Proceedings of the 4th International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/2026.33103

Research on Excitation Waveform Classification and Loss Characteristics of Magnetic Components Integrating Time-Delay System Theory and Data-Driven Method

Dandan Wang

University of Shanghai for Science and Technology, Shanghai, China
18821206786@163.com

Abstract. With the extensive application of power electronics technology in the fields of new energy and information communication, the existing core material loss models can hardly meet the requirements of practical applications. This paper deeply analyzes the loss of magnetic components through mathematical modeling, providing theoretical support for the design of high-performance magnetic components. The study finds that the change of magnetic flux density in the excitation waveform can be characterized by only a small number of time series points. Combined with the idea of Fast Fourier Transform, a few key features can be identified for waveform classification. The morphological features of the magnetic flux density time series are extracted by principal component analysis, and then the random forest model is used to realize the effective identification of waveform categories. The validity of the model is verified on the given test set, which well reflects the rationality and effectiveness of the model.

Keywords: Waveform classification, Feature extraction, Principal component analysis, Random forest

1. Introduction

With the development of the third-generation power semiconductor technology, high frequency, high power density and high reliability have become the development directions of power converters. As a key device, the volume, weight, loss and cost of magnetic components directly affect the overall performance of the system. To achieve high-efficiency and high-power density design, in addition to meeting electrical parameters, the loss of magnetic components must be reduced.

A large number of studies have been carried out by scholars at home and abroad on the modeling of core loss. Bertotti [1] proposed the loss separation theory, which decomposes the total core loss into three parts: hysteresis loss, classical eddy current loss and anomalous loss. The three parts are proportional to the 1st power, 2nd power and 1.5th power of frequency respectively and can be linearly superimposed. Steinmetz [2] proposed the Steinmetz empirical formula through experiments, pointing out that the hysteresis loss is proportional to the 1.6th power of the amplitude of magnetic flux density, and the microstructure of the material is a key factor affecting the loss. Roshen [3] proposed a core loss calculation model suitable for non-sinusoidal waveforms,

DOI: 10.54254/2753-8818/2026.33103

overcoming the problem of limited application scope of the traditional Steinmetz formula. Ye Jianying et al. [4] constructed a loss model through dynamic hysteresis loop equivalence and differential loss measurement. Experimental verification shows that the model can effectively evaluate and predict the core loss of magnetic components in power converters.

Winding loss mainly originates from the resistance of copper conductors and can be accurately obtained by finite element simulation of electromagnetic fields. Core loss is the power loss generated by magnetic materials under high-frequency alternating magnetic flux [5]. Due to the complex microstructure of high-frequency magnetic materials themselves, and their loss is related to many factors such as operating frequency, magnetic flux density, excitation waveform, operating temperature and core material, showing complex nonlinearity and intercorrelation, there is a big difference between the existing core material loss models and the needs of practical applications. Therefore, based on waveform feature analysis, this paper constructs a classification and recognition method for magnetic flux density waveforms to realize the effective characterization of excitation waveforms.

2. Problem analysis and scheme

Excitation waveform is a key factor affecting core loss. Different waveforms lead to different rates of change of magnetic flux density, resulting in significant differences in loss characteristics. The existing models have deficiencies in accuracy and applicability, and it is difficult to accurately characterize the dynamic loss behavior. The main difficulties are as follows: various waveform shapes and complex change laws; loss is affected by the coupling of multiple factors; most existing models are based on sinusoidal excitation and have poor adaptability to non-sinusoidal scenarios.

To solve the above difficulties, this paper proposes the following solutions: firstly, visualize and analyze the waveforms according to material types to preliminarily reveal the shape differences of different categories of waveforms. Secondly, use principal component analysis to reduce the dimension of the sampled values of magnetic flux density to solve the computational burden and overfitting problems caused by high-dimensional data. Finally, construct a random forest classification model based on the reduced-dimensional feature variables and verify the classification effect of the model.

3. Model establishment

Facing the multi-classification problem of excitation waveforms, this paper uses the feature variables extracted above and selects two ensemble learning models, random forest and XGBoost, for comparative analysis to deal with the problems of feature overlap and large data volume.

3.1. Modeling idea

Firstly, data preprocessing is carried out, including normalization, missing value and outlier processing. Secondly, visual analysis is carried out to extract the distribution and shape features of magnetic flux density. Then, the applicability of the data is tested through factor analysis, and principal component analysis is used for dimensionality reduction to extract feature variables that can characterize the waveform shape. Finally, a random forest classification model is constructed based on the extracted feature variables, and the XGBoost model is introduced for comparative analysis to test the classification effect.

DOI: 10.54254/2753-8818/2026.33103

3.2. Random forest model

Random forest is a non-parametric statistical method based on ensemble learning [6]. By integrating multiple decision trees and introducing randomness, it has good generalization ability and adaptability to high-dimensional data. Its core idea is to obtain stable and high-precision prediction results through "multi-model voting/averaging" by constructing and integrating multiple decision trees. Its expression is as follows:

$$\hat{y} = \frac{1}{B} \sum_{b=1}^{B} h_b(x) \tag{1}$$

where $h_b(x)$ is the prediction of the b-th tree.

3.3. XGBoost model

XGBoost [7] adopts the gradient boosting framework. By iteratively optimizing the loss function and introducing regularization terms, it performs excellently in both classification accuracy and computational efficiency. In the XGBoost algorithm, first input the training set samples $I = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$, the maximum number of iterations T, the loss function L, and the regularization coefficient, and finally output the strong learner f(x). The loss function expression of the XGBoost algorithm in this paper is:

$$L_t = \sum_{i=1}^{m} L\left(y_i, f_{t-1}(x_i) + h_t(x_i)\right) + \gamma J + \frac{\lambda}{2} \sum_{j=1}^{J} \omega_{tj}^2 \tag{2}$$

where J is the number of leaf nodes, and ω_{tj} is the optimal value of the J -th leaf node.

4. Experiments and results

4.1. Visual analysis of the distribution and shape characteristics of magnetic flux density

The four divided materials are recorded as Material 1, Material 2, Material 3 and Material 4, and data visualization is carried out. The magnetic flux density of different waveforms is shown in the following figures:

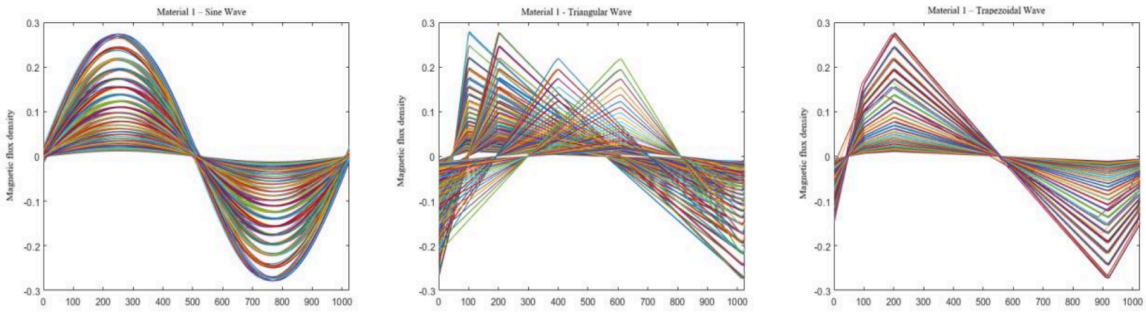

Figure 1. Excitation waveform of Material 1

DOI: 10.54254/2753-8818/2026.33103

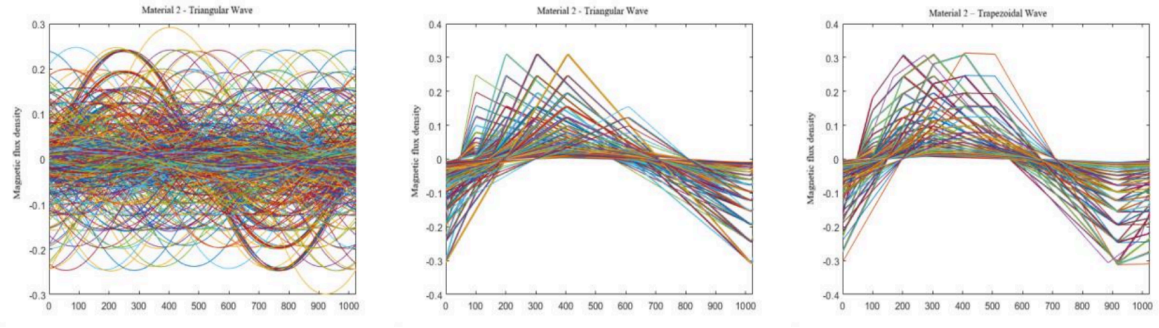

Figure 2. Excitation waveform of Material 2

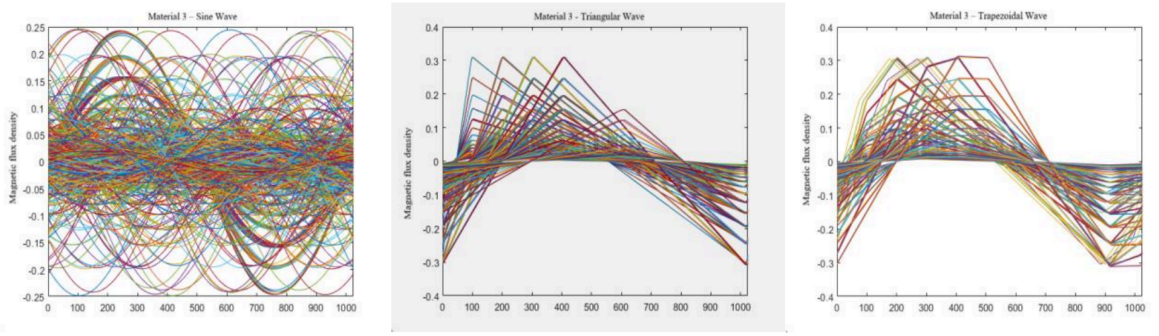

Figure 3. Excitation waveform of Material 3

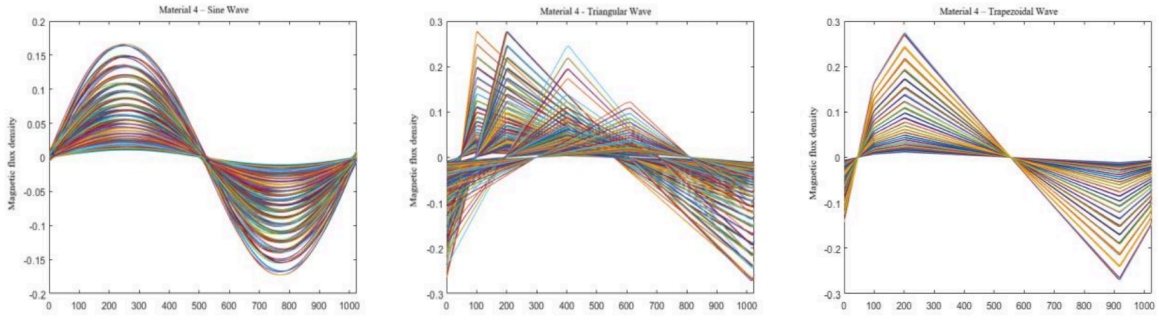

Figure 4. Excitation waveform of Material 4

Through the observation of the excitation waveform images from Figure 1 to Figure 4, it is found that the amplitudes of magnetic flux density under the three excitation waveforms have little difference.

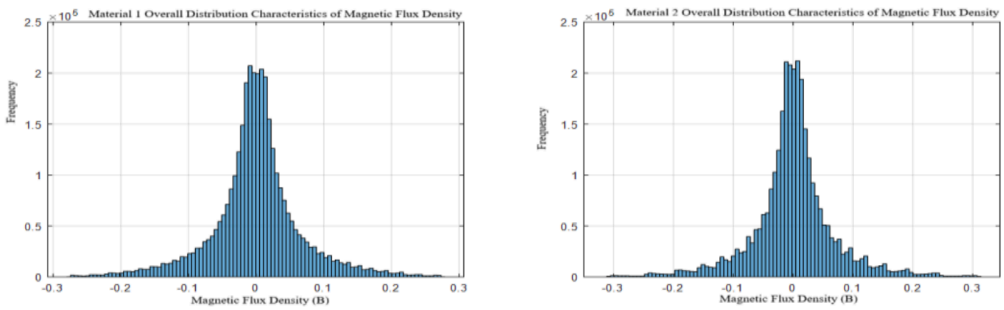

Figure 5. Overall distribution characteristics of magnetic flux density of Material 1 and Material 2

DOI: 10.54254/2753-8818/2026.33103

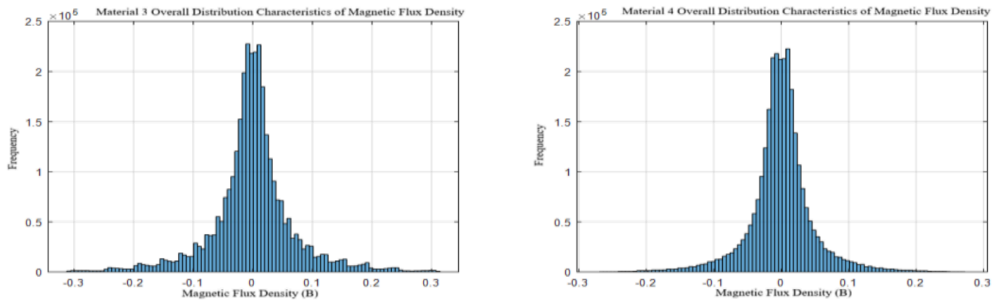

Figure 6. Overall distribution characteristics of magnetic flux density of Material 3 and Material 4

Through the observation of the magnetic flux density distribution images of Material 1, Material 2, Material 3 and Material 4 in Figure 5 and Figure 6, a series of significant features are obtained: 1. It presents the classic normal distribution characteristics; 2. The left and right sides of the image are perfectly mirror-symmetric, reflecting the balance of data distribution; 3. The data set is concentrated at the center point of the curve, that is, where the mean value is located, indicating that most observations are clustered around this central trend, in line with the 3σ principle; 4. The tail of the curve extends infinitely to both ends, gradually approaching but never touching the horizontal axis. It can be concluded that the magnetic flux density of these materials generally follows the characteristics of normal distribution.

4.2. Extraction of feature variables

To ensure the numerical stability and efficiency of model training, the magnetic flux density is first normalized to scale the features to a unified range, so as to speed up the learning speed, avoid weight deviation caused by feature scale differences, ensure that the model learns each feature fairly, and improve the convergence speed and final accuracy of the optimization algorithm. Part of the normalization results are shown in the following table:

Table 1. Normalization results (partial)

Serial number	0 (Magnetic flux density)	1	2	3	4
1	0.503614706	0.506678334	0.50973931	0.51280247	0.51586076
2	0.491299098	0.494369517	0.497435194	0.500508573	0.503574377
3	0.446646753	0.449683267	0.452707	0.455729072	0.458759801
4	0.493677553	0.496729379	0.499779855	0.502834867	0.505887006
5	0.505398131	0.508446966	0.511488966	0.514533685	0.517580023
...					

To determine the number of key features in the time series, factor analysis is used to identify the potential structure in the data, and the correlation between multiple observed variables is explained through unobservable factors, so as to realize dimensionality reduction and feature extraction. The factor analysis results are shown in the following table:

DOI: 10.54254/2753-8818/2026.33103

Table 2. Factor analysis results

Component	Total	Initial eigenvalue Variance percentage	Cumulative %	Total	Extraction sums of squared loadings Variance percentage	Cumulative %	Total	Rotation sums of squared loadings Variance percentage	Cumulative %
1	679.385	66.411	66.411	679.385	66.411	66.411	548.323	53.600	53.600
2	266.289	26.030	92.441	266.289	26.030	92.441	270.387	26.431	80.030
3	57.222	5.594	98.035	57.222	5.594	98.035	181.322	17.725	97.755
4	12.858	1.257	99.292	12.858	1.257	99.292	11.948	1.168	98.923
5	1.001	0.603	99.895	6.170	0.603	99.895	9.317	991	99.834
6	1.001	0.098	99.993	1.001	0.098	99.993	1.628	159	99.993
7	0.047	0.005	99.997						
8	0.012	0.001	99.998						
9	0.007	0.001	99.999						
10	0.004	0.000	100.000						
11	0.002	0.000	100.000						

As shown in Table 2, when the number of variables is 6, 99% of the variance of magnetic flux density can be explained, which indicates that most of the information in the data is captured by a few main factors. This result verifies that the data set has high correlation and is suitable for using PCA to extract key feature variables, so as to reduce the data dimension while retaining most of the variation information. Furthermore, principal component analysis is performed on the data, and a loading matrix is obtained, which shows the weight of each original variable on each principal component. The loading matrix is the key to understanding the relationship between original variables and principal components, and it reveals which variables contribute the most to the principal components. In the obtained loading matrix, part of the values are shown in the following table:

Table 3. Loading matrix (partial)

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6
0.026400164	0.043161285	0.063738658	0.016969436	0.075865162	0.028557559
0.026620908	0.042949099	0.062320932	0.017565952	0.074463474	0.02856395
0.026841771	0.042737087	0.060917015	0.018158474	0.07308077	0.028573644
0.027063034	0.042524885	0.059525674	0.018744437	0.071721226	0.02858321
0.027284473	0.042312789	0.058142731	0.019325371	0.070379016	0.028591098
0.027506245	0.042100172	0.056765269	0.019903514	0.06904763	0.028595014
0.027727973	0.041887177	0.05539165	0.020481227	0.067723891	0.028593233
0.027949642	0.041674202	0.054020547	0.021058799	0.066405081	0.028587622
0.028171295	0.041460476	0.052651258	0.021637927	0.065090797	0.028576097
0.028392869	0.041246474	0.051283107	0.022217876	0.063779336	0.028560966

The loading matrix obtained by principal component analysis can effectively reduce the dimension of the sampled values of magnetic flux density in the training set and the set to be

DOI: 10.54254/2753-8818/2026.33103

classified. This step is to reduce the data dimension by projecting the original data onto the principal components, while retaining the variation information in the original data as much as possible. The dimension-reduced data not only simplifies the model complexity, but also helps to improve the computational efficiency and generalization ability of the model. The feature variables of the dimension-reduced training set and the set to be classified are obtained as follows:

Table 4. Feature variables of the dimension-reduced training set (partial)

Serial number	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6
1	9.484302866	-6.094564289	1.545888836	0.20230056	0.916616129	-0.16878469
2	9.326183261	-6.331729305	1.548505399	0.236809881	0.935585458	-0.175726159
3	8.7174211	-7.129300656	1.555991782	0.3381735	0.996205756	-0.2015362
4	9.346310921	-6.273606494	1.542768651	0.224933761	0.941586846	-0.184621562
5	9.485095763	-6.058325725	1.541698632	0.201808565	0.92679924	-0.182707514
...						

Table 5. Feature variables of the dimension-reduced set to be classified (partial)

Serial number	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6
1	5.297304889	-6.10685015	-3.41093203	0.562503265	0.313661947	-0.164624802
2	5.308489364	-6.09101399	-3.429813277	0.576840261	0.329242252	-0.148236817
3	5.436318276	-6.020496422	-3.462587861	0.725123205	0.35152288	-0.064770164
4	5.488626758	-6.032071626	-3.431690863	0.751415118	0.406519043	-0.037052587
5	5.376558166	-6.070149155	-3.445401712	0.650927275	0.331803465	-0.10533069
...						

4.3. Analysis of classification results of random forest and XGBoost

To construct an appropriate classification model, some parameters need to be set, such as the maximum depth of the tree for decision trees and the learning rate for neural networks. The model is trained and the classification accuracy of the model is tested by dividing the data set into training set and test set, and the brand-new data for the training set is provided through the training set, so as to verify the generalization ability of the model.

4.3.1. Random forest classification

When conducting data analysis and model training in MATLAB, the data set is divided into 70% training set and 30% test set to balance the model fitting degree and generalization ability. A random forest model containing 10 decision trees is constructed using the training set and tested on the test set. The obtained confusion matrix is as follows:

DOI: 10.54254/2753-8818/2026.33103

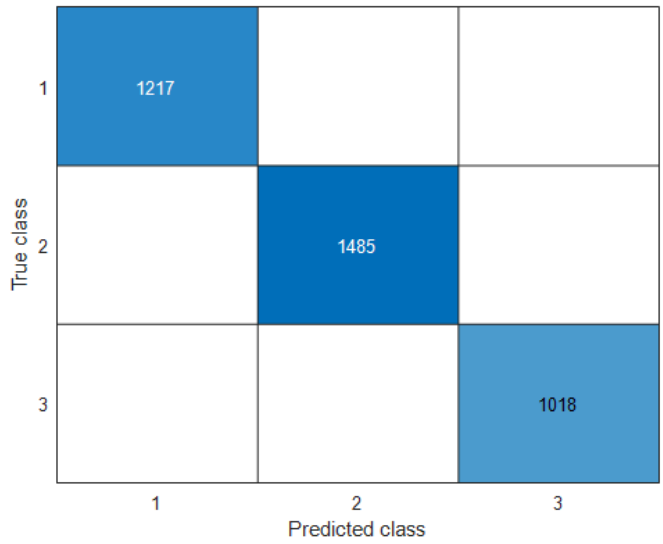

Figure 7. Confusion matrix

As shown in the figure, the accuracy of the random model on the test set reaches 100%, which proves the effectiveness of the random forest model in classifying excitation waveforms.

4.3.2. XGBoost classification model

XGBoost can improve the computational efficiency of large-scale data sets. In this paper, after normalizing the training set and test set in Python, the XGBoost parameters are determined through cross-validation: the maximum depth of the tree is 100, the learning rate is 0.1, and the number of iterations is 500. The evaluation indicators of each model are shown in the following table:

Table 6. Model indicator data

Model	Random forest	XGBoost
F1	1	0.85
ACC	1	0.82

4.4. Classification of excitation wave types

The trained model is applied to the set to be classified of 80 samples, and the number of each of the three waveforms in the samples is counted. The results are shown in Table 7 below:

Table 7 Classification results of excitation waveforms of 80 samples

Table 7. Classification results of excitation waveforms of 80 samples

Excitation waveform	Quantity
1	20
2	44
3	16

DOI: 10.54254/2753-8818/2026.33103

Table 8. Classification results of this sample

Serial number	Classification results of excitation waveforms (1~35 samples)	Serial number	Classification results of excitation waveforms (36~80 samples)
1	2	45	3
5	2	55	2
15	1	65	2
25	2	75	2
35	3	80	1

5. Model evaluation

The study finds that among the two models in this paper, the prediction performance of XGBoost is worse than that of random forest. It is speculated that this is because random forest randomly selects a part of features for splitting when constructing decision trees. This mechanism helps to reduce overfitting and makes the model less sensitive to feature selection. Therefore, when there are redundant or noisy features in the data, random forest can still maintain good performance.

6. Conclusion

This paper uses principal component analysis to reduce the dimension of a large number of sampled values of magnetic flux density, retaining the differences between various data as much as possible. Precisely because the feature variables hardly damage the information retained in the original data, nearly perfect results can be achieved in the recognition of excitation waveforms. Thus, it lays a crucial solid foundation for in-depth understanding of the core loss mechanism and optimizing core design. The random forest model is very convenient in both principle and implementation. The model not only has strong generalization ability, but also has high accuracy under the data provided in this problem. The training results also prove that the classification accuracy of the model reaches 100%.

References

[1] G. Bertotti. (1988). General properties of power losses in soft ferromagnetic materials. IEEE Transactions On Magnetics, 621-630.
[2] C. P. Steinmetz. (1892). On the law of hysteresis. AIEE Transactions, vol. 9, 3–64.
[3] Rudy Severns. (1991). HF core loss for non-sinsoidal waveforms. Proc. HFPC’91, 140-148.
[4] Ye Jianying, Chen Wei, Wang Jinghui. (2015). Research on Core Loss Model Under PWM Wave and DC Bias Excitation [J]. Proceedings of the CSEE, 35(10): 2601-2606.
[5] Chen Lin. 2012. Research on Core Loss of Magnetic Components in Power Electronic Converters [D]. Huazhong University of Science and Technology.
[6] Zhang Yi, Tang Yichen. (2024). Prediction Model of Calcium Ion Concentration in Seawater Circulating Cooling System Based on Random Forest Algorithm [J]. Salt Science and Chemical Industry, 53(07): 19-22+26.
[7] Chen T, Guestrin C. 2016. Xgboost: A scalable tree boosting system [C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 785-794.