of artificial intelligent decision systems, the challenge of learning to construct effective decisions in problems, scenarios and environments characterized by significant uncertainty is encountered commonly. An extensive body of research has been devoted to the development of learning processes and methods which are able to operate within the constraints of uncertainty often described by the shortage of prior information about the problem distribution while having the ability to produce effective decisions and improve their quality in the process. Due to the nature of the constraints in this type of problem, the necessary core capacity of such methods is the ability to extract maximum information from the problem data, including in the raw form and utilize it effectively for the construction of correct, i.e., empirically successful decisions. In this work, we propose and demonstrate an intelligent process of analysis and construction of the conceptual structure of problem data in the “constrained prior” context that requires effective learning with minimal prior data, based on the determination of a structure of probabilistic, “fuzzy” prototype classes/regions. The process and application of iterative learning, starting with minimal sets of problem data is demonstrated with a model dataset of images of basic geometric shapes. The proposed approach demonstrated an effective ability to learn the conceptual structure of problem data with minimal samplings and improve the quality of learning and associated decisions over learning iterations.

Keywords

Intelligent decision systems, concept learning, prototype learning, clustering, fuzzy clustering

1. Introduction

Intelligent decision systems which find applications in many functions and domains of modern society and technology depend on correct interpretation of inputs, expressed in certain measurable parameters that described the data space of the problem. It is logical that similar inputs or observations should induce similar decisions: this logic allows a decision that has been verified as correct and effective for one input sample to be applied to the class of inputs that are essentially similar to it, allowing for effective (the decisions produced by the system are consistent) and efficient (completely new decisions do not need to be constructed for every new input) process and models of constructing decisions.

However, the problem of how the relationship of essential similarity between inputs in general data spaces can be determined appears to be not so trivial. Particular challenge is presented by the cases and scenarios where the knowledge or information about the distribution of data points in the problem space is not available “at prior” that is, to a system in the training regime before it can be set in operation as is the case with conventional methods of supervised classification [1]. In such

Intelitsis'25: The 6th International Workshop on Intelligent Information Technologies & Systems of Information Security, April 04, 2025, Khmelnytskyi, Ukraine

* Corresponding author.

† These authors contributed equally.

✉ sdolgikh@kai.edu.ua (S. Dolgikh); oksana.mulesa@unipo.sk (O. Mulesa); vsabadosh@gmail.com (V. Sabadosh)

ORCID 0000-0001-5929-8954 (S. Dolgikh); 0000-0002-6117-5846 (O. Mulesa); 0009-0006-9933-444X (V. Sabadosh)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

cases that can be designated as “learning with constrained prior” intelligent systems must possess the ability to bring out, determine or calculate the relationship of similarity directly from samplings of data in the problem space and without massive prior information about characteristics of its distribution [2]. Developing approaches to deal with this type of problems is the subject of this work.

2. Prior work

The subject of this study lies at the conjunction of several actively researched directions and fields in data science and theory and practice of intelligent systems. The region of the problems we approach here can be defined as learning with constrained prior information, “learning with constrained prior” problem, that is, data with limited and/or insufficient prior knowledge about the distribution of the problem to employ conventional methods of machine intelligence. In this regard, methods of self-supervised, unsupervised and generative learning and dimensionality reduction were proven effective in their ability to identify and determine characteristic structures of types or patterns of similarity with a wide array of data of realistic complex types.

To reduce the dimensionality, improve the interpretability of data and, as a result, reduce the computational complexity of subsequent methods of analysis, a wide range of methods, both linear and non-linear were researched, including PCA, SNE [3, 4], prototype learning including with generative neural models [5, 6], dimensionality reduction and manifold learning [7, 8] and others. These methods compress data of large dimensions while preserving its essential information content, not in the least, in the context of this study, the relationship of similarity in the spaces (embeddings) of informative factors/features. It was shown that these methods can be used in combination with fuzzy approaches such as fuzzy C-means [9].

This observation brings us to the field of fuzzy sets that have been applied extensively to problems in multifactorial data spaces. Papers [10, 11] demonstrated successful applications of fuzzy pattern recognition and fuzzy modeling to evaluate, compare, select, prioritize, and/or organize alternative decision options. Such approaches have been shown to simplify decision-making processes and reduce their complexity.

Fuzzy models are effectively used in intelligent decision-making systems. Such integration allows to work effectively with multidimensional data and ensure the accuracy of decisions in complex conditions of uncertainty.

For example, [12] presented a novel neuro-fuzzy diagnostic system based on a non-iterative ANN and an original fuzzy information model. In [13] a novel effective decision-making method aimed to assist in clinical practice based on integrated fuzzy information models and data mining was proposed.

An interesting approach demonstrated recently in [14] is a combination of prototype models and fuzzy methods such as fuzzy C-means in the analysis of conceptual data structures. The method demonstrated the potential for accurate classification while managing data uncertainty. A related study [15] presented a multi-objective optimization method that showed potential in confident determination of the optimal prototype structures via simultaneous minimization of the training error on historical training data and minimization of the intra-cluster variance.

In this work we attempted to advance and refine methods of determination of the prototype structure, including fuzzy prototypes to the case and essential limitations in the problem of constrained prior by incorporating optimization and iterative improvement of conceptual models of problem data.

In approaching the problem of learning with constrained prior for optimal decisions produced by intelligent decision systems with arbitrary type of problem data space, we first examine the process of construction of fuzzy-prototype models with arbitrary samplings of the problem data, attempting to avoid essential assumptions and/or constraints such as pre-known content of prototype classes. Next, observing that in the constrained setting, general data of the problem, not necessarily associated with verified outcomes, can be accumulated in the active process of decision making, we examine quality characteristics of fuzzy prototype models, such as precision/resolution relative to

the size of the data samplings. Formulating these methods and processes allows us to test the hypothesis that using a combination of iterative accumulation of problem data and construction of fuzzy-prototype models with more representative and detailed samplings produces more precise models of the problem data distribution leading to improve quality of decisions based on such models.

3. Methods

3.1. Problem formulation

In this work we will deal with data that is obtained as a sampling of an unknown distribution D , $W = \{P, F\}$ where $P = \{p\}$: the individual points of observation that describe the domain, $F = \{f\}$, the observable factors recorded in the sampling. Accordingly, each data point $p \in P$ is described by a set of observable factors $f_p = F(p)$.

In the task of interpretation of the data D for making decisions the critical challenge is to establish an association between an observation $x \in P$ and a class of similar observations $K(x)$ that can be associated with a correct decision $M(K, a)$, where a : the parameters of the decision function.

An additional challenge we will be discussing in this work relates to the case/scenarios where the information about the association, the factorization relationship R :

$$K(p) = R(F(p)). \quad (1)$$

is limited or absent at the outset of the study. In such cases, one cannot rely on a known function or logical sequence to connect an observed instance of the phenomena of interest to the correct decision. This range of problems/scenarios will be referred to as “learning with constrained prior” problem.

In approaching this problem, one is faced with the challenge of determination of classes of similarity in complex data described by a large set of observable factors F without sufficient prior information about the factorization relationship (1), often inferred from sets of known associations $(p, K(p))$ known as annotated or labeled data.

3.2. Fuzzy prototype analysis

Prototype analysis is a well-known approach in problems and scenarios where prior information about a given type of data or distribution, D is limited or not available, confounding or precluding resolution of the factorization relationship (1) from known associations observation, class [5, 16].

These methods work particularly well with data that can be described or expressed by a large number of observable factors with a possibility of a strong redundancy in the observation points described by them (multiple homogeneous observable factors).

Following the well-researched process of unsupervised analysis and determination of conceptual structure that is in data science that commonly involves strong reduction of dimensionality, a structure of natural prototypes or concepts can be derived from a general representative sampling of the distribution D : $P(D) = \{p_k\}$ that can be interpreted as the basic framework of the essential types of similarity p_k that approximate the distribution:

$$x \in D \rightarrow \exists k, p_k: t(x) \in p_k, \quad (2)$$

where $t(x)$ is the image of the observation x in the informative prototype space (the embedding) commonly of reduced functionality [5, 6]. In many specific models and methods of unsupervised embedding, the inverse association from the prototype to the observable space exists as well $y(t) = G(t)$, $y \in D$, where $G(t)$: the generative transformation [17].

In this work, we propose an extension of the prototype analysis based on the observation that the association between an observation and its prototype class (2) in most practical cases with many methods is not categorical but rather, probabilistic, i.e. the probability of association of an

observation $x \in D$ to the prototype class p_k is described by the prototype probability distribution $\rho(x, p_k)$:

$$\rho(x, p_k) = W(E(x)) \in p_k, \sum_k \rho(x, p_k) = 1 \quad (3)$$

where $E(x)$: the embedding transformation $D \rightarrow P$, W : the probability of a point in the prototype space belonging to a specific prototype class. The relationship in (3) defines probabilistic or “fuzzy” association between the observations in the observable data space D and the prototype classes P resolved with methods of prototype analysis.

The advantage of the methods just described in application to the learning with constrained prior problem comes from the observation that whereas prior knowledge of the problem can be limited or constrained, it may not necessarily be the case for the general raw data in the problem data space.

Moreover, this data can be accumulated in the process of the interactions of the learning system with the problem data space, with a possibility of iterative, progressive learning from the experience, based on the results of the earlier observations and learning iterations.

Then, with the data accumulated through this process, methods of analysis of its informative structure can be applied, including, as discussed earlier, neural generative learning, prototype learning, deep dimensionality reduction with preservation of the information content and many others. This approach can offer essential insights into the conceptual composition of the distribution of the problem that can be used for determination of the prototype structure and subsequent applications in intelligent decision systems.

3.3. Integration with decision systems

Application of the fuzzy prototype analysis in intelligent decision systems operating in the context of constrained prior problems can be proposed straightforwardly via the construction of decisions based on the natural, intrinsic similarity of the observations, as:

$$t(x), t(y) \in p_k \rightarrow D(x) \cong D(y) \quad (4)$$

where $D(x), D(y)$: decisions produced (constructed) for the observations x, y , $t(x), t(y)$: their images in the informative embedding space of the problem.

In other words, once the structure of fuzzy prototypes in the problem data D has been determined via application of the fuzzy prototype analysis as described in the preceding sections, observations that belong in the same prototype class with high confidence can be associated with similar decisions. Fuzzy prototype models can as well provide some informative insights for observations with less confident association to prototype classes, as will be discussed further in the results section.

3.4. A demonstration of fuzzy prototype analysis

In this work we illustrate the methods and workflow of the fuzzy prototype analysis for intelligent decision systems with a dataset that models a case of a learning with constrained prior problem.

3.4.1. Model dataset

For an illustration of the fuzzy prototype method, we will consider here an example of an observable distribution is described by a large number of numerical parameters of the same type such as images. We will use the dataset of images of basic geometric shapes that was described in [18].

The images in the dataset were of three basic types: circles, triangles and backgrounds, of variable size and grayscale contrast. The resolution of the images was 64×64 pixels i.e. each data point corresponding to a single observation was expressed in 4,096 numerical factors in the range $[0, 1]$.

3.4.2. Informative dimensionality reduction

An informative embedding the model dataset was obtained by applying a generative neural network model with the architecture of a convolutional encoder, as described in [18]. This class of neural architectures, being of the type of self-supervised learning, does not require annotated datasets for training [17]. It is trained by reducing the error of reproduction (generation) of the samples in a set of observables samples that can represent a sampling of the distribution of the problem.

Thus, it is essential to note that informative embedding spaces constructed by application of generative models to the problem dataset do not depend on any prior information about the distribution and fully satisfy the constraints of the problem. Examples of distributions of data samples in the embedding spaces of trained generative models are shown in Figure 1.

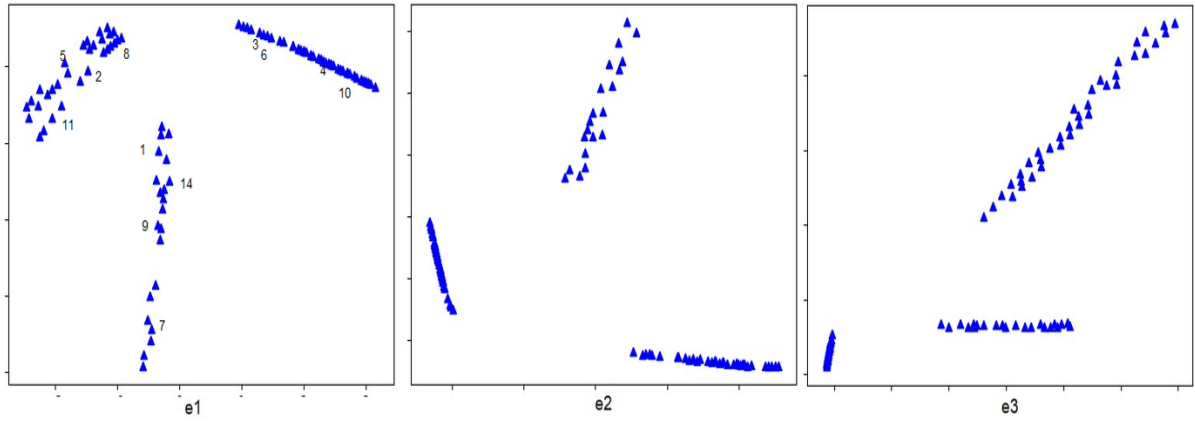


Figure 1: Distributions of multifactorial homogeneous data (images of geometric shapes, [18]) in low-dimensional informative embedding spaces.

Another point that is essential for the analysis and discussion here is that successful learning could be achieved with relatively small samplings of the distribution, in the model example, as small as dozens or even single samples per class. Granted, this observation needs to be taken with caution and may not be readily extendable to significantly more complex problem data. Still, it shows that meaningful initial learning of problem data of significant complexity as described earlier can be achieved with limited samplings and moreover, as noted earlier, do not require annotations with known types or classes, or any other form of prior knowledge about the problem distribution.

Finally, it is worth noting that the method of construction of informative low-dimensional embeddings of the problem data used here is not unique and a wide selection of methods of self-supervised, unsupervised learning and dimensionality reduction has been studied and applied successfully. A more detailed discussion of the types of the methods, their differences, etc., would fall beyond the scope of this work.

3.4.3. Fuzzy-prototype structure of problem data

In the example that we use in this work, groups of samples of the unknown distribution of the problem in the high dimensional space of observable factors characterized by essential similarity are modeled by the types of the geometric shape in the dataset of images.

This type of model corresponds to problems and scenarios where an unknown distribution of the problem is described by a large number of the numerical factors with approximately equal significance in the effect of interest in the distribution (homogeneous multifactorial data).

It was shown [18] that in some such cases, the prototype/conceptual structure of the data $P(D)$ discussed in the preceding sections can be determined or resolved with sufficient confidence by application of methods of unsupervised ensemble learning and clustering that do not depend on significant or, in fact, any prior knowledge about the unknown distribution.

As a result of application of such methods, the distribution of data points that correspond to a certain sampling of the original data S in the informative embedding space $E(S)$ can be represented by the fuzzy-prototype structure $P_f(D)$:

$$P_f(D): \{ P(D), \rho(x, p_k) \} \quad (4)$$

where $P(D) = P(E(S))$, the sequence of the prototypes derived from the distribution $E(S)$ in the derived informative embedding space, ρ : the prototype probability distribution.

Again, one can observe that the fuzzy-prototype structure of the problem data (4) effectively approximates the observable distribution by providing a probabilistic model of distribution of arbitrary representative of the problem data D between the conceptual prototypes.

3.4.4. Iterative learning

One can observe that the process of the resolution or construction of the fuzzy-prototype structure of the problem data described in the preceding sections was in effect, static with respect to the basic sampling of the problem data $S(D)$ that was used for the derivation of the structure/model of the fuzzy prototypes.

As we noted earlier, an essential advantage of the methods of unsupervised learning with constrained prior problems is the potential to accumulate general, non-annotated data in the course of the learning process. Such more extensive and detailed samplings can in their turn provide additional, more detailed information about the conceptual structure of the distribution. Then, repeating the described process iteratively with a sequence of extended samplings of data S_1, S_2, \dots, S_k in the problem space it can be possible to produce more precise, “sharper” fuzzy prototype models with diminishing uncertainty in the probability distribution of the prototype classes. This iterative process is illustrated in Figure 2.

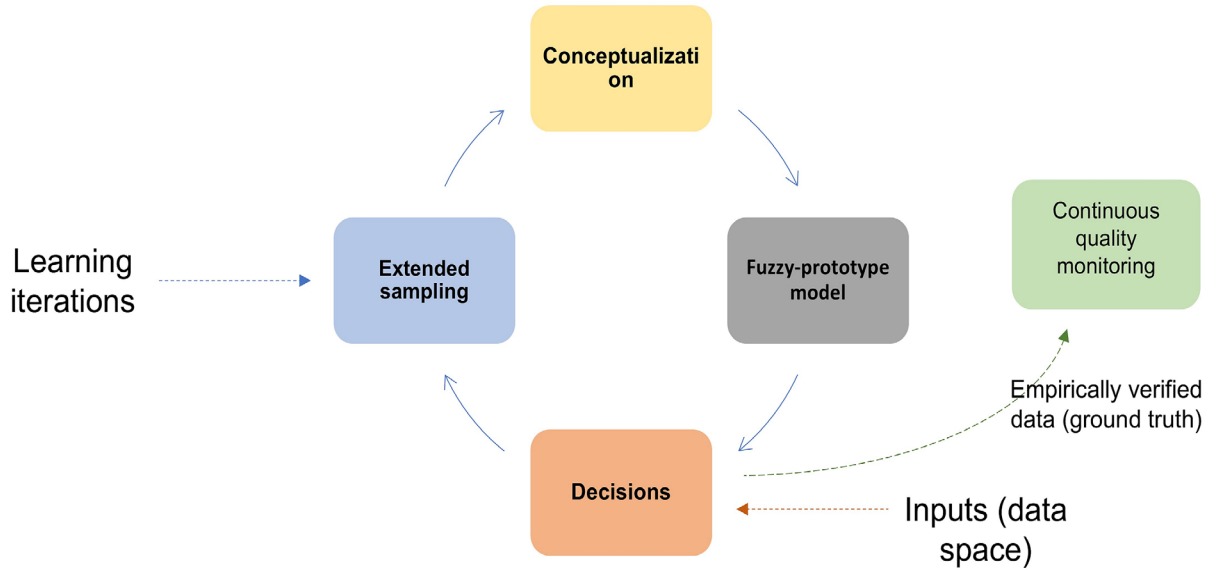


Figure 2: Iterative learning of fuzzy prototype structure via extended samplings.

Thus, one can expect that the iterations of the fuzzy-prototype models $P_f(D)_k$ obtained with more descriptive samplings S_k would produce more precise models of the prototype probability distribution, reducing the uncertainty in the distribution of the observations between the prototype classes. In the next section we attempt to verify this hypothesis with the model dataset of images.

4. Results

In this section we illustrate the method of construction of fuzzy-prototype models of problem data in constrained prior problems with the dataset of images as described in Section 3.3.1. To examine the relationship between the precision or resolution of the fuzzy-prototype model on the size of sampling, we used samples of three different sizes: S-90, having 30 images per type (i.e., geometric shape); S-150, 50 images per type; and S-300, with 100 images per type.

4.1. Samplings and construction of fuzzy-prototype models

Iterative learning of fuzzy-prototype models of problem data in the constrained prior setting can face another challenge in the early stages of the process: that of instability of learning with small data. This limitation applies only to some problems, where both known and general data are constrained, whereas in other cases, samplings of general data that is not associated with prior knowledge would not be limited or constrained. Still, in this work we chose to address the case where general data is limited along with the annotated one, and the processes of learning of prototype structure and collection of general samplings proceed alongside each other. For this reason, the initial, starting sample was chosen to be of a rather small size, about two dozen instances per conceptual class, that is in our case, the type of geometric shape.

One can note before proceeding to further analysis that the minimum threshold of the size of samplings that is necessary for initial learning of the prototype structure in the data is not an obvious choice; it depends on several factors such as conceptual complexity of data, characteristics of the variation in the informative embedding factors of the prototype classes and others. Addressing this question in full detail would be a challenging problem of its own that merits another study. With the data used in this work it was found by trial that the chosen size of the initial sampling was sufficient for the purposes of the study.

To construct fuzzy-prototype models with samplings of the problem data modeled by the model dataset of images, the process described in [18] was used. To address the challenge of stability in learning with small data for smaller-size samplings [19, 20], an ensemble [21, 22] of generative neural models was used. To outline it briefly, after construction of informative embeddings with neural models of self-supervised learning, clustering in the embedding space was applied to identify characteristic regions/clusters of samples that were associated with the concept/prototype classes. In this work we used clustering by a visual observation method, but the demonstrated approach can be extended straightforwardly to use known methods of unsupervised clustering such as DbScan, MeanShift and others [23-25].

4.2. Construction and resolution/precision of fuzzy-prototype models

Based on the structure of characteristic types/concepts/prototypes associated with clusters in the informative embedding space of the problem data, one can derive the fuzzy-prototype model of the data (in the iteration of the sampling S_k used to calculate the structure) $P_f(S_k)$ via the process illustrated in Figure 3. Many specific implementations of the process are possible with arbitrary number of clusters and without limitations for the dimensionality of the informative embedding space nor the methods of constructing it.

At a given iteration characterized by a general sampling of the problem distribution D , S_k the precision or resolution of the fuzzy prototype model $P_f(S_k)$ can be characterized by the confidence factor t_c indicating the minimal confidence threshold for an association of an observation x to a prototype class P_k , and the confusion matrix $M(P(x), P_{\text{true}})$:

$$\rho(x, p_k) \geq t_c \rightarrow P(x) = p_k P_f(D): \{ P(D), \rho(x, p_k) \}, \quad (5)$$

where $P(x)$: the prototype class associated with an observation x by the fuzzy-prototype model at the confidence threshold t_c , P_{true} : the true, known class associated with the observation.

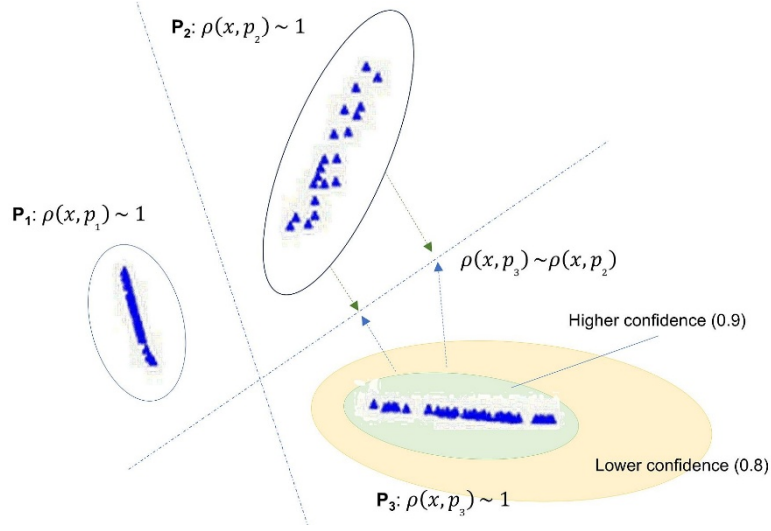


Figure 3: Fuzzy-prototype model from informative embedding of data sampling.

An example of the confusion matrix produced by a fuzzy prototype model for the model data of geometric images is shown below (from [9]).

Table 1

Confusion Matrix, S-150 Fuzzy Prototype Model, $t_c = 0.95$

Shape, Cluster	Cluster 0	Cluster 1	Cluster 2
Circle	1.0	0.	0.
Triangle	0.25	0.75	0.
Background	0.	0.15	0.85

4.3. Iterative learning with fuzzy-prototype models

To demonstrate the usability of the fuzzy-prototype approach with real data, we applied the process of construction of fuzzy-prototype models with three samplings of the model problem data of progressively larger size, as described earlier in this section: S-90, S-150, and S-300.

As can be seen in Table 2, where the precision characteristics of fuzzy-prototype models constructed with the samplings at two confidence thresholds, $t_c = 0.8, 0.9$ are given, fuzzy-prototype models obtained with iteratively extended samplings via the process shown in Figure 2 showed progressive improvement in the precision/resolution of the prototype classes.

The precision/resolution used to measure the performance of the models was calculated from the confusion matrix M of the model as a pair (tuple) (a, c) of:

1. The accuracy, a , measured as the sum of the diagonal (correct) predictions of the prototype class, divided by the number of classes: $a = \frac{1}{N} \sum_i M_{ii}$.
2. The confusion, c , measured as the sum of the non-diagonal (incorrect) predictions of the prototype class, divided by the number of combinations of classes: $c = \frac{1}{N(N-1)} \sum_{i \neq j} M_{ij}$.

Table 2

Precision/resolution in Iterative Learning with Model Data, Accuracy/Confusion Metrics

Sampling	Size (per type)	Precision, $t_c=0.8$	Precision, $t_c=0.9$
S-90	30	0.811/ 0.082	0.773/ 0.092
S-150	50	0.867/ 0.067	0.824/ 0.075
S-300	100	0.912/ 0.054	0.885/ 0.633

It can be noted in conclusion of this section that the stability of the structure and content of clusters in the embedding space is a key dependency and a requirement for confident determination of the fuzzy prototype structure. The use of an ensemble of generative neural models assured that the cluster/prototype structure resolved by the method described characteristic patterns in the model problem data.

The results of this experiment show that the precision of association of observations to prototype classes determined via the process described in this work improves steadily with accumulation of new data, even without the dependency on the known association between the observations and their true, externally known class. This knowledge was used in this work only for verification of the performance of the models.

4.4. Application of fuzzy-prototype models in intelligent decisions systems in the constrained prior context

The method of construction of fuzzy-prototype models developed in this study has shown the potential of progressive improvement in the precision of association of the data points in the problem data space with characteristic types, concepts or prototypes. Importantly, the fuzzy-prototype structure can be initially resolved with smaller samplings and the method allows to accumulated data in the process of learning/operation with progressively improving precision, resulting in higher effectiveness of the decisions associated with prototypes. A word of caution that needs to be said here is that data used to demonstrate the method was relatively simple with respect to its conceptual content (i.e., the number of distinct types) and more complex types/spaces of problem data may require samplings of larger size for confident determination of the prototype structure.

Another interesting potential of the proposed method can be pointed out for observations (i.e., data points in the problem data space) that produce less confident, “confused” associations to prototype classes. In such cases, if the decision space and process allow so, the constructed decision that is a mixture of the “pure” decisions associated with prototype classes can be more effective than any of the pure decisions. Let consider an example with a prototype structure of three classes $p_1 - p_3$, and an observation x_m : $\rho(x_m, p_1) = 0.4, \rho(x_m, p_2) = 0.3, \rho(x_m, p_3) = 0.3$. Then, the decision created as a combination of the pure decisions $d_1 - d_3$ associated with the prototype classes, for example as:

$$d(x_m) = \sum_k \rho(x_m, p_k) d_k,$$

if the decision space/process allows such kind of mixing, can be more effective than any of the pure decisions $d_i(p_i)$ associated with the prototype classes.

5. Conclusions

In this work, we attempted to approach the problem of learning with constrained prior for intelligent decision making by combining the results from several thriving fields of data science and the research in intelligent systems: methods of unsupervised and/or self-supervised learning, generative learning, prototype/concept learning, unsupervised, including density clustering and fuzzy clustering with iterative aggregation of samplings of the problem data space to outline the direction in which intelligent systems can begin learning complex data spaces with minimal data improving the quality of the decisions in the process of learning. As outlined in the review section, these results connect with reported results in these fields and are supported by them. Several concluding comments need to be added to this summary of the findings of this study.

First, while one can expect the outlined approach to be sufficiently general to accommodate a broad range of realistic complex data, specific characteristics and implementations of the method suitable for different types of data/problem spaces can vary. This can include methods of producing informative embeddings, characteristics of the embeddings including dimensionality, methods of

clustering/fuzzy conceptualization, minimal sampling, the number and pace of learning iterations and other essential characteristics.

Another essential point that was mostly left out of the scope of our discussion is verification of the quality and effectiveness of the learning process and the decisions produced by the learning system. A verification loop, based on batches of samples verified empirically, with known outcomes can be collected in each learning iteration and integrated into the learning process to evaluate the quality and progress of learning. This essential function of the system will be examined in a future work.

A close and seemingly natural connection between the processes of learning the conceptual structure in general problem data and the ability of intelligent systems to construct effective decisions and responses based on the relationship of essential similarity in the input data space was noted. In that way, a decision learned, and verified once can be applied to a class of inputs in the problem data space, improving the generality, effectiveness and efficiency of the learning process. An interesting direction of research that can be pursued in another work is the potential to construct fuzzy or “hybrid” decisions for observations with less certain association to concept/prototype classes, as discussed briefly in Section 3.4.

The problem of learning with constrained prior data described and examined here, where sufficiently large volumes of empirically verified knowledge about the problem distribution for conventional approaches in learning have not been accumulated emerge on a regular basis in modern science and technology. We expect that the methods proposed and examined in this work will be instrumental and find application in the theory and practice of development of effective and efficient intelligent decision systems capable of working with strongly constrained problems and environments.

Declaration on Generative AI

The authors have not employed any Generative AI tools in preparation of this work.

References

- [1] J. E. Van Engelen and H. H. Hoos, A survey on semi-supervised learning, *Machine Learning*, vol. 109, no. 2, pp. 373–440, 2020. doi:10.1007/s10994-019-05855-6.
- [2] L. Von Rueden, S. Mayer, K. Beck, B. Georgiev, S. Giesselbach, R. Heese, et al., Informed machine learning – A taxonomy and survey of integrating prior knowledge into learning systems, *IEEE Trans. Knowl. Data Eng.*, 2021. doi:10.1109/TKDE.2021.3079836.
- [3] M. Rabasovic, D. M. Pavlovic, and D. Sevic, Analysis of laser ablation spectral data using dimensionality reduction techniques: PCA, t-SNE and UMAP, *Contrib. Astron. Obs. Skalnaté Pleso*, vol. 53, 2023. doi:10.31577/caosp.2023.53.3.51.
- [4] D. Kracker, J. Garcke, and A. Schumacher, Automatic analysis of crash simulations with dimensionality reduction algorithms such as PCA and t-SNE, in *Proc. 16th Int. LS-DYNA Forum*, 2020. URL: <https://lsdyna.ansys.com/wp-content/uploads/attachments/t1-1-a-automotive-011.pdf>.
- [5] T. Isaiev and T. Kysil, Method of Creating Custom Dataset to Train Convolutional Neural Network, *Computer Systems and Information Technologies*, no. 4, pp. 37–44, Dec. 2024. doi:10.31891/csit-2024-4-5.
- [6] S. Dolgikh, Topology of conceptual representations in unsupervised generative models, in *Proc. 26th Int. Conf. Information Society and University Studies (IVUS-2021)*, Kaunas, Lithuania, CEUR-WS.org, vol. 2915, pp. 150–157, 2021.
- [7] M. Meilä and H. Zhang, Manifold learning: What, how, and why, *Annu. Rev. Stat. Appl.*, vol. 11, pp. 263–290, 2024. doi:10.1146/annurev-statistics-040522-115238.
- [8] Y. Wang, H. Huang, C. Rudin, and Y. Shaposhnik, Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data

- visualization, *J. Mach. Learn. Res.*, vol. 22, pp. 1–73, 2021. URL: <https://doi.org/10.48550/arXiv.2012.04456>.
- [9] Y. Shen, T. Chen, Z. Xiao, B. Liu, and Y. Chen, High-dimensional data clustering with fuzzy C-means: Problem, reason, and solution, in *Lecture Notes in Computer Science*, vol. 12861, I. Rojas, G. Joya, and A. Català, Eds. Cham: Springer, 2021, pp. 89–100.
 - [10] N. A. Setiawan, P. A. Venkatachalam, and A. F. M. Hani, Diagnosis of Coronary Artery Disease using Artificial Intelligence based decision support system, 2020, Arxiv 2007.02854.
 - [11] G. Marín Díaz, R. Gómez Medina, and J. A. Aijón Jiménez, Integrating fuzzy C-Means clustering and Explainable AI for robust galaxy classification, *Mathematics*, vol. 12, no. 18, p. 2797, 2024. doi:10.3390/math12182797.
 - [12] I. Izonin, R. Tkachenko, I. Dronuyk et al., Predictive modeling based on small data in clinical medicine: RBF-based additive input-doubling method, *Mathematical Bioscience Engineering*, 18 (3) (2021) 2599–2613.
 - [13] C. Ylenia, D. L. Chiara, I. Giovanni, R. Lucia A Clinical Decision Support System based on fuzzy rules and classification algorithms for monitoring the physiological parameters of type-2 diabetic patients. *Mathematical Biosciences and Engineering*, 18(3) (2021) 2687–2708. URL: <https://doi.org/10.3934/mbe.2021135>.
 - [14] W. Mao and K. Xu, Enhancement of the classification performance of fuzzy C-means through uncertainty reduction with cloud model interpolation, *Mathematics*, vol. 12, no. 7, p. 975, 2024.
 - [15] X. Gu, M. Li, L. Shen, G. Tang, Q. Ni et al., Multiobjective evolutionary optimization for prototype-based fuzzy classifiers, *IEEE Transactions on Fuzzy Systems*, 31 (5) (2023), 1703–1715.
 - [16] L. Makarova and M. Tatarenko, Comparative Analysis of Methods and Models for Estimating Size and Effort in Designing Mobile Applications, *Computer Systems and Information Technologies*, no. 3, pp. 26–33, 2024. doi:10.31891/csit-2024-3-4.
 - [17] M. Welling, D. Kingma, An introduction to variational autoencoders, *Foundations and Trends in Machine Learning*, 12 (4) (2019), 307–392.
 - [18] S. Dolgikh, Modeling of small data with unsupervised generative ensemble learning, in: *Proceedings of the 5th International Conference on Informatics and Data-Driven Medicine (IDDM-2022) Lyon France 2022*, CEUR-WS.org volume 3302, pp. 35–45.
 - [19] L. Gao, D. Wang, L. Zhuang, X. Sun, M. Huang, and A. Plaza, BS3LNet: A new blind-spot self-supervised learning network for hyperspectral anomaly detection, *IEEE Trans. Geosci. Remote Sens.*, vol. 61, art. no. 5504218, pp. 1–18, 2023. doi:10.1109/TGRS.2023.3246565.
 - [20] R.N. D’souza, P.Y. Huang, F.C. Yeh, Structural analysis and optimization of Convolutional Neural Networks with a small sample size, *Scientific Reports* 10 (2020) 834.
 - [21] A. Wu, W. Ge, and W.-S. Zheng, Rewarded Semi-Supervised Re-Identification on Identities Rarely Crossing Camera Views, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 15512–15529, Dec. 2023. doi:10.1109/TPAMI.2023.3292936.
 - [22] S. Kumar, P. Kaur, A. Gosain, A comprehensive survey on ensemble methods, in: *Proceedings of the 2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, Mumbai, India, 2022, pp. 1-7.
 - [23] J. Yan and X. Wang, Unsupervised and semi-supervised learning: The next frontier in machine learning for plant systems biology, *Plant J.*, vol. 111, no. 2, pp. 301–314, 2022. doi:10.1111/tpj.15905.
 - [24] P. Bhattacharjee, P. Mitra, A survey of density based clustering algorithms. *Frontiers of Computer Science* 15 (2021) 151308.
 - [25] D.R. Hunter, Unsupervised clustering using nonparametric finite mixture models, *WIREs Computational Statistics* 16 (1) (2024) e1632.