IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 2, February 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

An Intelligent Patent Recommender System Based on Machine Learning

Primrosetina.M*, N. Santhi*

*Noorul Islam Centre for Higher Education, Kumaracoil, India

Associate Professor, Noorul Islam Centre for Higher Education, Kumaracoil, India

ABSTRACT: Recommendation methods are frequently used in a variety of industries, including online product customization searches and customer-focused advertising. The mechanism for a patent recommendation system to find semantically pertinent patents for additional trend analysis and technology mining is developed in this study. The suggested recommender uses Machine Learning (ML) techniques like doc2vec for Natural Language Processing to encode patent papers in the vector space with sentence transformers to make it possible to do semantic analysis on the patent files. The doc2vec and Sentence transformer algorithms are used in the newly created patent recommender to train the model for document vectorization employing the relevant patent's context. The vectors allow for numerical processes like patent document's clustering and similarity assessment. There are two primary phases in the referral process. The process of extracting key phrases involves first employing doc2vec to process the original patent's words and then using sentence transformer algorithm to find the original patents' central phrase. Lastly, the key phrases that is nearest to the centre are chosen using cosine similarity. The research tests the recommender system and methodology's improved accuracy and effectiveness using a few drone technology sub-domains as case scenarios.

KEYWORDS: Natural Language Processing; Patent Recommender; Patent analysis;Sentence Transformers; Siamese architecture

I. INTRODUCTION

Today,Companies rely on intellectual property (IP), such as patents, because these intangible assets make up a larger portion of a company's net worth. Because of this, large amount of literature are being produced by not only companies but also educational universities at a soaring rate.The innovation, originality, and utility of an invention are all described in the patent paperwork, along with any distinctions from previous art. Businesses frequently discover relevant patents or published literatures that introduce them to current technologies when they create or adopt new inventive technologies, which help them avoid violating the intellectual property rights (IPRs) of others. Thepatent literatureare thus,further analysed for even more focused tasks like understanding the technological field better, finding solutions to problems, doing literature survey to generate ideas for fresh initiatives and comprehend cutting-edge methodologies [1]. In general, investigators spend a great deal of time looking for relevant related study[2]. Finding suitable articles using keyword-based searching methods is difficult because they call for some background understanding in the domain. People typically don't precisely state their data demands, that makes this process much more difficult [3]. Numerous document recommendation algorithms have lately been suggested to solve this problem[4]. These algorithms should serve a number of purposes. For beginner researchers, methods should suggest a wide range of articles, while for senior researchers, suggestions ought to be more in line with their pre-existing interests or guide them toward pertinent research activities [5]. The ability to access relevant material more quickly and simply owing to patent recommendation systems often benefits business and academics' professionals[6].The fundamental objective of a recommender system is to include or omit pertinent items utilizing information. Content-based sorting, collaborative sorting, and the combination method are some of the various types of recommender system. The primary aim of conventional recommenders' in different fields is to locate and rank possibilities for products or services that users may be fascinated by based on their internet search history, the likings and preferences and ratings they have set, and their previous purchase behaviour. In order to increase the precision of the recommenders, many parameters are evaluated and employed as sources to build the references.As input elements for recommender system, the user's contributions on social networks or on a preferred fans site, for example, are collected.

Recommendation systems were developed for various applications in the government and other libraries. But in terms of intellectual property, the drawn-out process demands extra effort to read and interpret the massive amounts of patents. The World Intellectual Property Organization (WIPO), the European Patent Office (EPO), the Chinese State Intellectual Property Office (SIPO), and the United States Patent and Trademark Office are just a few of the patent

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

datasets from which a vast variety of patent applications are gathered after conducting basic patent searches.The purpose of this proposal is to develop an intelligent patent recommendation system that would focus on the much more accurate suggestion of sub-domain patents for future investigations into the details of technology developments and gaining new insights, without duplicating the comprehensive patent searches that so many patent datasets currently provide. In order to demonstrate how the patent recommender works and locate more precise sub-field patents linked to the initially found patents, the research uses a few cases of sub-domain patentdocuments for the drone field as case studies. The project's goals are to categorize specific papers according to their subjects and sub-domains in drone technology. Text mining approaches using natural language processing (NLP) are essential for creating text and content-based recommenders when the buried information is contained in patent papers. Therefore, the focus of this research is on the deliberate selection of patent papers that are semantically significant. The relevant literature is evaluated in Part 2 with an emphasis on text analysis and recommender systems. The suggested technique and methods for recommendations are formulated and discussed in Part 3. The system structure, sample execution, and outcome study based on a patent search for NLP technology with recommendation results are all described in Part 4. Part 5 concludes by summarising the findings and recommendations of the study and offering suggestions for future studies.

II. RELATED WORK

Classification of Patents according to Industry 4.0 Pillars using Machine Learning Algorithms is proposed in this research. The ability of stakeholders to adapt to the changing industrial development must therefore be assessed by analysing patterns and trends in intellectual property (IP) information. Because the data in patent bibliographic papers is both organised and unstructured, text mining or machine learning must be used to analyse the data. In order to determine if the institution is prepared to deal with the fourth industrial revolution, this article established a patent trend by evaluating the patent data of the Intellectual Property Corporations of Malaysia (MyIPO). To achieve this aim, MyIPO patent data was categorised using a patent classification system according to the tenets of Industry 4.0. Additionally, the datasets for this study were created using data from the MyIPO Online Search and Filing System for patents. The dataset, however, just includes the name of the patent and the year it was published. The title's short text data presented a barrier for this investigation because it contains little semantic information and is quite sparse. Five popular classifiers were employed in this study to categorise the text. The Support Vector Machine (SVM) was found to be the machine learning classifier that classified the training and testing datasets with the best degree of accuracy. According to the research presented in this paper, the pillar of Industry 4.0 with the highest level of innovation in terms of patents is the autonomous robot [7].

The multi-class categorization methodology for science and technology policy utilizing patent information is suggested in this research. Technological exchange is the process of transferring a technology's excellence to its current market value. IP-driven Research & Development and corporate decision-making are utilized to assist effective study and development. Firms establish their patent collections and construct technological management approach through IP-driven Research &Development. They safeguard key rights and make usage of those patents to collaborate with other businesses. The rapid development of conversion systems in current civilization made it challenging to adapt earlier technology policy methods to innovation. This is as a result of their employment of quantitative methodologies that were expert-based. Using qualitative findings to ensure neutrality is challenging. To overcome these restrictions, numerous earlier researches have put forth models for assessing technique based on patents information. However, these methods risk losing contextual data while processing reference information, necessitating the utilization of a lexical analyser suited for analysing patent terms. To circumvent this restriction, this study makes utilization of a lexical analyser built utilising a deep learning architecture. Additionally, the suggested strategy divides the technologies into several groups and utilizes statistical results and bibliographic data from patents as independent variable. Sequentially assessing a technology's worth is how the multi-class categorization is done. This algorithm returns several classes in chronological order, allowing for class comparisons. Additionally, it is model-agnostic, allowing for the implementation of a variety of techniques. To test the viability of the suggested technique, it ran tests employing real patent data. The suggested technique was able to categorise actual patents into an ordered multi-class based on the outcomes of the study. Additionally, the results' objectivity could be ensured. This occurs as a consequence of the data in the patent specification being utilized in the models. Additionally, the system that combined qualitative and bibliography data performed better for categorization than the system that solely used quantifiable information. Consequently, by grouping the value of technology into more precise groups, the suggested technique could aid in the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

effective maintenance of businesses. However, Technological policy must leverage patents from multiple nations in order to advance effectively in accordance with market creation [8].

Thehybrid strategy that takes into accounts both professional judgement and patent data to identify developing potential ideas. It is now crucial for businesses to spot emerging, promising technologies that can be used to start new enterprises or enhance existing ones. This is due to the speed at which technology and industrial value chains are changing. Patent study is one of the methods most regularly used to find new, promising inventions. Many earlier studies have used patents to characterise emerging technologies since patents.Though, the majority of earlier research heavily relied on patent data when evaluating prospective innovations, even if promisingness is determined by a number of additional criteria that are not explained by patent data. This study suggests a hybrid technique that takes into account both expert opinions and patent data to discover upcoming potential innovations in order to get beyond the limitations of earlier approaches. In order to analyse the data, they first created a set of criteria for evaluating possibly prized patents, then it had experts assess a subset of the patents from a bigger collection of patents using the conditions, and lastly it used the assessment results to find additional possibly prized patents among the remaining patents in the collection of patents. At this point, an active semi-supervised learning technique was used, combining a sizable amount of unlabelled data with a modest amount of labelled data. Patents and patent attributes are the two layers that make up an analysis model, and it uses patent features such technology characteristics to categorise patents as promising or unpromising. The study results showed that semi-supervised learning integrated with active learning has plausible effects in successfully searching for evolving brilliant technologies or straining out non-promising technologies with less humanoid input. The suggested method was applied to the automobile industry sector, and its usage was proved. A huge set of patents could be labelled using a small number of labelled patents, saving time and effort when experts assess patents. This is a pioneering methodological effort to use active semi-supervised learning to patent analysis. Practically speaking, the research results make it possible to balance expert and data-driven decision-making by effectively using expert opinions to discover promising innovations and anticipate a future innovation ecosystem[9].

As the nation actively fosters the advancement of technologies and scientific and works to strengthen its capacity for autonomous innovation, the security of technologies assets is receiving increasing consideration. China has advanced quickly in a variety of technical and scientific disciplines recently, and each year sees a rise in the amount of patent registrations. Moreover, a number of issues with patent reliability have emerged, such as underdeveloped patent technology and a low rate of patent issuance. In this article, the parameters of patent quantity and assessment processes are examined. In order to create a model for evaluating patents, it first measure the indications used to evaluate the validity of patents and merge that information with the language of the patent. Technologies for multitasking learning training is based on US patents with patent level labelling. Secondly, the recent study identified is transferred from the English patents to the Chinese patents, where the task of physical labelling is minimised through the utilization of effective learning technologies and transferable learning innovation. Therefore, a collaborative training-based methodology for Chinese patent quality assessment was developed and put into practice. The study's techniques significantly enhanced the model's predictive power and enhanced the migratory impact. Numerous empirical outcomes demonstrate the effectiveness of the Chinese patent integrity assessment method. In order to conduct studies on patent quality assessment designs from various angles, offer patent decision support for associated companies, and identify future directions of investigation for research facilities and patent innovators, this study makes utilisation deep learning and natural-language processing innovation. Given the limits of machine translation strategies, the writers argue that employing machine translated Chinese information reduces the overall outcome of the methodology when utilizing source systems trained on U.S. patents and degrades the preliminary modeling impact [10].

Utilizing blockchains method to forecast text mining and patent analytics. Information technologies (ITs) have contributed significantly to the advancement of our civilization, and the rapid pace of their development fosters competition among areas as well as businesses. As a result, predicting the direction of technology in the future can help with investment and technology choice decisions. Researchers, investors, and government organisations have all paid close attention to blockchain technology due to its wide range of impressive uses. The objective of this study is to assess blockchain technology and examine its tendencies in light of how the World Intellectual Property Organization (WIPO) database classifies them.Additionally, we give special consideration to registered patents in some of the most well-known patent databases in the world, like the USA patent database. By using the text mining and clustering approach, we were able to identify the most recent technological developments in blockchain patents. The findings show that the growth phase has been attained for the registered patents in the USA patent database. This indicates that the blockchain is receiving more attention now, with the majority of patents focusing on crypto currencies and their uses in the financial sector. Blockchain technology is still in its infancy, and researchers and creators are continually refining it[11].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

In societal structure, industry, and technology, digitizing is a significant force. These advancements in digitization are also influencing patent administration, particularly the discovery for and evaluation of patents. This overview, in contrast to earlier research, concentrates on the fundamental digital business patterns that have emerged as commonplace in the scenery of prescribers' patent knowledge bases and interrogation techniques. Previous studies have primarily concentrated on the specific purpose and methodologies utilised for patent analysis and search. The analysis of 7 governmental and twenty private suppliers resulted in the identification of fifteen distinct digitalization developments. Although commercial suppliers prefer to concentrate on more complex trends for patent research, such as predictive analytics, public providers concentrate exclusively on developments for patent searching, such as machine translation. The hybridized coding method was utilized to identify the Fifteen tendencies, and they all lead to the 4 digitalization areas of cloud services, information management, information processing, and artificial intelligence. Finally, conflicts brought on by the development of these patterns are examined, such as the smooth changeover from searching to evaluation versus the capacity to describe lesser. As of September 2021, all browser providers and Adobe have stopped supporting the software plug-in Adobe Flash Player, which is necessary for the analysis to be displayed graphically[12].

III. PROBLEM STATEMENT

Although many researches have been focused on analyzing text from the patent documents for patent classification and trend analyzing, very few have focused on building an efficient patent recommendation. Latent semantic analysis and latent Dirichlet allocation are the algorithms most commonly applied in recent research. LSA use the matrix decomposition and LDA are based on probabilistic model. LSA/LSI outperforms when the training data is little, but not when there is a large amount of data. Recommendation methods are frequently used in a variety of industries, including online product customization searches and customer-focused advertising. Hence, it is important to develop a mechanism for patent recommendation system to find semantically pertinent patents for additional trend analysis and technology mining. In order to analyze technical publications like patents employing Natural Language Processing-based approaches, the study presents a computer-assisted intelligent recommender.

IV. PROPOSED SYSTEM

1. Text Analysis

Text analysis is the practise of reading and comprehending human-written text using computer tools to gain business insights. Software for text analysis can autonomously categorise, analyse, and extract data from texts to find trends, connections, sentiments, and other useful information. Text analysis can be used to quickly and accurately evaluate a variety of text-based sources, including emails, documents, posts on social media, and product reviews, just like a human would.

The stages in text analysis involve:

1. Data Gathering: A sizable number of patents in the drone domain and its subdomains are searched and gathered for the case studies from the USPTO patent and Lens database.

2. Data Preparation
    a. Tokenization : Tokenization is the procedure of separating the original text into pieces that make semantic sense.
    b. Part-of-speech tagging: The tokenized text is given grammatical tags by part-of-speech tagging. Applying this step, for instance, to the tokens stated before yields the following text: Benefits: Verb; analytics: Noun; businesses: Noun.
    c. Parsing: The tokenized words are connected in a meaningful way via parsing using English grammar. It facilitates the visual representation of word relationships by text analysis tools
    d. Lemmatization: Lemmatization, a linguistic procedure, reduces words to their lemma, or dictionary form. The word "visualise" is the dictionary's definition of "visualising."
    e. Stop words removal: Words like and, or, and for are examples of stop words because they provide little to no semantic context for a sentence. The programme may take them out of the structured text depending on the use case.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

2. Text Vectorization

The research approach provides a new patent recommendation system that employs the doc2vec technique to train a NN model for text vectorization utilising the context of domain patents. The vectors allow for numerical operations like grouping and similarity analysis of patent documents. There are two primary phases in the referral process. First, word-embedding is used to identify the significant terms in the patents that the user has read. A sizable number of patents in the drone domain and its subdomains are searched and gathered for the case studies from the USPTO patent and Lens database. The data in these patents are then prepared with preprocessing as mentioned above (Tokenization, Part-of-speech tagging, Parsing, Lemmatization, Stop words removal). Then the training set for the doc2vec NN model is comprised of these patents. Once vectorization is done, sentence transformers are applied to the patents. The proposed research methodology's workflow is broken down as follows. First, the domain patents are gathered in accordance with the user's interests for the initial patents selection. To find pertinent patents, you may search patent databases and platforms like Derwent Innovation, the USPTO, and Google Patents. A good number of patents in the sub-domains are chosen for case studies based on the titles, abstracts, and claims of the patent papers. The recommendation mechanism for further computation includes these early patents. The process of extracting key phrases involves first utilising doc2vec to vectorize the words of the primary patents and then averaging all word vectors to get the initial patents' central phrase for each sub-domain. Finally, the key words that are closest to the centre are chosen using cosine similarity.Figure 1 illustrates the suggested system's schematic diagram and flow from the system concept.

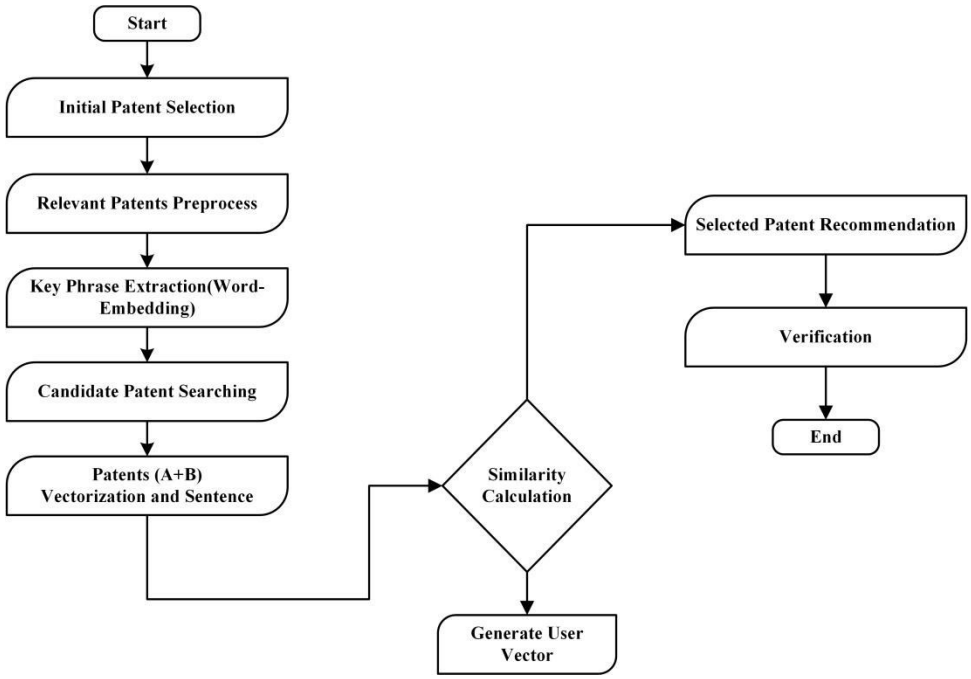


Figure 1: **Workflow of the Proposed System**

Importing the key words that were previously extracted in order to choose additional candidate patents from the USPTO/Google Patents database is necessary for the identification of the suggested patents. The structure uses a web crawler and USPTO API to choose potential patents from the patent pool in the USPTO/Google Patents database. Itsuggests patents that are clearly connected to the original patents based on these prospective patents. Altogetherthe patents, including all early patents and potential patents, are vectorized using Doc2vec. Doc2vec is similar to word2vec in idea, but differs from it in that it includes a vector for the patent ID. Doc2vec's design is broken into two groups based on the Distributed Memory Model of Paragraph Vectors.Comparing the capacity of TF-IDF, Latent Dirichlet Allocation, Word2vec, and Doc2vec to detect document similarity the overall finding is that Doc2vec produces a superior result when comparing document similarity. The research uses doc2vec as a result to provide a clever patent recommendation system. Three patents and their titles are utilised as training text as an example. After doc2vec application, the initial patents also is used with sentence transformer for further analysis.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

3. Sentence Transformers

SBERT is fine-tuned on sentence pairs using a Siamese architecture. Essentially, it just uses one BERT modelling. However, it is simpler to conceive of this as two models with coupled weights as we interpret sentences A and B as pairs during training.Figure 2 from the system architecture displays the sentence transformers' schematic diagram.

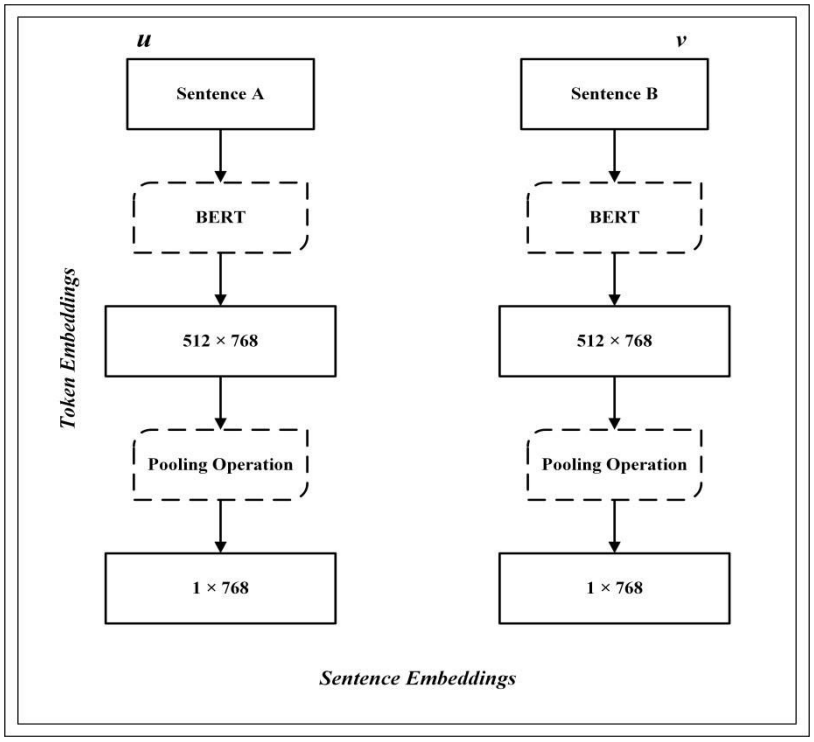


Figure 2:**Sentence Transformers Architecture**

Sentence-Transformers are a state-of-the-art Python framework for embedding sentences, text and images. These embedding are can then be used for classification or clustering. To train a Sentence Transformers model, one needs to enlighten the computer in some way that two sentences have a definite degree of similarity. Hence, each example in the data needs a label or structure that permits the model to appreciate whether two sentences are alike or different. Sentence Transformers are similar to Cosine Similarity. Given that the vector was trained using the domain document context, the cosine similarity between these 2 document vectors is a useful indicator of the semantic similarities between the two articles. The vectors A and B's similarity measure expression.

Patent Model Tree

Semi-structured information instances include patent records. The identities of the scientists, the filing dates, the International Patent Classification code, and other structural information are included. Textual data also includes an abstract, the entire text of a copyright, the assertions, as well as some unorganized drawings. To compute patent similarity, patent information must be cleanly organised according to its features. In computing similarity based on the analysis of copyright information and variations in data structures, the PMT concept was used. Different kinds of information are processed by PMT in various ways. PMT, a tree - based data architecture, represents the copyright information structure. The origin of PMT denotes the entire patent information collection that could then be partitioned into a number of kid nodes depending on the characteristics of the intellectual database. The leaf nodes are just data extracted from patent documents, such as the inventors' residences, the date of issuance, asserts, etc. The significance of a child node in respect to its parent node is determined by its load factor.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

Patent Analysis

Patent analysis is a primary kind of intellectual property that identifies technological and patenting patterns by using data from patent documents as well as other data from the patent lifecycle (patent data). Though there are difficulties and inconsistent terminology in diverse languages, containing French and Spanish, others like Russian, borrow terms from other languages. Patent landscape, mapping, and cartographic are other phrases that are employed interchangeably with patent statistics. Patent analytics is a burgeoning professional subject that includes data cleansing, text mining, machine learning, geographic mapping, statisticsvisualizations, and analysis of patent data and scientific literature.Patent analytics is a technique that is widely used in business and is being investigated more and more by the public sector to help make educated choices about, among other things, how to prioritize and invest in R&D, manage IP portfolios, commercialize technologies, and collaborate on research.

In the previous era, spreadsheet-based data analysis approaches have been employed to examine the tools and processes utilized in patent applications. Old-school patent analyses have recently been combined with data science, machine learning, semantics technology, and artificial intelligence in the domain of intellectual property, along with a rise in systems for patent visualizations. Furthermore, there has also been an increase in the utilization of open-source software and tools for copyright analysis, in addition to the implementation of different machine learning techniques. Several programs advocate creating statistics, displays, and visualization's in a semi-automated manner. The stages and tasks required in patent analytics are now only briefly described in a relatively small number of methodological resources. Patent analytics groups typically collaborate with R&D departments and patent attorneys, with the relevant data influencing choices about IP, commercial strategy, and business strategy.The following processes are generally included in creating a patent statistics study and a patent landscapes assessment:

- Defining the subject and proposal's parameters,
- The IP research, which facilitates getting patent information,
- The standardization and cleansing of data,
- Visualizing and analyzing data
- When writing the presentation, consider using narrative and storytelling, and
- Distributing and disseminating the analysis

Patent analytics is a repetitive procedure that frequently necessitates project recopying and adaption based on what is learned throughout the process. Different tools, including open-source and commercial ones, can be used for analytics. Some of these tools are incorporated in patent data collection, while others are more general data processing, visualization, and statistics software.

V. RESULTS AND DISCUSSION

The results of the proposed system here. Figure 3a and 3b shows the top rows of the patent documents given to the system as database for the system to execute the patent recommendation and find the relevant results from. Figure 4 shows the total number of patent documents provided to the system as database. Once the database is provided to the system, an abstract to given as input to the system based on which the system will text analyse the patent documents and find similar patent documents. Figure 5shows the outcome of the new abstract given to the system as input and also shows the patent recommendation system retrieved top-most similar/relevant 4 abstracts of the top relevant patents searched, analyses and found from the database by the patent recommendation system. Figure 6 shows the Headings & Cosine Similarity of Top n similar patents. This research develops patent recommendation system shown by the result which retrieves relevant patent documents.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

	ï»¿"#"	Jurisdiction	Kind	Display_Key	Lens_ID	Publication Date	Publication Year	Application Number	Application Date	Priority Numbers	...	Sequence Count	CPC_Classifications
0	1	KR	A	KR 20210049065 A	041-397-336-667-905	2021-05-04	2021	KR 20210052612 A	2021-04-22	KR 20210052612 A	...	0	Y02A50/2351;;B01D47/06;;B01D53/32;;B01D46/00;;...
1	2	CN	A	CN 106662313 A	145-386-375-019-078	2017-05-10	2017	CN 201680002265 A	2016-09-27	CN 2016100319 W	...	0	B64C2201/024;;B64C2201/024;;B64D47/02;;B64D47/...
2	3	JP	A	JP 2019093856 A	190-248-368-723-223	2019-06-20	2019	JP 2017223941 A	2017-11-21	JP 2017223941 A	...	0	NaN
3	4	US	B2	US 10973083 B2	070-234-967-727-017	2021-04-06	2021	US 201916582924 A	2019-09-25	US 201916582924 A;;US 201615352559 A	...	0	H04W84/18;;H04W84/18;;H04W4/08;;H04W4/08;;H04W...
4	5	JP	A	JP 2017052389 A	194-597-484-758-598	2017-03-16	2017	JP 2015177680 A	2015-09-09	JP 2015177680 A	...	0	NaN

5 rows × 33 columns

Figure 3a:**Head of First 5 Rows of the Given Dataset (Showing first 12 Columns)**

IPCR_Classifications	US Classifications	NPL Citation Count	NPL Resolved Citation Count	NPL Resolved Lens ID(s)	NPL Resolved External ID(s)	NPL Citations	Legal Status
B01D47/06;;B01D39/00;;B01D46/00;;B01D53/24;;B0...	NaN	0	0	NaN	NaN	NaN	DISCONTINUED
F21S8/10;;B64D47/02;;B64D47/06;;F21V8/00;;F21V...	NaN	0	0	NaN	NaN	NaN	DISCONTINUED
B64C39/02;;B64F1/36;;H01Q1/36;;H04B7/15	NaN	0	0	NaN	NaN	NaN	PENDING
H04W84/18;;H04W4/08;;H04W4/40;;H04W8/18	NaN	65	20	136-199-342-791-23X;;043-059-237-468-047;;045-...	10.1117/12.2014530;;10.1007/978-3-540-39850-9_...	U.S. Appl. No. 15/352,559, filed Nov. 15, 2016...	ACTIVE
B64C13/18;;B64C27/08;;B64C27/68;;B64C29/00;;B6...	NaN	0	0	NaN	NaN	NaN	ACTIVE

Figure 3b:**Head of First 5 Rows of the Given Dataset (Showing first 12 Columns)**

```
Out[34]: ï»¿#                          200
Jurisdiction                  200
Kind                          200
Display_Key                   200
Lens_ID                       200
Publication Date              200
Publication Year              200
Application Number            200
Application Date              200
Priority Numbers              200
Earliest Priority Date        200
Title                         200
Abstract                      200
Applicants                    200
Inventors                     200
Owners                        117
URL                           200
Document Type                 200
Has Full Text                 200
Cites Patent Count            200
Cited by Patent Count         200
Simple Family Size            200
Extended Family Size          200
Sequence Count                200
CPC_Classifications           200
IPCR_Classifications          198
US Classifications              5
NPL Citation Count            200
NPL Resolved Citation Count   200
NPL Resolved Lens ID(s)        19
NPL Resolved External ID(s)    19
NPL Citations                  49
Legal Status                  200
dtype: int64
```

Figure 4: **Total Patents**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

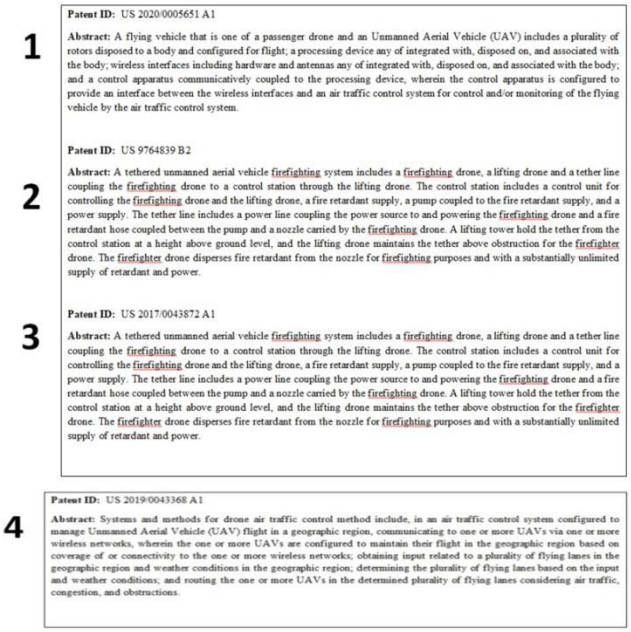
|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

Given Input Abstract:

To supply power to a plurality of drones by using a power supply cable for a drone group composed of a plurality of drones.

Solution: A drone group 1 includes a plurality of drones 10 which can perform flight attitude control by driving and controlling motor means, wired cables 12 for connecting adjacent drones 10 of the plurality of drones 10 with each other to connect the plurality of drones 10 like a network, a power supply cable 13 to which one end of at least one drone of the plurality of drones 10 is connected, and power supply means 14 connected to the other end side of the power supply cable 13 and arranged on the ground. In the drone group 1, power for driving the plurality of drones 10 is supplied from the power supply means 14 to each drone 10 through the power supply cable 13 and the wired cables 12 from the power supply means 14.



1

Patent ID: US 2020/0005651 A1

Abstract: A flying vehicle that is one of a passenger drone and an Unmanned Aerial Vehicle (UAV) includes a plurality of rotors disposed to a body and configured for flight; a processing device any of integrated with, disposed on, and associated with the body; wireless interfaces including hardware and antennas any of integrated with, disposed on, and associated with the body; and a control apparatus communicatively coupled to the processing device, wherein the control apparatus is configured to provide an interface between the wireless interfaces and an air traffic control system for control and/or monitoring of the flying vehicle by the air traffic control system.

2

Patent ID: US 9764839 B2

Abstract: A tethered unmanned aerial vehicle firefighting system includes a firefighting drone, a lifting drone and a tether line coupling the firefighting drone to a control station through the lifting drone. The control station includes a control unit for controlling the firefighting drone and the lifting drone, a fire retardant supply, a pump coupled to the fire retardant supply, and a power supply. The tether line includes a power line coupling the power source to and powering the firefighting drone and a fire retardant hose coupled between the pump and a nozzle carried by the firefighting drone. A lifting tower hold the tether from the control station at a height above ground level, and the lifting drone maintains the tether above obstruction for the firefighter drone. The firefighter drone disperses fire retardant from the nozzle for firefighting purposes and with a substantially unlimited supply of retardant and power.

3

Patent ID: US 2017/0043872 A1

Abstract: A tethered unmanned aerial vehicle firefighting system includes a firefighting drone, a lifting drone and a tether line coupling the firefighting drone to a control station through the lifting drone. The control station includes a control unit for controlling the firefighting drone and the lifting drone, a fire retardant supply, a pump coupled to the fire retardant supply, and a power supply. The tether line includes a power line coupling the power source to and powering the firefighting drone and a fire retardant hose coupled between the pump and a nozzle carried by the firefighting drone. A lifting tower hold the tether from the control station at a height above ground level, and the lifting drone maintains the tether above obstruction for the firefighter drone. The firefighter drone disperses fire retardant from the nozzle for firefighting purposes and with a substantially unlimited supply of retardant and power.

4

Patent ID: US 2019/0043368 A1

Abstract: Systems and methods for drone air traffic control method include, in an air traffic control system configured to manage Unmanned Aerial Vehicle (UAV) flight in a geographic region, communicating to one or more UAVs via one or more wireless networks, wherein the one or more UAVs are configured to maintain their flight in the geographic region based on coverage of or connectivity to the one or more wireless networks; obtaining input related to a plurality of flying lanes in the geographic region and weather conditions in the geographic region; determining the plurality of flying lanes based on the input and weather conditions; and routing the one or more UAVs in the determined plurality of flying lanes considering air traffic, congestion, and obstructions.

Figure 5: **Patent ID and Abstract of Top n Similar Patent Details**

	top_Patent_ids	cosine_similarity	Abstracts
0	US 2020/0005651 A1	0.7416	A flying vehicle that is one of a passenger dr...
1	US 9764839 B2	0.7170	A tethered unmanned aerial vehicle firefightin...
2	US 2017/0043872 A1	0.7170	A tethered unmanned aerial vehicle firefightin...
3	US 2019/0043368 A1	0.7099	Systems and methods for drone air traffic cont...
4	US 2019/0035285 A1	0.6909	Systems and methods for drone air traffic cont...

Figure 6: **Headings and Cosine Similarity of Top n Similar Patents**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Issue 2, February 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1102034 |

VI. CONCLUSION

The technique assists patent researchers in gathering extremely relevant patents with precision and effectiveness prior to subsequent technological trend studies in any domain. Patent examiners can quickly deploy the patent recommendation system as a pre-process module for pertinent patent retrieval for any domain-specific patent analytics. The suggested technique uses semantic analysis to intelligently choose a reduced group of semantically related patents from a large collection of potential patents. The accuracy of each of these activities is confirmed to match expert manual picks, and they are all done within the IPCs of the original patents. Itdiscovers that the technique greatly enhances recommendation performance through the authentication of the examples. This study's key contribution is to enhance patent recommendation outcomes without the intuitive judgments of domain experts. Comparing the recommended patents with the recommendations of the domain experts in the drone sector has provided a detailed description of the validation process, which is the challenging aspect of this study. Despite the fact that this research has substantially described the full process of developing a patent recommender, the application to different domains might still be further examined in other studies. However, the suggested doc2vec-based methodology's scalability and consistency, particularly when it's used to create patent recommenders for various domains needs to be thoroughly checked. For appropriate representation, doc2vec model trainings need a lot of training materials.

REFERENCES

[1] J. Liu, H. Zhang, and Z. Guo, “A defense mechanism of random routing mutation in SDN,” *IEICE Trans. Inf. Syst.*, vol. 100, no. 5, pp. 1046–1054, 2017.

[2] M. Amami, R. Faiz, F. Stella, and G. Pasi, “A graph based approach to scientific paper recommendation,” in *Proceedings of the international conference on web intelligence*, 2017, pp. 777–782.

[3] X. Li, Y. Chen, B. Pettit, and M. D. Rijke, “Personalised reranking of paper recommendations using paper content and user behavior,” *ACM Trans. Inf. Syst. TOIS*, vol. 37, no. 3, pp. 1–23, 2019.

[4] Q. Yang *et al.*, “A novel hybrid publication recommendation system using compound information,” *World Wide Web*, vol. 22, no. 6, pp. 2499–2517, 2019.

[5] K. Sugiyama and M.-Y. Kan, “Serendipitous recommendation for scholarly papers considering relations among researchers,” in *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, 2011, pp. 307–310.

[6] M. Le, S. Kayal, and A. Douglas, “The Impact of Recommenders on Scientific Article Discovery: The Case of Mendeley Suggest.,” in *ImpactRS@ RecSys*, 2019.

[7] W. A. Z. W. C. Jafery , M. S. S. Omar , N. A. Ahmad; H. Ithnin , Classification of Patents according to Industry 4.0 Pillars using Machine Learning Algorithms , 2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS).

[8] J. Lee, J. Kang, S. Park, D. Jang, and J. Lee, “A Multi-Class Classification Model for Technology Evaluation,” *Sustainability*, vol. 12, no. 15, p. 6153, 2020.

[9] L. Parmentier, “Mary-Morstan: a multi-objective modular framework to automatically configure machine learning algorithms,” PhD Thesis, Université de Lille, 2022.

[10] X. You, J. Liu, Z. Wang, Z. Liu, X. Lv, and J. Y. Kim, “Applying Deep Learning Technologies to Evaluate the Patent Quality with the Collaborative Training,” *Wirel. Commun. Mob. Comput.*, vol. 2021, 2021.

[11] S. M. H. Bamakan, A. B. Bondarti, P. B. Bondarti, and Q. Qu, “Blockchain technology forecasting by patent analytics and text mining,” *Blockchain Res. Appl.*, vol. 2, no. 2, p. 100019, 2021.

[12] L. Walter, N. M. Denter, and J. Kebel, “A review on digitalization trends in patent information databases and interrogation tools,” *World Pat. Inf.*, vol. 69, p. 102107, 2022.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details