Using Associative Petri Net with Over-Sampling Techniques to Construct an Aviation Incident Prediction Model

Hsiu-Sen Chiang, Tsung-Che Hsieh, Chia-Chen Chen, Neil Y. Yen, and Mu-Yen Chen

***Abstract*—Most current research applying data mining techniques to aviation safety focus on the acquisition and selection of data attribute fields, while the selection and analysis of classification algorithms has been largely overlooked. This study seeks to address this gap as a means of identifying key factors which result in aviation fatalities. The study data is taken from the Accident/Incident Data System (AIDS) flight accident database, which comprises an imbalanced dataset. The Synthetic Minority Oversampling Technique (SMOTE) is used to minimize the impact of data imbalance on classification results for categories with small datasets. Feature attribute selection was implemented through information gain (IG). This study also develops an Associative Petri Net (APN) model for comparison against five other classification methods: which are Naive Bayes, BayesNet, Support Vector Machine (SVM), Decision Tree (C4.5), and Radial Basis Function Network (RBFN). IG is used to identify 10 important factors: Event City, Event State, Flight Phase, Aircraft Model, Aircraft Series, Operator, Primary Flight Type, Flight Conduct Code, Flight Plan Filed Code, and Nbr of Engines. Results show that our proposed APN model had the highest overall accuracy rate, F-measure score, G-mean score among all algorithms. APN is based on the Apriori algorithm to provide a rule-based concept. Thus using the APN algorithm to build the model could produce an expert system for flight accident prediction.**

***Index Terms*—Aviation safety, attribute selection, imbalanced dataset, SMOTE, associative petri net.**

I. Introduction

According to the International Air Transport Association's (IATA) 2015 Safety Report, the air transport industry's average loss rate of passengers is 0.32 per million flights, equivalent to one serious accident every three million flights. In recent years it has become common practice to apply data mining techniques to historical flight accident data to identify measures which can potentially further reduce the risk of aviation accidents. For example, Shyur [1] to identify the most appropriate safety prediction model for aviation and the results of these explorations may lead to a different view from the traditional statistical analysis.

Manuscript received November 1, 2016; revised January 22, 2017. This work was supported by Ministry of Science and Technology of the Republic of China.

H. S. Chiang, T. C. Hsieh, and M. Y. Chen are with the National Taichung University of Science and Technology, Taichung, 40444 Taiwan (e-mail: hschiang@nutc.edu.tw, s4851798@gmail.com, mychen@nutc.edu.tw).

C. C. Chen is with National Chung Hsing University, Taichung, 402 Taiwan (e-mail: emily@nchu.edu.tw).

Neil Y. Yen is with University of Aizu, Aizu-Wakamatsu, Fukushima Pref. 965-8580 Japan (e-mail: neil219@gmail.com).

While flight safety prediction models have existed for some time, most research on such models has focused on the acquisition and selection of data attribute fields, while relatively few studies have addressed the selection and analysis of classification algorithms which are used to classify unknown data for prediction system construction. Classification algorithms are typically trained to identify characteristic patterns using a training dataset, and are then applied to classify testing data. This study uses different classification algorithms for data mining of flight accidents to identify a classification algorithm suitable for the aviation industry. Flight accident data is taken from the Accident/Incident Data System (AIDS) of the United States Federal Aviation Administration (FAA) to identify key factors leading to fatal flight accidents.

The dataset includes many different makes and models of aircraft, which might result in overfitting, thus we restricted our analysis to Boeing aircraft only. The dataset was also imbalanced in terms of a disparity between casualties and non-casualties. To reduce the influence of this data imbalance classification results, the Synthetic Minority Oversampling Technique (SMOTE) [2] method was used to increase the number of samples in categories with small datasets. Feature attribute selection was carried out by means of information gain (IG) [3]. As for classification models, our Associative Petri Net (APN) [4] was compared against other five classification methods: Naive Bayes [5], BayesNet [6], Support Vector Machine (SVM) [7], Decision Tree (C4.5) [8], and Radial Basis Function Network (RBFN) [9]. Finally, this study used the confusion matrix as the main performance indicator.

II. Literature Review

A. Impact Factors and Relevant Research on Flight Safety

The International Civil Aviation Organization (ICAO) defines an airplane accident as an accident occurring in the course of aircraft operation from the time anyone boards an aircraft for the purpose of flight, until all persons have left the aircraft, which directly causes death or injury, including to those not on board the aircraft, or causing substantial damage to the aircraft, such as the loss of a wing or engine failure, or causing the aircraft to go missing. The National Transportation Safety Board (NTSB) classifies causes of flight accidents as human factors, environment-related factors and machinery-related factors. Human factors entail a wide range of uncertainties, and this study focuses on factors that could be quantified and controlled (such as pilot

doi: 10.18178/ijfcc.2017.6.1.482

certification and other related attributes). Such factors are widely used as indicators for flight safety risk control and risk assessment and prediction.

Gürbüz *et al*. [10] proposed a method to reduce feature dimensionality by using a rough set and finding dependencies, with ID3 -based decision trees used for classification. Using data from the AIDS flight accident database, and Polyanalyst and Rosetta data mining tools, this study found that the two methods of attribute selection following classification had approximately equal overall effectiveness. In 2011, Gürbüz *et al*. [11] examined attribute selection and data preprocessing using 15 datasets provided by Turkish Airlines, including 19 different attribute dimensions. Data mining tools used included Polyanalyst, SPSS Clementine, Minitab and Rosetta, and attributes were selected using regression analysis, anomaly detection analysis, dependency identification and rough set. They found that dependency identification and rough sets were the two most effective attribute selection methods.

Shi and Luo [12] applied MATLAB to build flight safety models for China's Wuhan airport. Their proposed method combined the information gain with rough sets to screen and select attribute features, building a prediction model using Back-Propagation Neural Network (BPNN). Tanguy *et al*. [13] used Natural Language Processing (NLP) with n-gram [14] to mine texts taken from airplane accident reports provided by the Aviation Safety Report System (ASRS) and European Coordination Center for Accident and Incident Reporting Systems (ECCAIRS). The "N" is a contiguous sequence of a given number of items (e.g., words); an n-gram of size 1 is referred to as a "unigram"; size 2 is a "bigram"; size 3 is a "trigram", and so on. For example, in the phrase, "Occurrences involving air traffic management or communications, navigation, or surveillance service issues", the trigrams approach sequences the text into the following possible results: "Occurrences involving air", "Involving air traffic", "Air traffic management", and so on. In building the classification model, support vector machine (SVM) was used in the classification algorithm. Finally, Tanguy *et al*. (2015) concluded that text mining methods are more flexible than indicator prediction methods.

Another concern is the effectiveness of safety-related videos presented to airline passengers. Seneviratne and Molesworth [15] showed that using humorous pre-flight safety videos could effectively attract and maintain passenger attention and help them recall critical safety information. In addition, Vine *et al*. [16] simulated an "engine failure on take-off" scenario to test the pilot's response under pressure and the effect on subsequent flight control.

B. Synthetic Minority Oversampling Technique (SMOTE)

The Synthetic Minority Oversampling Technique synthesizes a small amount of data for oversampling [2]. Random oversampling simply increases a small number of samples from certain categories by duplication, which can result in overfitting. SMOTE rearranges and synthesizes features available in all samples of a few categories to obtain new samples which are then added to the dataset. The SMOTE variables (T, N, k) contain a sample T of categories with few samples increase N%. "k" refers to the nearest neighbor samples which need to be referenced when composing the samples.

El-Sayed *et al*. [17] used SMOTE to increase samples for imbalanced autism data and thus improve classification accuracy and reliability. Dong [18] applied SMOTE to improve a high dimension and imbalanced dataset of DNA features to predict prostate cancer. Sarakit *et al*. [19] conducted a textual emotion classification of Youtube user comments, using SMOTE to address dataset imbalance and effectively improve classification results.

III. Associative Petri Net (APN)

The APN model used in this study [4] made inferences based on Apriori approach and degree of truth. If the final node has a high degree of truth, then the degree of relationship establishment was large. APN was derived from Apriori operations, making it a rule-based system that could be opened to final interpretation. APN produced good results for ECG feature detection of cardiac anomalies [20] and E-mail viruses [21], and was used in this study to explore classification of flight accidents.

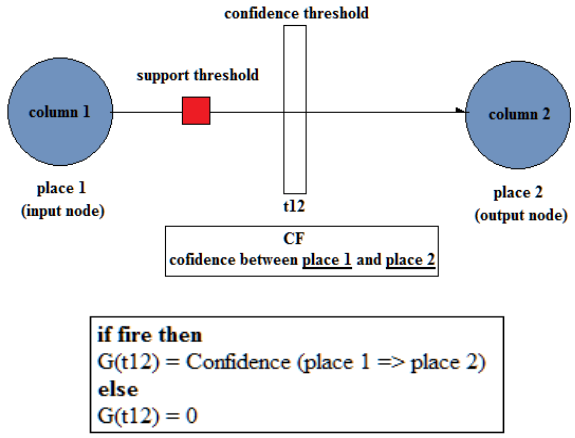

Fig. 1. A generalized APN structure.

The APN model contain three key node types: 1) place, which is circular; 2) support, which is square, representing the correlation threshold between the pre- and post-states; and 3) transition, which is a long strip. The APN model is mainly used to discuss dynamic processes, so the implementation rules focus on enabling and firing the change; that is, to fire enabled transition. During the state transition, inferences are added according to different firing conditions, further exploring changes in the influence and degree of correlation between the pre- and post-states. Assuming that N is an APN, N has three nodes: place, support, and transition. The arcs extend from one place to one transition or from one transition to one place, and are weighted in terms of positive integers. The support is used to evaluate the correlation strength between the pre- and post-states. When the strength of association exceeds the corresponding threshold value, it will fire an enabled transition. The transition contains a relational function G, which converts the various relationships between the pre-and post-states into appropriate representation values. For example, an APN structure is shown in the Fig. 1.

The value in column 1 for the testing data is 11, while the value in column 2 is 21). Before the testing data begins to travel through the APN structure, the support and confidence between the input and output nodes should be calculated and evaluated to determine whether each state has fired and met the threshold value. If the state is "fired" then the value of G(t12) is likely to be between the input and output nodes (G(t12) = CF). If the state is "not fired" then the value of G(t12) would be 0 (G(t12) = 0). After all the parameters in the APN structure are established, the degree of truth comes into play (as shown in Fig. 2). The APN structure has five connecting types, labeled 1 through 5 [4]. This study uses types 2 to select the minimum output value, and type 4 to select the maximum output value.

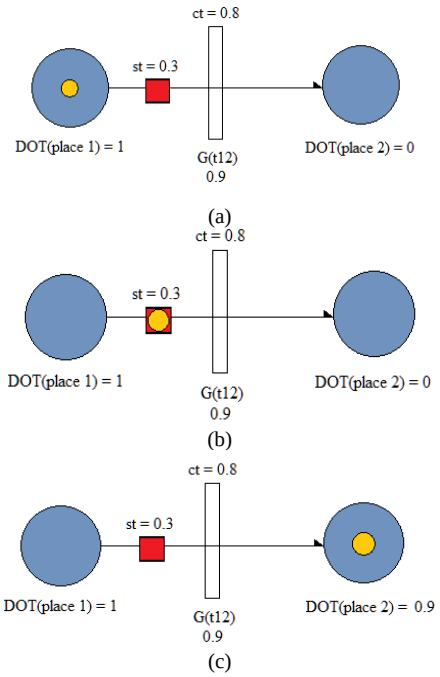

Fig. 2. Firing of a marked APN. (a) Before firing state. (b) Checking state. (c) After firing state.

IV. Experimental Design

A. Experimental Environment

In terms of experimental hardware, the operating system was WIN7 64-bit, 4-core CPU, 8G RAM. The system applied the Weka data mining software to the FAA Aviation Safety Information Analysis and Sharing (ASIAS). AIDS is a database of general civil aviation accidents established in 1978 by the United States Federal Aviation Administration (FAA) [22] based on data from the National Transportation Safety Board (NTSB). AIDS now includes over 100,000 records with 27 original data attributes. Weka was developed at New Zealand's University of Waikato, providing data mining and machine learning software written in Java.

B. Experimental Process

Step 1: Data collection

The data used in this study were obtained from the AIDS database. Data over-fitting could result from the use of datasets with multiple flight models, thus only incidents involving Boeing aircraft from January 2000 to April 2016 were considered for a total of 1209 incidents.

Step 2: Data Pre-Processing

This study used Java to develop the APN and import datasets, while Weka was used to implement and compare the other five algorithms. Symbols (e.g., %, ') that cannot be imported into Weka were deleted, as were records with incomplete data to produce a final dataset of 756 entries. Of these, 54 resulted in casualties (True) while the remaining 702 did not (False). In this study, information gain was used to evaluate the importance of attributes and to reduce the overall dimensions. The original data included 27 attribute types, which was then reduced to 26 by merging the two casualty types. SMOTE was then uesd to increase the original 54 casualty entries to 702, equal to the number of non-casualty entries, thus eliminating data imbalance.

Step 3: Using Classification Algorithms to Build Models

Classification models were established using six algorithms: APN, BayesNet, Naive Bayes, Support Vector Machine (SVM), Decision Tree (C4.5) and Radial Basis Function Network (RBFN). Prediction indicator (Overall classification accuracy, G-mean, F-measure and AUC) were used to test model performance. 10-fold cross validation was used to train and test samples.

Step 4: Result Comparison and Analysis

This study expected the comparison to differentiate the characteristics of each method. In addition, further discussion would be made based on the selected attribute characteristics.

C. Evaluation of Performance Indicators

This study used a confusion matrix (Table I) to calculate F-measure and G-mean. TP and TN were positive (which could be regarded as non-casualty samples) and negative samples (which could be regarded as casualty samples) that were correctly categorized. That is, positive and negative samples matched the original data classification after classification prediction. FP and FN are respectively the positive and negative samples that were wrongly predicted in the classification models.

TABLE I: Confusion Matrix

Prediction / Reality	Positive	Negative
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

The indicator $F-measure = \frac{2 \times \Pr ecision \times \mathrm{Re}\, call}{\Pr ecision + \mathrm{Re}\, call}$ used in this study combined precision=TP/(TP+FP) and recall=TP/(TP+FN). Precision refers to the accuracy of classification prediction after all data samples were predicted by a categorical model. For example, when predicting casualties (assuming a positive category), the number of all positively categorized data after prediction was used as the denominator base, divided by the number of data samples that were correctly predicted to be positive. Recall refers to the sample rate that was correctly predicted to be the same category in the actual data sample. For example, when the

majority of samples (assuming the negative categorization) predicted to be negative in the non-casualty category, it could be called a high recall rate. The F-measure range was 0 to 1, and the higher the score, the better.

The indicator, $G-mean=\sqrt{\Pr ecision\times \operatorname{Re} call}$ used precision and recall as the basis for calculation. The G-mean range was 0 to 1, and the higher the score, the better. AUC was the area under the ROC line; the higher the score, the closer the ROC to the left top which means it would have the better classification performance.

V. Experimental Results

A. Experimental Dataset

TABLE II: Attribute Introduction

No	Attributes	Description
1.	AIDS Report Number	The identification number of each flight accident incident
2.	Local Event Date	Time of Flight Accident
3.	Event City	The city where the accident occurred
4.	Event State	The state where the accident occurred
5.	Event Airport	The airport where the accident occurred
6.	Event Type	Types of flight accidents
7.	Aircraft Damage	Damage of aircraft
8.	Flight Phase	Flight phase
9.	Aircraft Make	Manufacturer of aircraft
10.	Aircraft Model	Aircraft model
11.	Aircraft Series	Aircraft serial number
12.	Operator	Operator of aircraft
13.	Primary Flight Type	Type of flight (such as non-scheduled flights, etc.)
14.	Flight Conduct Code	Flight code (e.g. short commute)
15.	Flight Plan Filed Code	Flight plan filed code (e.g. instrument flight)
16.	Aircraft Registration Nbr	Aircraft registration number
17.	Aircraft Engine Make	Aircraft engine manufacturer
18.	Aircraft Engine Model	Aircraft engine model
19.	Engine Group Code	Aircraft engine group code
20.	Nbr of Engines	Number of aircraft engine
21.	PIC Certificate Type	Certificates held by captain who flew the airplane involved in an accident
22.	PIC Flight Time Total Hrs	Captain's total flight hours
23.	PIC Flight Time Total Make-Model	The total number of flight hours of the captain flying the same aircraft model
24.	PIC Flight Time 90D Total Time	Captain 's flight time within 90 days
25.	PIC Flight Time Total Make-Model	Captain's flight time of the same aircraft model within 90 days
26.	IsFatal	Whether there are casualties of the Boolean (for the target classification attribute)

Table II describes the 26 attribute types after combining the casualty attributes. Blue shaded rows show the 10 attribute fields selected after dimension reduction. Red shaded rows show the target attributes of this study, regardless of whether the incident resulted in casualties or not.

B. Parameter Settings

APN requires a network graph, as shown in Fig. 3. Values for places must be put in the corresponding fields. These nodes were ranked using the information gain feature selection method, and nodes with stronger relationships are closer and closer to the final node (also called the goal place). We used the trial and error method to evaluate the degree of importance among these nodes. In Fig. 3, the main effects of the input nodes for the output are {P00}, {P01} and {P7, P8, P9, P10}, and the node connection used TYPE 2 to obtain higher performance. We set a threshold of 0.01 for support and confidence, because the classification results under the requirement was satisfied. After setting the necessary parameters, APN could obtain information from different nodes. Network access requires membership input, because the data attributes obtained for this study were mostly discrete values such as code and serial number. There was no continuous value to be separated, so the input places were all 1.0. In the goal place, APN would obtain the highest score between them and predict the result. This score means that the combination of such feature attributes and casualties had a high probability of occurring at the same subset concurrently. In addition, some special cases have identical scores, so the concept of the prior probability is used.

An example of traveling APN is provided as follows. First, the degree of truth in input places were all 1.0, {DOT(P1), DOT(P2), DOT(P3), DOT(P4), DOT(P5), DOT(P6)} = {1.0, 1.0, 1.0, 1.0, 1.0, 1.0}. Suppose the parameters in the APN structure were complete (e.g., transition CF) such that the structure could be traveling. The nodes {{P7}, {P8}, {P9}, {P10}} and nodes {{P1}, {P2}, {P3}, {P4}, {P5}, {P6}} were connecting by type 4, therefore, the maximum output is chosen. For example, the value of the transition for node {P7} and nodes {{P1}, {P2}, {P3}, {P4}, {P5}, {P6}} would be {G(t17), G(t27), G(t37), G(t47), G(t57), G(t67)} = {0.206, 0.410, 0.140, 0.132, 0.136, 0.126}, in the "fired" state it obtains DOT(P7) = Max{ DOT(Pn)* G(tn7), n=1,2,…,6.} = 0.410. The nodes {P8}, {P9}, {P10} were calculated by the same method, {DOT(P7), DOT(P8), DOT(P9), DOT(P10) } = {0.410, 0.229, 0.546, 0.605}. Next, nodes {{P00}, {P01}} and nodes {P7, P8, P9, P10} are calculated. Because the nodes are connected by type 2, the minimum output is chosen. Transitions were {G(t00), G(t01)} = {0.024, 0.975}; DOT(P00) = Min{DOT(Pn)* G(tn00), n=7,8,9,10.} = 0.006; DOT(P01) = Min{DOT(Pn)* G(tn01), n=7,8,9,10.} = 0.223. Eventually, G(P01)=0.223 > G(P00)=0.006. The score was the degree of truth, namely it was the relationship between the membership set. Therefore, this reasoning of the testing data was P01 (IsFatal = "true").

The other five classification algorithms set the parameters. In Weka, Naive Bayes had no special parameters. The BayesNet estimator was used to check the correlation of each attribute to create a graphic representation of the correlation network. The support vector machine core used a linear function. The smallest leaf of the decision tree (C4.5) was 5, and the radial basis function network was divided into 4 groups in the hidden layer.

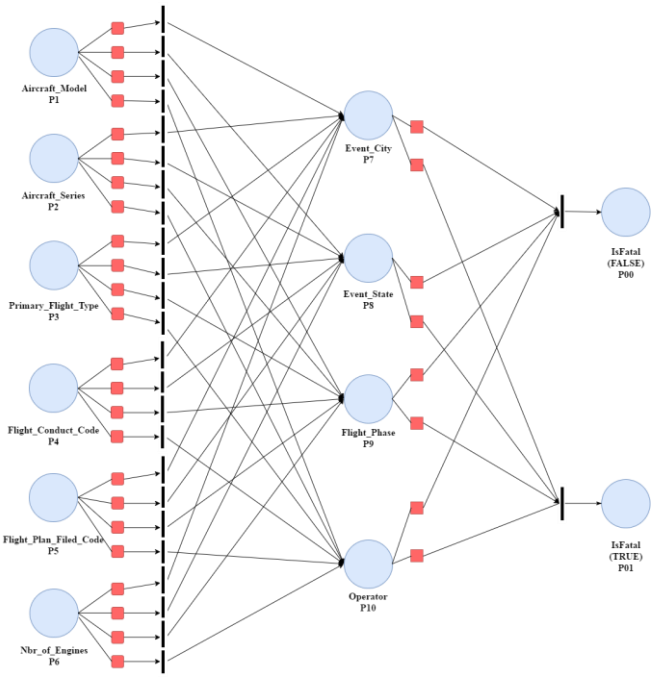

Fig. 3. APN network node setting diagram.

C. Performance Evaluation

This study compared the performance of six different classification algorithms using 11 important feature attributes (including IsFatal). Table III shows the results using 10-fold cross validation without SMOTE classification. Table IV showed the results of using 10-fold cross validation together with SMOTE. The positive items in Tables III and IV represent "IsFatal = "true"", because this study aims to predict aviation incidents based on this imbalanced dataset. Table 3 shows that, without SMOTE, positive categories with few samples provide poor classification results because the number of samples is insufficient to find suitable features for determination. Therefore, the SMOTE sample increase technique was used to inflate the number of positive samples to equal those of other categories.

TABLE III: RESULTS OF 10-FOLD CROSS VALIDATION WITHOUT SMOTE

Methods	TP	TN	FP	FN	Precision	Recall
APN	2	696	52	6	0.037	0.25
BayesNet	12	630	72	42	0.143	0.222
Naive Bayes	9	646	56	45	0.183	0.167
SVM	4	693	9	50	0.308	0.074
C4.5	0	702	0	54	0	0
RBFN	2	675	27	52	0.069	0.037
Methods	**F-Measure**		**G-Mean**		**AUC**	**Accuracy**
APN	0.065		0.096		0.59	0.923
BayesNet	0.174		0.178		0.643	0.849
Naive Bayes	0.151		0.151		0.666	0.866
SVM	0.119		0.151		0.531	0.921
C4.5	0		-		0.477	0.929
RBFN	0.048		0.051		0.676	0.896

Table IV shows that APN, SVM, BayesNet, C4.5 and RBFN achieved overall accuracy rates of 0.9. APN had the highest F-Measure and G-Mean indicators, but the BayesNet had the best AUC indicator. The overall results showed that SMOTE could effectively increase the classification performance of positive categories with few samples (True). These results were affected by the characteristics (positive item) of the training data, and SMOTE pre-processing clearly results in improved performance. SMOTE application can effectively resolve data imbalance, and efficiency can be further improved by applying the Apriori-based APN reasoning to the dataset.

TABLE IV: RESULTS OF 10-FOLD CROSS VALIDATION AND SMOTE

Methods	TP	TN	FP	FN	Precision	Recall
APN	657	687	45	15	0.936	0.978
BayesNet	678	591	111	24	0.859	0.966
Naive Bayes	657	545	154	45	0.81	0.936
SVM	663	661	41	39	0.942	0.944
C4.5	671	631	71	31	0.904	0.956
RBFN	671	623	79	31	0.895	0.956
Methods	**F-Measure**		**G-Mean**		**AUC**	**Accuracy**
APN	0.956		0.957		0.958	0.957
BayesNet	0.909		0.911		0.97	0.904
Naive Bayes	0.868		0.871		0.959	0.858
SVM	0.943		0.943		0.943	0.943
C4.5	0.929		0.930		0.953	0.927
RBFN	0.924		0.925		0.967	0.921

VI. CONCLUSION

A close relationship is found among aircraft type, engine type, flight stage and the degree of damage to the airplane in aviation accidents. Area-related attributes such as Event City and Event State might be related to density of flight routes. Other attributes include indirect human factors such as operator attribute which might be related to the corporation/organiation's operating goals. For Nbr of Engines, when the engine was damaged, if there were other engines to be used, it could reduce flight accidents. In this study, the dataset provided by the AIDS database was used to train and compare APN with the other five classification algorithms.

Finally, it was found that APN of this study had the highest overall accuracy rate 95.7%, F-measure score 95.6%, G-mean score 95.7% among all algorithms. These results indicated that the APN could be efficient due to the plentiful training data within clear and balanced characters. APN used Apriori algorithm as the base, so there was a rule-based concept. In other words, the final output could be interpreted. Therefore, if the APN algorithm is used to build the model, it is possible to develop an expert system for flight accident prediction.

ACKNOWLEDGMENT

The authors wish to thank the Ministry of Science and Technology of the Republic of China for financially supporting this research under Contract Grants No. MOST105-2410-H-025-015-MY2,MOST105-2511-S-005-001-MY3, and MOST105-2410-H-025-013.

REFERENCES

[1] H. J. Shyur. "A quantitative model for aviation safety risk assessment," *Computers & Industrial Engineering*, vol. 54, pp. 34-44, Feb. 2008.

[2] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
[3] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Proc. the Ninth International Workshop on Machine Learning*, pp. 249-256, 1992.
[4] D. H. Shih, H. S. Chiang, and B. Lin, "A generalized associative Petri net for reasoning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, pp. 1241-1251, Aug. 2007.
[5] I. Rish. "An empirical study of the naive Bayes classifier," *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, vol. 3, pp. 41-46, Jan. 2001.
[6] I. Ben-Gal, "Bayesian networks," *Encyclopedia of Statistics in Quality and Reliability*, 2007.
[7] C. C. Chang and C. J. Lin. "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
[8] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, 1993.
[9] D. S. Broomhead and D. Lowe. "Radial basis functions, multi-variable functional interpolation and adaptive networks," *Royal Signals and Radar Establishment Malvern (United Kingdom)*, March 1988.
[10] F. Gürbüz, L. Özbakir, and H. Yapici, "Classification rule discovery for the aviation incidents resulted in fatality," *Knowledge-Based Systems*, vol. 22, pp. 622-632, Dec. 2009.
[11] F. Gürbüz, L. Özbakir, and H. Yapici. "Data mining and preprocessing application on component reports of an airline company in Turkey," *Expert Systems with Applications*, vol. 38, pp. 6618-6626, June 2011.
[12] L. Shi and F. Luo, "Research on risk early-warning model in airport flight area based on information entropy attribute reduction and BP neural network," I*nternational Journal of Security and Its Applications*, vol. 9, pp. 313-322, 2015.
[13] L. Tanguy, N. Tulechki, A. Urieli, E. Hermann, and C. Raynal, "Natural language processing for aviation safety reports: from classification to interactive analysis," *Computers in Industry*, vol. 78, pp. 80-95, 2015.
[14] G. Sidorov, F. Velasquez, E. Stamatatos, A. Gelbukh, and L. Chanona-Hernández, "Syntactic n-grams as machine learning features for natural language processing," *Expert Systems with Applications*, vol. 41, pp. 853-860, Feb. 2014.
[15] D. Seneviratne and B. R. Molesworth, "Employing humour and celebrities to manipulate passengers' attention to pre-flight safety briefing videos in commercial aviation," *Safety Science*, vol. 75, pp. 130-135, June 2015.
[16] S. J. Vine, L. Uiga, A. Lavric, L. J. Moore, K. Tsaneva-Atanasova, and M. R. Wilson, "Individual reactions to stress predict performance during a critical aviation incident," *Anxiety, Stress, & Coping*, vol. 28, pp. 467-477, Dec. 2014.
[17] A. A. El-Sayed, M. A. M. Mahmood, N. A. Meguid, and H. A. Hefny, "Handling autism imbalanced data using synthetic minority over-sampling technique (SMOTE)," in *Proc. 2015 Third World Conference on Complex Systems (WCCS)*, pp. 1-5, Nov. 2015.
[18] S. Dong, "Exploring effects of various data preprocessing methods on the classification of DNA microarray data," in *Proc. the International Conference on Bioinformatics & Computational Biology (BIOCOMP)*, p. 74, 2015.
[19] P. Sarakit, T. Theeramunkong, and C. Haruechaiyasak, "Improving emotion classification in imbalanced YouTube dataset using SMOTE algorithm," *2015 2nd International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, pp. 1-5, 2015.
[20] H. S. Chiang, D. H. Shih, B. Lin, and M. H. Shih, "An APN model for arrhythmic beat classification," *Bioinformatics*, vol. 30, pp. 1739-1746, Feb. 2014.
[21] H. S. Chiang, D. H. Shih, M. H. Shih, and J. M. Chang, "APN model construction for malicious email detection," *Expert Systems with Applications*, vol. 42, pp. 5696-5704, Aug. 2015.
[22] Federal Aviation Administration, Accident and Incident Data System (AIDS). October 2016. [Online]. Available: http://www.asias.faa.gov/pls/apex/f?p=100:12:0::NO:::

Hsiu-Sen Chiang received his Ph.D. degree in Information Management from National Yunlin University of Science and Technology, Taiwan, in 2007. He is an associate professor in Department of Information Management, National Taichung University of Science and Technology, Taichung, Taiwan. His current interests include data mining, Petri net, biomedical science and internet marketing.

Tsung-Che Hsieh is the graduated student in Information Management from National Taichung University of Science and Technology, Taichung, Taiwan. His current interests include data mining, petri net, and sentiment analysis.

Chia-Chen Chen received her Ph.D. degree in information management from National Chiao-Tung University in Taiwan. She is an associate professor of Management Information Systems at National Chung Hsing University, Taiwan. Her current interests include e-learning, u-learning, data mining, and smart living.

Neil Y. Yen received his Ph.D. degree in the Department of Human Informatics and Cognitive Sciences from Waseda University, Japan. He is an associate professor of School of Computer Science and Engineering at University of Aizu, Japan. His current interests include human-centric computing, big data science, and anticipatory computing.

Mu-Yen Chen received his Ph.D. degree in information management from National Chiao-Tung University in Taiwan. He is a professor of information management at National Taichung University of Science and Technology, Taiwan. His current interests include artificial intelligent, data mining, and soft computing.