A combined Bayesian network method for predicting drive failure times from SMART attributes

Shuai Pang
Email: pangshuai@nbjl.nankai.edu.cn
Yuhan Jia
Email: jiayuhan@nbjl.nankai.edu.cn
Rebecca Stones
Email: rebecca.stones82@nbjl.nankai.edu.cn
Gang Wang
Email: wgzwp@nbjl.nankai.edu.cn
Xiaoguang Liu
Email: liuxg@nbjl.nankai.edu.cn
College of Computer and Control Engineering,
Nankai University, Tianjin, China

***Abstract*—Statistical and machine learning methods have been proposed to predict hard drive failure based on SMART attributes, and many achieve good performance. However, these models do not give a good indication as to when a drive will fail, only predicting that it will fail. To this end, we propose a new notion of a drive's health degree based on the remaining working time of hard drive before actual failure occurs. An ensemble learning method is implemented to predict these health degrees: four popular individual classifiers are individually trained and used in a Combined Bayesian Network (CBN). Experiments show that the CBN model can give a health assessment under the proposed definition where drives are predicted to fail no later than their actual failure time 70% or more of the time, while maintaining prediction performance standards at least approximately as good as the individual classifiers.**

***Index Terms*—Combined Bayesian Network, Ensemble Learning, SMART, Hard Drive Failure Prediction**

I. Introduction

A. Problem Description

Most Internet services are provided by data centers where hard drives serve as the main storage devices. Although hard drives are reliable in general, they are still the most commonly replaced hardware components in most data centers [1]. If a server crashes due to hard drive failure, it will not only decrease the overall user experience, but also cause a direct economic loss to the service provider. Therefore, predicting drive failures before they actually occur can enable operators to take actions in advance and avoid unnecessary losses.

Most major hard drives now support Self-Monitoring, Analysis, and Reporting Technology (SMART), which measures drive characteristics such as operating temperature, data error rates, spin-up time, etc. [2], [3]. Certain trends and sudden changes in these parameters are thought to be associated with an increased likelihood of drive failure. However, SMART parameters alone may not be sufficiently useful for predicting drive failures; combinations of SMART parameters can have an impact on failure probability. Hard drive manufacturers estimated that the threshold-based algorithm implemented in drives can only obtain a failure detection rate (FDR) of 3.10% with a low false alarm rate (FAR) of the order of 0.1% [3]. To improve predicting accuracy, some machine learning methods based on SMART attributes have been used. Li et al. explore building hard drive failure prediction models based on classification and regression trees [4]. Hamerly and Elkan employed two Bayesian approaches to predict hard drive failures based on SMART attributes [5]. Hughes et al. proposed two statistical methods to improve SMART prediction accuracy; they proposed two different strategies: a multivariate rank-sum test and an OR-ed single variate test [6]. Murray et al. proposed a new algorithm based on the multiple-instance learning framework and the naive Bayesian classifier [7]. Zhao et al. employed Hidden Markov Models and Hidden Semi-Markov Models to predict hard drive failures [8].

Despite hard drives deteriorating gradually in practice, most prediction approaches are binary classifiers which just predict whether a drive failure will occur in a certain period of time. As such, they do not give a detailed indication of a drive's health, which could be used to provide more response options and time to back up or migrate data and analyze the reasons for drive failure. Li et al. present a Regression Tree (RT) model to evaluate the health degree which can give a health assessment for a drive rather than a simple classification result [4]. They define a drive's health degree as the fault probability. A drive's health degree can be utilized to represent trends in drive failure. Therefore, the operator can respond to warnings raised by the prediction model in order of their health degrees. However, there are two drawbacks for RT model: (a) Health

978-1-5090-0620-5/16/$31.00

degrees are defined by probabilities, which is not intuitive. If we know the health degree of a drive is 50%, for example, the remaining working time of the drive is unknown and can still be unclear when to back up or migrate data. (b) The prediction accuracy of RT model is low. The result of RT model can be considered as a curve and it should decline as the health degree decreases under ideal conditions. However, most curves generated by RT model are clustered.

B. Our Contribution

In this paper, the notion of a drive's health degree is adapted to encapsulate failure trends and an ensemble learning method is implemented to predict these health degrees. Our main contributions are listed below:

1) We propose a new definition of health degree, defined by the remaining working time of a hard drive before actual failure occur. The longer the remaining working time, the healthier the hard drive.
2) To predict the health degree of hard drives, we implement a Combined Bayesian Network (CBN) model. We train four individual classifiers using Backpropagation Artificial Neural Networks (BP ANN), Evolutionary Neural Network (ENN), Support Vector Machine (SVM) and Classification Tree (CT) methods. To predict the remaining working time of hard drive with real-world data, we implement CBN to combine the learning results from these individual classifiers. Experiments show that this ensemble learning method achieves good performance on predicting health degrees based on given features.

II. Methods and Implementation

A. Definition of Health Degree

Previously, health degree is defined as a measure of the likelihood for hard drive failure: the higher the fault probability, the lower the health degree for hard drives. However, a probability-based definition may not be particularly useful to an operator. Here, we redefine a drive's health degree as the remaining working time of a drive before actual failure occurs: the longer remaining working time, the healthier for hard drives.

To obtain the health degree of a hard drive under this new definition, we need to predict its remaining working time. We do this using the drive's SMART attributes. We divide the remaining working time of hard drives into several intervals according to the degree of urgency, for example, we use 0 to 48 hours as a level of health degree. If we predict the remaining working time of a drive is within this interval, then we migrate data immediately to ensure the data security. If instead the remaining working time is more than 500 hours, the drive can continue operating for the data on the drive will be safe for a relative long period of time. Therefore, each interval represents a class label (i.e., each interval represents a level of health degree), and a classifier for predicting health degree can be constructed.

B. Combined Bayesian Networks

We first train BP ANN, ENN, SVM and CT as individual classifiers, which are combined using a Bayesian Network, to give the Combined Bayesian Network (CBN). A Bayesian Network (BN) is a probabilistic network model where nodes are variables and edges represent their probabilistic interdependencies. A BN can be used for classification and the nodes are divides into two kinds: the class variable and feature variables; we infer the class variable from feature variables. There are three motivating reasons for using a CBN model:

1) To construct a individual classifier with high classification accuracy is difficult, but it is straightforward to construct several simple classifiers and combine them, giving a higher classification accuracy than the individual classifiers [9], [10], [11], [12], [13].
2) We train BP ANN, ENN, SVM and CT as individual classifiers for they have been shown to perform well in detecting failed drives [3], [4], [14] and they have simple construction which are easy to train. However, BP ANN, ENN, SVM and CT are binary classifiers, and we usually need more than two health degrees, so they can not be used to predict health degree individually. BN give a classification that admits multiple class labels.
3) The outputs of BP ANN, ENN, SVM and CT contain some trend of stand or fall of drives [4]. For example, we set the class label of BP ANN for good drives as 1 and drives that fail as 0, and we set thresholding used to get the binary result as 0.5 (i.e., if the result of BP ANN is greater than 0.5, the corresponding drive will be classified as good), so if the output of BP ANN for drive A is 0.9, and 0.1 for drive B, we can conclude that compared with drive A, drive B is more likely to fail. In fact, the outputs of individual classifiers reflect the distances of samples to the optimal hyperplane constructed by classifiers and they serve as evidence which indicates how well drives work. So the outputs of individual classifiers can be treated as features in BN.

1) Structure of CBN: In the CBN model, we consider the predicted health degree of hard drives as the class variable and the outputs of BP ANN, ENN, SVM and CT as feature variables. To construct the model, we set the class variable node as the parent of all feature variable nodes, as in a Naive Bayesian Network. Dependency relationships between feature variables are based on the following consideration: for one record, the results of BP ANN, ENN, SVM and CT can reflect each other's outputs. Once we know that the prediction results of CT have a certain trend, the results of the other three classifiers are likely to display the same trend. For example, if the prediction result of a drive is good by CT, then ENN, SVM and BP ANN may get the same result. Moreover, conditioning on the prediction results of ENN and SVM can also increase the probability of BP ANN's prediction results to get certain value. This relationship is reflected by adding edges between variables in CBN and the direction of edges are determined by predicting performance of each individual classifiers. We

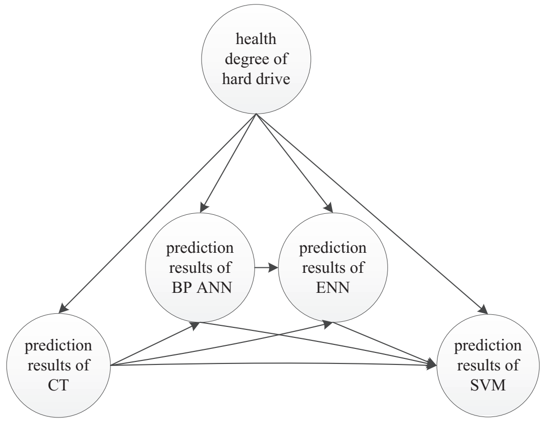

Fig. 1. Structure of the Combined Bayesian Network.

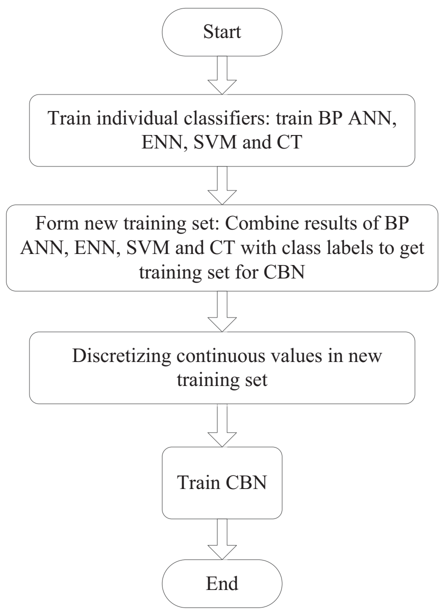

Fig. 2. Flowchart of the CBN training method.

make CT a parent node of ENN, SVM and BP ANN. We make BP ANN a parent node of ENN and SVM, and ENN a parent node of SVM. The structure of CBN is shown in Fig. 1.

2) Training and Prediction of CBN: To train the CBN model, we first train the four classifiers with the same training set. For each record in the training set, we combine the four results predicted by the individual classifiers with the health degree level of the record to get a new training record used to train the CBN model. A flowchart of the CBN training method is given in Fig. 2.

1) *Train individual classifiers.* We use records based on SMART attributes in training set to train BP ANN, ENN, SVM and CT models. The details will be introduced in the next subsection.
2) *Form new training set.* After the individual classifiers are constructed, the records (we call these records original records) in training set are fed into four individual classifiers again to get a new training set for CBN. Each record in the new training set consists of five components which are predicting results of four individual classifiers and corresponding health degree level of original record.
3) *Discretizing continuous values in new training set.* CBNs can be used with discrete and continuous data. However, processing continuous data requires the implementation of integral algorithms, which increases computational burden. The outputs of BP ANN, ENN and SVM are continuous and their value range is respectively $[0, 1]$, $[0, 1]$ and $[-1, 1]$; the output of CT are 1 or -1. Discrete data can highlight the correlation of the variables in a CBN and make learning and inference more reasonable. Therefore, we convert the continuous outputs to a discrete format. The value range is partitioned with equal width and the width is set to 0.002.
4) *Train CBN model.* To fully specify the CBN model and represent the joint probability distribution, it is necessary to specify the conditional probabilities of each node with records in the new training set. Since there are no unobserved variables in our network, we implement the maximum likelihood approach to estimate parameters. If P_{ijk} denotes the conditional probability of $X_i = k$, with parent X_j, the maximum likelihood approach is expressed as follows:

$$P_{ijk} = \frac{N_{ijk}}{N_{ij}}$$

where N_{ijk} is the number of instances in which $X_i = k$ and the parent of variable X_i is X_j, and N_{ij} is the number of instance which the parent of variable X_i is X_j. P_{ijk} represent conditional dependencies which give the probability of the node that takes a particular set of values for the node's parent variables in CBN.

To predict the health degree of drives, we collect attributes based on SMART attributes use them with BP ANN, ENN, SVM and CT to get intermediate results, which are then fed into the CBN to infer the health degree level. Given the results of BP ANN, ENN, SVM and CT, we can compute posterior probabilities of a given class, $\mathrm{P}(C|B, E, S, R)$, where C denotes the health degree, and B, E, S and R respectively denote the prediction result of BP ANN, ENN, SVM and CT. The probability of health degree class $C = i$ conditioned on given BP ANN, ENN, SVM and CT results can be computed as follows:

$$\begin{aligned}\mathrm{P}(C = i|B, E, S, R) &= \mathrm{P}(C|B, E, S, R) \\ &= \frac{\mathrm{P}(C, B, E, S, R)}{\mathrm{P}(B, E, S, R)} \\ &= \frac{\mathrm{P}(C, B, E, S, R)}{\sum_j \mathrm{P}(C = j, B, E, S, R)}\end{aligned}$$

TABLE I
THE INCORPORATED SMART DRIVE FEATURES.

ID	Attribute Name
1	Raw Read Error Rate
2	Spin Up Time
3	Reallocated Sectors Count
4	Seek Error Rate
5	Power On Hours
6	Reported Uncorrectable Errors
7	High Fly Writes
8	Temperature Celsius
9	Hardware ECC Recovered
10	Reallocated Sectors Count
11	Raw Read Error Rate
12	Hardware ECC Recovered
13	Reallocated Sectors Count (raw value)

This is computed by:

$$\begin{aligned}&\mathrm{P}(C=i,B,E,S,R)=\mathrm{P}(C,B,E,S,R)\\&=\mathrm{P}(R|B,E,S,C)\times\mathrm{P}(S|B,E,C)\times\mathrm{P}(E|B,C)\\&\quad\times\mathrm{P}(B|C)\times\mathrm{P}(C).\end{aligned}$$

The probability $\mathrm{P}(C)=\mathrm{P}(C=i)$ is the prior probability of class i and is estimated from the training set:

$$\mathrm{P}(C=i)=\frac{\text{number of class-}i\text{ samples}}{\text{total number of samples}}.$$

C. Individual Classifiers

1) Data collection and features : To construct our classifier models, we collect data from a real-world data center. There are a total of $23,201$ drives in the dataset, labeled "good" or "failed" [14]. During a period of eight weeks, each good drive was sampled every hour. Some records were missed because of sampling or storing errors. For drives which failed, samples in a period of 20 days before actual failure were recorded. Some of drives which failed lost some records if they had not survived 20 days of operation since we began to collect the data. Each drive that failed was marked with its remaining working time before actual failure. For good drives, they were marked with ">480 hours".

The individual classifiers use the same feature collection, which has 13 features selected by statistical methods mentioned in [4], [14], listed in Table I.

2) Backpropagation Artificial Neural Network : In this paper, we implement BP ANN as individual classifier to classify the stand or fall of a hard disk. The 13 features mentioned above serve as input layer nodes. There are one hidden layer with 25 nodes. The output layer has one node and its output value is a positive real number between 0 and 1. A drive would be predicted to be failed only if the output of the neural network is smaller than a threshold. Otherwise, the drive would be predicted to be healthy. The other parameters and training methods are the same as [14].

3) Evolutionary Neural Network : Artificial Neural Networks (ANNs) and Genetic Algorithms (GAs) are both techniques for obtaining basic functions of intelligent information by simulating models of life. ANNs are inspired by brain activity whereas GAs are inspired by the tenets of evolution. The two methods can be combined into an Evolutionary Neural Network (ENN), which utilizes a GA to optimize the weights of an ANN. More details can be found in [15]. In our implementation, a vector of weights is the "chromosome" (with each weight corresponding to a feature in Table I and the squared error between the expected output and the actual output is used as loss function. Then the GA searches for the best solutions (the weights that minimize the loss function) for the ANN. We set the population size (total number of "chromosome") 10 and max searching generations is 500.

4) Support Vector Machine : SVM constructs a set of hyperplanes in a high-dimensional space, which can be used for classification and has been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming. In the CBN model we select select parameters such as kernel function and soft margin parameter as [3] did. The result of SVM is a real number between -1 and 1. Through comparing the output of SVM with the threshold, it can be determined whether a drive is failed.

5) Classification Tree : CT model is a binary recursive partitioning procedure which is capable of processing continuous and nominal attributes as predictors. Each node in CT model represents a feature and is split into two children, then each of the children is in turn split into grandchildren. We use Information Gain as the splitting function in our models. The split procedure searches through all values of the input attributes to find out the best partition variable which maximizes the gain in information. The output of the CT will be a single value, 1 or -1 (healthy or failed). We adopt the practices in [4] for CT model.

The reason for incorporating the CT model into the proposed CBN is that it achieves good performance on detecting at-risk drives [4]. Detecting at-risk drives is a precondition of predicting health degrees which can increase prediction accuracy of healthy drives.

III. EXPERIMENTS AND RESULTS

A. Experimental Setup

1) Training and Test Sets: To evaluate the CBN model more practically, we divide the dataset into training and test sets according to time rather than randomly. There are total of $22,758$ healthy drives. For each healthy drive, we take the earlier 70% of the records within one week as training data, and the later 30% as test data. We randomly choose 6 records per healthy drive from its daily records set in the training set to train classifiers. In this way, we can eliminate the bias of a single drive's sample in a particular hour and provide enough information to describe the health condition of the drive. The test set for healthy drives contains $955,836$ records. Since there are far fewer drives that fail than healthy drives and the chronological order of them was not recorded, we include all

TABLE II
PERFORMANCE AT PREDICTING WHICH DRIVES WILL FAIL.

model:	CBN	BP ANN	ENN	SVM	CT
FAR (%)	0.079	0.286	0.301	0.092	0.070
FDR (%)	95	91	90	83	96

the drives that fail and divide them randomly into training and test sets. The training set has 343 drives and $242,473$ records, and the test set has 100 drives and $37,902$ records.

2) Partitions: The predicted remaining working time of hard drives is divided into intervals. We experiment with three partitions as indicated in Fig. 3. The levels are determined by the degree of urgency: the lower the health degree level, the more urgently a drive needs to be attended to. If two drives are predicted to fail, we can attend to the drive with a lower health degree first. It seems reasonable to suspect that greater separability would be more useful for lower health degrees, and less important for higher health degrees. So for the first and second partition, the interval lengths are unequal: the lower the level, the shorter interval length and less remaining working time. For comparison, the intervals of the third partition are chosen to be the same length (48 hours).

B. Experimental Results

When we predict the health degree of a drive, we adopt a voting-based detection algorithm. It may not be appropriate to predict health degree level of a drive only by one record as a single abnormal record might be due to e.g. measurement noise. When predicting the health degree of a drive, we check the last 5 consecutive records, and the drive's health degree is determined by the classification results obtained by most records, i.e., the mode. If the mode is not unique, we choose the one that results in the lowest health degree, erring on the side of caution.

1) Failure Prediction for Hard Drives : To measure the overall performance of our model, we firstly follow the common practice in previous research in hard drive failure prediction problems. We test the performance of the CBN model at predicting which drives fail in terms of their failure detection rate (FDR) and false alarm rate (FAR); the results are given in Table II. We consider the drives with the highest health degree as healthy drives, and the drives with lower health degree as at-risk ones. As control groups, BP ANN, ENN, SVM and CT are also tested. We measure the performance of these four individual classifiers when they obtain their lowest FAR. The CBN model achieved a FDR of 95% at the FAR of 0.079% with the second partition, outperforming BP ANN, ENN and SVM in terms of both FAR and FDR, but was marginally outperformed by CT. However, unlike the CBN model, the CT model is a binary classifier and does not give a health degree.

2) Health Degree Prediction for Hard Drives : To further test classification accuracy of the CBN model for health degree prediction, we compare the performance on healthy drives and drives that fail separately and adopt different evaluation criterion for these cases. This is motivated by the observations: (a) The classification accuracy of individual classifier is over 90% for predicting healthy drives but only 40% to 60% for predicting drives that will fail. (b) In the test set, the total number of healthy drives is $22,758$ compared with 100 drives that fail; if tested together, the prediction performance will be largely determined by the prediction results for healthy drives.

For healthy drives, to evaluate the CBN model under the three different partitions, we define the precision as

$$\frac{\text{no. healthy drives with highest health degree level}}{\text{no. healthy drives}}.$$

For drives that fail, to test the performance of CBN model, we consider two evaluation criteria. The first is classification precision which is defined as accuracy of the health degree prediction for all test drives. The second criterion is the conservative prediction rate (CPA) which is the proportion of failed drives that were predicted to fail no later than they actually failed, thereby giving sufficient time to attend to these drives.

Table III shows the classification performance of the CBN model. We see that if the first division method is adopted, prediction accuracy is over 60%. The four individual classifiers are binary classifiers which can only predict whether a drive is failed or not rather than its health degree individually, so we construct a multi BP ANN with multiple class labels as a control group. Multi BP ANN is an artificial neural network whose output layer contains more than one node and each node corresponds to a healthy level. The results in Table III show that CBN consistently outperforms the multi-class BP ANN model. The results are identical with expectation. For that, CBN model combines four individual classifiers to perform ensemble. The base task of prediction models can be commonly described as finding a suitable hypothesis by searching through a hypothesis space for a particular problem. Each model represents a single hypothesis. Even if the hypothesis space contains hypotheses which are very well-suited for a problem, it may be very difficult to find a good one. However, combining multiple hypotheses to form a better hypothesis can yield better performance than single model. Moreover, the more class labels, the harder to find a good hypothesis for a single model. As Table III showed, with increase of total number of partitions (healthy degree levels), CBN model yields performance much more better than multi-class BP ANN model.

Fig. 4 plots the classification accuracy of the CBN model under the three different partitions. Here, an accurate prediction is made if a drive is predicted as failing in a given interval and fails in that interval. By increasing of total number of classes, we decrease the interval sizes which means less samples for each interval with the total number of training samples unchanged. As we known, total number of samples has a strong influence on the predicting accuracy of a classifier: the more samples for each label, the higher the accuracy of the classifier. Predicting that a drive belongs to a smaller interval

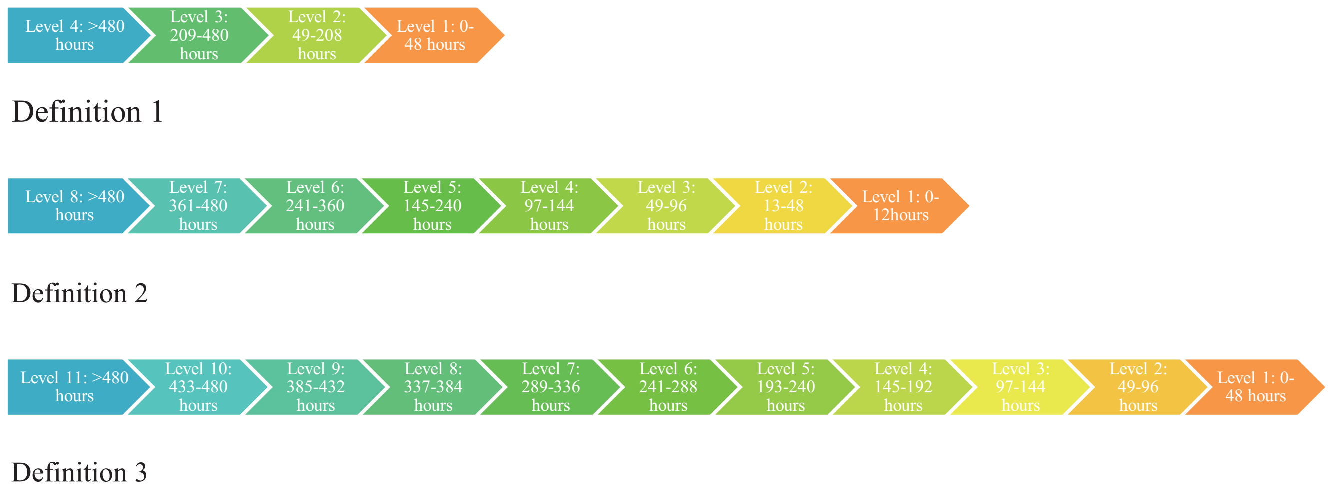

Fig. 3. The three partitions used for classifying a drive's health degree.

TABLE III
PERFORMANCE OF THE CBN MODEL VS. A MULTI-CLASS BP ANN MODEL.

	CBN			multi-class BP ANN		
	Good Drives (%)	Failed Drives (%)		Good Drives (%)	Failed Drives (%)	
	Precision	Precision	CPA	Precision	Precision	CPA
Partition 1	99.0	62.2	83.4	96.5	51.2	71.0
Partition 2	99.2	44.6	74.9	97.7	37.2	62.6
Partition 3	99.2	38.8	67.6	97.2	26.3	53.6

(i.e. less samples) is a stronger claim and consequently more difficult to satisfy, which is why we see an overall decrease in classification accuracy.

IV. CONCLUDING REMARKS

In this paper, we propose a new definition of "health degree" defined by the remaining working time of hard drive. We implement a Combined Bayesian Network model, an ensemble learning method, to estimate the health degree of hard drives. This is achieved by combining the results of four trained individual classifiers, BP ANN, ENN, SVM and CT, using SMART attributes. Experimental results indicate that the CBN model outperforms the BP ANN, ENN and SVM models, and performs comparably to the CT model in terms of FDR and FAR, while having the additional benefit of predicting when failure occurs (instead of just predicting that it will occur).

In future research, it may be worthwhile to try other statistical and machine learning methods to build more effective health degree prediction models, and incorporate them into the Combined Bayesian Network model.

REFERENCES

[1] K. V. Vishwanath and N. Nagappan, "Characterizing cloud computing hardware reliability," in *Proc. 1st ACM Symposium on Cloud Computing*, Indianapolis, IN, 2010, pp. 193–204.
[2] B. Allen, "Monitoring hard disks with SMART," *Linux Journal*, no. 117, Jan 2004.
[3] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado, "Machine learning methods for predicting failures in hard drives: A multiple-instance application," *J. Mach. Learn. Res.*, vol. 6, pp. 783–816, 2005.
[4] J. Li, X. Ji, Y. Jia, B. Zhu, G. Wang, Z. Li, and X. Liu, "Hard drive failure prediction using classification and regression trees," in *Proc. 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, Atlanta, Georgia, USA, 2014, pp. 1–12.
[5] G. Hamerly and C. Elkan, "Bayesian approaches to failure prediction for disk drives," in *Proc. 18th International Conference on Machine Learning*, 2001, pp. 202–209.
[6] G. F. Hughes, J. F. Murray, K. Kreutz-Delgado, and C. Elkan, "Improved disk-drive failure warnings," *IEEE Transactions on Reliability*, vol. 51, no. 3, pp. 350–357, 2002.
[7] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado, "Hard drive failure prediction using non-parametric statistical methods," in *Proc. International Conference on Artificial Neural Networks*, Istanbul, Turkey, 2003.
[8] Y. Zhao, X. Liu, S. Gan, and W. Zheng, "Predicting disk failures with HMM- and HSMM-based approaches," in *Proc. 10th Industrial Conference on Advances in Data Mining: Applications and Theoretical Aspects*, 2010, pp. 390–404.
[9] T. G. Dietterich, "Ensemble learning," in *The Handbook of Brain Theory and Neural Networks*. Cambridge, UK: MIT press, 2002.
[10] M. Keams and L. G. Valiant, "Cryptographic limitation on learning Boolean formulae and finite automata," *J. ACM*, vol. 41, no. 1, pp. 67–95, 1994.
[11] D. Optiz and R. Maclin, "Popular ensemble learning methods: An empirical study," *J. Artif. Intell. Res.*, vol. 11, pp. 169–198, 1999.
[12] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensemble," *Machine learning*, vol. 51, no. 2, pp. 181–207, 2003.
[13] P. Sollich and A. Krogh, "Learning with ensembles: How overfitting can be useful," *Adv. Neural Inf. Process Syst.*, vol. 8, pp. 190–196, 1996.
[14] B. Zhu, G. Wang, X. Liu, D. Hu, S. Lin, and J. Ma, "Proactive drive failure prediction for large scale storage systems," in *Proc. 29th IEEE Conference on Massive Storage Systems and Technologies (MSST)*, Long Beach, CA, 2013, pp. 1–5.

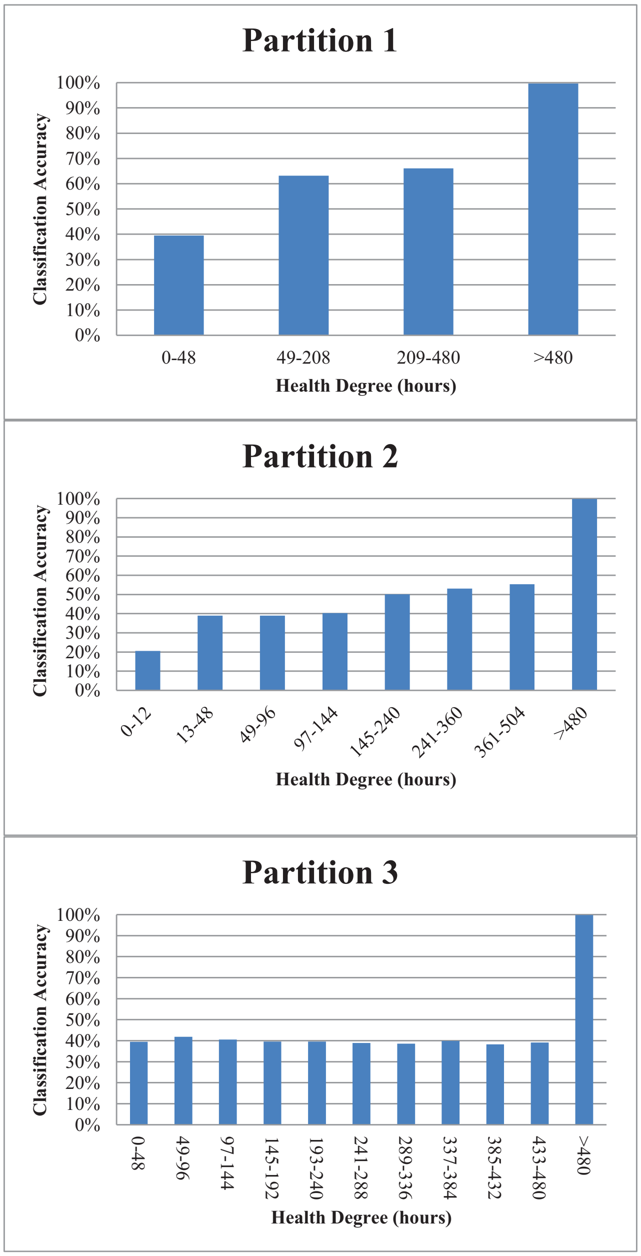

Fig. 4. Prediction accuracy of each health degree levels.

[15] W. Z. Lu, H. Y. Fan, and S. M. Lo, “Application of evolutionary neural network method in predicting pollutant levels in downtown area of hong kong,” *Neurocomputing*, vol. 51, no. 2, pp. 387–400, 2003.