Computers, Materials & Continua CMC, vol.63, no.2, pp.769-785, 2020

OTT Messages Modeling and Classification Based on Recurrent Neural Networks

Guangyong Yang[1], Jianqiu Zeng[1], Mengke Yang[2,*], Yifei Wei[3], Xiangqing Wang[3] and Zulfiqar Hussain Pathan[4]

Abstract: A vast amount of information has been produced in recent years, which brings a huge challenge to information management. The better usage of big data is of important theoretical and practical significance for effectively addressing and managing messages. In this paper, we propose a nine-rectangle-grid information model according to the information value and privacy, and then present information use policies based on the rough set theory. Recurrent neural networks were employed to classify OTT messages. The content of user interest is effectively incorporated into the classification process during the annotation of OTT messages, ending with a reliable trained classification model. Experimental results showed that the proposed method yielded an accurate classification performance and hence can be used for effective distribution and control of OTT messages.

Keywords: OTT messages, information privacy, nine-rectangle-grid, hierarchical classification, recurrent neural networks.

1 Introduction

In the era of big data, the quantity of the information increases rapidly. The risk and value of the information have attracted much research attention. In recent years, big data mining technology has been developing fast. However, existing methods still wrongly pass a lot of spam information, causing loss and distress to users and not providing the full usage value of big data. The risk and value need to be considered in order to develop the big data industry. Therefore, the tradeoff between information privacy and value becomes a big challenge [Pan, Qin, Yi et al. (2019)]. Since mobile phone messages affect the daily work and life of many people, the number of mobile phone short messages generated every year adds up to hundreds of billions. In recent years, with the development of Three Network Convergence, OTT (over the Top) messages have become more and more popular such as WeChat and Fetion in China. The vast amounts of OTT messages contain a large number

[1] School of Economics & Management, Beijing University of Posts and Telecommunications, Beijing, 100876, China.
[2] School of Automation, Beijing University of Posts and Telecommunications, Beijing, 100876, China.
[3] School of Electronic & Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China.
[4] Mehran University of Engineering & Technology, Jamshoro, Pakistan.
[*] Corresponding Author: Mengke Yang. Email: yangmengke@139.com.
Received: 31 May 2019; Accepted: 02 August 2019.

of spam information, which endangers the safety of users' information [Wang, Shen and Zhang (2015); Kim, Kim, Hwang et al. (2017); Wang (2018)].

At present, in the field of OTT message classification, compared with traditional long text messages, OTT messages are sparse in real time and have non-standardized characteristics. There are some problems like 'feature dimension disaster' and 'no semantic features' which dampens the accuracy of short text classification off traditional text classification models.

This paper focuses on WeChat, Fetion, and other OTT messages. According to the two dimensions of value and privacy, this paper presents a Nine-Cell hierarchical classification model and information utilization countermeasures based on hierarchical classification, and classifies OTT messages based on the rough set theory. After constructing the text model, we use recurrent neural networks to classify OTT messages. The classification process is integrated into the attention content and interest tags of the users. Experimental results show that the Long Short-term memory (LSTM) network classification algorithm is more reliable than the current K-Nearest Neighbor (KNN) classification algorithm and Convolutional Neural Networks (CNN) classification algorithm. LSTM provides more accurate classification for OTT messages and a more effective spam management way [Pan, Qin, Yi et al. (2019)].

2 Related research

Information Classification has recently become a hot research topic in data mining. A general process is to construct an accurate description or model for each type of data according to the characteristics of training data. Then, the model classifies unknown data. SVMs (Support Vector Machines), decision trees, Bayesian neural networks and so on are all commonly used techniques for information classification [Delany, Buckley and Greene (2012)]. The advantage of the decision tree and Bayesian neural networks is that the results can be expressed in an easy-to-understand rule form. The disadvantage is that only a few rules can be found and the classifier constructed by these rules cannot represent the data completely. Thus, the rules obtained by mining are not necessarily significant and the rules of significance are not necessarily well-tapped.

Unlike decision trees or Bayes neural networks, association classification is a rule-based classification technique that can mine all rules without the need of the minimum support. As the minimum support reduces, the number of mining rules increases so that the data can be completely described and the associated classification technology can be applied to generate better classifiers for making the message classification more accurate [Jin (2015)]. With the emotions disclosed by Twitter, Ahmed Esmin et al. proposed the hierarchical classification method used for automatic classification for three categories: neutral, positive, or negative emotions [Ahmed, Roberto and Stan (2012)]. Ghorab et al. [Ghorab, Zhou, O'Connor et al. (2013)] proposed personalized information services based on surveys and classifications.

Chen et al. [Chen, Liao and Shi (2014)] proposed the time, relationship, and type attributes of information as the key attributes of user information requirements. They proposed a user interest update and classification mechanism based on the key attributes and constructed a personalized user interest update model, a user interest management model, and a user classification model of a push service. Cui [Cui (2010)] discussed the classification of short messages in Weibo by mapping the keywords to the semantic concepts and implemented the message classification using the semantic KNN classification algorithm through the research of the interest mining of microblog authors. Li [Li (2015)] proposed adding one or more tags for each log then aggregating similar resources according to the tags. By introducing clustering analysis, they deduced the interest preferences of the users and built the user interest model to construct a tag-based user personalized information service process framework. Gao et al. [Gao and Tang (2017)] proposed a personalized information service strategy based on trajectory clustering, real-time access to user access and retrieval of trajectory data combined with a trajectory clustering algorithm to establish a dynamic user retrieval model and rapid real-time personalized information services. Xiong et al. [Xiong and Cao (2018)] effectively mined high-quality user interest phrases and identified their categories from the content of texts to create the interest mining of microblog users and to improve the accuracy of user mining interest with recall based on the phrase-based theme model and the automatically constructed user interest knowledge base. Qiu et al. [Qiu (2013); Xu and Hu (2014)] divided user interests into long-term interests and short-term interests, and built a user interest model according to the importance of users and the Weibo keywords.

Recently, the information classification has been attracting more research attention. Existing methods typically employ classification techniques and construct the user interest model with the dimensionality of keywords, user attributes, and user behavior. There is limited research on the classification theory of OTT messages, especially on the combination of management and technology. However, the dimensionality of the information evaluation is unclear and the user subjective characteristics of OTT messages were not sufficiently considered. So, a hierarchical model will be meaningful for a better OTT message classification. In the process of information hierarchical classification, the research on the combination of privacy and value is still limited [Wei, Yu, Richard et al. (2010); Wei, Yu, Richard et al. (2018)]. The research on user interest model is able to boost the development of information similarity analysis and precise recommendations.

3 Application of hierarchical classification method to big data

The hierarchical classification of information will benefit the healthy development and information utilization. The information classification will help cope with the potential problem of "one size fitting all".

3.1 Nine-rectangle-grid information hierarchical classification model

The Nine-rectangle-grid was created by Ouyang Xun, the calligrapher of the Tang Dynasty in China. Its main purpose was to let Chinese characters be written more easily. The Nine-rectangle-grid is the predecessor of the Sudoku, but it has a more complex characteristics than the Sudoku. The summation of three numbers in the vertical and horizontal is equal to 15. The Nine-rectangle-grid method has been widely used in the field of enterprise management, education, teaching, and algorithm analysis [Huang (2015); Wang (2016); Dong and Zhu (2015)]. Existing methods usually classify information along a single dimension, e.g., time, length, and its value. This paper proposes modelling two-dimensional information, value and privacy, with Nine-rectangle-grid.

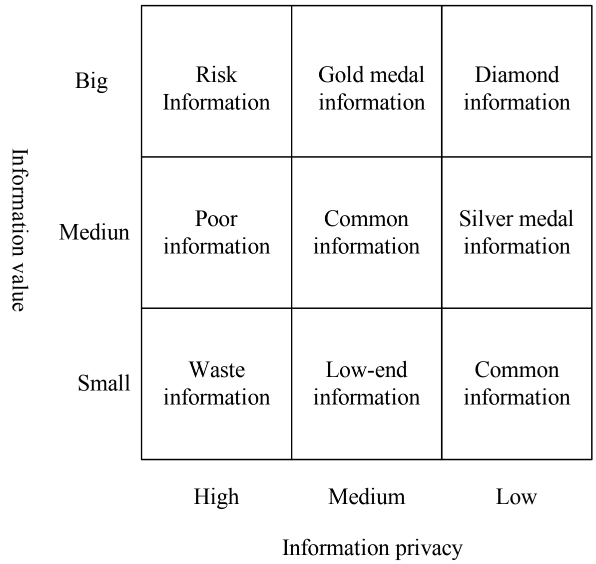


Figure 1: Nine-rectangle-grid information matrix chart

The Nine-rectangle-grid matrix has two dimensions (attributes), the information value and information privacy. The information privacy attribute takes the value of High, Medium, and Low, and the information value attribute takes the value of Big, Medium, and Small. The information value denotes the importance or the vital interest to the users or units, and the information privacy refers to whether an information encapsulates personal privacy or unit security information. The matrix is divided into nine blocks by distinguishing the information value to be small, medium, and large, and the information privacy to be high, medium, and low. The models were constructed containing the following elements: corresponding waste information, low-end information, common information, poor information, silver medal information, risk information, gold medal information, and diamond information.

Let X represent the information value dimension, and Y represent the information privacy dimension. The probability of every element of the Nine-rectangle-grid can be formulated as

$$P(X_iY_j) = P(X_i / Y_j)P(Y_j) = P(Y_j / X_i)P(X_i) \\ i, j = 1, 2, 3 \tag{1}$$

3.2 Information hierarchical classification based on the rough set theory

Assume that there are seven types of information that constitute set A={x1, x2, x3, x4, x5, x6, x7}, each type has a value attribute. In accordance with the different values, we can divide the information into W1={red, yellow, blue} as the three categories. For example, the Low-value information constitutes a set X1={x1, x2, x5}, the Medium-value information constitutes a set X2={x3, x6}, and the High-value information constitutes a set X3={x4, x7}.

According to the Nine-rectangle-grid structure in Fig. 1, the information has another privacy attribute as W2={strong, medium, weak} corresponding to the three sets: {x1, x7},{x3, x4},{x2, x5, x6}.

Adding the W1 attribute to set A:

A/W1={X1, X2, X3}={{x1, x2, x5},{x3, x6},{x4, x7}} (value classification) (2)

A/W2={Y1, Y2, Y3}={{x1, x7},{x3, x4},{x2, x5, x6}} (privacy classification) (3)

Through the intersection operation:

Law value & weak privacy {x1, x2, x5}∩{x2, x5, x6}={x2, x5}; law value & medium privacy {x1, x2, x5}∩{x3, x4}=0; law value & high privacy {x1, x2, x5}∩{x1, x7}={x1}; medium value & weak privacy {x3, x6}∩{x2, x5, x6}={x6}; medium value & medium privacy {x3, x6}∩{x3, x4}={x3}; medium value & high privacy {x3, x6}∩{x1, x7}=0 high value & weak privacy {x4, x7}∩{x2, x5, x6}=0; high value & medium privacy {x4, x7}∩{x3, x4}={x4}; high value & high privacy {x4, x7}∩{x1, x7}={x7}.

All of these concepts calculated by intersection and union comprise a knowledge system W=W1∩W2 and (A/W1, A/W2).

Next, we consider the information attributes and decision that are listed in Tab. 1.

Table 1: Information attributes and decision

Information	Value	Privacy	Usability
X_1	Low	Strong	No
X_2	Low	Weak	Free use
X_3	Medium	Medium	Strive for use
X_4	High	Medium	Rational use
X_5	Low	Weak	Free use
X_6	Medium	Weak	Active use
X_7	High	Strong	Use after processing

X1-X7 contain all the information types. The value and privacy are the condition attributes and the decision attribute are the availability attribute. In Tab. 1, we use each type of information in a different strategy, which has manifested the idea of the hierarchical classification of information [Luo, Li, Yi et al. (2016)].

4 OTT messages classification algorithm based on deep neural networks

4.1 Chinese short messages processing status

With the rapid development of mobile communication technology and wide distribution of mobile phones, the OTT message service has become an indispensable communication channel in common daily life. The main content of OTT messages is text, so the classification of OTT messages can be cast as the classification of short message texts in this paper. Text messages are divided into eight categories according to the value and the privacy of information. The eight classes are diamond information, gold information, silver information, common information, low-end information, poor information, waste information, and risk information. The Long Short-Term Memory (LSTM) has many advantages and achieved intriguing results in spam filtering within the text classification algorithm [Pan, Fan, Yu et al. (2017)]. This paper proposes using the LSTM to classify text messages according to the text content.

4.2 LSTM

LSTM is a special type of recurrent neural networks that can learn the long-term dependence of information. It was proposed in 1997 and has recently been refined and promoted by Graves et al. [Graves and Schmidhuber (2005)]. LSTM has achieved considerable success in various fields, especially in the area of short text processing [Graves (2012)].

4.3 Jieba segmentation

The Jieba segmentation algorithm is a Chinese word segmentation method based on the combination of rules and statistics [Jieba]. Jieba segmentation algorithm pre-stores some Chinese word probability values through a large number of experiments and supports three Chinese word segmentation modes: 1) accurate mode, which tries to cut the most accurate sentence suitable for text analysis; 2) full mode, in which the formed words of sentence are scanned out. This mode is very fast but does not deal well with the ambiguity; 3) search engine model, which is based on the accurate model and cuts the long word again to improve the recall rate, is suitable for search engine segmentation.

4.4 Preprocessing

In the preprocessing stage, we firstly use the Jieba segmentation tool to segment each Chinese text. We choose the accurate model in this paper because this work aims to classify OTT messages. Once the word segmentation is carried out, we remove repetition and sort all the word segmentation results and encode sorted words. The encoding rules are according to the number of the words.

5 Experimental analysis

5.1 Experiment setups

The 6 users included 3 men and 3 women, with an average age of 32 years old. They are employees of colleges or companies. There was a total of 10548 messages, as given in Tab. 2.

Table 2: OTT messages of eight classifications

Category	Diamond	Gold	Silver	Common	Low-end	Poor	Waste	Risk
Quantity	1035	1186	1061	1037	840	973	3655	761

After the preprocessing of the texts, they were classified by the LSTM network, KNN, and CNN. Experiments were conducted for the classification of two classes, four classes, and eight classes. In each experiment, we randomly select 10% of the data from all data as the test set, and the remaining data is used as the training set. The detail is given in Tab. 3.

Table 3: Division of training set & test set

Item	Training set	Test set	Total
Number	9489	1059	10548
Percentage (%)	89.96	10.04	100.00

5.2 Results

This section shows the experimental results of classification on the dataset. The proposed LSTM method is compared with the KNN method and the 1-D convolutional method (CNN-1D). Tab. 4 gives the result for two classes in which diamond, gold, silver, and common information are categorized to be one class, and low-end, poor, waste, and risk information are categorized to be the other class. From Tab. 4, we can see that LSTM provides the best classification. It outperforms the KNN method and the CNN-1D method, verifying the effectiveness of the proposed method. Fig. 2 sketches the experimental comparison.

Table 4: Classification comparison between LSTM, KNN, and CNN-1D for two classes. Class A includes diamond, gold, silver, and common, and class B includes low-end, poor, waste, and risk information. The number of messages is given for each classification method

Actual \ Prediction	Classifier	Class A	Class B	Total
Class A	LSTM	364	70	434
	KNN	285	149	
	CNN-1D	284	150	
Class B	LSTM	36	589	625
	KNN	169	456	
	CNN-1D	151	474	

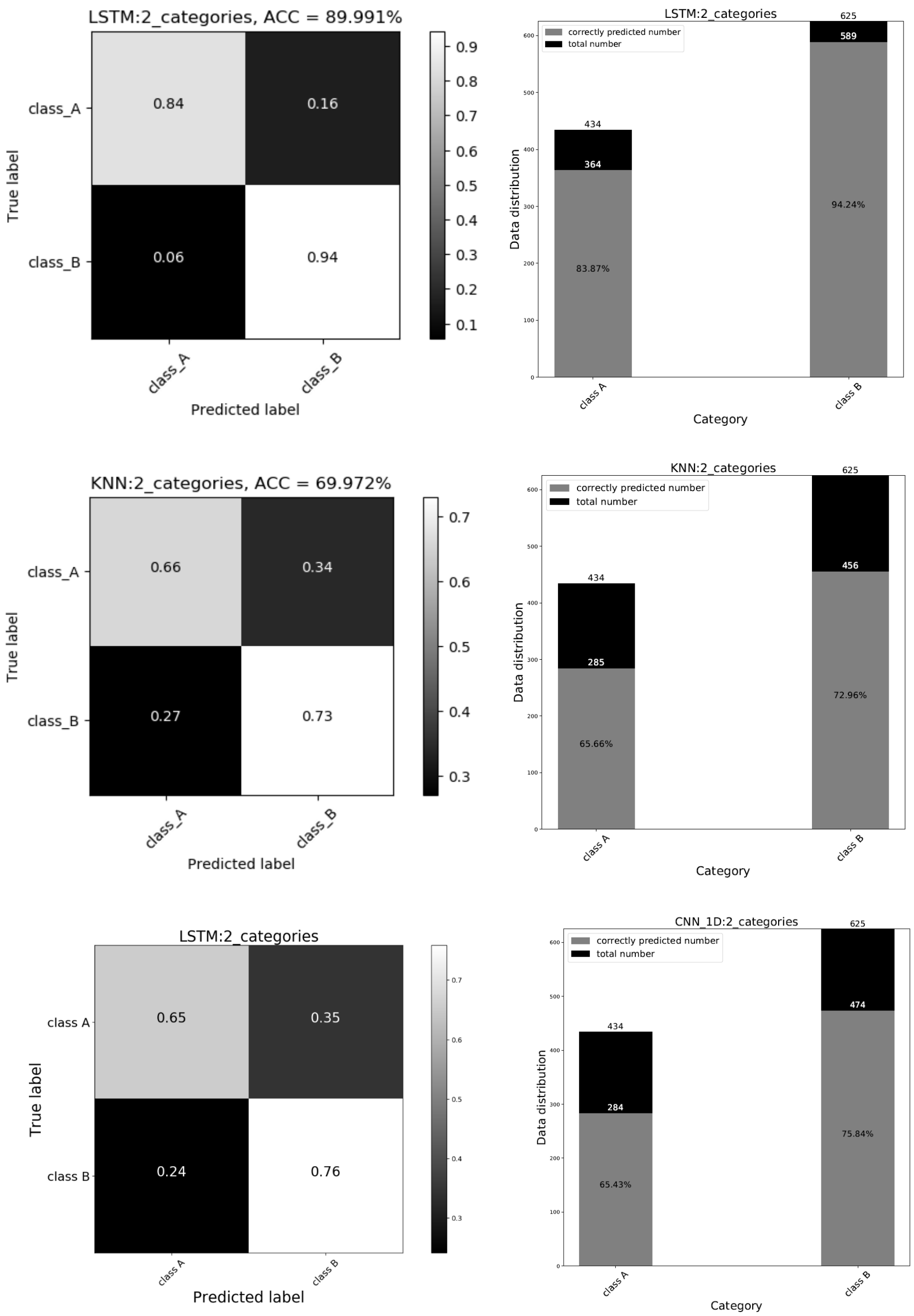
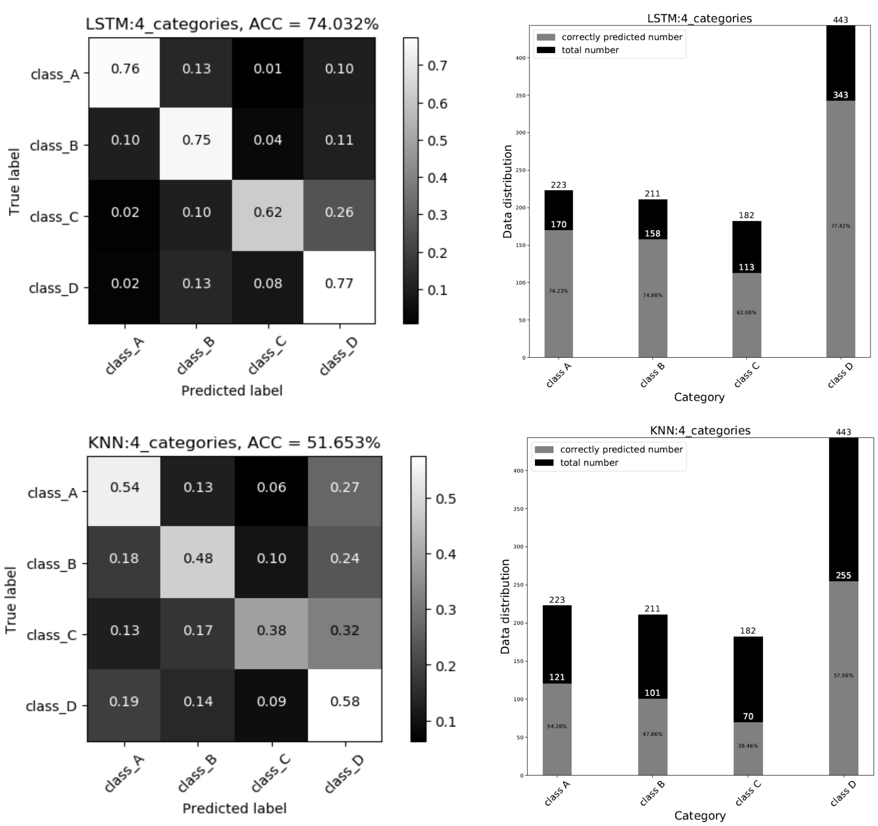


Figure 2: Classification comparison between LSTM, KNN, and CNN-1D for two classes

From Tab. 5, we can still see that LSTM provides the best classification, outperforming the KNN method and the CNN-1D method. Fig. 3 sketches the experimental comparison for the classification of four classes.

Table 5: Classification comparison between LSTM, KNN, and CNN-1D for four classes. Class A includes diamond and gold, Class B includes silver and common, Class C includes low-end and poor, and Class D includes waste and risk. The number of messages is given for each classification method

Actual \ Prediction	Classifier	Class A	Class B	Class C	Class D	Total
	LSTM	170	28	2	23	
Class A	KNN	121	28	14	60	223
	CNN-1D	118	27	18	60	
	LSTM	22	158	8	23	
Class B	KNN	38	101	22	50	211
	CNN-1D	34	96	25	56	
	LSTM	3	19	113	47	
Class C	KNN	23	31	70	58	182
	CNN-1D	23	20	85	54	
	LSTM	10	56	34	343	
Class D	KNN	83	64	41	255	443
	CNN-1D	66	56	57	264	



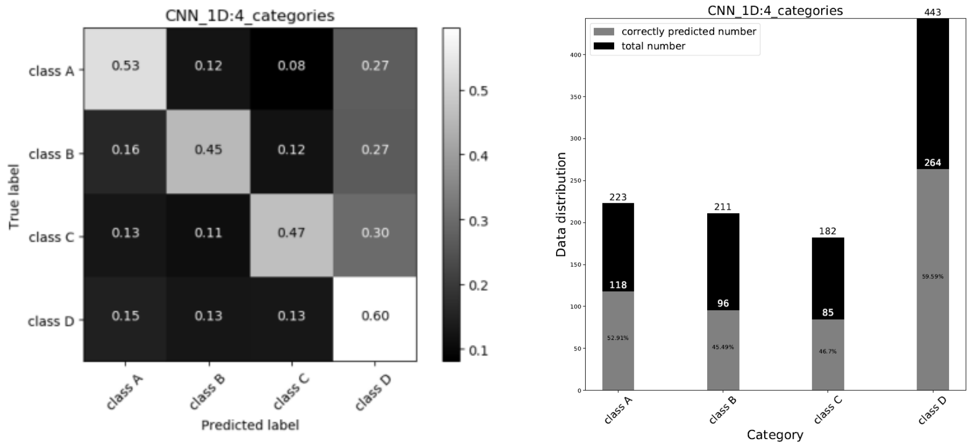


Figure 3: Classification comparison between LSTM, KNN, and CNN-1D for four classes

Tab. 6 gives the result for the whole eight classes. Tab. 6 reveals that the proposed LSTM method outperforms the KNN and CNN-1D method for the task of classifying each class. Fig. 4 sketches the visual comparison between the three methods. Also, Tabs. 4-6 show that KNN and CNN-1D give a comparable performance to each other.

To more finely evaluate the performance of the proposed algorithm, the commonly used sensitivity (SE), specificity (SP), positive rate (PP), and accuracy (ACC) as follows:

$$Sensitivity(SE) = \frac{TP}{TP + FN} \times 100\% \tag{4}$$

$$Specificity(SP) = \frac{TN}{TN + FP} \times 100\% \tag{5}$$

$$PositivePredictivity(PP) = \frac{TP}{TP + FP} \times 100\% \tag{6}$$

$$Accuracy(ACC) = \frac{TN + TP}{TP + TN + FP + FN} \times 100\% \tag{7}$$

Tab. 7 shows the comparison between LSTM, KNN, and CNN-1D in terms of SE, SP, PP, and ACC for the message classification of two categories. It can be seen that the proposed LSTM gives the best performance. For example, higher than 69.97% by KNN and 71.58% by CNN-1D. Fig. 5 illustrates the visual comparison of Tab. 7. In terms of SE, SP, and PP, the proposed LSTM also outperforms KNN and CNN-1D, verifying the effectiveness of the proposed method in classifying OTT messages.

Table 6: Classification comparison between LSTM, KNN, and CNN-1D for eight classes. The number of messages is given for each classification method

Prediction / Actual	Classifier	Diamond	Gold	Silver	Common	Low-end	Poor	Waste	Risk	Total
Diamond	LSTM	90	3	2	2	0	3	1	3	104
	KNN	61	5	6	2	3	2	24	1	
	CNN-1D	54	8	6	6	1	5	24	0	
Gold	LSTM	12	80	2	2	3	3	12	5	119
	KNN	9	47	10	6	7	2	37	1	
	CNN-1D	11	44	6	10	2	6	39	1	
Silver	LSTM	9	4	78	3	1	3	4	5	107
	KNN	12	12	49	2	1	2	29	0	
	CNN-1D	5	6	64	6	1	4	20	1	
Common	LSTM	6	7	5	49	16	6	13	2	104
	KNN	10	16	6	32	3	6	31	0	
	CNN-1D	4	8	4	47	3	6	30	2	
Low-end	LSTM	0	2	0	3	50	6	18	5	84
	KNN	9	6	1	5	23	4	32	4	
	CNN-1D	8	3	3	5	19	8	36	2	
Poor	LSTM	4	0	0	4	11	64	12	3	98
	KNN	5	8	5	4	3	50	23	0	
	CNN-1D	8	3	5	7	1	50	23	1	
Waste	LSTM	5	8	2	15	56	20	244	16	366
	KNN	31	37	19	19	26	23	202	9	
	CNN-1D	24	23	24	41	17	20	208	9	
Risk	LSTM	0	2	0	0	6	1	14	54	77
	KNN	4	4	4	2	4	4	26	29	
	CNN-1D	4	5	1	6	1	5	30	31	

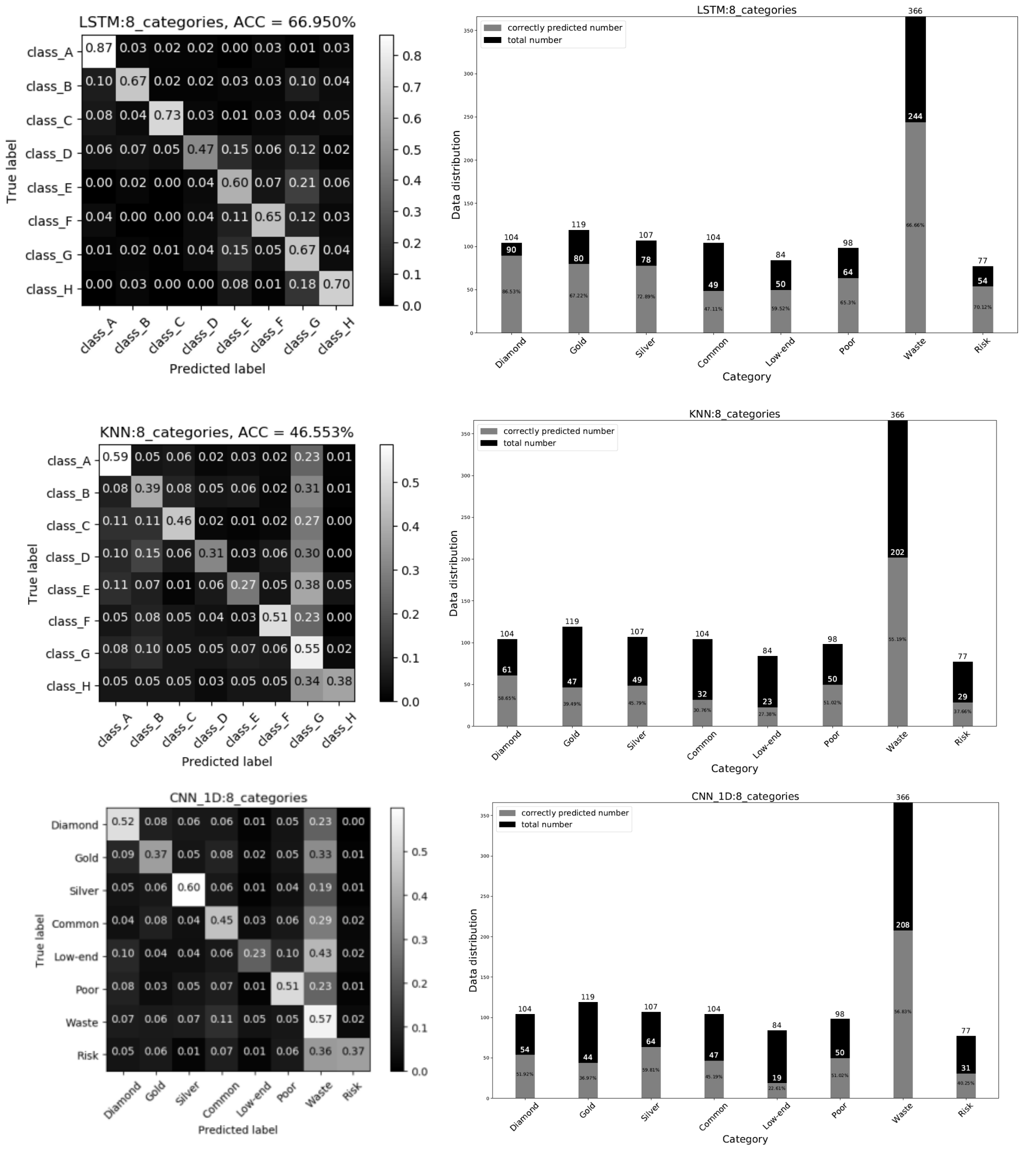


Figure 4: Classification comparison between LSTM, KNN, and CNN-1D for eight classes

Table 7: Classification comparison between LSTM, KNN, and CNN-1D for two classes

	A_LSTM (%)				A_KNN (%)				A_CNN-1D (%)			
Index	SE	SP	PP	ACC	SE	SP	PP	ACC	SE	SP	PP	ACC
Class A	83.87	94.24	91	89.99	62.78	75.37	65.67	69.97	65.44	75.84	65.29	71.58
Class B	94.24	83.87	89.38	89.99	75.37	62.78	72.96	69.97	75.84	65.44	75.96	71.58

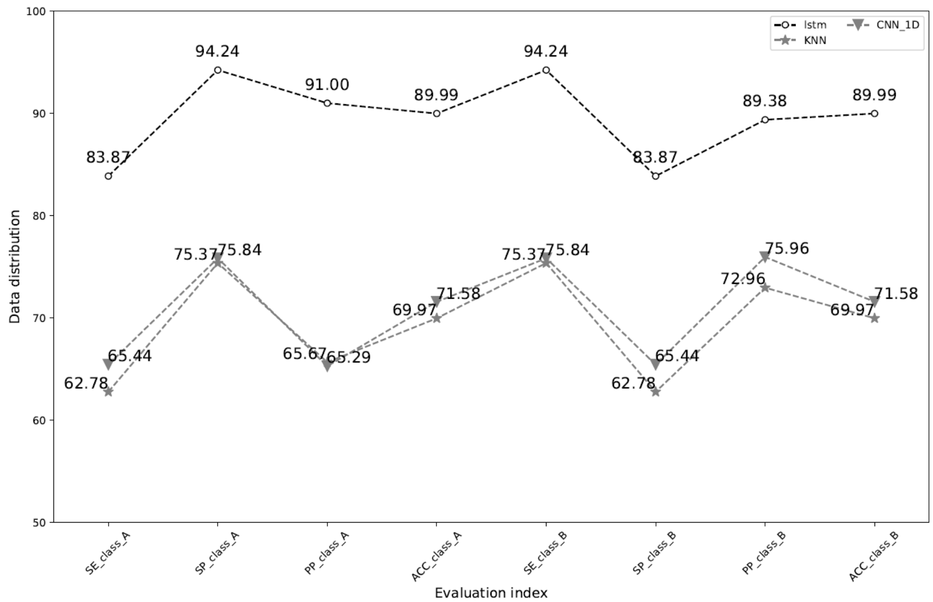


Figure 5: Classification comparison between LSTM, KNN, and CNN-1D for two classes

Tab. 8 shows the comparison between LSTM, KNN, and CNN-1D for the message classification of four categories. Again, the proposed LSTM gives the best performance. For example, a similar comparison can also be observed for Class B, Class C, and Class D. Fig. 6 gives the visual comparison of classification for Tab. 8.

Table 8: Classification comparison between LSTM, KNN, and CNN-1D for four classes

	B_LSTM (%)				B_KNN (%)				B_CNN-1D (%)			
Index	SE	SP	PP	ACC	SE	SP	PP	ACC	SE	SP	PP	ACC
Class A	76.23	95.81	82.93	91.69	54.26	82.78	45.66	76.77	52.92	85.29	48.96	78.47
Class B	74.88	87.85	60.54	85.27	47.87	85.50	45.09	78.00	45.50	87.85	48.24	79.42
Class C	62.09	94.98	71.97	89.33	38.46	91.22	47.62	82.15	46.70	88.60	45.95	74.79
Class D	77.43	84.90	78.67	81.78	57.56	72.73	60.28	66.38	59.59	72.40	46.98	67.04

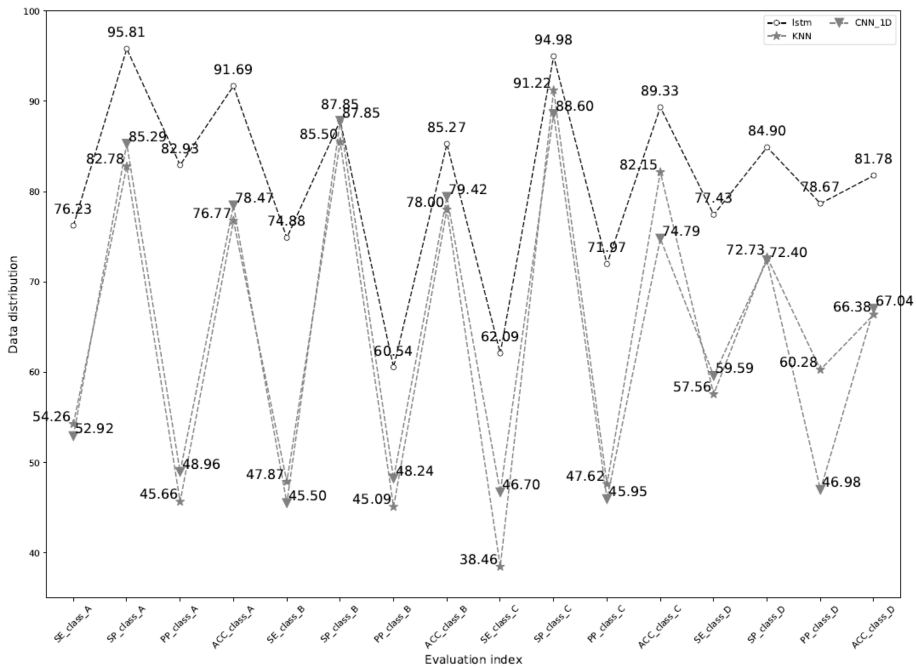


Figure 6: Classification comparison between LSTM, KNN, and CNN-1D for four classes

Tab. 9 shows the comparison between LSTM, KNN, and CNN-1D for the message classification of eight categories, i.e., the finest classification. Again, the proposed LSTM gives the best performance. For example, Fig. 7 illustrates the visual comparison of Tab. 9. A similar comparison can also be observed for the class Gold, Silver, Common, Low-end, Poor, Waste, and Risk.

Table 9: Classification comparison between LSTM, KNN, and CNN-1D for eight classes

	C_LSTM (%)				C_KNN (%)				C_CNN-1D (%)			
Index	SE	SP	PP	ACC	SE	SP	PP	ACC	SE	SP	PP	ACC
Diamond	86.54	96.23	71.43	95.28	58.65	91.62	43.26	88.39	49.04	93.30	45.76	89.24
Gold	67.23	97.23	75.47	93.86	39.50	90.64	34.82	84.89	36.98	94.04	44.00	87.63
Silver	72.90	98.85	87.64	96.22	45.79	94.64	49.00	89.71	59.81	94.85	56.64	91.31
Common	47.12	96.96	62.82	92.07	30.77	95.81	44.44	89.42	45.19	91.52	6.72	86.97
Low-end	59.52	90.46	34.97	88.01	27.38	95.18	32.86	89.80	22.62	97.33	42.22	91.41
Poor	65.31	95.63	60.38	92.82	51.02	95.53	53.76	91.41	51.20	94.38	48.08	90.37
Waste	66.67	89.32	76.73	81.49	55.19	70.85	50.00	65.44	56.83	70.85	50.73	66.01
Risk	70.13	96.03	58.07	94.15	37.66	98.47	65.91	94.05	40.26	98.37	65.96	94.33

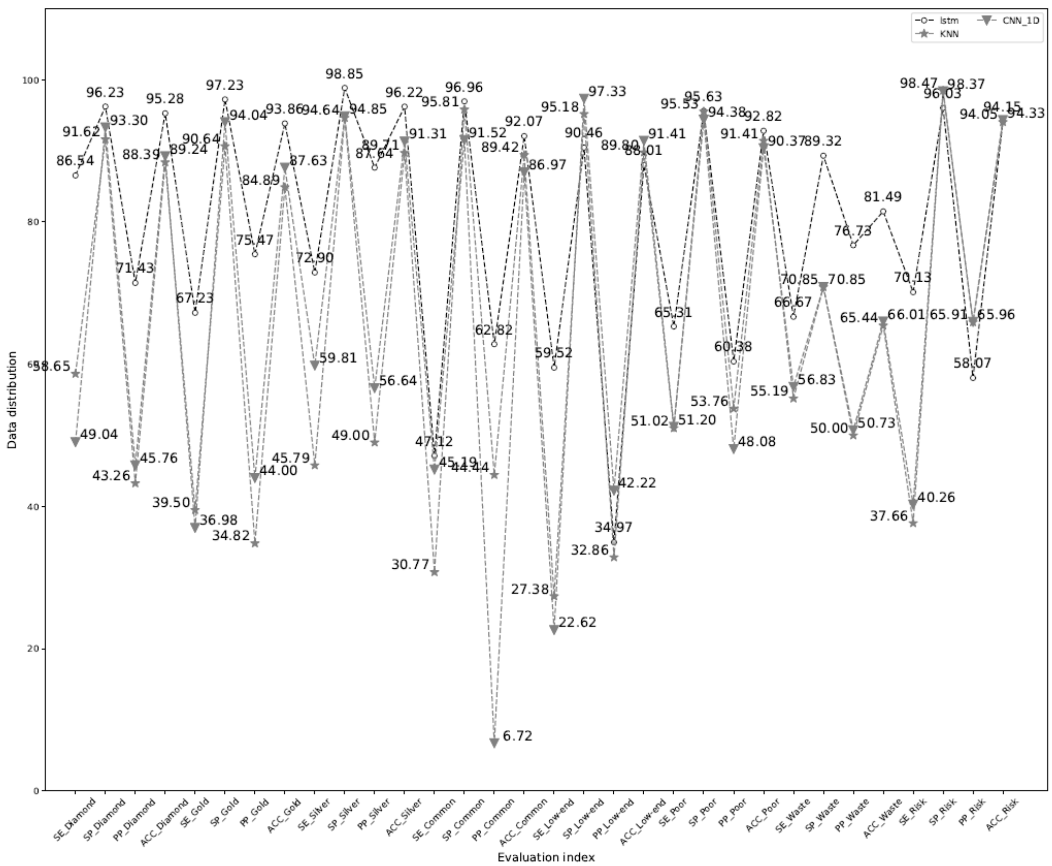


Figure 7: Classification comparison between LSTM, KNN, and CNN-1D for eight classes

From Tabs. 7 to 9 and Figs. 5 to 7, it can be seen that the proposed LSTM provides a better classification result than KNN and CNN-1D. It can effectively classify the meaningful and unmeaningful messages finely and help screen out many messages that are not worth reading and hence save time.

6 Conclusions and future work

This paper studied the classification of OTT messages. We constructed the Nine-rectangle-grid information matrix according to the privacy and value dimension of information and proposed the information decision table by using the rough set theory. According to the user preference of OTT messages, over 10,000 messages were labelled by using the exact pattern of Jieba's word segmentation. After preprocessing and labelling the OTT messages, we classify them with LSTM, KNN, and CNN-1D. Experiments were conducted on the classification of two, four, and eight categories. The results verified that the classification accuracy by LSTM is higher than those of KNN and CNN-1D. The proposed classification approach can be used for the accurate classification, delivery, and effective control of OTT messages.

Acknowledgment: This work is supported by the Research on Big Data in Application for Education of BUPT (No. 2018Y0403), Fundamental Research Funds of BUPT (No. 2018XKJC07, 2018RC27) and the National Natural Science Foundation of China (No. 61571059).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Ahmed, E.; Roberto. L.; Stan, M. (2012): Hierarchical classification approach to emotion recognition in Twitter. *Machine Learning and Applications*, vol. 2. pp. 381-385.

Chen, C.; Liao, G.; Shi, X. (2014): Model of user for personalized information push service. *Information Science*, vol. 11, no. 87, pp. 71-76.

Cui, Z. (2010): Short message classification of microblogging based on semantic. *Modern Computer*, vol. 8, no. 24, pp. 18-20.

Delany, S.; Buckley, M.; Greene, D. (2012): SMS spam filtering: methods and data. *Expert Systems with Applications*, vol. 39, no. 10, pp. 9899-9908.

Dong, Y.; Zhu, J. (2015): Raversal nine-rectangle-grid image search algorithm and its application. *Mathematical Theory and Applications*, vol. 3, no. 35, pp. 107-113.

Gao, Y.; Tang, S. (2017): Personalized information service strategy based on trajectory clustering. *Information Studies: Theory & Application*, vol. 6, pp. 87-90.

Ghorab, M. R.; Zhou, D.; O'Connor, A.; Wade, V. (2013): Personalised information retrieval: survey and classification. *User Modeling and User-Adapted Interaction*, vol. 4, pp. 381-443.

Graves, A.; Schmidhuber, J. (2005): Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, vol. 18, no. 5, pp. 602-610.

Graves, A. (2012): *Supervised Sequence Labelling with Recurrent Neural Networks*, pp. 1735-1780. Springer Berlin Heidelberg.

Huang, Y. (2015): General electric company: the differential management of the nine-rectangle-grid. *Enterprise Management*, vol. 3, pp. 55-57.

Jin, Y. (2015): *SMS Classification Based on Association Classification (Ph.D. Thesis)*. Beijing University of Technology, China.

Kim, M. S.; Kim, E.; Hwang, S. Y.; Kim, J.; Kim, S. (2017): Willingness to pay for over-the-top services in China and Korea. *Telecommunications Policy*, vol. 41, no 3, pp. 197-207.

Li, Z. (2015): Research on tags-based personalized information services. *Wireless Internet Technology*, vol. 5, pp. 116-117.

Luo, C.; Li, T.; Yi, Z; Fujita, H. (2016): Matrix approach to decision-theoretic rough sets for evolving data. *Knowledge-Based Systems*, vol. 99, pp. 123-134.

Pan, S.; Fan, J.; Yu S.; Wang, B. (2017): A method of behavior evaluation based on web browsing information. *Smart Grid and Electrical Automation*, pp. 697-700.

Pan, Z.; Qin, H.; Yi, X.; Zheng, Y.; Khan, A. (2019): Low complexity versatile video coding for traffic surveillance system. *International Journal of Sensor Networks*, vol. 30, no. 2, pp. 116-125.

Pan, Z.; Yu, W.; Yi, X.; Khan, A.; Yuan, F. et al. (2019): Recent progress on generative adversarial networks (GANs): a survey. *IEEE Access*, vol. 7, pp. 36322-36333.

Qiu, J. (2013): *Research on User Interest Modeling Based on Microblog Social Network (Ph.D. Thesis).* Shanghai Jiao Tong University, China.

Sun, J. (2012): Jieba, Chinese word segmentation tool. https://github.com/fxsjy/jieba/.

Wang, C. (2018): A ramble on "network neutrality" and OTT. *People's Posts and Telecommunications News*.

Wang, F.; Shen, L.; Zhang, L. (2015): Research and implementation of service connection between converged communications & OTT. *Telecom Engineering Technics and Standardization*, vol. 1, pp. 6-9.

Wang, Q. (2016): Research on the spiral sudoku teaching mode based on blended learning. *Modern Educational Technology*, vol. 9, pp. 46-52.

Wei, Y.; Yu, F.; Richard, F.; Mei, S; Han, Z. (2018): User scheduling and resource allocation in het nets with hybrid energy supply: an actor-critic reinforcement learning approach. *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 680-692.

Wei, Y.; Yu, F.; Richard, F.; Song, M. (2010): Distributed optimal relay selection in wireless cooperative networks with Finite State Markov channels. *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2149-2158.

Xiong, C.; Cao, Y. (2018): Research of microblog user interest mining based on microblog posts. *Journal of Computer Applications*, vol. 6.

Xu, X.; Hu, S. (2014): Research on user interest model of social network sites. *Science Technology and Industry*, vol. 14, no. 11, pp. 156-160.