

CiteOpinion: Evidence-based Evaluation Tool for Academic Contributions of Research Papers Based on Citing Sentences

Xiaoqiu Le^{1†}, Jingdan Chu², Siyi Deng², Qihang Jiao²,
Jingjing Pei², Liya Zhu¹, Junliang Yao²

¹National Science Library, Chinese Academy of Sciences, Beijing 100190, China

²University of Chinese Academy of Sciences, Beijing 100190, China

Citation: Xiaoqiu Le, Jingdan Chu, Siyi Deng, Qihang Jiao, Jingjing Pei, Liya Zhu, Junliang Yao (2019). CiteOpinion: Evidence-based evaluation tool for academic contributions of research papers based on citing sentences. *Journal of Data and Information Science*, 4(4), 26–41

DOI: 10.2478/jdis-2019-0019

Received: Nov. 8, 2019

Revised: Nov. 22, 2019

Accepted: Nov. 28, 2019

Abstract

Purpose: To uncover the evaluation information on the academic contribution of research papers cited by peers based on the content cited by citing papers, and to provide an evidence-based tool for evaluating the academic value of cited papers.

Design/methodology/approach: CiteOpinion uses a deep learning model to automatically extract citing sentences from representative citing papers; it starts with an analysis on the citing sentences, then it identifies major academic contribution points of the cited paper, positive/negative evaluations from citing authors and the changes in the subjects of subsequent citing authors by means of Recognizing Categories of Moves (problems, methods, conclusions, etc.), and sentiment analysis and topic clustering.

Findings: Citing sentences in a citing paper contain substantial evidences useful for academic evaluation. They can also be used to objectively and authentically reveal the nature and degree of contribution of the cited paper reflected by citation, beyond simple citation statistics.

Practical implications: The evidence-based evaluation tool CiteOpinion can provide an objective and in-depth academic value evaluation basis for the representative papers of scientific researchers, research teams, and institutions.

Originality/value: No other similar practical tool is found in papers retrieved.

Research limitations: There are difficulties in acquiring full text of citing papers. There is a need to refine the calculation based on the sentiment scores of citing sentences. Currently, the tool is only used for academic contribution evaluation, while its value in policy studies, technical application, and promotion of science is not yet tested.

Keywords Cited paper; Citing paper; Citing sentence; Citation motive; Citation sentiment; Academic contribution; Evaluation



1 Introduction

San Francisco Declaration on Research Assessment^① states in 2013 that the Journal Impact Factor has a number of well-documented deficiencies as a tool for research assessment, so it calls for a broad range of measures to objectively evaluate the academic value of research papers.

Number of citation received by a paper is often used to represent the academic value of the paper, but it is widely known that citing behaviors vary greatly between different fields, in addition to the inherent problems with using citations to evaluate the contributions focusing on policy study, or technology application, or promotion of science. Many institutions then utilize some normalized citation impact such as CNCI (Category Normalized Citation Impact) of Clarivate Analytics^② and SNIP (Source Normalized Impact per Paper) of Scopus^③, where citation impact is represented by the ratio of the citation frequency of the author's articles to the citation frequency of articles in the specific research field. Some use h-index to evaluate the research impact of an author or a team. However, there are still fundamental challenges in these citation-based indicators: A citation may be made because of different motives, and it can be positive or negative. In most cases, number of citations alone will be difficult to accurately measure the academic contribution of a research paper.

Some institutions today have begun adopting a peer review mechanism based on representative works in person, team, and project evaluation. There, representative works (for example, representative papers provided by the authors or selected by third parties) are thoroughly analyzed and evaluated by peer specialists. This method works in avoiding the biases inherent in citation indicators, but it is time-intensive, and often, is difficult to find relevant specialists who are not related to those being evaluated. Besides, peer reviewers are prone to rely on, knowingly or implicitly, the most accessible but maybe misleading statistical indicators under time pressure. Therefore, it becomes an important theoretic and practical need to explore the feasibility to extract and analyze the rich information hidden in the citing behaviors and citing context to be used objectively and conveniently in evaluation for the academic value of the research papers and the authors or the team authored the papers.

^① San Francisco Declaration on Research Assessment <https://sfдора.org/read/>

^② Category Normalized Citation Impact. <http://help.incites.clarivate.com/inCites2Live/indicatorsGroup/aboutHandbook/usingCitationIndicatorsWisely/normalizedCitationImpact.html>

^③ How is SNIP (Source Normalized Impact per Paper) used in Scopus? https://service.elsevier.com/app/answers/detail/a_id/14884/supporthub/scopus/kw/snip/



In addition, given the rich dynamics of research, the complexity of academic contribution, and the invaluable insights of the experts, it is necessary to combine expert peer review with objective evidence-based analysis. When peer reviewers find difficulty in accurately deciding or reaching consensus on the impact of an academic result, it will be beneficial to use an evidence-based analysis tool to present factual evidence or analytic summaries out of the large pool of citing information, thus providing the reviewers with evidence-based references.

2 Theoretic framework for citation content-targeted evaluation

Citation content refers to the sentences citing the cited paper in a citing paper (Liu et al., 2015). This paper considers citation content as the anchor text in a citing paper linking to the cited paper (Fig. 1). It is called citing sentence or citation context (Lei et al., 2016).

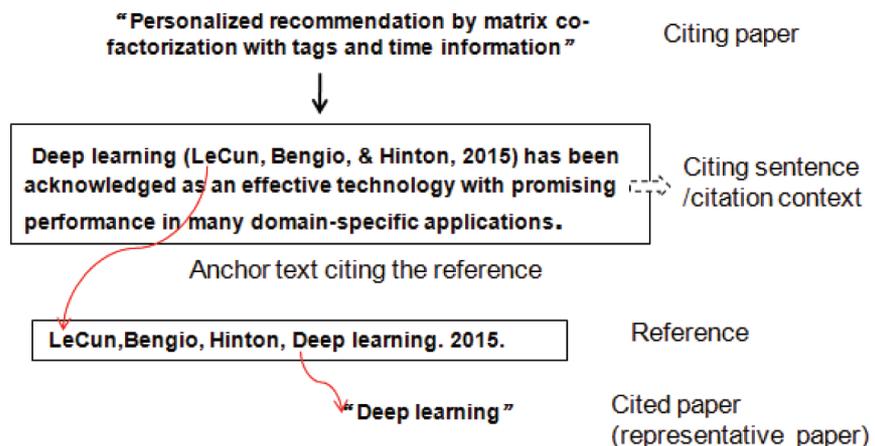


Figure 1. Diagram for Concept of Citation Content.

There are studies on conducting academic evaluation from the perspective of citation content. They can be roughly summarized into three approaches:

First is to evaluate the actual impact on the citing paper from the cited paper based on citation motive analysis. There are various reasons that scientific papers cite other papers. In 1964, Garfield (1964) systematically illustrated 15 citation reasons when studying Autonomous Citation Indexing. Brooks (1985) proposed 7 types of citation motives: Currency scale, negative credit, operational information, persuasiveness, positive credit, reader alert and social consensus. There are many other similar categorizations. For example, Some (Ma et al., 2009) think that citation motive is a citation intention hidden in the citing sentence and can be only confirmed



by the original author, thus it is necessary to obtain feedback data from the citing authors to evaluate the cited paper. However, the method is inoperable in large-scale evaluation. In this case, automatic identification becomes another important approach, which uses natural language processing technology to identify the category of the citing sentence based on clue words and sentence structural rules.

The second approach is to conduct sentiment classification on citing sentences. That means to assess the peer reviewers' evaluation on the paper according to the distribution of citing sentences in different sentiment categories. Small (2011) proposed to analyze citation content by expanding the concept of sentiment. He defines the sentiment as the attitude and inclination towards the cited paper and classifies citation content into 11 categories including Importance, Discovered, Achieved, Improved and Applied. Teufel et al. (2006) classified citing sentences into 4 categories by reference to the classification result of Spiegel-RÖsing (1977): ① Explicit statement of weakness; ② Contrast or comparison with other work; ③ Agreement/usage/compatibility with other work; ④ A neutral category. Each category is identified based on clue words. Liu et al. (2015) divided citation content into positive, neutral and negative categories based on Spiegel-RÖsing and Teufel and conducted sentiment classification on citing sentences according to the subjects and clue words of citing sentences.

The third is to study how much the cited paper's points are used by peers in perspective of topic analysis between the citing sentences and the cited paper. Mohammad (2009) studied how to find that citing sentences can play a key role and citation content can reflect the application value of the cited paper by extracting summarized information from the original author's abstract and the citing sentences. Liu et al. (2014) studied the conformity between citing sentences and the abstract of the cited paper and subject terms. They found that a good conformity between citing sentences and the abstract of the cited paper can characterize the content of the cited paper well. Liu et al. (2013) compared the potential topic differences between the abstract of a citing paper and the citing sentences. They found there are clear overlaps in the topics between the two parts and that the topic of the abstract contains more professional terms.

In comparison with the expert-dependent peer review, citation content-targeted academic evaluation has the following strengths:

- 1) The authors of the citing sentences are all citing authors. No matter what their citation motives are, their research areas are mostly relevant to the research area of the cited author. Thus, "peer reviewers in the same or related specialties" are easy to be found;
- 2) The content in citing sentences to be evaluated is published and relatively objective, thus reducing risks on subjective judgment without evidences;



- 3) Citing sentences can support multi-perspective evaluation by peers from various aspects including motives, sentiment and topic evolution. Potential prejudice from a single perspective is thus avoided.

Academic evaluation based on citation content has made some progress, but few analysis tools are yet seen to help peer reviewers make decision in the evaluation process. Hence, this paper developed a series of method explorations and engineering practices. Targeting the given citing papers, the paper achieved the analysis and calculation of key evaluation factors including automatic machine recognition and academic contribution point extraction from citing sentences, citing sentence evaluation analysis, contribution point knowledge distribution and calculation for citing sentence statistical indicators. Based on that, the paper established CiteOpinion, an evidence-based analysis tool for academic evaluation.

3 Technical approach for citation content-targeted evaluation

3.1 Basic analysis logic

The analysis logic of CiteOpinion based on the following assumptions: (1) The contribution and contribution degree of an article can be objectively and qualitatively determined by peer reviewers in the same field who have cited the article; (2) the contribution (representing the academic value to some extent) of an article to its research field can be quantified by analyzing the diffusion network of its core innovation points in the citation process. Therefore, CiteOpinion design targets the citing sentences in citing papers. It reveals the impact of a cited paper in academia by identifying the citation content and citation evaluation of the citing authors, the diffusion pattern and degree of the cited content (contribution points) in the citation network, etc. Thus, the tool provides evidence-based evaluation for the academic contribution of academic papers and assists peer reviewers in making objective decisions.

In comparison with typical quantitative analysis tools (Table 1), CiteOpinion targets citing sentences as the processing object. It starts with a finer aspect that focuses on text content mining, and the analysis result is presented in the form of knowledge units (evaluation evidence) with quantitative indicators.

Table 1. Comparison between CiteOpinion and Conventional Quantitative Analysis Tools.

	Quantitative Analysis Tool	CiteOpinion
Data object	Metadata of paper	Citing sentences
Data granularity	Article	Sentence
Analysis focus	Statistical indicators	Text content mining
Result form	Relationship diagrams and data sheets	Evaluation evidence text, relationship diagrams and data sheets



3.2 Measuring academic contribution of cited papers based on citing sentences

The measurement of the academic contribution of cited papers mainly involves two questions: How to identify the contribution points of a paper and how to measure the contribution degree of a paper. CiteOpinion is designed with an academic contribution measurement system for cited papers based on content analysis and quantitative analysis on citing sentences (Fig. 2). For the identification of contribution points, the tool focuses on the contribution points proposed by the author and those mentioned by peers in citing sentences. By contrastive analysis, the tool can objectively disclose how well the contribution points of a paper are recognized by peers. With regard to academic contribution degree, the tool conducts quantitative analysis and calculation on citing sentence evaluation, citing sentence quantitative analysis and knowledge diffusion.

Citing sentence evaluation mirrors the citing authors' recognition for the achievements of the representative works they cited. Liu et al. (2015) classified citing sentences into 6 types under the three categories (positive, neutral, and negative ones) based on the citation character, but they did not differentiate citation motives from sentiment. This paper divides citing sentences into positive, neutral and negative (Table 2) based on sentiment analysis from the perspective of evaluation: A positive citation refers to citing sentences holding a commendatory, approving and admiring attitude towards a paper's results; a neutral citation refers to citing sentences briefly stating or rephrasing a paper's results without obvious expression of sentiment; a negative citation refers to citing sentences suggesting the defects, shortcomings or mistakes of a paper's results. Citing sentences are herein classified by calculating sentiment scores. The calculation formula for the sentiment score is:

Table 2. Sentiment categorization of citing sentences.

Sentiment Category	Definition	Sentiment Score Range (E)
Positive	Holding commendatory, approving and admiring attitude	1>E>0
Neutral	Brief statement or rephrasing, without obvious expression of sentiment	E=0
Negative	Describing defects, shortcomings or mistakes	0>E>-1

$$E = r \left(W_s + \sum_{i=1}^m (W_{li}) \right) \tag{1}$$

Wherein W_s represents the weight of the sentence structure, while W_l refers to the weight of sentiment words. Each weight has positive and negative scores at the range of 0-1, and $|W_s| > |W_l|$. m represents the number of sentiment words, and r is the adjustment coefficient.



The quantitative analysis indicators for citing sentences correspond to conventional article-based quantitative indicators and cover the total number of citing sentences, the number of non-self-citing sentences (at home/abroad), the number of self-citing sentences, the number of citing sentences in JCR ranking (Q1/Q2/Q3/Q4), etc. Besides, this paper adds in the number of core citing papers (i.e. the number of citing sentences in a single citing paper is ≥ 2) and the number of citing sentences in Cell, Nature and Science (CNS). The former is used to measure major following studies based on the results of Representative Works. Although the latter contains Q1 ranking, in practice, many authors or evaluators of Representative Works still pay more attention to the evidences that a paper is cited by the CNS. These indexes indirectly measure the academic contributions of Representative Works from different perspectives and can be used to indirectly measure the recognition degree among peers and impact of papers.

In order to measure the impact of the academic contribution points of cited papers on the discipline advancement, CiteOpinion conducts calculation and disclosure on intra-disciplinary knowledge transfer and inter-disciplinary knowledge transfer. Intra-disciplinary knowledge transfer mainly analyzes the transfer of contribution points to other research areas in the citation network of the current discipline and the evolution of the newly generated study topics; inter-disciplinary knowledge transfer reveals the situation where the contribution points are introduced and put in extensive application by other disciplines. They are measured by disciplinary knowledge transfer rates T_i and T_m , respectively which are calculated as per:

$$T_i = N_{cl}/N_{cl} \times 100\%$$

$$T_m = N_{cs}/N_{ts} \times 100\%$$

Wherein N_{cl} represents the number of the leaf nodes of the citing papers that are in the same discipline category as the cited paper; N_{cl} is the total number of the leaf nodes of the cited paper's discipline category; N_{cs} is the number of the first-level disciplines of the citing paper; and N_{ts} is the total number of first-level disciplines.

Based on the calculation and analysis results for the above measuring indexes, CiteOpinion established a series of evidence-based services used for academic contribution to provide evidence for manual work on academic evaluation. The major evidences are: Representative Works' academic contribution points, highlight citing sentences, negative citing sentences, originality evaluation, contribution points cited by peers, statistical indicators for citing sentences, diagram for disciplinary knowledge transfer of contribution points, etc.



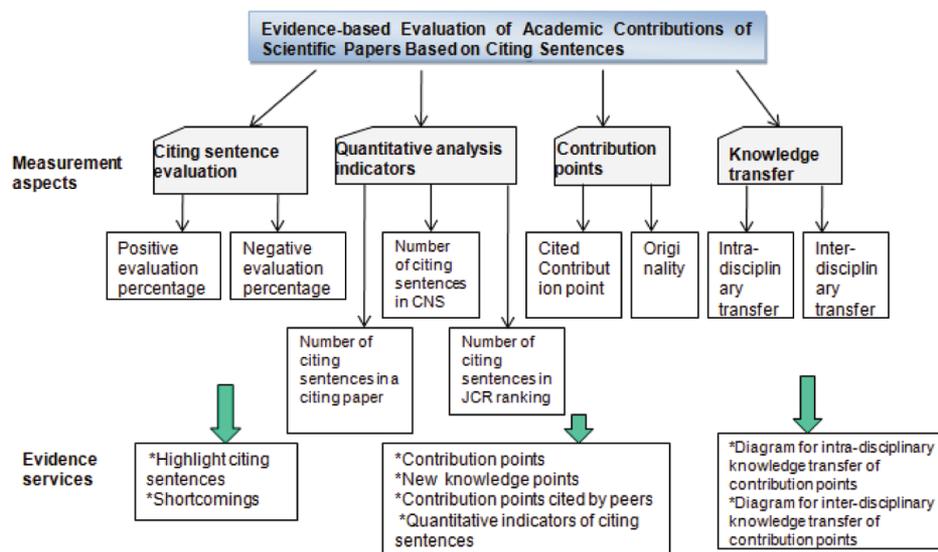


Figure 2. Measurement framework for academic contribution of representative paper.

4 Implementation scheme of citation content-targeted evaluation

4.1 System framework

CiteOpinion focuses on analysis and recognition of citing sentences from representative papers. The system comprises the recognition of citing sentences, academic evaluation knowledge units mining, indicators computation generation of evidence-based analysis reports (Fig. 3). Recognition of citing sentences is based on the full text of citing papers. It will build a citing sentence database from representative works by recognizing citing sentences from citing papers, and then mining the knowledge units in evaluation of academic contributions based on the citing sentences and full text. Knowledge units mainly include academic contribution points, problems, method sentences, concluding sentences, original sentences and subject terms, which can be used for subsequently calculating evaluation indexes of academic contributions and analyzing knowledge transfer. In terms of evidence-based evaluation, the system provides automatic and manual selection of analysis results and finally generates analysis reports on evidence-based evaluation of academic contributions of representative works.

4.2 Automatic recognition of citing sentences

Considering the mode of citation, citing sentences in journal articles often have normative marks, involving author-year or superscript numbers at the end of



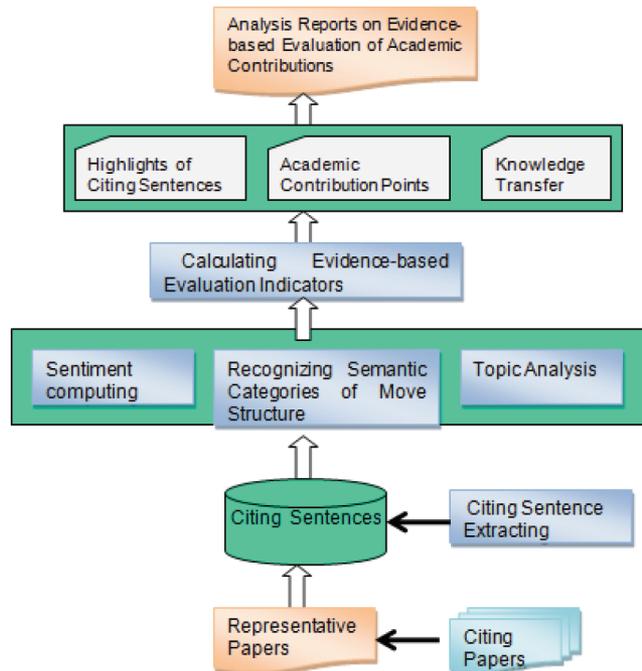


Figure 3. CiteOpinion system framework.

sentences. It is easy to recognize citing sentences from journal articles stored in structured or semi-structured forms (such as web pages). However, it is difficult to recognize citing sentences from scientific papers given that most of them are now saved in unstructured PDF files.

In the preliminary study, this paper designed a recognition model for text blocks with parallel relationship (Pei et al., 2019) to recognize text blocks in PDF files sharing parallel relationship. This can be used to identify references in citing papers and then recognize citing sentences from the full text. The model utilizes visual information of the page layout for vector representation of paragraphs. It relies on one-layer Autoencoder for dimensionality reduction, combines with character vectors and style vectors of paragraphs, and uses the CNN (Convolutional Neural Network) classification model to get text blocks of references. The precision of reference recognition is 90.1% and that of citing sentence recognition is 92%.

4.3 Calculating sentiment scores of citing sentences

The calculation formula of the traditional dictionary-based sentiment analysis for citing sentences is as follows. The process is mainly backed by pattern matching via syntactic parsing results and sentiment dictionaries. Sentiment dictionaries have



two levels: sentences and words, of which each is given a respective weight of w_s and w_i . Basic algorithm is:

- 1) perform data cleaning against all citing sentences of the representative works to remove disturbance terms (such as inaccurate segmentation and unreadable characters);
- 2) utilize the Stanford CoreNLP Tools to tokenize, Part-of-speech tagging and parsing;
- 3) perform sentence pattern matching against the marked citing sentences. If it succeeds, then $w_s > E$;
- 4) match sentiment words and calculate the frequency (f). If it succeeds, then $E = E + fw_i$;
- 5) normalize the sentiment score E.

4.4 Recognizing academic contribution points of papers

Academic contribution points refer to the innovative research results stated in scientific papers, such as new knowledge, resolved difficult problems and new application models. Contribution point recognition is to identify paragraphs containing them. CiteOpinion can complete this process under the following two circumstances:

- 1) If journal articles have standard descriptions about contribution points, first, the above-mentioned model (Pei et al., 2019) can be used to identify text blocks with parallel relationship and next, some rules will be used to identify whether such text blocks are contribution points.
- 2) For papers having no standard description about contribution points, semantic annotation for move categories will be performed against all sentences. Currently, three move categories, namely problems, method sentences and concluding sentences, can be recognized by multiple dichotomous models (Fig. 4). Based on ULMFiT (Howard et al., 2018) pre-trained language model, the dichotomous models can get word embeddings suitable for scientific papers by fine-tuning on a great deal of scientific text for language model training. Based on the word embeddings, input few labeled data for dichotomous supervised learning. The learning process adopts the fine-tuning mechanism: the structure and parameters of language model are followed; training parameters is unfreezed layer by layer; the learning rate follows triangle changing. Recognition results (positive precision) are 89% for problems, 88% for method sentences and 91% for concluding sentences. Contribution points are identified on the basis of the three types of sentences.



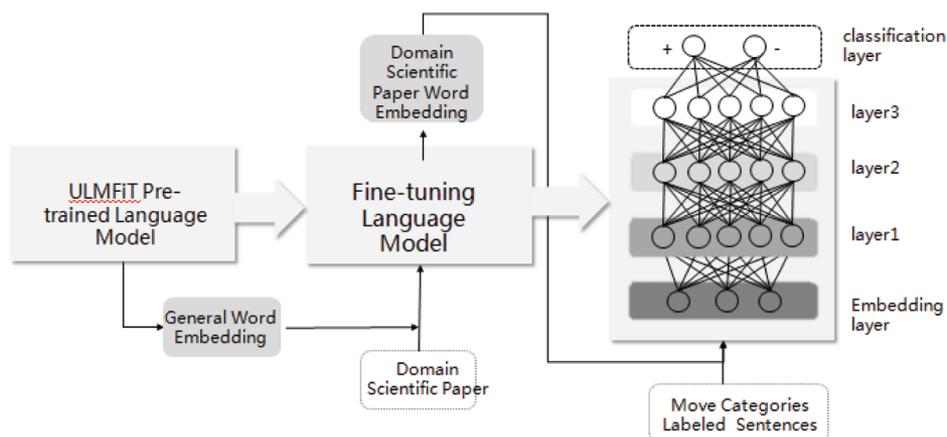


Figure 4. Move structure category recognition model.

4.5 Knowledge transfer of contribution points

To reveal content changes when contribution points transfer among disciplines, CiteOpinion extracts keywords of papers based on metadata (titles, abstracts, and keywords) of the representative works' citing papers and calculate changes of subject terms within a certain period. Integrated with indexes such as core citing papers and move categories of citing sentences, researchers can find proofs that achievements of the representative works stimulate follow-up researches.

5 A case study—analysis on citing sentences in representative works of A.M. Turing Award laureates

5.1 Selection of representative works

Take academic papers published by Geoffrey Hinton, 2018 A.M. Turing Award laureate, to explain processing flow and results of CiteOpinion. Search for all papers of Geoffrey Hinton in the Web of Science (WoS) and Scopus; display these papers in reverse order based on the cited counts; select five out of the top 10 papers that match with contribution points stated in the citation (Table 3). The following explanation uses Learning Representations by Back-propagating Errors as an example.

Processing Flow:

- 1) Acquire citing papers of representative works in WoS or Scopus;
- 2) Acquire full texts of citing papers;
- 3) Analyze the full texts and recognize citing sentences from them;
- 4) Analyze the contents of citing sentences;
- 5) Develop the evidence-based report.



Table 3. Representative papers of Geoffrey Hinton.

#	Author	Representative Papers	Citing Papers
1	Hinton Geoffrey	Deep learning	7,630
2	Hinton Geoffrey	Reducing the dimensionality of data with neural networks	4,509
3	Hinton Geoffrey	A fast learning algorithm for deep belief nets	4,538
4	Hinton Geoffrey	Learning representations by back propagating errors	6,598
5	Hinton Geoffrey	Dropout: A Simple Way to Prevent Neural Networks from Overfitting	4,142

5.2 Evidence-based results analysis

1) Highlights of citing sentences

In Web of Science Core Collection, this representative work of Hinton has 6,981 citing papers and we select 3,159 citing sentences from 2,789 pieces of full texts. After calculating the respective sentiment score of them, we get 998 (32%) positive, 2,076 (65%) neutral and 85 (3%) negative citing sentences. 282 out of 998 positive ones score greater than 0.9, showing the tendency of obvious approval and high praise. CiteOpinion selects highlight evaluation sentences from these citing sentences that enjoy high sentiment scores. Follow-up researchers speak highly of the improved BP algorithm proposed by Hinton when citing this representative work. Six citing papers regard it as a breakthrough (Table 4); four think it is a major improvement/new finding; 92 praise it as being important/popular/widely used.

Table 4. Highlights of citing sentences praising the contribution point as a breakthrough.

	Highlights of Citing Sentence	Titles of Citing Paper
1	The steepest descent algorithm, also known as the error backpropagation (EBP) algorithm [8,9], dispersed the dark clouds on the field of artificial neural networks and could be regarded as one of the most significant breakthroughs for training neural networks.	Application of Neural Networks to Automatic Load Frequency Control
2	In addition to the development of new ANN algorithms that were more neural-inspired (e.g. Hopfield networks), another major breakthrough that helped lead to a resurgence in neural network research was the rediscovery of the backpropagation technique (LeCun, 1985, Rumelhart et al., 1986, Werbos, 1990).	A historical survey of algorithms and hardware architectures for neural-inspired and neuromorphic computing applications
3	The next major breakthrough happened in late 80s with the invention of back-propagation and a gradient-based optimization algorithm to train a neural network with one or two hidden layers with any desired number of nodes (Rumelhart et al., 1986).	Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines



2) Academic contribution points

Recognize the contribution points of the original representative work's abstract and that of the citing sentences. The results are shown in Table 5. The contribution

point of the original paper is “we describe a new learning procedure, back-propagation, for networks of neurone-like units.” The citing author developed his/her contribution points as ① the learning rate of BP algorithm in neural network training, which accounts for 44% of the total citing sentences. ② Central mechanism including activation function and hidden layer(s) of BP algorithm, which accounts for about 42% of the total citing sentences. ③ The structure of neural network and feedforward multilayer perceptron, which accounts for about 11% of the total citing sentences. After comparing the contribution points of the representative work and citing sentences, we find that follow-up researchers pay more attention to central mechanisms of the representative work’s contribution points, efficiency in application and basic structure. They will deeply explore the essence, instead of simply copying such contribution points. Owing to that, peer reviewers can also understand the actual contributions of the representative work recognized by counterparts.

Table 5. Comparison of contribution points of representative work and citing sentences.

Contribution points mentioned by the author in the original representative work	We describe a new learning procedure, back-propagation, for networks of neurone-like units.
Contribution points mentioned in the citing sentences	<ol style="list-style-type: none"> 1. In MBGD, the learning rate is very important to the convergence speed and quality in training. Many different schemes, e.g., momentum [6], averaging [15], AdaGrad [16], RMSProp [17], Adam [18], etc., have been proposed to optimize the learning rate in neural network training. Adam may be the most popular one among them. [43.84%] 2. As for the extrapolation, a smooth activation function that only acts on the hidden layer(s) is recommended. Back-propagation is the second part of the algorithm [37]. This is the central mechanism that allows neural network methods to “learn.” [42.36%] 3. The feedforward multilayer perceptron is one of the most popular types of ANNs; it was developed by Rumelhart et al. [23], and it is presented in Supplementary I. This network also consists of an input layer, one or more hidden layers, and one output layer. [10.95%]

3) Knowledge transfer of contribution points

Changes in research topics of follow-up citing authors since 1987 when Hinton published the paper are shown in Fig. 5. Specifically, persistent ones include propagate error and representation. Afterward, the number of persistent research topics increases every five years and the research scope involves many new AI-based fields, such as pattern recognition, classification, prediction, computer vision, image analysis, knowledge extraction, sentiment analysis, speech recognition, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), language model and reinforcement learning (RL). Meanwhile, the research result quickly transferred to other disciplines (WoS Discipline Category, Fig. 6): four first-level disciplines, including Life Sciences, Physical Sciences, Social Sciences and Arts & Humanities,



realizing 100% cross-discipline. It's worth noting that 42 second-level disciplines enjoy the citing frequency that is greater than five. It is safe to conclude that the representative work's research result not only greatly impacts on its own field, but also is widely used in many fields of other disciplines. Objectively, this reflects the actual academic value of the representative work.

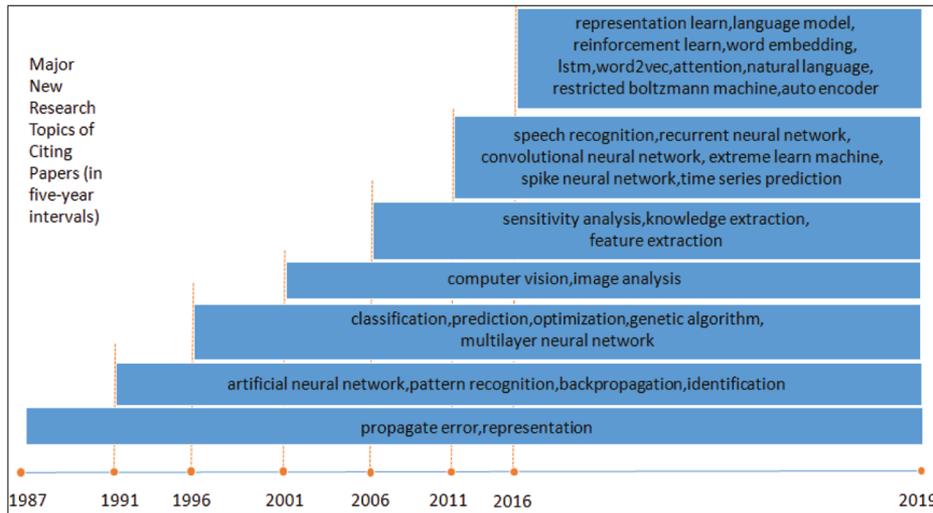


Figure 5. Changes in major new research topics of representative paper's citing papers.

6 Conclusion

Evaluating academic contributions of research papers, researchers, or research teams is real and serious need and has been a focus of research and practice. To meet the challenge represented by lack of evidence-based and multi-dimensional analysis tools, CiteOpinion integrates sentiment analysis, deep learning, clustering analysis and other methods to reach a rich and objective evaluation of research paper which in turn can be used to assess the academic contribution of researchers and research teams. It provides evidence-based analytic results, including highlight citing sentences, negative citing sentences, contrastive analysis of contribution points, analysis of contributing theme's knowledge diffusion and citing sentence statistical indexes. It works as an online analysis tool that can be used for quasi-real time analysis to support decision making in talent discovery, performance benchmarking, and competition analysis.

In practical sense, need for machine-readable full-text data is obvious because the difficulty of acquiring the full text of the citing and cited papers lead to less-than-expected level of accurate extraction of citing sentences especially for larger



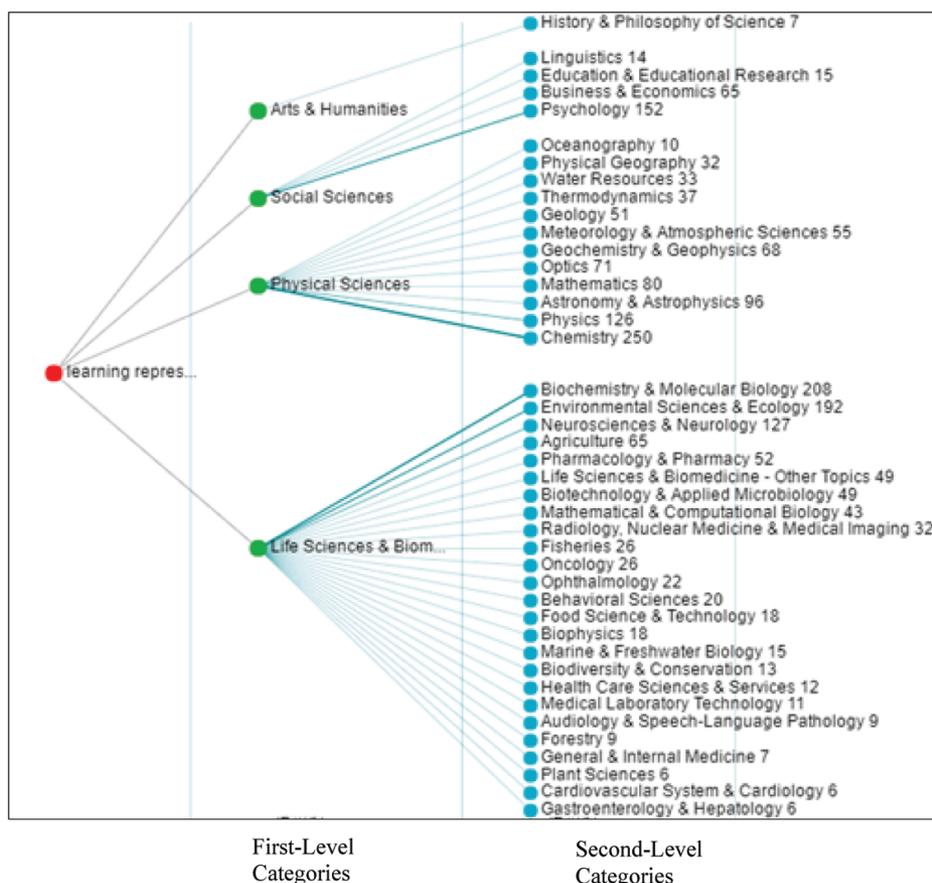


Figure 6. Distribution of representative work's citing sentences in different disciplines.

scale and real time application, to reach the fullest potential of the theoretic and technical capabilities of CiteOpinion. However, based on the samples, it is proved to be effective in reflecting the contribution and recognition by counterparts of representative works. Additionally, because of the limitations of sentiment dictionaries, fine tuning of sentiment scores of a few citing sentences are need to reveal their real sentiment in more granular ways, which needs further research and improvement.

Acknowledgment

Thanks to Professor Xiaolin Zhang, *JDIS*, for his meticulous comments for revision. Also, gratitude is expressed to Lin He, National Science Library, for his proposal for the project construction of which CiteOpinion was developed to address such needs.

Author Contributions

Xiaoqiu Le (lexq@mail.las.ac.cn) designed the study. Jingdan Chu (chujingdan@mail.las.ac.cn) designed the software structure, Siyi Deng (dengsiyi@mail.las.ac.cn), Qihang Jiao (jiaoqihang@mail.las.ac.cn), Jingjing Pei (peijj0926@foxmail.com), Liya Zhu (zhuly@mail.las.ac.cn), Junliang Yao (yaojunliang@mail.las.ac.cn) performed the program. Xiaoqiu Le, Jingdan Chu, and Siyi Deng participated in writing up the manuscript.

References

- Brooks, T.A. (1985). Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science*, 36(4), 223–229.
- Garfield, E. (1964). Can citation indexing be automated? In the Proceedings of the Symposium on Statistical Association, Washington DC, 8490. <http://garfield.library.upenn.edu/essays/V1p084y1962-73.pdf>
- Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. In the Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. <https://arxiv.org/pdf/1801.06146.pdf>
- Lei, S.W., Chen, H.H., Huang, Y., & Lu, W. (2016). Research on automatic recognition of academic citation context. *Library and Information Service*, 60(17), 78–87. DOI:10.13266/j.issn.0252-3116.2016.17.012
- Liu, S.B., & Chen, C.M. (2013). The differences between latent topics in abstracts and citation contexts of citing papers. *Journal of the American Society for Information Science and Technology*, 64(3), 627–639.
- Liu, S.B., Wang, B., Tang, D.L., Ma, X., & Ding, K. (2015). Research on paper influence based on citation context: A case study of the Nobel Prize winner's paper. *Library and Information Service*, 59(24), 109–114.
- Liu Y., & Cui L. (2014). The information value of citation context in paper content analysis. *Library and Information Service*, 58(6), 101–104.
- Ma, F., & Wu, Y.S. (2009). A survey study on motivations for citation: A case study on periodicals research and library and information science community in China. *Chinese Journal of Library and Information Science*, 2(3), 28–43.
- Mohammad, S., Dorr, B., Egan, M., Hassan, A., & Zajic, D.M. (2009). Using citations to generate surveys of scientific paradigms. In the Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Boulder: Association for Computational Linguistics, 584–592.
- Pei, J.J., & Le, X.Q. (2019). Identifying coordinate text blocks in discourses. *Data Analysis and Knowledge Discovery*, 3(5), 51–56. DOI:10.11925/infotech.2096-3467.2018.1380
- Small, H. (2011). Interpreting maps of science using citation context sentiments: A preliminary investigation. *Scientometrics*, 87(2), 373–388.
- Spiegel-Rösing, I. (1977). Science studies: Bibliometric and content analysis. *Social Studies of Science*, 7(1), 97–113.
- Teufel, S., Siddharthan, A., & Tidhar, D. (2006). Automatic classification of citation function. In the Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. Sydney: Association for Computational Linguistics, 103–110.

