# DISCRIMINATIVE PATTERN WITH MULTIPLE SUPPORT VECTOR MACHINE (MSVM) PREDICTION FRAMEWORK (DPMSvmPF)

B.Sabarish[1],B.Prem Kumar[2],C. Daniel NesaKumar[3]

*[1]MCA Student, Hindusthan College of Arts and Science, Coimbatore*

*[2]MCA Student, Hindusthan College of Arts and Science, Coimbatore*

*[3]Asst Professor, Department of MCA, Hindusthan College of Arts and Science,Coimbatore*

*([1]sabarishdbs@gmail.com,[2]prempkkumar007@gmail.com,[3]danielnesakumar@gmail.com)*

***Abstract-***In the review, two series of methods have been introduced to solve prediction issues which consist of classification and regression. Straightforward methods such as generalized linear methods have normal accuracy however well-built interpretability on a set of straightforward attributes. The other series, consists of tree-based methods, organize mathematical, unconditional and high dimensional attributes into a complete formation by means of prosperous interpretable information in the data. In this work proposed a new Discriminative Pattern with Multiple Support Vector Machine (MSVM) Prediction Framework (DPMSvmPF) in the direction of achieve the prediction steps with attractive their merits of together efficiency and interpretability. Specifically, DPMSvmPF make use of the brief discriminative patterns with the purpose of are on the prefix hyperplances in the classifier. Furthermore, DPMSvmPF chooses a restricted no. of the useful discriminative patterns with searching designed for the mainly successful pattern combination in the direction of fit universal linear classifiers. At last compress the no. of discriminative patterns with choosing the mainly effective pattern combinations with the purpose of fit addicted to a widespread classifier model with increased prediction results. This part of fast and successful pattern extraction authorizes the well-built preventability and interpretability of DPMSvmPF.

***Index terms-****Discriminative Pattern;Generalized Linear Model;Classification;Regression;Multiple Support Vector Machine(MSVM);Discriminative Pattern with Multiple Support Vector Machine(MSVM) Prediction Framework (DPMSvmPF).*

## I. INTRODUCTION

For informational indexes with class labels, association patterns [1] with the purpose of happen with irregular recurrence in a few classes versus others can be of significant value. Additionally allude to them as discriminative examples [2-3] in this work, in spite of the fact that these examples have likewise been explored under different names, for example, rising examples [4], differentiate sets [4] and directed descriptive standards [5]. Discriminative examples have been appeared to be valuable for enhancing the prediction results [6] and for circle covering test subgroups [5]. The calculations for finding discriminative examples normally utilize a measure for the discriminative intensity of an example. Such measures are by and large characterized as a component of the example's relative help in the two classes, and can be characterized both essentially as the proportion of the different varieties, for example, information gain, and Gini record and so forth.

Frequent design based characterization has been investigated lately and its capacity was shown by numerous examinations in a few spaces, including (1) association classification [7] on straight out information, where a classifier is fabricated in view of high- support, high-confidence rules; and (2) frequent pattern based classification [8] in light of content or information with complex structures, for example, successions and diagrams, where discriminative frequent examples are taken as highlights to high quality classifiers. A successive itemset (pattern) is an arrangement of things that happen in a dataset no not as much as a user determined least support (minsup). Visit designs have been investigated generally in grouping undertakings. These examinations accomplish promising grouping precision and show the achievement of successive examples in classification.All the more imperatively, the grouping assignments connect extraordinary significance to the successive itemsets that are profoundly discriminative as for the class labels. Since visit itemsets are produced exclusively in view of help data, not founded on discriminative power, countless itemsets can be created among the mining step. At the point when the total mining comes about are restrictively substantial, yet just the exceedingly discriminative ones are of genuine intrigue, it is wasteful to sit tight perpetually for the mining calculation to complete and after that apply highlight choice to post-process the tremendous measured mining comes about. Notwithstanding for a feature selection with straight difficulty, it could be extremely costly to process a hazardous number, for example, millions, of highlights which is a typical scale in frequent patterns.In this work proposed a new Discriminative Pattern with Multiple Support Vector Machine (MSVM) Prediction Framework (DPMSvmPF) in the direction of achieve the prediction steps with attractive their merits of together efficiency and interpretability. Furthermore, DPMSvmPF chooses a restricted no. of the useful discriminative patterns with searching designed for the mainly successful pattern combination in the direction of fit universal linear classifiers.

## II. LITERATURE REVIEW

Fang et al [9] introduced a far reaching discourse that characterizes these four example composes and researches their properties and their relationship to each other. Furthermore, these thoughts are investigated for an assortment of datasets (ten UCI datasets, one gen e articulation dataset and two hereditary variety datasets). Classify discriminative examples into four gatherings in view of the accompanying kinds of cooperations: (I) driver-traveler, (ii) coherent, (iii) autonomous added substance and (iv) synergistic past independent option. The outcomes show the presence, attributes and measurable noteworthiness of the distinctive kinds of examples. They likewise delineate how design portrayal can give novel bits of knowledge into discriminative example mining and the discriminative structure of various datasets.Li et al [10] proposed a way to deal with remove the discriminative examples for effective human activity acknowledgment. Each activity is considered to comprise of a progression of unit activities, every one of which is spoken to by an example. Given a skeleton grouping, first naturally separate the key-outlines for unit activities, and afterward mark them as various examples. Likewise further utilize a factual metric to assess the discriminative capacity of each example, and characterize the sack of solid examples as neighborhood highlights for activity acknowledgment. Test comes about demonstrate with the purpose of the separated nearby descriptors could give high exactness in the activity acknowledgment, which show the productivity of strategy in extricating discriminative examples.He et al [11] considered a novel thought named restrictive discriminative example to address this issue. To mine restrictive discriminative examples, proposed a Conditional Discriminative Patterns Mining (CDPM) calculation to produce an arrangement of non-repetitive discriminative examples. Test comes about on genuine informational indexes show that CDPM has great execution on expelling repetitive examples with the purpose of gotten from noteworthy sub-designs to create a compact arrangement of important discriminative examples.Cheng et al [12] proposed a few all inclusive models to screen cardiopulmonary conditions, including DPClass, a novel learning approach. Likewise precisely plan movement dataset covering status from GOLD 0 (solid), GOLD 1 (mellow), GOLD 2 (direct), the distance to GOLD 3 (serious). Sixty-six subjects take an interest in this investigation. After de-distinguishing proof, their strolling information is connected to prepare the prescient models. The Radial Bias Function (RBF) with Support Vector Machine (SVM) display yields the most exactness while the DPClass demonstrate gives better understanding of the model systems. Not just give promising answers for screen wellbeing status by essentially conveying a cell phone, yet in addition exhibit how socioeconomics impacts prescient models of cardiopulmonary illness.

Yan et al [13] proposed a novel ordering model in view of discriminative successive structures that are distinguished through a diagram mining process. The proposed approach not just gives an exquisite answer for the graph indexing issue, yet in addition shows how database ordering and inquiry preparing is able to profit by information mining, particularly Frequent Pattern Mining (FPM). Results demonstrate that the minimal list worked under this model can accomplish better execution in preparing graph inquiries. Since discriminative successive structures catch the inherent qualities of the information, they are moderately steady to database refreshes, in this way encouraging examining based element extraction and incremental file upkeep. Besides, the ideas created here can be summed up and connected to ordering groupings, trees, and other muddled structures too.Fang et al [14] tended to the need of exchanging off the fulfillment of discriminative example disclosure with the effective revelation of low-support discriminative examples from such informational indexes. Likewise proposed a group of antimonotonic measures named SupMaxK with the reason for sort out the arrangement of discriminative examples into settled layers of subsets, which are logically more total in their scope, however require progressively more calculation. Specifically, the individual from SupMaxK with K = 2, named SupMaxPair, is appropriate for thick and high-dimensional informational. Examinations on both manufactured informational indexes and a malignancy quality articulation informational collection show that there are low-bolster designs that can be found utilizing SupMaxPair however not by existing methodologies. Besides, we demonstrate that the low- support discriminative examples that are just found utilizing SupMaxPair from the tumor quality articulation informational collection are factually critical and organically important. This delineates the complementarity of SupMaxPairXo existing methodologies for discriminative pattern discovery.

## III. PROPOSED WORK

In this work proposed a new Discriminative Pattern with Multiple Support Vector Machine (MSVM) Prediction Framework (DPMSvmPF) in the direction of achieve the prediction steps with attractive their merits of together efficiency and interpretability. Specifically, DPMSvmPF make use of the brief discriminative patterns with the purpose of are on the prefix hyperplances in the classifier.

Furthermore, DPMSvmPF chooses a restricted no. of the useful discriminative patterns with searching designed for the mainly successful pattern combination in the direction of fit universal linear classifiers. At last compress the no. of discriminative patterns with choosing the mainly effective pattern combinations with the purpose of fit addicted to a widespread classifier model with increased prediction results. This part of fast and successful pattern extraction authorizes the well-built preventability and interpretability of DPMSvmPF.

3.1. Problem Formulation

For an prediction task, the information is an arrangement of n cases in a d-dimensional component space together with their labels $(x_1, y_1)(x_2, y_2), \ldots.(x_n, y_n)$, for $\forall i\ (1 \le i \le n), x_i \in R^d$. It is significant that the qualities in the illustration xi can be either ceaseless (numerical) or discrete (clear cut). As absolute highlights can be changed into a few double sham pointers are can accept xi without the loss of consensus. The class $y_i$ is either a class (type) marker or a genuine number contingent upon the particular errand. In past example based techniques, designs are extricated from unmitigated qualities and in this manner they are just ready to deal with the nonstop factors after cautious manual discretization, which is precarious and frequently requires earlier learning about the information. Proposed DPMSvmPF structure is to take in a brief model that comprises of a little arrangement of discriminative examples from the preparation information, which learns and predicts the cases as precisely as would be prudent, i.e., anticipate the right class marker in order assignments and foresee near the genuine number in regression tasks.Condition is a thresholding boolean capacity on a particular element measurement. The condition is as $(x_{\cdot j} < v)$ or $(x_{\cdot j} \ge v)$, where j shows the particular measurement and v is the limit value. The relational operator in a condition is either < or ≥. For any measurement j in highlights comparing to paired pointers, we limit v to be 0:5. Example is a conjunction clause of conditions on particular element measurements. Discriminative Patterns allude to those examples which have solid flags on the learning undertakings, given the names of information. For instance, an example with high information gain on the characterization preparing information, or an example with little mean square mistake on the relapse preparing information, is a discriminative example. Discriminative examples have covered prescient impacts. In particular, a couple of discriminative examples are unique instances of different examples. Be that as it may, the second example just encodes a subset of information focuses that the primary example encodes, and along these lines, it doesn't give additional data to the learning procedure. This normal wonder demonstrates that generally taking the best discriminative examples in view of free heuristics squanders the financial plan of the quantity of examples, when the straight blend of these examples are not synergistic. Top-k Patterns are formalized as a size-k subset of discriminative examples, which has the best execution in light of the preparation information.

A. Discriminative Pattern Generation

The primary segment in the DPMSvmPF system is the age of superb discriminative examples. Utilize hyperplane to elude the arrangement of cases falling into a particular hub in the machine learning classifier. SVM multi-class order is most likely the one-against-all technique. It builds k SVM models where k is the quantity of classes. The $m^{th}$ SVM is prepared with the greater part of the cases in the $m^{th}$ class with positive labels, and every single other case with negative labels. Consequently known l training information $(x_1, y_1), \ldots, (x_l, y_l)$, where $x_i \in R^n, i = 1, \ldots l$ and $y_i \in (1, \ldots k)$ is the class of $x_i$, the $m^{th}$ SVM resolve the subsequent issue:

$$\min_{w^m, b^m, \xi^m} \frac{1}{2} (w^m)^T w^m + C \sum_{i=1}^{l} \xi_i^m$$

$$(w^m)^T \phi(x_i) + b^m \ge 1 - \xi_i^m, if\ y_i = m,$$

$$(w^m)^T \phi(x_i) + b^m \le -1 + \xi_i^m, if\ y_i \ne m$$

$$\xi_i^m \ge 0, i = 1, \ldots l$$

where the training data $x_i$ are mapped in the direction of a higher dimensional space with the function $\phi$ and C is the fine factor. Reducing $\frac{1}{2} (w^m)^T w^m$ means with the purpose of we would like in the direction of maximize $2/||w^m||$, the margin among two clusters of data. When data are not linear distinguishable, there is a fine factor $C \sum_{i=1}^{l} \xi_i^m$ which be able to decrease the no. of training errors. The basic idea following SVM is in the direction of search meant for a balance among the regularization period $\frac{1}{2} (w^m)^T w^m$and the training errors. Following solving (1), there are k conclusion task:

$$(w^1)^T \phi(x) + b^1,$$
$$\vdots$$
$$(w^k)^T \phi(x) + b^k$$

Articulate x is in the group which have the main value of the choice function:

$$class\ of\ x \equiv \arg\max_{m=1,\ldots k} ((w^m)^T \phi(x) + b^m)$$

For all intents and purposes we take care of the double issue of (1) whose number of factors is the same as the quantity of information in (1). Consequently k l-variable quadratic programming issues are solved. Another significant strategy is known as the one-against-one technique. It was presented in [16], and the main utilization of this procedure on SVM was in [17]. This technique develops k(k - 1)=2 classifiers where everyone is prepared on information from two classes. For preparing information from the $i^{th}$ and the $j^{th}$ classes, we fathom the accompanying double characterization problem:

$$\min_{w^{ij},b^{ij},\xi^{ij}} \frac{1}{2}(w^{ij})^T w^{ij} + C\sum_{i=1}^{l} \xi_t^{ij}$$

$$(w^{ij})^T \phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, if\ y_t = i,$$

$$(w^{ij})^T \phi(x_t) + b^{ij} \leq -1 + \xi_i^{ij}, if\ y_t = j$$

$$\xi_t^{ij} \geq 0, i = 1, \ldots l$$

There are distinctive techniques for doing the future testing after all k(k-1)=2 classifiers are built. After a few tests, we choose to utilize the accompanying voting procedure : if sign(($w^{ij}$)T ϕ (x) + $b^{ij}$)) says x is in the $i^{th}$ class, at that point the vote in favor of the $i^{th}$ class is included by one. Something else, the $j^{th}$ is expanded by one. At that point foresee x is in the class with the biggest vote. The voting approach portrayed above is likewise called the "Maximum Wins" technique. In the event that those two classes have indistinguishable votes, figured it may not be a decent technique, now we essentially select the one with the smaller index. For each discriminative example, there is one comparing twofold measurement depicting whether the occurrences fulfill the example or not. Since the measurement of the example space is equivalent to the quantity of discriminative examples which is an expansive number after the age stage, we have to additionally choose a set number of examples and therefore influence the example to space little and productive. It is additionally justified regardless of a say that this mapping procedure can be completely parallelized for speedup.

B. Top-k Pattern Selection

After a substantial pool of discriminative examples is created, additionally top-k determination should be done to recognize the most informative and interpretable examples. A guileless path is to utilize heuristic capacities, for example, information gain and gini index, to assess the noteworthiness of various examples on the prediction and pick the best positioned designs. In any case, the impacts of best ranked patterns in view of the basic heuristic scores may have a huge segment of covers and in this manner their blend does not work ideally. Consequently, to accomplish the best execution and discover integral examples, propose two effective arrangements: forward choice and LASSO, which settle on choices in light of the impacts of the example mixes as opposed to considering distinctive pattern autonomously.

## IV. EXPERIMENTATION RESULTS

In this area, direct broad trials to exhibit the interpretability, proficiency and adequacy of proposed DPMSvmPF and Discriminative Pattern-based Prediction structure (DPPred) [18]. Additionally first present test settings, examine the effectiveness and interpretability, and afterward give the outcomes on grouping and relapse undertakings and also parameter examination. A few characterization datasets from University of California, Irvine (UCI) Machine Learning Repository are utilized as a part of the analyses, as appeared in Table 1 with measurements of the quantity of cases and the quantity of highlights. In the datasets adult, hypo and sick, the proportion of standard prepare/test part is 2:1. Along these lines, for the other grouping and relapse datasets, we separate the datasets into prepare/test (2: 1) by unprejudiced examining as preprocessing. For classification tasks, to contrast and DDPMine, utilize the same datasets including adult, hypo, sick, crx, sonar, chess, waveform, and mushroom.. Since both DDPMine and DPPred accomplish relatively close precision on the datasets waveform and mushroom, these two datasets are overlooked. What's more, the execution of DPPred and DPMSvmPF on high-dimensional datasets (nomao, musk and madelon datasets) is additionally explored, since DDPMine performs ineffectively on high-dimensional information. The metric is the precision on the testing information: higher exactness implies better execution. For relapse datasets, we pick general datasets, for example, bike and crime, and additionally clinical datasets where designs will probably be available, for example, parkinsons. Moreover, to make the mistakes in various datasets practically identical, min-max standardization is embraced to scale the consistent names into [0; 1]. The metric is the Rooted Mean Square Error (RMSE) on the testing information: a lower RMSE implies better execution.

*Table.1 The statistics of real-world datasets from UCI Machine Learning Repository*

| Type | Dataset | # instances | # dimensions | Variable type |
|---|---|---|---|---|
| **Classification** | Adult | 45222 | 14 | Mixed |
| | Hypo | 3772 | 19 | Mixed |
| | Sick | 3772 | 19 | Mixed |
| | Chess | 28056 | 6 | Mixed |
| | Crx | 690 | 15 | Mixed |
| | Sonar | 208 | 60 | Numeric |

| | | | | |
|---|---|---|---|---|
| **High dimension** | Nomao | 29104 | 120 | Mixed |
| | Musk | 7074 | 166 | Numeric |
| | Madelon | 1300 | 500 | Numeric |
| **Regression** | Bike | 17379 | 10 | Mixed |
| | Parkinsons | 5875 | 16 | Numeric |
| | Crime | 1994 | 99 | Numeric |

Think about proposed DPMSvmPF classifier on the same datasets utilized as a part of DPPred-F and DPPred-L are appeared in the figure 1 and figure 2. The outcomes are appeared in Table 2 and table 3. DPPred-F and DPPred-L dependably have higher precision over other existing techniques. An essential reason of this preferred standpoint is that the applicant designs created by tree-based models in DPPred and DPMSvmPF are significantly more discriminative and hence more compelling on the particular characterization undertaking than those successive however less valuable examples. All the more shockingly, DPMSvmPF shows far better execution than the complex DPPred display on a few datasets, while its correctnesses on different datasets are as yet similar with RF, which is because of the viability of the example determination module where we select the ideal example mix as opposed to choosing designs autonomously. This demonstrates the proposed show is extremely compelling in arrangement undertakings while it is profoundly brief and interpretable.

*Table.2 Test accuracy on classification datasets*

| **Dataset /classifiers** | **Accuracy(%)** | | | | | |
|---|---|---|---|---|---|---|
| | **Adult** | **Hypo** | **Sick** | **Crx** | **Sonar** | **Chess** |
| **DPMSvmPF** | 87.52 | 99.25 | 98.63 | 89.93 | 88.36 | 93.51 |
| **DPPred-F** | 85.93 | 99.06 | 98.12 | 88.21 | 86.14 | 91.54 |
| **DPPred-L** | 84.02 | 99.21 | 97.53 | 88.12 | 84.41 | 91.91 |
| **DT** | 83.69 | 92.05 | 92.36 | 76.69 | 72.69 | 89.05 |
| **DDPMine** | 81.25 | 90.56 | 90.58 | 86.36 | 75.79 | 90.08 |

*Table.3 Test error rate on classification datasets*

| **Dataset /classifiers** | **Error Rate (%)** | | | | | |
|---|---|---|---|---|---|---|
| | **Adult** | **Hypo** | **Sick** | **Crx** | **Sonar** | **Chess** |
| **DPMSvmPF** | 12.48 | 0.75 | 1.37 | 10.07 | 11.64 | 6.49 |
| **DPPred-F** | 14.07 | 0.94 | 1.88 | 11.79 | 13.86 | 8.46 |
| **DPPred-L** | 15.98 | 0.79 | 2.47 | 11.88 | 15.59 | 8.09 |
| **DT** | 16.31 | 7.95 | 7.64 | 23.31 | 27.31 | 10.95 |
| **DDPMine** | 18.75 | 9.44 | 9.42 | 13.64 | 24.21 | 9.92 |

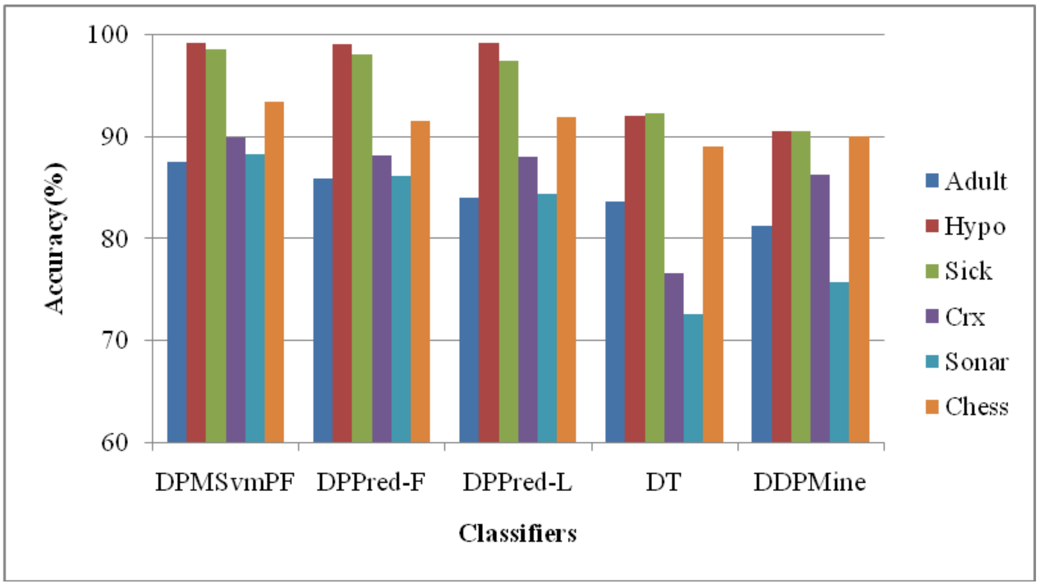


Fig.1: Classification accuracy results comparison

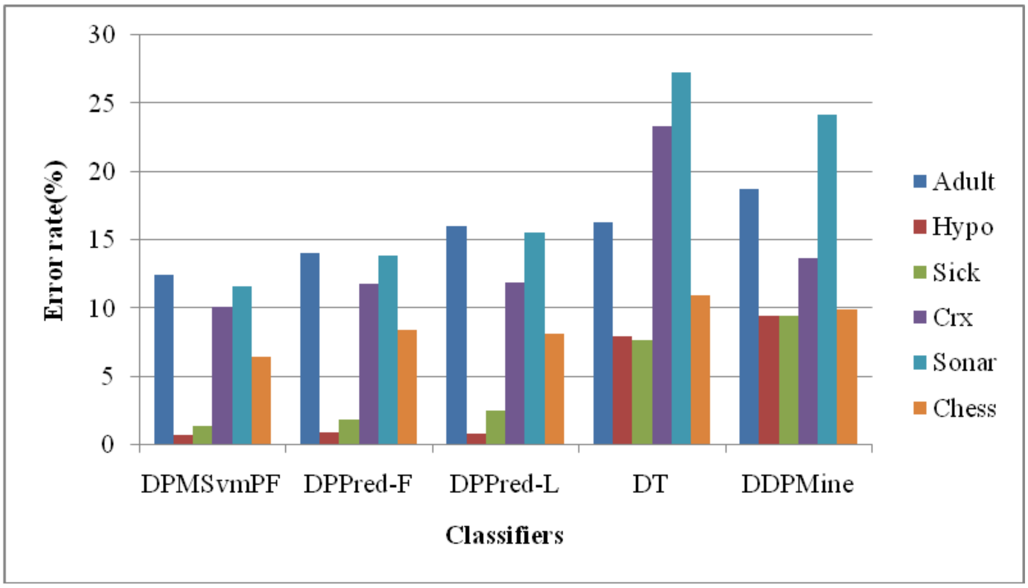


*Fig.2:Error rate results comparison*

## V. CONCLUSION

In this paper, we proposes an new Discriminative Pattern with Multiple Support Vector Machine (MSVM) Prediction Framework (DPMSvmPF) to address the classification issues and give high results with a lesser amount of discriminative patterns. Furthermore, DPMSvmPF selects a restricted no. of the useful discriminative patterns with searching designed for the mainly successful pattern combination in the direction of fit universal linear classifiers. The size of discriminative patterns is compacted by choosing the generally successful pattern combinations related in the direction of their analytical results in a widespread classifier. Extensive results shows with the purpose of DPMSvmPF is capable in the direction of classifier high-order interactions and presented a small number of interpretable patterns in the direction of assist human experts recognize the data. DPMSvmPF gives comparable or even better performance than the state-of-the-art methods DPMSvmPF and random forest model in classification.

## REFERENCES


[1]. P.-N. Tan, M. Steinbach, and V. Kumar. Introduction to data mining. Addison-Wesley ,2005.

[2]. G. Fang, R. Kuang, G. Pandey, M. Steinbach, C. Myers, and V. Kumar. Subspace differential coexpression analysis: problem definition and a general approach. Pacific Symposium on Biocomputing , 15:145–56, 2010.

[3]. G. Fang, G. Pandey, W. Wang, M. Gupta, M. Steinbach, and V. Kumar. Mining low-support discriminative patterns from dense and high-dimensional data. IEEE TKDE, 2011.

[4]. G. Dong and J. Li. Efficient mining of emerging paterns: Discovering trends and differences. In Proc. of SIGKDD , pp. 43–52, 1999.

[5]. P. Novak, N. Lavrac, and G. Webb. Supervised Descriptive Rule Discovery: A Unifying Survey of Contrast Set, Emerging Pattern and Subgroup Mining. Journal of Machine Learning Research, 10:377–403, 2009.

[6]. H. Cheng, X. Yan, J. Han, and C.-W. Hsu. Discriminative frequent pattern analysis for effective classification. In ICDE , pp. 716–725, 2007.

[7]. A. Veloso, W. M. Jr., and M. Zaki, "Lazy associative classification," in Proc. of ICDM , 2006, pp. 645–654.

[8]. M. Deshpande, M. Kuramochi, N. Wale, and G. Karypis, "Frequent substructure-based approaches for classifying chemical compounds," IEEE Trans. Knowl. and Data Eng. , vol. 17, no. 8, pp. 1036–1050, 2005.

[9]. Fang, G., Wang, W., Oatley, B., Van Ness, B., Steinbach, M. and Kumar, V., 2011. Characterizing discriminative patterns. arXiv preprint arXiv:1102.4104.

[10]. Li, J., Liu, G. and Wong, L., 2007, Mining statistically important equivalence classes and delta-discriminative emerging patterns. In Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 430-439). ACM.

[11]. He, Z., Gu, F., Zhao, C., Liu, X., Wu, J. and Wang, J., 2017. Conditional discriminative pattern mining: Concepts and algorithms. Information Sciences, 375, pp.1-15.

[12]. Cheng, Q., Shang, J., Juen, J., Han, J. and Schatz, B., 2016, Mining discriminative patterns to predict health status for cardiopulmonary patients. In Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, pp. 41-49.

[13]. Yan, X., Yu, P.S. and Han, J., 2005. Graph indexing based on discriminative frequent structure analysis. ACM Transactions on Database Systems (TODS), 30(4), pp.960-993.

[14]. Fang, G., Pandey, G., Wang, W., Gupta, M., Steinbach, M. and Kumar, V., 2012. Mining low-support discriminative patterns from dense and high-dimensional data. IEEE Transactions on Knowledge and Data Engineering, 24(2), pp.279-294.

[15]. Lo, D., Cheng, H., Han, J., Khoo, S.C. and Sun, C., 2009, Classification of software behaviors for failure detection: a discriminative pattern mining approach. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 557-566.

[16]. S. Knerr, L. Personnaz, and G. Dreyfus. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In J. Fogelman, editor, Neurocomputing: Algorithms, Architectures and Applications. Springer-Verlag, 1990.

[17]. U. H.-G. Kressel. Pairwise classification and support vector machines. In B. Sch¨olkopf, C. J. C. Burges, and A. J. Smola, editors, Advances in Kernel Methods – Support Vector Learning, pp. 255–268, Cambridge, MA, 1998. MIT Press.

[18]. Shang, J., Jiang, M., Tong, W., Xiao, J., Peng, J. and Han, J., 2018. DPPred: An Effective Prediction Framework with Concise Discriminative Patterns. IEEE Transactions on Knowledge and Data Engineering, 30(7), pp.1226-1239.