

Credit Risk Rating Using State Machines and Machine Learning

Behnam Sabeti, Hossein Abedi Firouzjaee, Reza Fahmi, Saeid Safavi, Wenwu Wang, and Mark D. Plumbley

Abstract—Credit risk is the possibility of a loss resulting from a borrower’s failure to repay a loan or meet contractual obligations. With the growing number of customers and expansion of businesses, it’s not possible or at least feasible for banks to assess each customer individually in order to minimize this risk. Machine learning can leverage available user data to model a behavior and automatically estimate a credit score for each customer. In this research, we propose a novel approach based on state machines to model this problem into a classical supervised machine learning task. The proposed state machine is used to convert historical user data to a credit score which generates a data-set for training supervised models. We have explored several classification models in our experiments and illustrated the effectiveness of our modeling approach.

Index Terms—State machine, machine learning, classification, credit risk, financial regulation.

I. INTRODUCTION

Credit risk refers to the possibility of loss due to a borrower’s failure to make payments on any type of debt. The goal of credit risk management is to maximize a bank’s risk-adjusted rate of return by maintaining credit risk exposure within acceptable parameters. More specifically, by measuring customers’ credit scores, banks and financial institutes monitor the expected rate of return for any debt and manage their portfolio by rejecting or adjusting high risk applications.

There are several configurations for a credit risk management system based on its features and expected results. These features include:

- **Type:** Credit scoring refers to a situation where the Credit Risk Management (CRM) system produces a score for each customer whereas in Credit rating a category of credit,

e.g. slightly risky, is produced for each customer.

- **Decision Rules:** Specifies whether it’s possible to investigate the reasons behind a produced result or not.
- **Mode:** Specifies whether the system procedure is updated online or not.
- **Loan amount:** This feature states whether the loan amount is considered in producing the credit score or not.
- **History:** This feature specifies if the CRM system uses customer’s history to generate a credit score.
- **Fraud:** This feature indicates whether the CRM system is paired with a Fraud Detection module. If so, the output of fraud detectors coupled with customer information can enhance the performance of CRM systems.

Configuration features for a credit risk management system are summarized in Fig. 1.

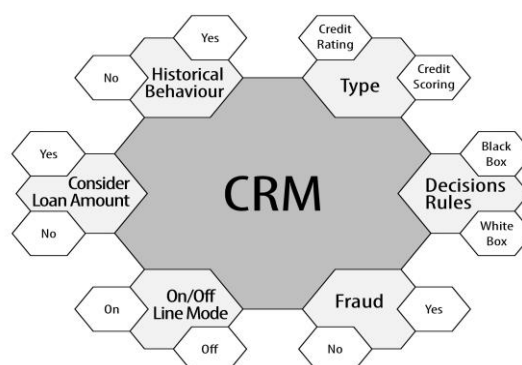


Fig. 1. Credit risk management types.

While credit risk management is a critical process for financial institutes, the growing number of customers and expansion of businesses have made it impossible or at least not feasible for bank to maintain their risk within acceptable parameters with traditional regulations. The increase in loan applications and also huge information available for customers, are indicators that this process is a good candidate for automation. There are also privacy matters, where banks need to make sure that their customers' data is not accessed by any unauthorized entity. This issue can also be addressed by an automation process where customer data is not exposed at any step and is always processed by computers.

Machine learning techniques have been employed for credit risk management in many settings since user information, transactions and historical data are gathered and stored by banks [1], [2]. The information available for users can serve as features for *Regression* and *Classification* models. Several models including *Support Vector Machine (SVM)*, *Decision Tree* and *Neural Network* have been utilized for this task. These previous approaches are briefly described

Manuscript received February 15, 2020; revised September 21, 2020. This work was partially supported by the EPSRC Projects entitled “Making Sense of Sounds” with grant reference number EP/N014111/1 and “Audio-Visual Media Research Platform” with grant reference number EP/P022529/1. We would also like to appreciate everyone at Miras Technologies International for their technical and moral support in this project.

Behnam Sabeti, Hossein Abedi Firouzjaee, Reza Fahmi are with Miras Technologies International, Number 92, Movahed Danesh St., Tehran, Iran (email: behnam@miras-tech.com, hossein@miras-tech.com, reza@miras-tech.com).

Saeid Safavi, Wenwu Wang, and Mark D. Plumbley are with CVSSP, University of Surrey, Guildford, Surrey GU2 7XH, UK (email: s.safavi@surrey.ac.uk, w.wang@surrey.ac.uk, m.plumbley@surrey.ac.uk).

in Section II.

In this paper we propose a classification model for credit risk rating, where we use available customer information, transactions and historical data to estimate new customers' credit risk category. The proposed approach uses a state machine to generate credit scores for previous customers who had been granted a loan. These scores together with customer information generates a data-set for training classification models. *Support Vector Machine (SVM)* and *Decision Tree* have been employed for classification where they showed promising results. The main contribution of this paper is the proposed state machine that can be leveraged to convert previous behavior in making payments to credit scores. In this way, a unified and consistent credit score can be computed across all previous customers, making it easy to employ machine learning models for credit risk management.

II. RELATED WORK

In order to approach credit risk management, more accurate and robust systems have been employed to drive expert decisions in recent years, exploring new techniques especially from the field of machine and deep learning. Recently, several approaches have been developed to address the problem of modelling the credit quality of a company or a customer, using both quantitative and qualitative information. In this section we briefly describe these methods.

Several studies employed *Support Vector Machines (SVM)* for credit risk scoring where rather than estimating a credit score, a category of credit is considered as the representative of one's credibility [3], [4]. SVM classifiers have been specifically found useful for feature selection and optimization [3]. To overcome the generalization issue of this approach, *Yu* proposed to integrate the concepts of fuzzy set, while adopting least square method to reduce the computational complexity of the model [5]. In order to further improve the computational complexity of SVM especially on large data-sets, *Harris* introduced the use of *Clustered Support Vector Machine* [6].

Tree based methods have also been employed for credit risk management [2], [7]. *Addo et al.* focused on credit risk scoring where they examined the impact of the choice of different machine learning and deep learning models [8]. They observed that the tree-based models are more stable than neural network models. Dimensionality reduction has also been leveraged in combination with *Decision Trees* to boost their accuracy [9]. *Random Forest* is another tree based model that improves the performance of other methods for credit risk management with a probabilistic approach [1].

Neural network is another approach to perform discriminant analysis in business research [10]. Using bank's default data, *Tam et al.* compared the neural network approach with linear classifiers, where empirical results show that neural model is a promising method of evaluating customer conditions [11]. Ensemble learning can leverage multiple Weak learners to boost their performance. This approach has been employed to further improve the performance of neural networks for credit risk assessment [12].

III. CREDIT RISK MANAGEMENT

A. Problem Definition

As mentioned in Section I, *Credit Risk Management* aims at minimizing the risk of customer failure to meet contractual obligations. Consider a set of customers for which we have personal, account and historical information. Using this information, we want to build a machine learning model that can automatically infer credit scores for new customers. To this end, we need to find a set of features for each customer and define a credit score associated with it to build a training data-set.

We have gained access to a bank database through our partner¹, which we are going to use to build our model. After investigating the data, we extracted potentially useful information which are described in Table I. In addition to his information, the customer behavior for repaying a loan is also available. We can use customer information to build a feature vector for each customer and leverage behavior data to generate an associated credit score. In the following section, the procedure for converting user repaying behavior to a credit score is described.

B. Credit Score Modeling

According to the bank's data, a customer is assigned to a state based on his adherence to contractual obligations:

- Normal: The customer has repaid all of the payments up until current date, based on the contract.
- Usance: The customer has failed to repay the last payment.
- Deferred: The customer has failed to repay the last k payments.
- Suspicious: the customer has failed to repay the last n payments ($n > k$).

TABLE I: AVAILABLE CUSTOMER INFORMATION

Type	Features
Personal	Age, Gender, Education
Account	number of active, inactive and closed accounts
Balance	min, max, average and variance of account balance
Transaction	min, max, average and variance of deposits and withdrawals
Previous Loans	min, max, average and variance of previous credit scores

These state are automatically extracted by the bank and are available in the database for each (*customer, loan*) pair (note that multiple loans might be granted to a single customer). In this step we need to define a strategy for converting a state sequence to a credit score. After consulting with domain experts, we defined a state machine which describes the customer transition between defined states. The proposed state machine is illustrated in Fig. 2. Note that the state machine is not complete-edged since, for instance, it's not possible by definition for a customer to go straight from the *Normal state* to the *Deferred state*.

¹ We cannot disclose the bank information due to an Non-Disclosure Agreement.

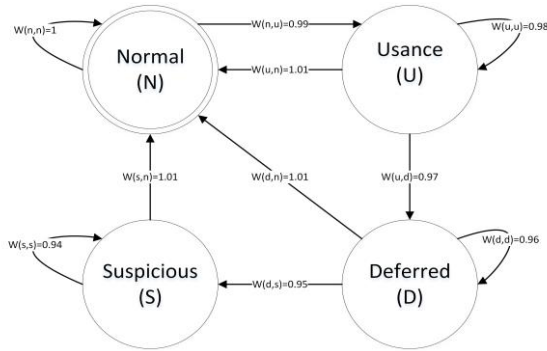


Fig. 2. The state machine for credit score generation.

Each edge in the state machine is associated with a weight, which specifies the penalty/reward for making the transition. At the beginning of repaying period, each customer has a credit score of one and is in the Normal state. Each transition in the state machine alters the credit score based on the associated weight. For instance, a transition from *Normal* to *Usance* state reduces the credibility of the customer due to the failure of last payment reimbursement. Thus the initial credit score is multiplied by the transition weights which are specified by domain experts. The final result is considered as the credit score for the given (*customer, loan*) pair. Algorithm 1 specifies the procedure for credit score computation.

Algorithm 1: The procedure for converting a sequence of states to a credit score.

Input: *S*: State sequence, *W*: Transition weights
Output: Credit score
 Score=1
i = 0
while *i* < *sequence_length* **do**
 Score = Score * *W*[*S*_{*i*}, *S*_{*i*+1}]
end

Using the proposed state machine we can extract a credit score for each (*customer, loan*) pair. Also, user information discussed in Section III. A can serve as feature vectors for a pair of customer and loan. This modeling scheme converts the initial problem to a supervised machine learning task, more specifically a Regression task.

Fig. 3 illustrates the histogram of computed credit scores for all users. A considerable proportion of users have been assigned with a credit score of one which is feasible since most users obey the obligations enforced by the contracts. This shows the asymmetric nature of the problem which should be considered in the classification phase.

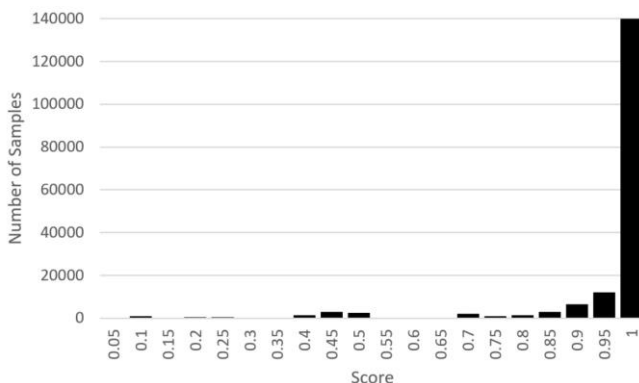


Fig. 3. Histogram of the computed scores across all data.

IV. MODEL SPECIFICATIONS

So far we have introduced the problem modeling approach. In this section, some extra details are presented to complete the modeling scheme and propose the overall architecture of the proposed framework.

A. Quantization

Instead of producing a credit score for each customer, we decided to categorize the credit scores and estimate the credit class. This approach turns our problem into a straightforward classification task. The quantization required for this step has been done by domain experts. We considered two types of quantization, namely 2-class and 5-class quantization, which are described in Table II.

TABLE II: QUANTIZATION POLICY FOR COMPUTED CREDIT SCORES

	Class Name	Criteria
2-class	No Risk	Score = 1
	Risky	Score ≠ 1
5-class	No Risk	0.99 ≤ Score
	With default history	0.95 ≤ Score < 0.99
	Slightly risky	0.8 ≤ Score < 0.95
	Middle risky	0.5 ≤ Score < 0.8
	Highly risky	Score < 0.5

B. Parameters

The parameters of the proposed state machine are critical to the performance of the overall modeling scheme, since they specify the credit score for previous customers and construct the training data. We have asked our partner’s domain experts to set these parameters in a way that the produced scores would be interpretable for the bank. Table III shows the specified parameters.

TABLE III: STATE MACHINE PARAMETERS ASSIGNED BY DOMAIN EXPERTS

<i>W</i>	Value	<i>W</i>	Value
<i>W</i> (n, n)	1	<i>W</i> (d, d)	0.96
<i>W</i> (n, u)	0.99	<i>W</i> (d, n)	1.01
<i>W</i> (u, u)	0.98	<i>W</i> (d, s)	0.95
<i>W</i> (u, n)	1.01	<i>W</i> (s, s)	0.94
<i>W</i> (u, d)	0.97	<i>W</i> (s, n)	1.01

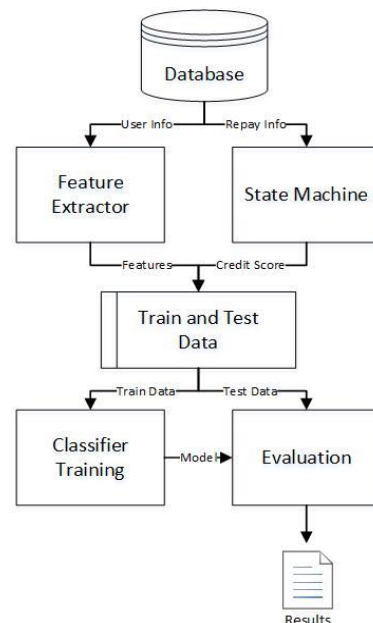


Fig. 4. The overall process for data-set generation, modeling.

C. Classifiers

We have used *Support Vector Machine (SVM)* [13] and *Decision Tree* as baselines for classification and evaluating the proposed problem modeling. In order to train the Decision Tree we leveraged *Gradient Boosting* [14] which will enhance the performance of classification through an ensemble of models. More specifically, we used *CatBoost* which is a library for gradient boosting on decision trees [15]. Fig. 4 illustrates the overall process for *Credit Risk Rating*.

V. EXPERIMENTS AND RESULTS

In this section, the proposed approach for credit risk score computation is employed to construct the training data-set. After presenting some insights about the generated data-set, SVM and Decision Tree classifiers are trained and evaluated.

The available information is not consistent for all customers. For instance, some customers are new to the bank and do not have any transaction data, while others might not have been granted a loan before. Thus it's not possible to extract all features for all customers. In order to extract the features in a coherent way, we have split the users:

- Group I: Customers who had never been granted a loan before and had less than two months of transaction history in time of their application.
- Group II: Customers who had never been granted a loan before and had more than two months of transaction history in time of their application.
- Group III: Customers who have been granted a loan before and had more than two months of transaction history in time of their current application.

Based on these categories, we can extract certain features for each group. For instance, users in *group I* do not have enough transaction history to extract related features, thus we ignore transactions for this group entirely. Table IV specifies the available features for each group of users.

TABLE IV: AVAILABLE FEATURES FOR EACH CUSTOMER GROUP

Feature Type	Group I	Group II	Group III
Personal	■	■	■
Account	■	■	■
Balance		■	■
Transactions		■	■
Previous Loans			■

So far we have introduced the procedure for credit score computation which generates the targets. Furthermore, User Splitting produces a consistent approach for feature extraction which combined with extracted targets generates the required data-set for training classification models. The obtained data-sets are described in Table V.

TABLE V: DATA-SET STATISTICS

		Group I	Group II	Group III
2-Class	No Risk	33580	99109	6917
	Risky	9844	23952	776
5-Class	No Risk	28440	83241	6125
	Default History	8744	25220	1168
	Slightly Risky	2412	6148	208
	Middle Risky	2348	4696	88
	Highly Risky	1480	3756	104
Total		43424	123061	7693

In order to evaluate the classifiers, we split 20% of the data as test set. *Accuracy (Acc)*, *Precision (P)*, *Recall (R)* and *F-Score (F1)* are the metrics used for evaluation. Table VI specifies these evaluation metrics for SVM and Decision Tree classifiers trained in two configurations, namely 2-class and 5-class classification.

The results described in Table VI show the effect of using more features on the classification performance. As described in Table IV *Group III* has all the features, while *Group II* lacks previous loans features. The same condition holds for *Group I* where no balance, transaction and previous loans data is available. Both classifiers reach their maximum performance on the third data group that show the effectiveness of previous loans information. The same holds for the second data group compared to the first data group, where adding transaction data improves the performance of both classifiers.

Comparing the 2-class and 5-class configurations, we can see that in the fine-grained case the classifiers are producing smaller *F-scores*, which makes sense since increasing the number of classes makes it more complicated for the classifier to model the data. On another note, *SVM* classifier outperforms *CatBoost* in *Group I* and *Group II*, while *CatBoost* shows better performance than *SVM* on the third data group. Overall the best performance belongs to the *CatBoost* classifier in the 2-class configuration and trained and evaluated on the third data group.

TABLE VI: SVM AND CATBOOST PERFORMANCE ON TEST SET

	Group	2-Class			5-Class		
		I	II	III	I	II	III
SVM	Acc	0.53	0.79	0.88	0.26	0.44	0.79
	P	0.68	0.71	0.79	0.38	0.55	0.61
	R	0.54	0.79	0.89	0.27	0.45	0.79
	F1	0.53	0.73	0.84	0.32	0.5	0.68
CatBoost	Acc	0.41	0.59	0.83	0.21	0.37	0.68
	P	0.7	0.76	0.9	0.39	0.57	0.72
	R	0.41	0.59	0.83	0.21	0.37	0.68
	F1	0.42	0.63	0.85	0.26	0.44	0.69

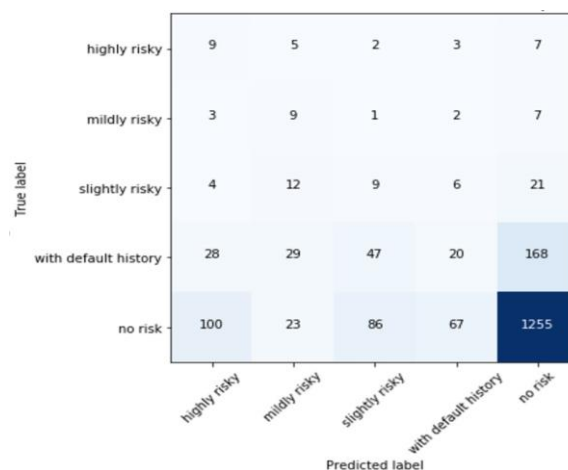


Fig. 5. The confusion matrix for *CatBoost* model trained and evaluated on *Group III* data.

In order to further investigate the performance of classifiers in the fine-grained situation, the confusion matrix for *CatBoost* classifier trained and evaluated on the third data group, a setup which reaches the best performance in 5-class configuration, is illustrated in Fig. 5. The confusion matrix shows the inability of the model to differentiate between *Risky* classes.

VI. CONCLUSION AND FUTURE WORK

Credit risk refers to the possibility of loss due to a borrower's failure to make payments on any type of debt. The goal of credit risk management is to maximize a bank's risk-adjusted rate of return by maintaining credit risk exposure within acceptable parameters. The growing number of customers and increasing amount of stored information for each customer makes it impossible to assess loan applications using traditional methods, while making the problem specifically fit for machine learning methods. In this paper, we proposed a state machine for converting previous repayments of each user to a credit score. The generated credit scores alongside some engineered features construct a supervised classification task. We then employed *SVM* and *Decision Tree* models for classification and evaluated the models with *F-Score*. Our experiments showed that *SVM* outperforms *Decision Tree* in some cases, while *Decision Tree* classifier reaches the best performance overall. We plan to further improve the proposed approach using feature selection and domain knowledge for feature engineering. Also after gathering more data samples, Deep Learning models will be evaluated for the given task.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

B. Sabeti and R. Fahmi conducted the research and modeled the problem. B. Sabeti and H. Abedi then analyzed the data and implemented the models. B. Sabeti wrote the paper. The research was conducted under the supervision of S. Safavi, W. Wang and M. Plumbley. All authors reviewed and confirmed the final version of the paper.

REFERENCES

- [1] J. Kruppa, A. Schwarz, G. Armingier, and A. Ziegler, "Consumer credit risk: Individual probability estimates using machine learning," *Expert Systems with Applications*, vol. 40, no. 13, pp. 5125–5131, 2013.
- [1] A. E. Khandani, A. J. Kim, and A. W. Lo, "Consumer credit-risk models via machine-learning algorithms," *Journal of Banking & Finance*, vol. 34, no. 11, pp. 2767–2787, 2010.
- [2] T. Bellotti and J. Crook, "Support vector machines for credit scoring and discovery of significant features," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3302–3308, 2009.
- [3] P. Danenas, G. Garsva, and S. Gudas, "Credit risk evaluation model development using support vector based classifiers," *Procedia Computer Science*, vol. 4, pp. 1699–1707, 2011.
- [4] L. Yu, "Credit risk evaluation with a least squares fuzzy support vector machines classifier," *Discrete Dynamics in Nature and Society*, vol. 2014, 2014.
- [5] Terry Harris, "Credit scoring using the clustered support vector machine," *Expert Systems with Applications*, vol. 42, no. 2, pp. 741–750, 2015.

- [6] J. Galindo and P. Tamayo, "Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications," *Computational Economics*, vol. 15, no. 1-2, pp. 107–143, 2000.
- [7] P. Addo, D. Guegan, and B. Hassani, "Credit risk analysis using machine and deep learning models," *Risks*, vol. 6, no. 2, pp. 38, 2018.
- [8] A. Petropoulos, V. Siakoulis, E. Stavroulakis, A. Klamargias et al., "A robust machine learning approach for credit risk analysis of large loan level datasets using deep learning and extreme gradient boosting," *The Use of Big Data Analytics and Artificial Intelligence in Central Banking*.
- [9] V. Pacelli and M. Azzollini, "An artificial neural network approach for credit risk management," *Journal of Intelligent Learning Systems and Applications*, vol. 3, no. 2, p. 103, 2011.
- [10] K. Y. Tam and M. Y. Kiang, "Managerial applications of neural networks: the case of bank failure predictions," *Management Science*, vol. 38, no. 7, pp. 926–947, 1992.
- [11] L. Yu, S. Y. Wang, and K. K. Lai, "Credit risk assessment with a multistage neural network ensemble learning approach," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1434–1444, 2008.
- [12] C. Saunders, M. O. Stitson, J. Weston, L. Bottou, A. Smola et al., *Support Vector Machine Reference Manual*, 1998.
- [13] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, pp. 1189–1232, 2001.
- [14] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: Unbiased boosting with categorical features," *Advances in Neural Information Processing Systems*, pp. 6638–6648, 2018.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Behnam Sabeti is a Ph.D candidate at Sharif University of Technology where his research mainly is focused on representation learning for Natural Language Processing. He received his master degree in artificial intelligence from SUT in 2015. He is also a project manager at Miras Technologies International, where he is in charge of managing and developing machine learning based products.



Hossein Abedi Firouzjaee received his bachelor's degree in mining engineering (mineral extraction) from University of Tehran and his master's degree in artificial intelligence from Amirkabir University of Technology in 2013 and 2016, respectively. His research interests include evolutionary computation, nature inspired algorithms and machine learning applications in real world problems.



Reza Fahmi is the director of artificial intelligence at Miras Technologies International. His main role is to manage and design software solutions based on AI, and also to lead Miras data science team. He has been working at Miras for 4 years, starting as a data scientist, CTO, and currently the director of AI. He graduated with a master's degree in artificial intelligence from Sharif University of Technology in 2016, where his research field mainly focused on 3D object tracking in computer vision.



Saeid Safavi, is a research fellow within CVSSP, where he joined in 2018. He received an MEng in 2009 and a PhD in 2014, all in electrical engineering, from University of Birmingham, UK, where he also worked as a post-doctorate researcher for a year in 2015. He then joined a H2020 project (OCTAVE) as a research fellow for two years at the University of Hertfordshire. He has served as a secretary of ISCA for the special interest on children computer interactions and co-organized four international workshops. He is also a reviewer for many internationally

recognised journals and conferences, e.g. CSL journal, IEEE transactions, Interspeech, and Speccom.



Wenwu Wang was born in Anhui, China. He received the B.Sc. degree in 1997, the M.E. degree in 2000, and the Ph.D. degree in 2002 from Harbin Engineering University, Harbin, China. He then worked with King's College London, Cardiff University, Tao Group Ltd. (now Antix Labs Ltd.), and Creative Labs, before joining University of Surrey, U.K, in May 2007, where he is currently a reader in signal processing, and a co-director of the

Machine Audition Lab within the Centre for Vision Speech and Signal Processing. Since 2018, he has been a guest professor with Qingdao University of Science and Technology, Qingdao, China. His current research interests include blind signal processing, sparse signal processing, audio-visual signal processing, machine learning and perception, machine audition (listening), and statistical anomaly detection. He has co-authored more than 200 publications in these areas. He served as an associate editor for the IEEE Transactions on Signal Processing from 2014 to 2018. He is also publication co-chair for ICASSP 2019, Brighton, U.K.



Mark D. Plumbley received the B.A. (Hons.) degree in electrical sciences and the Ph.D. degree in neural networks from the University of Cambridge, Cambridge, U.K., in 1984 and 1991, respectively. Following his Ph.D., he became a lecturer with King's College London, before moving to Queen Mary University of London in 2002. He subsequently became a professor and the director of the Centre for Digital Music, before joining the University of Surrey

in 2015 as a professor of signal processing. He is known for his work on analysis and processing of audio and music, using a wide range of signal processing techniques, including matrix factorization, sparse representations, and deep learning. He is a co-editor of the recent book on Computational Analysis of Sound Scenes and Events, and co-chair of the recent DCASE 2018 Workshop on Detection and Classifications of Acoustic Scenes and Events. He is a member of the IEEE Signal Processing Society Technical Committee on Signal Processing Theory and Methods, and a fellow of the IET.