Latent-Variable Classifiers Based on Total Correlation for Dynamic Environments

Cheol Ho Kim
Superintelligence Creative Research Laboratory
Eletronics and Telecommunications Research Institute
Daejeon, Republic of Korea
kimcheolho@etri.re.kr

Jung-Hoon Lee
Superintelligence Creative Research Laboratory
Eletronics and Telecommunications Research Institute
Daejeon, Republic of Korea
jhlee0914@etri.re.kr

Byounghwa Lee
Superintelligence Creative Research Laboratory
Eletronics and Telecommunications Research Institute
Daejeon, Republic of Korea
byounghwa.lee@etri.re.kr

Ock Kee Baek
Superintelligence Creative Research Laboratory
Eletronics and Telecommunications Research Institute
Daejeon, Republic of Korea
ok.baek@etri.re.kr

***Abstract*— This paper introduces a novel approach to latent-variable classifiers that employs the correlation explanation technique, an information-theoretic latent-variable clustering based on total correlation. The proposed classifier improves batch-mode classification accuracy in comparison to traditional training approaches. Additionally, it demonstrates the ability to incrementally learn from additional datasets involving new features, without accessing or reprocessing past data. These findings underscore the method's potential to serve as a robust machine learning model suitable for dynamic real-world environments.**

Keywords— classifiers, latent variables, total correlation, mutual information, incremental learning

I. Introduction

Latent-variable representation has facilitated the discovery of meaningful and succinct hidden variables within various real-world applications, including machine learning, statistics, econometrics, psychometrics, social sciences, and biology [1–2]. Information-theoretic latent-variable clustering techniques based on total correlations leverage not only pairwise relationships but also the holistic interplay of all variables, enabling a more comprehensive extraction of latent variables [3–5].

Integrating supervised learning classifiers with these latent variables holds the potential for developing enhanced classifiers. Moreover, latent variables extracted through unsupervised learning, independent of class information, retain their invariant properties even in the presence of additional training data or the introduction of new variables. This resilience renders them valuable for incremental learning in dynamic environments.

In this paper, we propose a novel approach to latent-variable classifiers by integrating classification with the correlation explanation technique [3–4], one of the total correlation-based latent-variable clustering methods. We validate the potential of our approach to achieve improved performance compared to conventional training methods in batch-mode classification tasks. Furthermore, our proposed method supports incremental learning in dynamic environments, accommodating situations where some existing features may be missing or new features are introduced. We demonstrate an upwardly increasing incremental curve as the feature space continues to expand.

II. Total Correlations

A. Information Entropy and Mutual Information

The concept of total correlation [3–5] is rooted in information theory. Information entropy, a foundational notion in information theory, quantifies the degree of uncertainty inherent in a probability distribution. For a discrete random variable (or feature) X, information entropy is defined as $H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$ where x_i indicates a value that X takes.

When considering two variables X_1 and X_2, the conditional entropy of X_1 given X_2 is defined as $H(X_1|X_2) = -\sum_{x_1,x_2} p(x_1,x_2) \log \frac{p(x_1,x_2)}{p(x_2)}$. The significance of $H(X_1|X_2)$ lies in representing the amount of uncertainty that remains regarding variable X_1 after the value of X_2 is known. The shared information entropy between these variables is referred to as mutual information, denoted as $I(X_1{:}X_2)$ and calculated using $I(X_1{:}X_2) = H(X_1) - H(X_1|X_2) = H(X_2) - H(X_2|X_1)$. Mutual information can be interpreted as the amount of uncertainty in variable X_1 that is jointly captured by variable X_2.

B. Total Correaltion for Multiple Variables

Mutual information for three or more variables can assume various forms due to the existence of diverse interrelationships. Among the commonly used measures, one is total correlation, also referred to as multivariate mutual information or multi-information. For a given set of n random variables $\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$, it is defined as $TC(\boldsymbol{X}) = \sum_{i=1}^{n} H(X_i) - H(\boldsymbol{X})$. The conditional total correlation for a random variable Z can be defined as $TC(\boldsymbol{X}|Z) = \sum_{i=1}^{n} H(X_i|Z) - H(\boldsymbol{X}|Z)$. When $TC(\boldsymbol{X}|Z)$ equals 0, as indicated by the factorization of X_i's conditioned on Z, it implies that Z contains complete information about all the common causes among the X_i's.

The shared information between $\boldsymbol{X}$ and Z, referred to as correlation explanation and denoted as $TC(\boldsymbol{X};Z)$, can be defined and calculated using the equation $TC(\boldsymbol{X};Z) = TC(\boldsymbol{X}) - TC(\boldsymbol{X}|Z) = \sum_{i=1}^{n} I(X_i{:}Z) - I(\boldsymbol{X}{:}Z)$. This equation implies that when $TC(\boldsymbol{X}|Z)$ is minimized, $TC(\boldsymbol{X};Z)$ is maximized. When $TC(\boldsymbol{X}|Z)$ equals 0, it signifies that Z explains all the correlations in $\boldsymbol{X}$. In other words, the correlation information among the variables constituting $\boldsymbol{X}$ is encapsulated within the variable Z.

C. Hierarchical Latent-Variable Clustering Based on Correlation Explanation

Leveraging the correlation explanation technique facilitates the establishment of hierarchical latent-variable clustering for multiple variables $\boldsymbol{X}$. Given a set of m random variables $\boldsymbol{Z} = \{Z_1, Z_2, \dots, Z_m\}$, $TC(\boldsymbol{X}; \boldsymbol{Z})$ can be computed as

$$TC(\boldsymbol{X}; \boldsymbol{Z}) = \sum_{i=1}^{n} I(X_i : Z) - \sum_{j=1}^{m} I(Z_i : X). \quad (1)$$

By maximizing $TC(\boldsymbol{X}; \boldsymbol{Z})$, it becomes possible to obtain latent variables $\boldsymbol{Z}$ that provide the most comprehensive explanation of $\boldsymbol{X}$. Applying this process iteratively enables the realization of hierarchical latent-variable clustering, as depicted in Fig. 1. Therefore, given $\boldsymbol{X}$, the objective function is to find the latent variables $\boldsymbol{Z}$ at each layer l that maximize $TC(\boldsymbol{X}; \boldsymbol{Z})$, i.e., $\max_{p(z|x)} TC(\boldsymbol{X}; \boldsymbol{Z})$.

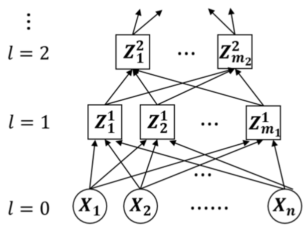

Fig. 1. Hierarchical latent-variable clustering based on the correlation explanation.

III. Latent-Variable Classification Based on Total Correlations

A. A Proposed Classification Method Based on Single-Layer Correlation Explanation

The correlation explanation technique facilitates concise and efficient feature transformations based on the mutual relationships among numerous variables through unsupervised learning. Leveraging this capability, we aim to integrate it into supervised learning, specifically classification. Figure 2 illustrates a classification approach wherein n features $X_1, X_2, \dots, X_n$ are provided. This approach employs single-layer correlation explanation to extract m latent variables $Z_1, Z_2, \dots, Z_m$, which are subsequently utilized as inputs to predict the class variable Y using a classifier. Given that these latent variables can function as meaningful abstractions of the input features, they can mitigate the training complexity of the classifier, potentially resulting in enhanced classification accuracy. The key hyperparameters encompass the number of latent variables m (typically 5–20)

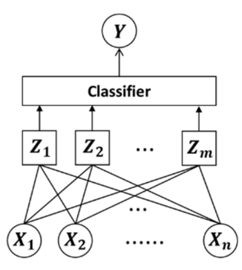

Fig. 2. Proposed classification method utilizing latent-variable representation based on correlation explanation.

and the cardinality k (typically 5–15) associated with each latent variable.

B. Datasets

The datasets chosen for experimentation encompass Mice-behavior, Mice-treatment, Mice-genotype, Spambase, Phishing, and Parkinson from the UCI Machine Learning Repository. Furthermore, the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset has been incorporated. A summary of the dataset details is presented in Table I.

TABLE I. DATASETS

Dataset	**Feature**	**Instance**	**Class**
Mice-behavior	77	1080	2
Mice-treatment	77	1080	2
Mice-genotype	77	1080	2
Spambase	57	4601	2
Phishing	68	11055	2
Parkinson	753	756	2
ADNI	44	12936	4

C. Batch-Mode Classification Results

The outcomes of the batch-mode classification are presented in Table II. We have compared the proposed latent-variable classifiers with the conventional training methods. For instance, in the case of Mice-behavior, the conventional method utilized naïve Bayes (NB) with raw features, whereas the proposed method trained an NB with the latent variables extracted from the raw features. In most cases, the latent-variable approach demonstrates superior performance over conventional training in the majority of cases. These results highlight the beneficial impact of utilizing the latent-variable representation in augmenting classification accuracy.

IV. Feature-Incremental Learning Based on the Latent-Variable Classifiers

A. Incremental Learning Stratege for Missing and New Features Utilizing Latent-Variable Representation

In real-world applications, batch-mode learning is often impractical, requiring models to learn progressively from

TABLE II. BACH-MODE CLASSIFICATION ACCURACY

Dataset	**Classifier**	**Accuracy by the proposed latent-variable method (%)**	**Accuracy by conventional training method (%)**
Mice-behavior	NB[a]	**100**	99.5
Mice-treatment	NB	**82.0**	76.0
Mice-genotype	NB	**85.0**	80.7
Spambase	NB	**90.0**	89.9
Phishing	TAN[b]	**93.4**	92.7
Parkinson	TAN	**87.3**	84.1
	BN[c]	**85.2**	78.3
	RF[d]	**88.4**	77.8
	XGB[e]	**88.4**	86.8
ADNI	BN	**91.1**	89.9
	RF	91.6	**94.0**
	XGB	92.4	**93.2**

[a] naïve Bayes, [b] tree-augmented naïve Bayes, [c] Bayesian network, [d] random forest, [e] extreme gradient boosting

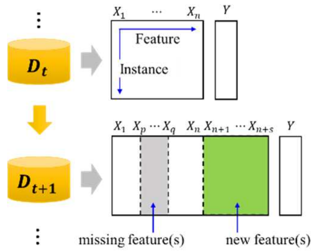

Fig. 3. An additional dataset with missing and new features.

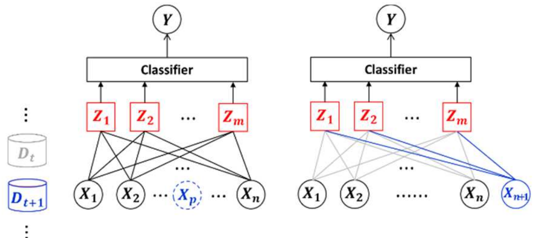

Fig. 4. Incremental updates of latent variables for missing and new features in the additional dataset.

sequentially provided segregated datasets. This scenario frequently involves encountering missing features or introducing new features, as depicted in Fig. 3. In the context of our proposed approach, latent variables capture the mutual relationships among existing variables independent of class information. This enables the preservation of a robust representation even when features undergo dynamic changes.

As illustrated in Fig. 4, the utilization of the proposed latent-variable classifier enables adaptive and incremental learning on a new dataset D_{t+1} with missing and new features. This is achieved through updating only the latent variables, even in the absence of access to the previous dataset D_t, thereby facilitating the incremental learning process.

B. Feasibility Test of Feature-Incremental Learning Using the Proposed Latent-Variable Classifiers

We conducted a feasibility test to assess the viability of incremental learning in scenarios involving the ongoing introduction of new features. We divided a single dataset into multiple mini datasets and incrementally increased the number of features, as shown in Fig. 5, through artificial preprocessing for each mini dataset. The proposed latent-variable classifier learns the mini datasets sequentially, ensuring that during the learning of the current mini dataset, there is no access to or reprocessing of the previous dataset. Furthermore, we employed the NB and tree-augmented NB (TAN) classifiers because they facilitate incremental learning.

Fig. 5. A scenario of feature incrementation where new features are continually introduced in additional datasets.

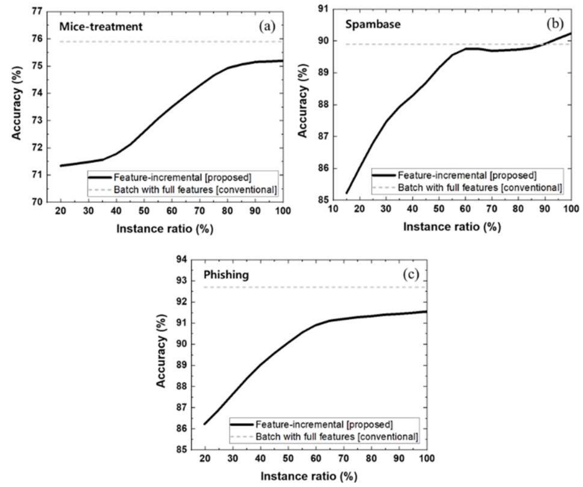

Fig. 6. The learning curves from feature-incremental learning. (a) Mice-treatment with NB, (b) Spambase with NB, (c) Phishing with TAN.

The results of feature-incremental learning for the three selected datasets are presented in Fig. 6. In each graph, the horizontal axis represents the ratio of the dataset that the classifier has learned, and the gray dashed line indicates the accuracy achieved by batch-mode conventional learning on the full-feature dataset before preprocessing. The proposed classifier demonstrates successful incremental learning curves, closely approaching the batch-mode accuracy, despite learning with much less information. This outcome indicates that the proposed latent-variable representation is capable of robust learning in the feature-incremental learning task.

V. Conclusion

We proposed a latent-variable classifier based on total correlation and demonstrated its ability to enhance the performance of batch-mode classification. Furthermore, we experimentally showed that the proposed method enables incremental learning on new datasets containing new features, even without access to past data. Our approach signifies its potential to serve as a robust machine learning model in dynamically changing real-world environments.

Acknowledgment

This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [23ZS1100, Core Technology Research for Self-Improving Integrated Artificial Intelligence System]

References

[1] G. Fang and P. Li, "On estimation in latent variable models," Proc. of the 38th Int. Conf. Mach. Learn. (PMLR) 139, 2021.

[2] X. Bing, F. Bunea, M. Royer, and J. Das, "Latent model-based clustering for biological discovery," iScience, vol. 14, pp. 125–135, 2019.

[3] G. V. Steeg and A. Galstyan, "Discovering structure in high-dimensional data through correlation explanation," Proc. of the 28th Conf. on Neural Inf. Process. Syst. (NeurIPS), 2014.

[4] G. V. Steeg and A. Galstyan, "Maximally informative hierarchical representations of high-dimensional data," Proc. of the 18th Int. Conf. on Artificial Intell. and Stat. (AISTATS), 2015.

[5] G. V. Steeg, "Unsupervised learning via total correlation explanation," Proc. of the 26th Int. Joint Conf. on Artificial Intell. (IJCAI), 2017.