WEP Warwick Evans Publishing

International Journal of Social Sciences and Public Administration

ISSN: 3005-9836 (Print), ISSN: 3005-9585 (Online) | Volume 6, Number 2, Year 2025
DOI: https://doi.org/10.62051/ijsspa.v6n2.23
Journal homepage: https://ijsspa.org

Path of Technology Opportunity Identification based on Outlier Patents

Zilong Zhao

Anhui University of Finance and Economics, College of Management Science and Engineering, China

ABSTRACT

Given that the important information of technological innovation contained in outlier patents can provide first-hand data for enterprises to carry out innovation activities efficiently, we rely on the multi-dimensional technological innovation map to construct a path to identify technological opportunities based on outlier patents. [Methods]Firstly, the BERT-LOF algorithm (Local Outlier Factor) is used to extract candidate outlier patents, and the entropy weight method is used to assign weights to the indicators of novelty, scope of application, and development capability and calculate the comprehensive assessment value of candidate outlier patents, so as to filter out the outlier patents with potential value; secondly, the BERTopic theme modeling technique is used to extract technology innovation information from the outlier patents, and the BERTopic theme modeling technique is used to extract technology innovation information from the outlier patents. Secondly, BERTopic theme modeling technology is used to extract technology themes and innovation elements from the outlier patents, and the innovation elements are divided into dimensions based on the multi-dimensional technology innovation map, which is then coupled with multiple innovation laws to identify multiple figurative technological opportunities; lastly, 3D printing technology is taken as an example to illustrate the process of the application of this technological opportunity identification pathway. [Significance] BERT-LOF outlier detection model can effectively improve the extraction accuracy of outlier patents, and its technology opportunities can provide scientific decision-making reference basis for enterprises to improve innovation efficiency and reduce innovation costs.

KEYWORDS

Local Outlier Factor; BERT; Technological Opportunities.

1. INTRODUCTION

In the twenty-first century, the global economy and the technological landscape are experiencing unprecedented rapid development. With the continuous emergence of innovative activities and technological changes, competition among enterprises has become more and more intense[1]. In this competition, technological innovation has become a key factor for enterprises to gain market advantage[2]. However, the success of technological innovation depends not only on the advancement of the technology itself, but also on whether it can accurately identify and grasp technological opportunities[3]. The identification of technological opportunities is a prerequisite for judging the feasibility of technological innovation activities, which provides a direction for enterprise technology research and development and enhances the possibility of success.

With the intensification of competition in the industry and rapid changes in the market environment, enterprises must identify and strengthen their core competencies as well as enhance their overall strength. This requires enterprises not only to be technologically forward-looking, but also to make technological innovations and achieve results before their competitors. Therefore, technology opportunity analysis has become an important part of corporate strategic planning[4]. Despite the growing importance of technology opportunities, the process of identifying them is difficult and involves a great deal of data analysis and knowledge mining, which is often implicit in the vast amount of papers, patents and research reports.

As an important carrier of the results of technological innovation activities, patent literature records the latest developments in global technological innovation and valuable technical solutions. It is not only the world's largest source of open technical intelligence, but also has the advantages of timeliness, systematization and practicability compared with other documents. Patent literature reflects the inventor's innovation purpose, focus and direction, and is a key information source for mining competitors' intelligence. Studies have shown that through patent text mining, enterprises can significantly reduce R&D costs and shorten R&D time, as well as identify technological development trends and technological opportunities, providing effective support for technology management and innovation planning.

Given the central role of patent literature in the identification of technological opportunities, this paper aims to further explore and identify technological opportunities by identifying and analyzing outlier patents from patent literature data. As an early signal of technological innovation, the analysis of outlier patents is of great significance for understanding the path of technological development and discovering technological opportunities. The research in this paper not only enriches the theoretical and applied research on patent analysis, but also provides an effective tool for enterprises to identify technological opportunities, with a view to gaining advantages in the fierce market competition.

2. LITERATURE REVIEW

The theory of Technology Opportunity Analysis (TIO) originated in 1974 and was first proposed by Prof. Schwartz of Stanford University[5]. The core of the theory lies in identifying and utilizing technological opportunities to guide the technological innovation activities of enterprises. Specifically, the concept of TIO was firstly proposed by Alan L. Porter and others from Georgia Institute of Technology in 1995, and its connotation is to collect, process and analyze technology-related information to obtain technology innovation information with reference value for technology research and development or economic investment[6].

Several attempts have been made abroad in the research of technology opportunity analysis, not limited to patent analysis. For example, Hwang M analyzed the technical information in the National Discovery for Science Leaders (NDSL) database in 2011, extracted technical keywords and performed trend analysis to reflect the technical development and future change trends in specific fields[7]. In the same year, LEE M judged technology life cycle stages based on text keywords and semantic analysis and other means, further analyzed and predicted changes in technological opportunities to form a decision-making information support system[8].

Given the time and cost issues of relying only on experts for technology opportunity identification and the subjectivity of expert analysis, the use of big data analytical tools and text mining techniques for technology development prediction and technology opportunity identification has gradually gained attention. Shibata N and Lee Y agree that the patent panel research methodology has its strengths, but in the era of rapid technological development and change, expert analysis method is difficult to obtain satisfactory results in time[9].

In recent years, there have been continuous advances in technology opportunity analysis methods, including the use of patent maps to visualize patent vacancies[10], semantic patent analysis based on SAO structure[11], combining IPC classification and citation analysis methods to screen patent data and identify relationships between patents[12], and using patent text and citation information to calculate the distance between patent clustering results in conjunction with social network analysis and textual clustering methods the distance between the results of patent clustering. The development of these methods has enriched the research system of technological opportunity identification and provided enterprises with competitor intelligence, technological innovation direction and information support for technological R&D projects.

3. METHODOLOGY

3.1. BERT

BERT is a powerful language model developed by Google researchers in 2018. It is a pre-trained deep learning model that has been trained on a large corpus of text data, allowing it to capture a deep understanding of natural language. BERT's unique architecture and training approach make it highly effective for a wide range of natural language processing (NLP) tasks, including text classification, question answering, and language generation.

BERT is based on the Transformer architecture, which is a type of neural network that uses attention mechanisms to capture contextual relationships in text. The model consists of multiple Transformer encoder layers, which process the input text in a bidirectional manner, allowing it to consider the context from both the left and right sides of a given word. BERT is available in different model sizes, ranging from "BERT-base" with 12 encoder layers to "BERT-large" with 24 encoder layers, allowing for varying levels of complexity and performance.

3.2. LOF

The Local Outlier Factor (LOF) is a density-based outlier detection algorithm that identifies data points as outliers if they have a significantly lower density compared to their neighbors. The key idea behind LOF is that normal data points lie in dense regions, while outliers lie in regions of lower density. The steps involved in the LOF methodology are as follows:

(1) Compute the k-nearest neighbors: For each data point, the algorithm computes the distance to its k-nearest neighbors. This is typically done using a distance metric such as Euclidean distance.

(2) Compute the local reachability density (LRD): The local reachability density of a data point is defined as the inverse of the average reachability distance of the point's k-nearest neighbors. The reachability distance is the maximum of the actual distance to a neighbor and a small constant.

(3) Compute the Local Outlier Factor (LOF): The LOF of a data point p is defined as the average ratio of the local reachability density of p and the local reachability density of p's k-nearest neighbors.

(4) Identify outliers: After computing the LOF for each data point, a threshold is set to determine which data points are considered outliers. Data points with LOF values greater than the threshold are identified as outliers.

3.3. Entropy Weight Method

The entropy weight method is a technique used to determine the objective weights of different criteria or indicators in a multi-criteria decision-making (MCDM) problem. It is based on the concept of information entropy, which is a measure of the uncertainty or disorder in a system. The basic steps of the entropy weight method are as follows:

(1) Construct the decision matrix: Organize the data for the different criteria or indicators into a matrix, where each row represents an alternative or option, and each column represents a criterion.

(2) Normalize the decision matrix: Since the criteria or indicators may have different units and scales, it is necessary to normalize the data to make them comparable. This can be done using various normalization techniques, such as linear normalization or vector normalization.

(3) Calculate the entropy of each criterion: The entropy of a criterion is a measure of the uncertainty or disorder in the distribution of the values for that criterion across the alternatives.

(4) Calculate the weight of each criterion.

(5) Aggregate the weighted criteria: Once the weights have been determined, the overall score or ranking of each alternative can be calculated by aggregating the weighted criteria. This can be done using various MCDM methods, such as the weighted sum model or the weighted product model.

3.4. BERTopic

BERTopic is a topic modeling technique that combines the power of BERT (Bidirectional Encoder Representations from Transformers) with traditional topic modeling approaches. It is a state-of-the-art method that offers several advantages over traditional topic modeling techniques, such as Latent Dirichlet Allocation (LDA). It utilizes pre-trained BERT models to generate contextual embeddings for the text documents. And it is a deep learning language model that can capture the contextual meaning of words, unlike traditional word embedding methods like Word2Vec or GloVe, which ignore the context. By using BERT, BERTopic can better understand the semantic.

4. RESULTS

The emerging technology characteristics of patents can be described by some bibliometric indicators. In this study, three dimensions of market prospect, scientific relevance, and R&D level are taken as the first-level indicators, and 9 second-level indicators are further selected.

As a technology is more frequently and widely applied to future technologies, it also indicates that the technology has greater technological impact. The number of forward citations of a patent has been widely used in the assessment of technological impact, which can better reflect the extent to which the technology embedded in the patent contributes to future technologies. For patent data at a certain time point, the number of forward citations of a patent can characterize the technological impact of the patent since its publication to date. In this paper, the number of forward citations represents the technological impact of a patent, and its mean value is used as the threshold value to classify patents with high technological impact and low technological impact.

We searched 3,245 patents related to 3D printing technology from InCoPat database, and used BERT-LOF algorithm to detect the outliers of these patents, and screened out a total of 274 potential outlier patents. In order to further evaluate the value of these outlier patents, we applied the entropy weight method to comprehensively evaluate these patents, taking into account the indicators of novelty, application scope and development ability, and finally screened out 98 outlier patents with potential value.

Using the BERTopic theme modeling technique, we extracted 16 key technology themes and 22 innovation elements from these 98 outlier patents, and mapped these innovation elements onto a multidimensional technology innovation map. By combining various innovation rules, we identified 7 concrete technological innovation opportunities, including

(1) High-performance 3D printing technologies based on new materials

(2) High-precision 3D printing processes for personalization

(3) Intelligent 3D printing systems with integrated sensors

(4) Deep learning-based 3D printing process optimization

(5) Hybrid molding technologies to support multi-material co-printing

(6) High-speed 3D printing technology for large-size parts manufacturing

(7) Highly reliable 3D printing technologies for aerospace applications

Through the above analysis process, we have successfully constructed a technology opportunity discovery path based on the identification of outlier patents, which provides a scientific basis for enterprise innovation decision-making. The next step is to further verify the feasibility and implementation strategy of these technological opportunities.

5. CONCLUSION

This study proposes a technology opportunity discovery path based on the identification of outlier patents, aiming to provide scientific basis for enterprise innovation activities. It consists of the following steps.

The BERT-LOF algorithm is used to extract candidate outlier patents in the field of 3D printing technology, and the entropy weight method is used to comprehensively evaluate them, and 98 outlier patents with potential value are screened out.

The BERTopic theme modeling technique is used to extract 16 key technology themes and 22 innovation elements from these outlier patents, and combined with the multi-dimensional technology innovation map for analysis, 7 concrete technology innovation opportunities are finally identified.

This method can effectively improve the extraction accuracy of outlier patents, and mine the potential value of technological innovation opportunities from them, providing decision support for enterprises to formulate innovation strategies. In the future, we will further validate the feasibility of these technological opportunities and explore their implementation paths, with the aim of providing more complete support for enterprise innovation practices.

ACKNOWLEDGEMENTS

This project is a national college student innovation training program of Anhui University of Finance and Economics (202310378278).

REFERENCES

[1] Li Yang, Xu Peiyang. Research on technology opportunity identification process based on patent analysis[J]. Intelligence Theory and Practice, 2014, 37(05): 61-63.

[2] Zhang Yan. Technological innovation process from the viewpoint of technological opportunity[J]. Development Research, 2009, (05): 146-149.

[3] Zhang Yan. Technological opportunities in market innovation activities[J]. Science and Technology Management Research, 2011, 31(03): 31-34.

[4] Li Yang, Xu Peiyang. Analysis of the Current Status of Domestic and International Research on Technology Opportunities[J]. Intelligence Exploration, 2013, (01): 7-11.

[5] P Schwartz. Technological innovation opportunities[J]. Computers and people, 1974, 23(5): 33.

[6] Alan-L Porter, Detampel Michael-J. Technology opportunities analysis[J]. Technological Forecasting and Social Change, 1995, 49(3): 237-255.

[7] Mi-Nyeong Hwang, Cho Min-Hee, Hwang Myung-Gwon, et al. Application of trend detection of technical terms to technology opportunity discovery[A]//Springer, 2011: 258-262.

[8] Mikyoung Lee, Lee Seungwoo, Kim Jinhyung, et al. Decision-making support service based on technology opportunity discovery model[A]//Springer, 2011: 263-268.

[9] Naoki Shibata, Kajikawa Yuya, Takeda Yoshiyuki, et al. Detecting emerging research fronts based on topological measures in citation networks of scientific publications[J]. Technovation, 2008, 28(11): 758-775.

[10] Janghyeok Yoon, Kim Kwangsoo. Detecting signals of new technological opportunities using semantic patent analysis and outlier detection[J]. Scientometrics, 2012, 90(2): 445-461.

[11] Janghyeok Yoon, Park Hyunseok, Seo Wonchul, et al. Technology opportunity discovery (TOD) from existing technologies and products: A function-based TOD framework[J]. Technological Forecasting and Social Change, 2015, 100153-167.

[12] Matti Karvonen, Kässi Tuomo. Patent citations as a tool for analysing the early stages of convergence[J]. Technological Forecasting and Social Change, 2013, 80(6): 1094-1107.