Article 作者為 HoTommy

提交日期: 2020年05月14日 07:29下午 (UTC+0800)

作業提交代碼: 1324062733

文檔名稱: Class_Disambiguation_of_Cancers_Type-v2.2.pdf (497.93K)

文字總數: 2926 字符總數: 16078

Cancer Literature Classification Methods Performance

3 heng-Yuan Ho
Dept. of Computer Science and
Information Engineering,
Asia University,
Taichung City, Taiwan
tommyho@asia.edu.tw

1 Mariana Syamsudin
Dept. of Computer Science and
Information Engineering,
Asia University, Taichung City, Taiwan
Dept. of Information Technology,
Pontianak State Polytechnic, Indonesia
marianasyamsudin@gmail.com

3 ueh-Chun Shen
Dept. of Computer Science and
Information Engineering,
Asia University,
Taichung City, Taiwan
107121006@live.asia.edu.tw

Abstract— The literary classification system is the best solution to improve the data search process. In terms of the need, its goal is to compare the relevant biomedical papers and discover novel knowledge to identify potential research issues. This paper will present cancer literature classification performance by comparing three approaches, Naïve Bayes, Neural Network and Linear Classifier with SGD training. The propose approaches classify biomedical literature in five classes of cancer literature type namely, bone cancer, gastric cancer, kidney cancer, skin cancer and papillary thyroid cancer by using 9259 documents. General steps for building classification refer to the classification of scientific literature. The result shows that all algorithms successfully can be used to classify cancer literature. However, for the best performance, it is strongly recommended to use Naïve Bayes and Neural Network.

Keywords—classification performance, naïve Bayes, neural network, linear classification

I. INTRODUCTION

Regard to the database of the World Health Organization (WHO) annual death registration, cancer is the second leading cause of death globally, 9.6 million deaths in 2018 [12]. Thus, cancer becomes an essecond leading area for biomedical research. The vast number of biomedical literature has proliferated, provide a rich source of knowledge to improve the development of biomedical research. Figure 1 shows the number of literature on cancer research for the past ten years. The data collected from PubMed using "cancer" as a keyword in the searching of title or abstract.



Figure 1. The Number of Cancer Publication in PubMed

Until the beginning of 2020, there are 845.908 cancer publications retrieved from PubMed. Abundant of cancer literature become an essential problem when researchers need to compare the relevant topic papers and discover novel

The work is supported by the Ministry of Science and Technology of Taiwan, u 21 Grant no. MOST 108-2221-E-468-010 -, and partially supported by the Ministry of Education of Taiwan, the Artificial Intelligence Talent Cultivation Project, and by Asia University, under Grant no. 107-ASIA-UNAIR-09.

knowledge to identify potential research issues. This study address to help researchers to overcome these difficulties by classifying the cancer literature using text mining methods. Realizing the superiority of the classification of text mining is expected to help cancer researchers find the literature easily and quickly. Sang-Woon Kim [8] used the TF-IDF and LDA scheme to support the paper classification system. This research produces a classification system that has two objectives. The fift objective is classifying research papers using keywords and topics with the support of high-performance com-22 ing techniques. Then, the papers that have been classified will be applied to search papers in the field of research expeditiously and efficiently. In addition, the classification performance is another concern, and bias can occur if inaccurate discipli 8 ry assignments exist in the scientific classification system especially if there is a significant proportion of multidisciplinary journals in the reference lists [10]. More consideration should be given to the robust and accuracy of classification 8 hemes by consideration of the importance of classification in the construction and analysis of bibliometric indicators

Different from the methods as mentioned above, in the following, we propose to specify literature classification in cancer type that will be remarkably effective to collect and analyze information in biomedical research. This study will also compare the classification performance of three approaches, Naïve Bayes, Neural Network, and Linear Classifier with SGD training. The Naïve Bayes approach was successfully applied to classify scientific literature into predetermined categories, according to the needs of the researchers. Furthermore, this approach increases the effectiveness of work in identifying and solving potential research proble 2 and dramatically facilitates scientific research. [8]. In recent years, neural network-based approaches become potential and dominant methodology in biomedical relation classification. The advantage of Neural network-based approaches that is can effectively and automatically learn the underlying feature represent 5 on from the labelled training data [11]. Stochastic Gradient Descent (SGD) has been successfully applied to large-scale and sparse machine learning problems often encountered in text classification and 7 ural language processing. Furthermore, through SGD convex loss functions, such as (linear) Support Vector Machines an 7 Logistic Regression, can be learned in a profoundly efficient approach to discriminative learning of linear classifiers [7].

There are five types of cancers literature cor 12 ered as a sample of classification problem, namely, Gastric cancer, Skin cancer, Bone Cancer, Kidney cancer, and Papillary Thyroid

Cancer. All corpus in the form of title and abstracts, in total 9,259 documents.

This project has six sections. It starts from converting data from xml format to csv format, performi 17 vectorization, removing stop words, stemming, doing term frequency-inverse document frequency (TF*IDF), and training as well as testing all data in Jupyter Notebook and Python environment.

Training and testing process will be carried out in three approaches. Respectivel 19 esting for 80% of training dataset and 20% of the testing dataset, 70% of training dataset and 30% of testing dataset, and 60% of training dataset and 40% of the testing dataset.

II. THE PROCEDURES FOR CLASSIFICATION

Referring to the classification of scientific literature, general steps for building a classification model as presented in Figure 2 are [9]:

- A. Determining a classification scheme
- B. Reading text documents
- C. Text preprocessing
- D. Classification

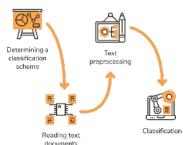


Figure 2. General Steps of Classification Process

A. Determining Type of Cancer Literature

To guarantee the results validity of the design, it is essential to do stages in the design shown in Figure 3.

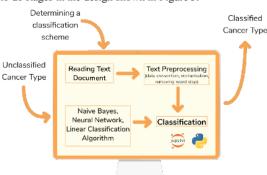


Figure 3. Model Architecture for Cancer Literature Classification

The stage determining a classification scheme is related to the development of a classifier of cancer literature. They are formed according to the five types of 12 cer literature that will be analyzed in this study are: Bone Cancer, Gastric Cancer, Kidney Cancer, Papillary Thyroid Cancer and Skin Cancer.

The following is a general explanation of each type of cancer. The first type is bounded the cancer. Many kinds of cancer can begin in the bones. The most common types of bone cancer in

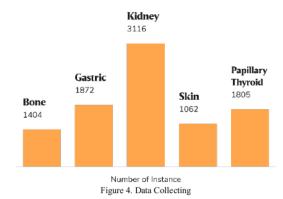
cildren and adolescents are osteosarcoma and Ewing's tumor. Bone cancer is divided into two types, namely primary and secondary bone cancer. Primary bone cancer exists and grows in bone ce 20 while secondary bone cancer starts from elsewhere and spreads to the bone. The second type of cancer is gastric cancer, which is a type of disease in the stomach lining. The symptoms of gastric cancer can be in the form of intagestion and stomach discomfort or pain. The third type is kidney cancer. It is one of the most common types of cancer in adults in their 60s or 70s. This cancer is also known as kidney cancer since it first appeared in a small tube lining in the kidney. The fourth type is papillary thyroid cancer. It is recognized as an asymptomatic disease with a mass in the neck as the most common symptom. The last type is skin cancer, which occurs because of long-term skin in the sun. There are three main types of diseases such as basal cell carcinoma, squamous cell carcinoma, and myeloma caused by abnormal growth of skin cells [1].

B. Text Documents

This phase is used to present the text documents in a 16 ar word format. All text document resources extracted from PubMed.gov, https://www.ncbi.nlm.nih.gov/pubmed/. The advanced search feature is used to collect tittle and abstract of literature from a set of massive papers efficiently as well as to get specific data by selecting Date — Publication, Title/Abstract, and name of cancer in the searching box.

In this paper, we utilized the title and abstract data of research papers, since these two parts of papers are the most 1 rt that users read to catch the main idea of the research before reading other contents in the body of a paper [8].

The number of data collected varies according to availability on PubMed servers and result of data cleaning processing. The data mapping can be seen in Figure 4.



The data in the form of titles and abstracts were collected, starting from the smallest to the most significant number in the sequence are as follows: 1,062 documents of skin cancer, 1,404 documents of bone cancer, 1,805 documents of papillary thyroid cancer, 1,872 documents of gastric cancer, and 3,116 documents of kidney cancer.

C. Text Preprocessing

The stage of text pre-processing starts from extracting Abstract Title and Abstract Text, then converting xml files into csv files. The sample code and result are presented in Figures 5 and 6 below:



Recall

It less Donoregolation of Codd inhibits town-rigonesis of networcome in view and in views.

Retreat: Ortonorcome is the most frequent engligant home cancer in terrupors. In this study, coll division cycle 6 (Cod on a spregulated in the tune tissues from pattern with networcome according to the resists of DM separating. Cod on a separation by such food frequent the supposts and coll (yellow arms) in colorocome remain architect.

This is considered to the suppost and coll (yellow arms) in colorocomes remain architect.

This (food you mulgasis effect and mechanism of closhofugin on rest with home concer paid).

unting with (CC-D), fine cromeric and transmull assay, respectively.

If the (CC-D), fine cromeric and transmull assay, respectively.

If the (CC-D) is the control of the

Figure 5. The Process of Data Extracting

Data in CSV Format		•	
A CAN DE LO COMPANION DE LA CO	1 - 1 - 1	Y.	(8)
1 title attetract carcertage			
Downings/ation of Citis inhibits tumorigenesis of estessarcona in vive and in vitre, "Our findings indicated that Citis may serve as an oncogene in extensarcona. Thus, inhibiting Citis may be:	a therapes	PER NO	med
1 Study on analigesis effect and mechanism of circolutagin on rats with bone cancer paintbone			
[Relationship between expression of high mobility group box 1 and inflammatory systemes in patients with bone concerpantbone			
5 Incidence and mortality of bone cancer among children, adolescents and young adults of Brazil, "Octeopartoma and living Sarcoma are the most incident bone cancers in all Brazilian regions. But			
 Suppression of WAX3-SPAX/DSE, attenuates bone cancer pain by regulating WXX21 and KCC2, Cur preliminary experiment indicated the activation of with noticinal kinases 1 (WWXI) in bone or 			
A faile for NP IRE in Organ Specific Cancer and Cancer Item Cells, "Concer dem cells (CSCS) account for sumor institution, invasiveness, metactator, and recurrence in a broad range of human can	DWIL Altho	sigh be	ogs.
1250 - Meditsplitsphosphosets unked Devandrion Bone Targeted Strategy for Treatment of Ostecuarcoma, "To reply to as wit unnert medical needs to treat extrementary, a form of primary	bone cano	65, 100	oros
Engineering oracidismic metaltiases to delineate the adactive response of andropen-deprived produte cancer in the bone metaltatic responsement. While strongly interactions are examined	stal in can	cerado	oten
Currynin Researcy (ggpfell Derived Carbonaled Apatite Nanocemen for Combined Anti-Caroer, Anti-Inflammators and Bone Regenerative Therapy, Some pancer or cotectarcome is an aggre-	tive carce	refer	Sept.
Therein to not maturation and part (RCMX) reduces bending of anti-RCMX and book to multiple requires and in, "Book" instruction and gar (RCMX), a former record factor records (RCMX) from the	nember is	named.	-
Mesparany superprinting the frequency of the present study of the present study. The present study is an use to develop a n	amato re	MATERIAL PROPERTY.	
Triple threat activity of PEOF in bone tumors. Tumor inhibition, tissue presentation and cardiopertection against absocubate, "Eigenest epithelium derived factor (PEOF) is known for its option	manus menne	erfies.	200
4 Transcript this one of the most common growth factors in the blood 6M-CSI is also involved in bone parez bone	-		
Simple exocution large area maging of nationals observe and mineralization of a sciencing intercurrence in lunar time. "Oderoacons in the most connect primary bone cancer type in hur	tant First	-	السند
# The Endocumathroot; Endocumiled Section in Bone: From Consequence to Orthogoroom, "Bone is a dynamic Stock, whose homeostasis is maintained by a fine Islance between caterolast IS			
T After 1955 amening they been passed past in rate via ministring pill mediated ghall not activation and required partnerships in the spinal domail from Jones pass (ACP) professories compromise			
I moderno and mortality of primary store cancer in China, 2016. This population based duly assert this most recently available estimates on primary born operand in China, 2016. This population based duly assert this most recently available estimates on primary born operand in China, asserting that it is not recently available estimates on primary born operand in China.			
2 Zetfilly optimization in some with career from premaration to contract the same part of the same and the same received and the same and the same received and the same receive			
 employ operation in some vert cards in the preferation to consideration, including the same deposit and statement of same values and expenses and improve patients. Distributional production in the same possible to promoting MOR and operation of local operation to both power patients and interesting to the promoting and interesting to the production of the produc			
(I) (I) this continue that we continue the continue to the con			
(2) Hydrophobicity may enthance membrane althinky and with usonal effects of Schill base remoduliny; observables complemes, factor at 4 alters of senseburic compounds have generated not	en etere	200	me eg.

Figure 6. The Data in CSV Format

The output documents are then provided for the next phase in text classification that includes:

Tokenization: Befo 4 partitioning into the token list, the document is treated as a string in advance.

Removing stop words: The process to remove the insignificant words or stop words such as "the", "a", "and", etc.

Stemming word: This sold is the process of conflating tokens to their root form. Converts different word forms into a similar canonical form can be done by applying the stemming algorithm, e.g. change the word "Classification" and "Classified" into the word "Classify".

The related program codes of this step are depicted in Figure 7. The next step can be seen in Figure 8, 11d it is responsible for inversing term/document freq 11ncy (TF*IDF). The TF-IDF has been widely employed to evaluate 11e relationship of each word in the collection of documents in the fields of information retrieval and text mining.

D. Classification

The classification will be applied in three treatments with three methods: Neural Network, Naïve Bayes, and Linear Classifier with SGD training. Moreover, in the following subsections, we used a diverse number to indexes each cancer literature, (1) for Bone Cancer, (2) for Gastric Cancer, (3) for Kidney Cancer, (4) for Papillary Thyroid Cancer, and (5) for Skin Cancer, respectively.



Figure 7. Code of Data Removing



Figure 8. TF*IDF

At the stage of classification, the researchers divided the data into two parts with different proportions. Specifically, the proportions of training dataset and testing data are 80% and 20% (called 0.2 for short), 70% and 30% (0.3), and 60% and 40% (0.4), respectively. Figures 9 \sim 14 show the results for each method's model accuracy and confusion matrix form in 0.3 proportion, while the complete result of accuracy is presented in Table 1.

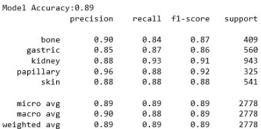


Figure 9. Model Accuracy of Neural Network with Proportion 0.3

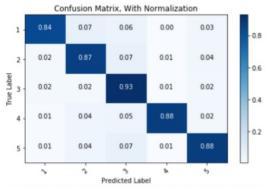


Figure 10. Confusion Matrix of Neural Network with Proportion 0.3

y:0.89			
precision	recall	f1-score	support
0.89	0.83	0.86	409
0.86	0.86	0.86	560
0.88	0.93	0.91	943
0.97	0.87	0.92	325
0.87	0.88	0.88	541
0.89	0.89	0.89	2778
0.89	0.88	0.88	2778
0.89	0.89	0.89	2778
	0.89 0.86 0.88 0.97 0.87	precision recall 0.89 0.83 0.86 0.86 0.88 0.93 0.97 0.87 0.87 0.88 0.89 0.89 0.89 0.89	precision recall f1-score 0.89 0.83 0.86 0.86 0.86 0.86 0.88 0.93 0.91 0.97 0.87 0.92 0.87 0.88 0.88 0.89 0.89 0.89 0.89 0.88 0.88

Figure 11. Model Accuracy of Naïve Bayes with Proportion 0.3

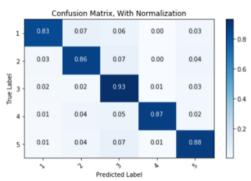


Figure 12. Confusion Matrix of Naïve Bayes with Proportion 0.3

Model Accurac	y:0.89			
	precision	recall	f1-score	support
bone	0.91	0.83	0.87	409
gastric	0.86	0.87	0.86	560
kidney	0.88	0.93	0.90	943
papillary	0.96	0.88	0.92	325
skin	0.88	0.88	0.88	541
micro avg	0.89	0.89	0.89	2778
macro avg	0.90	0.88	0.89	2778
weighted avg	0.89	0.89	0.89	2778

Figure 13. Model Accuracy of Linear Classifier with Proportion 0.3

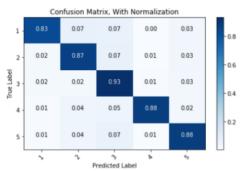


Figure 14. Confusion Matrix of Linear Classifier with Proportion 0.3

TABLE I. RESULT OF ACCURACY FOR 9,259 INSTANCES

Markend		Testing Dataset	
Mehtod Naive Beyes Neural Network	0.2	0.3	0.4
Naive Beyes	0.89	0.89	0.89
Neural Network	0.89	0.89	0.89
Linear Classification	0.89	0.89	0.89

III. EXPERIMENTAL RESULTS

A. The Analysis of Accuracy Result

After going through the stages of determining the classification scheme, reading text documents, text preprocessing and classification, the results for each condition, i.e., 0.2, 0.3, and 0.4, are obtained as shown in Table 2 below:

TABLE II. RESULT OF ACCURACY FOR HUNDREDS INSTANCES

Mr. Land	instang Dataset					
Mehtod	0.2	0.3	0.4			
Naive Beyes	0.90	0.97	0.93			
Neural Network	0.90	0.97	0.93			
Linear Classification	0.90	0.93	0.90			

Each method manifests the same performance for an accuracy result of 0.89/1 but produces different confusion matrix that will explain further in the analysis of confusion matric section. This condition presents the initial hypothesis that more number of instances result in more stable accuracy. The researchers also reclassified documents with fewer data to examine the performance of each algorithm. As shown in Table 2, the result of accuracy shows more vary for each method when using hundreds of data instances.

The data describe all methods have the same accuracy performance when using 20% testing dataset, each 0.9/1. While Naïve Bayes and Neural Network present higher performance than Linear Classification when using 30% and 40% testing dataset, each respectively at 0.97/1 and 0.93/1.

B. The Analysis of Confusion Matrix

The result of the confusion matrix is presented in Table 3 below. In this confusion matrix, the system manages to classify each cancer literature with average accuracy 0.84/1 for Bone cancer, Gastric cancer, Papillary cancer and Skin cancer. In contrast, Kidney cancer gets the highest accuracy on average 0.93/1. This condition is highly possible because of a more significant amount of kidney dataset than other cancer literature. For reference, Kidney cancer literature has 3116 datasets, whereas four other types of cancer literature have approximately only a thousand dataset.

TABLE III. RESULT OF CONFUSION MATRIX Cancer 0.2 0.3 0.4 NB NB LC Type NN NN NN 0.83 0.83 0.83 0.83 0.84 0.83 0.83 0.83 0.83 0.87 0.84 0.86 0.88 0.89 0.87 0.86 0.93 0.87 0.87 0.89 0.87 88.0 0.89 0.87 0.88 0.89 Papillary 0.88 0.87 0.89 0.88 0.87 0.89 88.0 0.87

Figure 15 shows a comparison of the confusion matrix of three types of the algorithm more precisely with the proportion of 0.3 testing data.

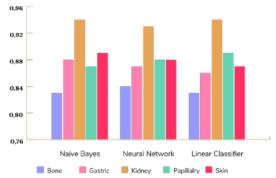


Figure 15. Comparison of Accuracy Performance (0.3)

C. The Analysis of Classification Result

Table 4 depicts information about the mapping of correct and incorrect instances. The data shows incorrect instances for the literature of Gastric cancer, Papillary Thyroid Cancer and Skin Cancer tend to kidney cancer literature in the number of 0.07/1. In addition, bone cancer literature tends to map inaccurate cases for Gastric cancer as well as Kidney cancer in the number of 0.07/1. On the other hand, incorrect instances of Kidney cancer literature itself tend to skin cancer with the highest missed classification in the number of 0.03/1.

TABLE IV. RESULT OF INCORRECTLY INSTANCES

	Euro	Gestric	Iddney	Papillary	Skin	Bone	Gentre	Knincy	Papillary	Skin	Bone	Capting	Kidney	Papillary	Skin
Bone	0.83	0.07	0.06	0.00	0.03	0.83	0.07	0.07	0.00	0:03	0.83	0.06	80.0	0.00	0.04
Gastric	0.02	0.88	0.06	0.00	0.03	0.02	0.87	0.07	0.01	0.03	0.02	0.86	0.07	0.01	0.04
Kidney	0.02	0.01	0.94	0.00	0.02	0.02	0.02	0.93	0.01	0:03	0.02	0.01	0.93	0.01	0.02
Paillary	0.00	0.06	0.05	0.87	0.02	0.01	0.04	0.05	0.88	0.02	0.00	0.03	0.05	0.89	0.02
Skin	0.00	0.04	0.07	0.00	0.89	0.01	0.04	0.07	0.01	0.88	0.01	0.04	0.07	0.00	0.87
	Correc	tiy Instan	ces												
	The M	ost incom	octly Inst	arves											

The data above were obtained from the confusion matrix summary of the neural network, with sequential proportions of 0.2, 0.3 and 0.4. Since it is considered as the best algorithm for classification of cancer literature.

IV. CONCLUSION

Based on confusion matrix analysis, the result indicates that type of cancer literature with the most significant number of instances will show the highest accuracy, that is Kidney Cancer literature and followed by other types of cancer literature based on their number of instances. Meanwhile, the mapping of incorrect instances confirms that the highest missed classification of each cancer literature tend to Kidney cancer while Papillary Cancer literature becomes the lowest number of missed classification.

Incorrect instances can be prompted by the relationship between two different cancer types, for example, incorrect literature interpretation of Bone Cancer and Gastric Cancer caused by relational of the research topic of both cancer, that usually Gastric Cancer initially presenting as bone metastasis.

Furthermore, the disambiguation classification of cancer type can also be provoked by strong association from each cancer type, as shown by Kidney Cancer. Considering the treatment for cancer survivor will use a considerable amount of medicine that will harm the patient's liver, Kidney Cancer will be the most common synchronous clinical effect in patients. This case vice versa with Papillary Thyroid cancer, wherein only $0.00 \sim 0.01$ will likely to be wrong mapped as Papillary Thyroid because of a weak relationship between

papillary cancer with other cancer topics, in this case, Bone Cancer, Gastric Cancer, Kidney Cancer and Skin Cancer.

The results of this study showed that all algorithms successfully could be used to classify cancer literature. However, for the best performance, it is strongly recommended to use Naïve Bayes and Neural Network methods by dividing 70% training dataset and 30% testing dataset.

REFERENCES

- American Cancer Society, https://www.cancer.org/cancer/bone-cancer.html
- [2] Colas, F. and Brazdil, P. (2006). Comparison of SVM and some older classification algorithms in text classification tasks. In: Bramer, M. (Ed.), Artificial Intelligence in Theory and Practice. Springer, Boston, MA, pp. 169-178.
- [3] Fei Shu, Charles-Antoine Julien, Lin Zhang, Junping Qiu, and Vincent Larivière. (2019). Comparing journal and paper level classifications of science, <u>Journal of Informetrics</u>Volume 13, Issue 1 February 2019 Pages 202-225
- [4] Fleuren, W.W.M. and Alkema, W., (2015). Application of text mining in the biomedical domain. Methods 74, 97-106. https://doi.org/10.1016/j.ymeth.2015.01.015
- [5] Mowafy M, Rezk A, and El-bakry HM, An Efficient Classification Model for Unstructured Text Document, American Journal of Computer Science and Information Technology ISSN 2349-3917
- [6] Peng Jin, Yue Zhang, Xingyuan Chen, and Yunqing Xia, Bag-of-Embeddings for Text Classification, Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (JICAI-16)
- [7] PubMed.gov, https://www.ncbi.nlm.nih.gov/pubmed/
- [8] Sang-Woon Kim and Joon-Min Gil, (2019), Research Paper Classification System Based on TF-IDF and LDA Scheme, Human-centric Computing and Information Sciences volume 9, Article number: 30.
- [9] Scikit Learn, https://scikit-learn.org/stable/modules/sgd.html
- [10] S, Snezhana, Using Text Mining To Classify Research Paper, https://www.researchgate.net/publication/319688772_Using text mining to classify research papers.
- [11] StackOverflow, https://stackoverflow.com/questions/
- [12] World Health Organization https://www.who.int/health-topics/cancer#tab=tab_1
- [13] Zhang, L., Janssens, F., Liang, L., and Glänzel, W. (2010). Journal cross-citation analysis for validation and improvement of journalbased subject classification in bibliometric research. Scientometrics, 82(3), 687-706.
- [14] Zhang, Y., Lin, H., Yang, Z., Wang, J., Sun, Y., Xu, B., and Zhao, Z. (2019). Neural network-based approaches for biomedical relation classification: A review. Journal of Biomedical Informatics, 103294. doi:10.1016/j.jbi.2019.103294

Article

原	創性報告				
	15% 似度指數	9%網際網絡來源	12 % 出版物	8 % 學生文稿	
主	要來源				
	hcis-jou 網際網絡來源	urnal.springerop ^ந	en.com		3%
	Wang, "Neural	nang, Hongfei Li Yuanyuan Sun, I network-based lical relation clas I of Biomedical I	Bo Xu, Zhehu approaches f ssification: A re	ian Zhao. or eview",	2%
	3 WWW.CC 網際網絡來源	omsis.org			1%
	4 Submit _{學生文稿}	ted to Macquari	e University		1%
	Zhang, Human	Ye, Xin Li, Xiar Wu Chen. "Dee Activity Real-Ti rian Navigation"	ep Learning-Baime Recognition	ased on for	1%

6 www.azherbs.com 網際網絡來源

1 %

7	ogrisel.github.io 網際網絡來源	1%
8	Fei Shu, Charles-Antoine Julien, Lin Zhang, Junping Qiu, Jing Zhang, Vincent Larivière. "Comparing journal and paper level classifications of science", Journal of Informetrics, 2019	1%
9	www.mdpi.com 網際網絡來源	1%
10	ijsetr.org 網際網絡來源	1%
11	repository.tudelft.nl 網際網絡來源	1%
12	Pang, Clifford. "Hyperthermia in oncology and nontoxic integrative treatments", Hyperthermia in Oncology, 2015.	1%
13	Submitted to Universiti Teknologi Petronas 學生文稿	<1%
14	Submitted to American Public University System 學生文稿	<1%
15	"Sustainable Agriculture Reviews 43", Springer Science and Business Media LLC, 2020 ^{出版物}	<1%

16	methadonenearme.com 網際網絡來源	<1%
17	trepo.tuni.fi 網際網絡來源	<1%
18	Submitted to University of Southampton 學生文稿	<1%
19	Submitted to Mahidol University _{學生文稿}	<1%
20	Submitted to Sunway Education Group 學生文稿	<1%
21	Lecture Notes in Computer Science, 2009.	<1%
22	Sang-Woon Kim, Joon-Min Gil. "Research paper classification systems based on TF-IDF and LDA schemes", Human-centric Computing and Information Sciences, 2019	<1%

 排除引述
 開

 排除參考書目
 開

排除相符處

關閉