Advanced Engineering Informatics 43 (2020) 101027

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

Intelligent compilation of patent summaries using machine learning and natural language processing techniques

Amy J.C. Trappey[a], Charles V. Trappey[b], Jheng-Long Wu[c,*], Jack W.C. Wang[a]

[a] Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan
[b] Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan
[c] School of Big Data Management, Soochow University, Taipei, Taiwan

ARTICLE INFO

Keywords:
Artificial intelligence
Machine learning
Natural language processing
Deep learning
Patent analysis

ABSTRACT

Patents are a type of intellectual property with ownership and monopolistic rights that are publicly accessible published documents, often with illustrations, registered by governments and international organizations. The registration allows people familiar with the domain to understand how to re-create the new and useful invention but restricts the manufacturing unless the owner licenses or enters into a legal agreement to sell ownership of the patent. Patents reward the costly research and development efforts of inventors while spreading new knowledge and accelerating innovation. This research uses artificial intelligence natural language processing, deep learning techniques and machine learning algorithms to extract the essential knowledge of patent documents within a given domain as a means to evaluate their worth and technical advantage. Manual patent abstraction is a time consuming, labor intensive, and subjective process which becomes cost and outcome ineffective as the size of the patent knowledge domain increases. This research develops an intelligent patent summarization methodology using artificial intelligence machine learning approaches to allow patent domains of extremely large sizes to be effectively and objectively summarized, especially for cases where the cost and time requirements of manual summarization is infeasible. The system learns to automatically summarize patent documents with natural language texts for any given technical domain. The machine learning solution identifies technical key terminologies (words, phrases, and sentences) in the context of the semantic relationships among training patents and corresponding summaries as the core of the summarization system. To ensure the high performance of the proposed methodology, ROUGE metrics are used to evaluate precision, recall, accuracy, and consistency of knowledge generated by the summarization system. The Smart machinery technologies domain, under the subdomains of control intelligence, sensor intelligence and intelligent decision-making provide the case studies for the patent summarization system training. The cases use 1708 training pairs of patents and summaries while testing uses 30 randomly selected patents. The case implementation and verification have shown the summary reports achieve 90% and 84% average precision and recall ratios respectively.

1. Introduction

Intellectual property (IP) is becoming more complex for companies to manage given the overwhelming number of patents being registered in multiple countries and are integrated and incorporated into new products and services as standard essential licensed sets used to effectively keep up with the increasingly short product lifecycles of industrial and consumer goods. Mismanagement of a company's IP and patents can prevent new products being launched on time or cause disputes over IP rights. Thus, during the entire product lifecycle (design and development, manufacturing, sales and marketing, disposal, or revitalizing), continuous patent search and analysis becomes critical tasks for sustaining business. Simple patent queries yield surprisingly large amounts of unstructured documents that are difficult for even the most competent teams of engineers to abstract, summarize and comprehend the essential knowledge efficiently, accurately, and consistently [1]. In Artificial Intelligence (AI) domain, Machine Learning (ML), Deep Learning (DL) and Natural Language Processing (NLP) techniques developed for text document processing and understanding help to organize the syntactical relations and structure of language for automatic processing and learning. These AI-based techniques require validation and extensive reliability testing in the fields of applications. For example, legal e-discovery and patent document analysis is still questioned when used in a court of law as the basis for litigation and

* Corresponding author.
E-mail addresses: trappey@ie.nthu.edu.tw (A.J.C. Trappey), trappey@faculty.nctu.edu.tw (C.V. Trappey), jlwu@gm.scu.edu.tw (J.-L. Wu).

https://doi.org/10.1016/j.aei.2019.101027
Received 26 November 2018; Received in revised form 16 June 2019; Accepted 25 November 2019

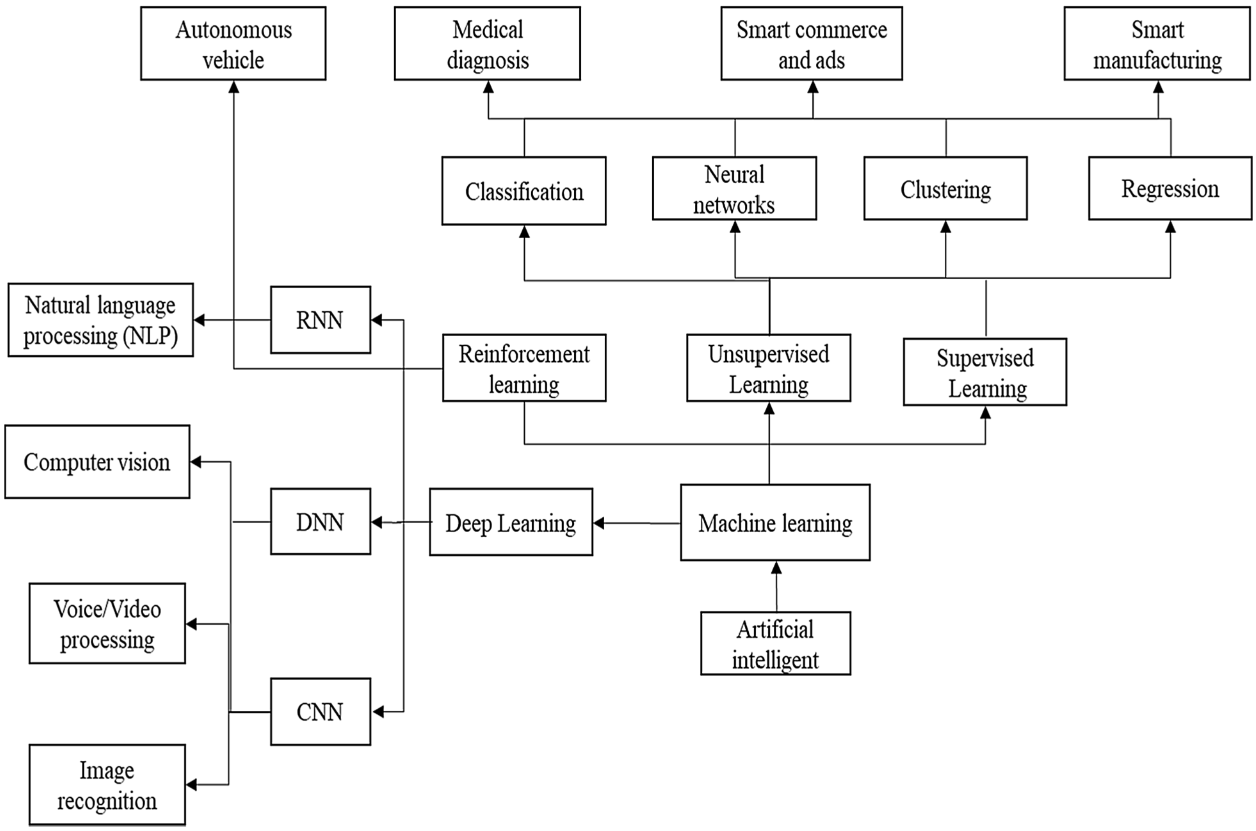

Fig. 1. Domain ontology of machine learning and deep learning methodologies.

evidence [2]. Because there is a lack of efficient patent knowledge extraction and understanding, enterprises that adopt licensed IPs for developing competitive short life cycle products, often fail to understand licensing limitations, outstanding arguments over ownership, usage and compliance limitations, and market access. As the most typical example, smart phones are incrementally developed through patent licensing since new models may consist of a collection (stack) of over a quarter of a million separate patents. New consumer communication products depend on the incremental extension and advancement of working platforms (instead of re-inventing every technology used) to meet the demands of the short product life cycles of these goods and services. The ability to verify the rightful claims of IP rights requires scrupulous attention because short life cycle technologies require a greater understanding of what is being offered in the licensed patents. The technological terms and claims, which potentially could sustain profitability of the new product in a globally and aggressively competitive marketplace are commonly misunderstood and lead to legal disputes [3]. Thus, summarization of the key patent claims and terms of patent documents is an indispensable capability for companies that need to abstract large number of patents (e.g., standard essential patents for mobile phone designs) into concise and structured summary reports in a limited time frame [4,5].

In order to automatically summarize large numbers of patents, each document must be decomposed to obtain the features embedded in the document using soft computing algorithms. The algorithms generate a summary for given patent or a set of patent documents. The semantic meaning, key features, and relationships are described and presented as a structured summary report for humans to review and reference [6,7]. The automatic summarization approach may be extractive or abstractive from a given document. The extractive method assembles the most meaningful paragraphs, sentences or phrases taken from the document (s). The abstractive method generates a summary using sentences and phrases which may or may not be entirely extracted from the original document [8,9]. The goal of automatic summarization is that the knowledge content of the original document is automatically extracted without significant loss of meaning. A patent summary report should include a categorizable name, the patent's claims, and an easily comprehensible description of the novelty, advantage, and use. This research (1) provides the AI-based methodology for a system to integrate abstractive and extractive summarization solutions (2) uses automatic summarization techniques to reduce the size of patent documents without the loss of key meanings, and (3) compiles Chinese or English patents into summary reports to test and validate that the summarization system is capable of handling multi-language patent documents.

2. Literature review

In this section, we review the literature of the development and application trends in artificial intelligence, natural language processing, and automatic text summarization. These methodologies and algorithms are the key enabling technologies used for intelligent patent summarization.

2.1. The new era of artificial intelligence (AI 2.0)

The term artificial intelligence, AI, was first established by scholars around 1956. The definition of AI at that time was that machines can perform certain functions, such as understand, learn, and make complex decisions, just as well as intelligent humans [10]. More than thirty years passed and the results of AI scholars, incorporating human expertise into machines to mimic human knowledge construction in general fields like the encyclopedia of knowledge are far from ideal [11–13]. At the turn of 21st century, the rapid development and integration of information technologies including big data analytics, Internet collective intelligence, human-machine and machine-machine hybrid intelligence, and autonomous and ubiquitous intelligent agents, has dramatically transformed AI into a promising evolutionary field called AI 2.0 [14].

After the early AI development era, a popular AI method was neural network modeling as a form of machine learning (ML). These are models of consciousness, which mimic human neuron and synaptic activities for concept understanding and learning. Particularly after the 1980's, many neural network model types and modeling algorithms were developed for general and specific purposes and applications with success [15]. Computing power has grown exponentially, enabling the less expensive use of Deep Learning (DL) algorithms for large numbers of applications that evolved from fundamental neural network principles after the year 2010. DL is able to process more complex tasks which generally were impossible to process using conventional ML. DL simulates human brain neurons with multiple layers to process complicated predictions and decision making [16–18]. Fig. 1 is a brief ML ontological schema, constructed using the cited ML and DL literature in the references. The focus of this research is in patent document summarization. Thus, the ML and DL literature relevant to solving natural language understanding, text mining, and text summarization are reviewed in the following two sub-sections.

2.2. Natural language processing

The goal of natural language processing (NLP) is to convert human language into a formal and machine-readable representation. NLP operates in two ways. One way uses inputs from human's natural language expressions to create machine-readable representations. The other way is for a machine to generate human language expression using internal computer codes. NLP applications include information extraction, machine translation, text summarization, keyword search and human-computer interfaces using natural languages. There are numerous traditional NLP applications in the literature [19,20]. The entity recognition method detects and extracts valuable concepts in patent documents that are commonly searched for among multiple users of the invention. Experimental results shown that the NLP technique is an efficient approach [21]. Milanez et al. [22] proposed new patent indicators using text mining of the patent claims. The experimental results shown that patent claim indicators have outperform patent titles and patent abstracts. Clercq et al. [23] applied Latent Dirichlet Allocation (LDA) and term frequency – inverse document frequency to successfully identify topics which signal the emergence of technical concepts in the domain of food waste, anaerobic digestion, and biogas patents.

Many scholars focus their NLP research and adopt different advanced AI techniques as needed, particularly deep learning algorithms, to improve functional capabilities in different knowledge domains. In order to capture the semantics from a big collection of text documents, neural network modeling incorporating word embedding vectorization is a proven machine learning technique in NLP [24]. Word embedding considers the dimensionality of all words from the original high-dimensional space document to construct a lower-dimension document using continuous vector space to construct a set of features for each word. Distributed representations of the set of features are made using a neural network-based language model that enables machines to learn the distributed representation of words and achieves the purpose of reducing the word space dimension. This breakthrough application was reported by Whitehead and Johnson [25] using word embedding to analyze semantic similarity associated with cancer therapies among US patent documents. The most popular modules for producing word vectors are word2vec by Google, FastText by Facebook, and Glove by Stanford.

2.3. Automatic text summarization techniques

Automatic text summarization can be divided into types and categories. For example, based on how a summary is formed for given text document(s), the approach used may be extractive or abstractive [26]. Extractive summarization approaches, e.g., word, sentence, and paragraph extractions, are used to extract and organize important words, sentences, and paragraphs from the original document(s) to form a summary. Abstractive summarization generates a summary to briefly describe the key elements of the original document(s) [27]. Regardless of the summarization type, the trend is toward embedding automatic text summarization into machine learning algorithms [28]. Severyn and Moschitti [29] proposed the Convolution Neural Network (CNN) model to rank and rematch the sentences. CNN applies a series of methods to input sentence vectors including convolution, non-linearity and reorganization operations. The purpose of the convolutional layer is to extract the relationships between the word order of sentences so that reorganization of individual words with new input materials create a meaningful summary. Yao et al. [30] presented a new Deep Neural Network (DNN) method for processing documents into abstracts. The DNN trains the sentence using a descending gradient method to optimize the sentence output weight. The K-Nearest Neighbor (KNN) method classifies the relevance of sentences and generates the final summary result.

A patent document consists of critical technical and business intelligence pertaining to the advances, uniqueness, and application values of the legally protected invention. The data, information, and knowledge extracted from relevant patents, such as assignees, titles, chronological issue dates, independent and dependent claims, and the prior-arts (patents and references) can be analyzed and synthesized by R&D and management teams to access a company's competitive market position and IP strategies [31]. An intelligent patent analysis and summarization has been proposed to generate summarization from processes which include entity recognition, lexical chain identification, segmentation description alignment, and composition detection. The system is a standalone engine to summarize single language patents for English language patents [32]. Girthana and Swamynathan [33] proposed a semantic query-based summarization system which applied a restricted Boltzmann machine method. The system can generate a summary when give the domain knowledge and the patents retrieved and related are provided by a patent analyst.

There are large amounts of information and knowledge embedded in patent documents which are an overwhelmingly challenge for quick knowledge discovery and learning. Thus, patent text summarization is a valuable research endeavor for IP and patent analytics which depends on automating and abstracting critical summaries from large collections of text documents.

3. Methodology

The patent summary method of this research is shown in Fig. 2. For the initial step, a set of the specific patents in a given topic (sub-domain) are collected from a large, globally available, archived patent collection. A set of specific patents are input to generate the final summary report related for a given topic. The summary report includes input patents, the summary result, and key phrases. The gold target of system is abstracts provided by Derwent Innovation (DI) [34] including novelty, use, and advantage. In order to train a smart summarization model from training examples, the research collects raw patent documents from the DI database. The title, abstract, and claims from the original patent document acts as the model input, and the patent novelty, use and advantage provided by the DI abstract are the model output are used for model training. The summary report generated by our proposed summary system includes novelty, usability, and technology advantage.

In this research, we propose a novel patent summarization model that includes NLP, sequence to sequence with attention, and word embedding algorithms. The proposed model includes four parts: raw data collection, text pre-processing, smart summarization model training, and model evaluation. The flowchart of the proposed model is

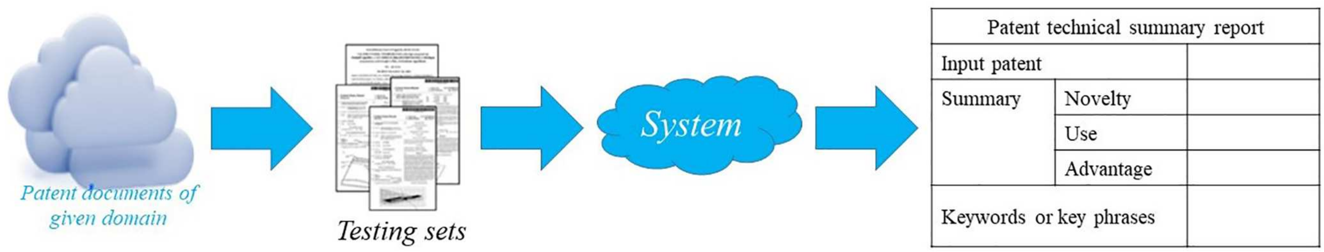

Fig. 2. Flowchart for generating the summary report.

shown in Fig. 3 and overall processes are described as follows:

(1) Raw and domain data collection: Collecting a set of the patent documents which include Chinese or English patents and is based on a given topic domain. These specific domain patent documents are split into the training set and testing set for model training and performance evaluation.
(2) Text pre-processing: The text in each patent document undergoes standard reductive cleaning to reduce ambiguity and duplication. Tokenization splits the sentence into a set of the token words according to space for English words and the Jieba[1] tool for Chinese. These token words are converted to lowercase and the stop words are removed.
(3) Smart summarization model training: The Sequence to Sequence with Attention (SSWA) model is used to encode the text information and then decode the encoded information to form a sentence in natural language. These two steps (encoding and decoding) of SSWA learn from the patent documents that are used as inputs. SSWA uses the input data (token words) and target data (summary words) to develop the model. The target data used as summary words are reviewed by professional patent engineers that have access to the Derwent Innovation platform for clarification. The input token words include the title, abstract and claim from the original patent document. The model uses a gradient-based optimization method to train the data and generate the best summarization model. After model training, the output is used to discover the attention words using SSWA and a sentence with the most attention words is retrieved as an important sentence.
(4) Model evaluation: The testing set is used to evaluate the summary generation model. This research uses the Recall-Oriented Understudy for Gisting Evaluation based (ROUGE-based) [35] metrics to evaluate the quality and accuracy of the summary results of the proposed model.

3.1. Raw and domain data collection

In this step, large sets of patent documents are collected from the DI platform. These patents include Chinese or English language documents. The sub-domain specific patent set is collected from raw patent documents in a given topic (sub-domain) based on intelligent algorithms such as LDA, Doc2vec, and classification. This research uses Doc2vec to collect several sub-domain patent sets. A conceptual view of the raw data collected is shown in Fig. 4. Doc2Vec is used to collect specific domain patents using the following steps:

(1) All patents of the sub-domain are used to learn the patent (document) terms that are embedded in sentences using the Doc2Vec algorithm.
(2) The algorithm computes the central term embedded in the patents.
(3) The most similar and related patents are identified using the cosine similarity.

[1] Jieba: Chinese word segmentation module. https://github.com/fxsjy/jieba.

3.2. Text pre-processing

For the sub-domain patents, the text is split into token words. This paper uses the space symbol for English documents, and the Jieba tool for Chinese documents. These token words are used for additional processes including conversion to lowercase and to remove stop words and punctuation. Fig. 5 depicts an example of text pre-processing for Chinese and English patents.

3.3. Smart summarization model training

This research separates each sub-domain patent into input data and target data. The input data includes the title, abstract, and claim from an original patent document. The target data includes the Derwent World Patents Index (DWPI)'s patent abstract (called patent **quick view**) collected from the DI platform. The DWPI patent abstracts are manually written by a team of domain experts in the DWPI writing center. The DWPI abstracts are published with a set reliable editorial procedure to ensure accuracy in summarizing the patent novelty, use, and advances [34]. The detailed training algorithms for smart summarization are described in Sections 3.3.1 and 3.3.2.

3.3.1. Sequence to sequence with attention (SSWA)

The SSWA model in this research includes an encoder and a decoder. Input sequences such as title, abstract and claim are converted to a fixed length using a zero-padding approach before input to the encoder. The encoder uses data as a vector to represent the patents. The decoder uses prior representation by the encoder as well as the target sequence (summary) to predict future summary words. The attention mechanism is used to focus attention on sequences represented by the encoder for each decoded word in the decoder. The overall process in the SSWA is shown in Fig. 6.

- Encoder

For the encoder step, the sequence vectors are used as input data to represent hidden vectors unique to the RNN model. For example, a sentence selected from the original patent content is a sequence of words $X = [e(x_1), e(x_2)\cdots, e(x_n)]$, where e denotes the word embedding matrix for each word x_j. The word embeddings matrix e is also jointly trained with neural networks. In order to accelerate training, this paper uses the zero-padding approach to insure the sequence of words of each document are of fixed length for accelerate training. The research uses a bi-directional long and short-term memory model to extract hidden features from the word embedding of a sentence. There are two directions to encode the hidden features including a forward-LSTM model $\overrightarrow{f_e}$ and a backward-LSTM model $\overleftarrow{f_e}$. Two hidden layers in the model map sequence the word vector hidden states so each word vector x_j is mapped into each hidden vector h_j^{en}, and is defined as:

$$h_j^{en} = [\overrightarrow{h_j^{en}}; \overleftarrow{h_j^{en}}] \tag{1}$$

$$\overrightarrow{h_j^{en}} = \overrightarrow{f_e}((e(x_j), \overrightarrow{h_{j-1}^{en}})) \tag{2}$$

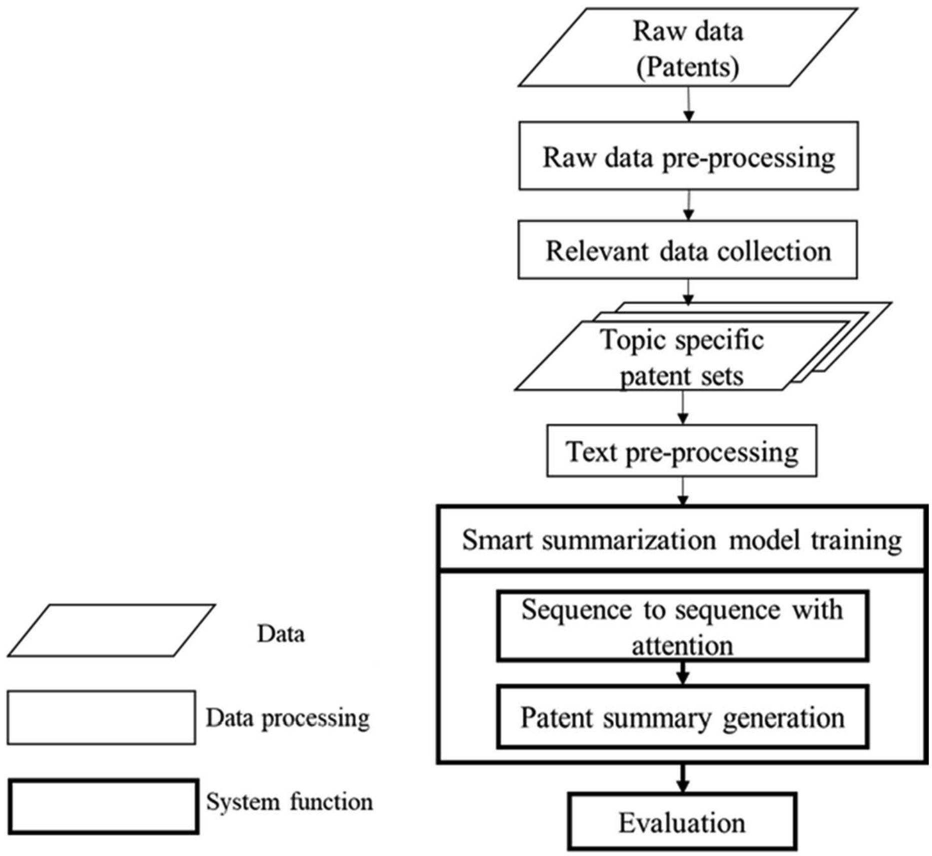

Fig. 3. Flowchart of smart summarization model.

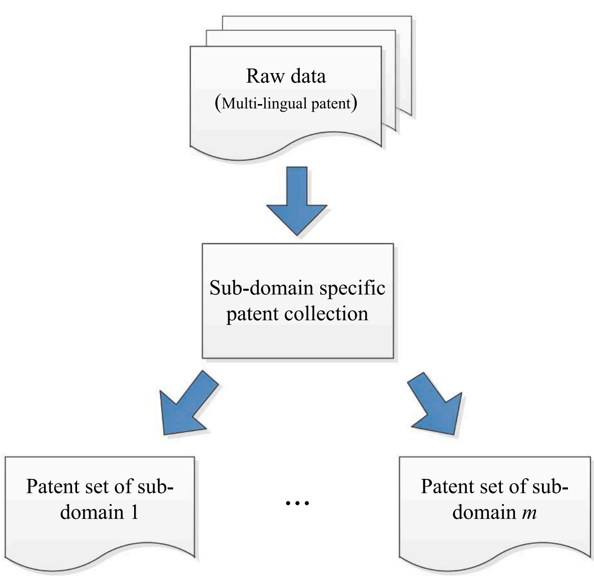

Fig. 4. Flowchart of sub-domain patent collection.

$$\overleftarrow{h}_j^{en} = \overleftarrow{f}_e((e(x_j), \overleftarrow{h}_{j-1}^{en})) \tag{3}$$

where $[\cdot]$ denotes the concatenation operation. The $\overrightarrow{h}_j^{en}$ and $\overleftarrow{h}_j^{en}$ are the forward hidden features and backward hidden features, respectively. The $j-1$ denotes the previous hidden state (vector) of the LSTM encoder at the jth step.

- Decoder

For the decoder step, the attention mechanism captures the hidden states before applying word generation to provide more weight for each decoded word. An LSTM model is used to predict output wordy_i. From the attention mechanism, each h_j^{en}assigns to its attention weight a_j for the ith output word of decoder so that the vector c_i is multiplied by the weighted sum as defined by:

$$c_i = \sum_j h_j^{en} \times a_j \tag{4}$$

$$a_j = \frac{\exp(\mu_j)}{\sum_j \exp(\mu_j)} \tag{5}$$

$$\mu_j = f_a(h_j^{en};\ h_{i-1}^{de}) \tag{6}$$

where the d_{i-1} denotes the hidden state of the LSTM model for the predicted word y_i. The variable f_a is a linear layer used to estimate the weight of attention assigned to the output layer. The LSTM model variable f_d predicts the hidden state of the output word y_i. The ith hidden state of the decoder is defined as:

$$h_i^{de} = f_d([c_i;\ e(y_{i-1})],\quad h_{i-1}^{de}) \tag{7}$$

where the $e(y_{i-1})$ denotes the word embeddings of the previous predicted word i-1. The output vector h_i^o and the selected word y_i for each output of decoder are defined as

$$h_i^o = f_o(h_i^{de}) \tag{8}$$

$$y_i = argmax(h_i^o) \tag{9}$$

where f_o is a model parameter used to predict for predicting terms of input. The notation y_i is the output of the SSWA model and all y_i for $i = 1, \ldots, m$ are the label of key terms highlighted in Fig. 7. The decoder is trained using the content vector c_i with word embedding of the previously predicted words y_{i-1} to predict the hidden vector of the

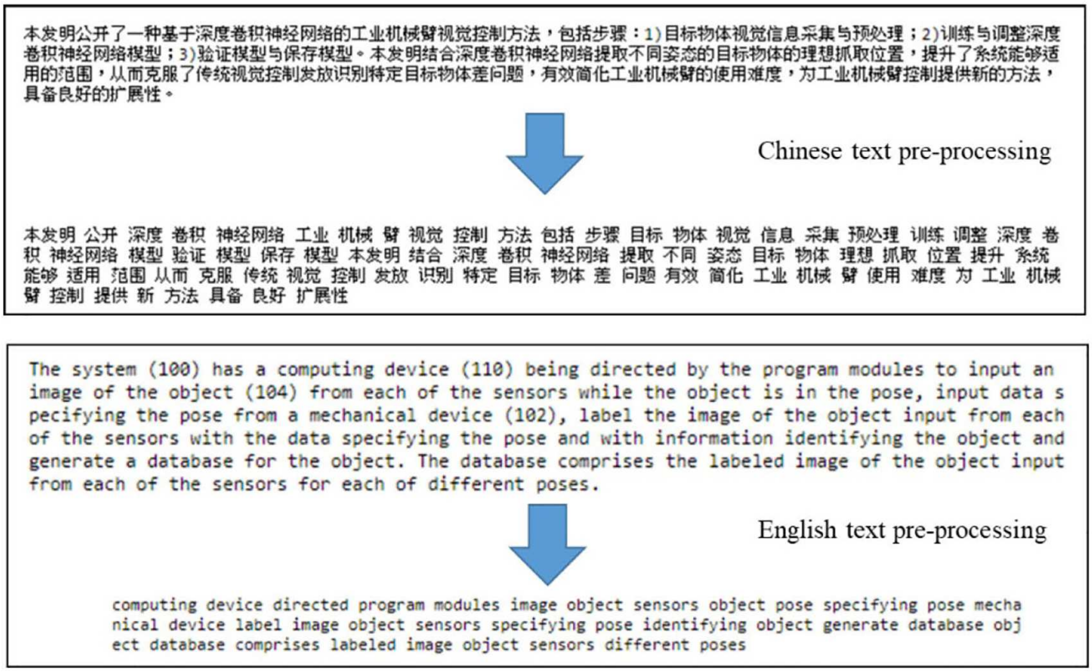

Fig. 5. An example of text pre-processing for Chinese and English patents.

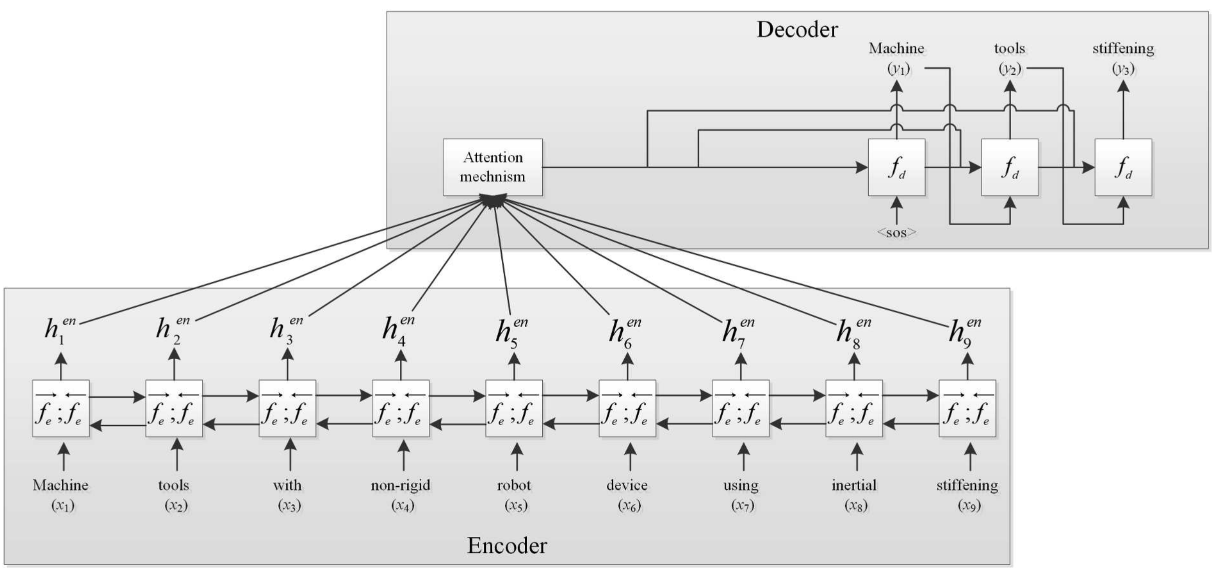

Fig. 6. Overall SSWA processes.

current output word y_t. The hidden vector of each output word is used to predict the current output word. The model input sequence sentence and the output sequence of the target sentence is used to define the relationship needed to generate a summary report. Prior to training the model, hyper-parameters such as number epoch, batch size, hidden size, and learning rate require definition. In this research, we propose the gradient descent method with the maximum likelihood loss function to converge model parameters to find an optimized patent summary model.

3.3.2. Summary generation

After training the SSWA model, the output predicts the best performing output words. These output words are not used to form the patent summary. A new algorithm for patent summary construction is proposed which combines abstractive and extractive summarization techniques. The process of summary generation follows:

(1) All words are split into several sentences according to end of the original sentence punctuation.
(2) The attention word aw^* which has the highest attention weight in the input sequence for each output word (decoded word) identified. The highest attention weight is defined by the $aw^* = \arg_j a_j$ in ith output decoder word.
(3) A count of the number of attention words in each sentence is assigned.
(4) The attention words with the highest weights in a sentence form the output summary.

The concept of summary generation is shown in Fig. 7. There are 20 target words for prediction. Therefore, the number of attention words is equal to 20 because each output word has the highest attention word weights of the input words. The input words are reduced to 16 attention

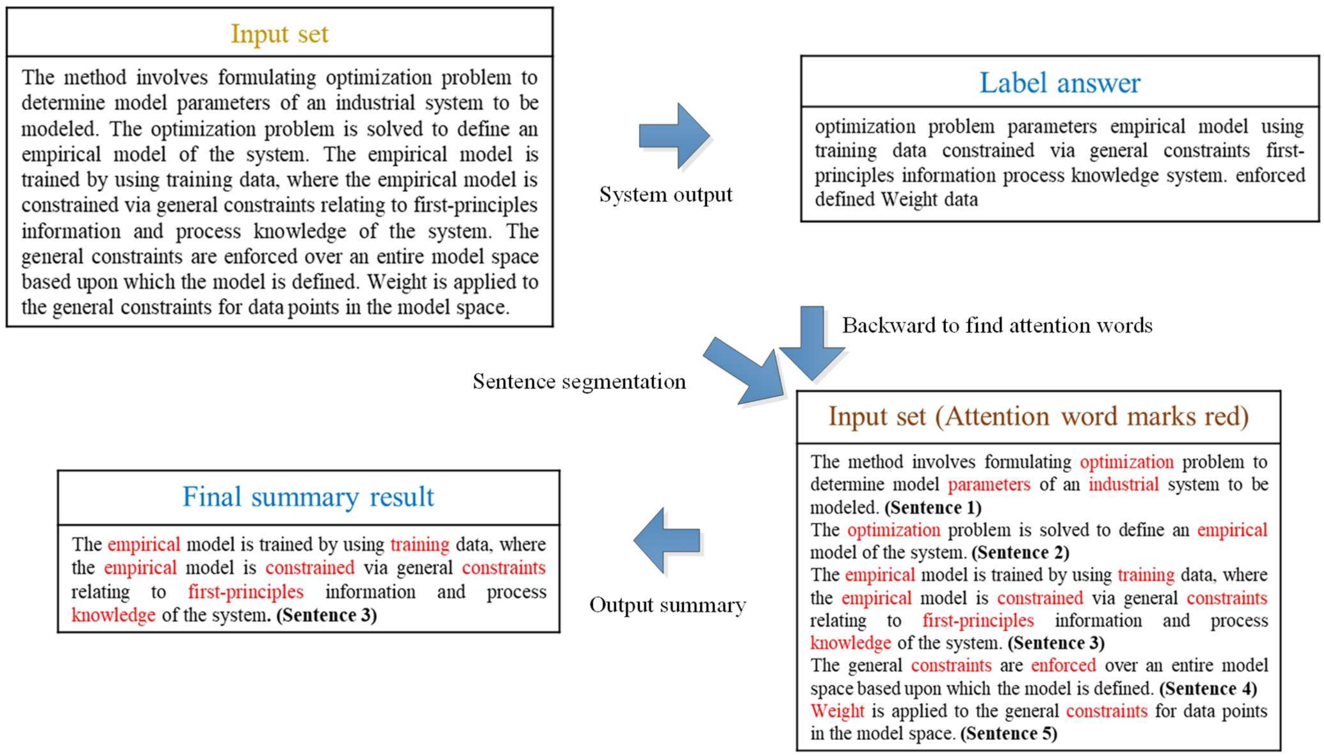

Fig. 7. An example of summary generation on smart machinery case.

words since some of the attention words are the same words. The third sentence at end of right side represents the output summary because it has eight attention words compared to the other sentences with fewer attention words.

3.4. Model evaluation

The ROUGE metrics are frequently used to measure the text generation quality of natural language such as text summaries and machine translation [35,36]. The output summary by our proposed model is compared with the target summary of human compilation such as n-grams and word sequences, and word matching has been used to verify the similarity of system results to human natural language. In this research, the raw patents or sub-domain patents are collected from the DI platform. The Quick View sample obtained provides information such as patent novelty, use, and advantage. The Derwent patent information is used as training material for the summary system and provides a reference for evaluating the performance between the target summary and the output summary. The recall indicator of the ROUGE metric measures how much of the target summary is captured or recovered by the output summary. The Recall indicator calculation is shown in Eq. (10).

$$\text{Recall} = \frac{\text{Number of overlapping sentences}}{\text{Total sentences in reference summary}} \tag{10}$$

A machine generated summary report can be extremely long, capturing all words in the target summary. Superfluous words may exist in the output summary words which makes the summary unnecessarily verbose. Therefore, this research considers the precision indicator metric of ROUGE to measure the summary generation quality. Precision measures how many words of the output summary are relevant or needed. The Precision indicator calculation is shown in Eq. (11).

$$\text{Precision} = \frac{\text{Number of overlapping sentences}}{\text{Total sentences in system output summary}} \tag{11}$$

4. Case study

The case study is based on AI techniques found in smart machinery patents. The Derwent patent index provides the technical summaries for the case corpus. Section 4.1 introduces the knowledge ontologies of smart machinery technologies and, specifically, the AI-based techniques deployed for smart machinery. The knowledge ontologies, based on literature and verified with domain engineers, are constructed in Section 4.1. Section 4.2 describes DI patent search queries in the field of smart machinery to select the domain patents for summarization system training and testing. Section 4.3 presents the technical summary compilation results and evaluates the results of each technical summary. Finally, Section 4.4 provides a brief discussion in regards to the outcome and implication of the case study.

4.1. Smart machinery ontology construction

The concept of smart machinery was first proposed in the 1950s to describe machines that have the ability to monitor and control multiple functions. Smart machinery provides higher productivity and quality with near-zero failure during the operating process [37]. Smart means that the machinery can be program controlled, such as the Computer Numerical Control (CNC) systems which use a computer to monitor and control the manufacturing process.

Dai and Wang [38] have defined advanced smart machinery as a machine which has automatic monitoring and analysis of processes, environment and the affiliated information. There are five attributes of smart machinery: information perception, behavioral decision making, quality assessment, self-learning, and network communication. Cao et al. [39] compare and contrast smart machinery with traditional machine tools and note that smart machines have the features of autonomy, self-learning and compatibility. The design and manufacture of smart machinery requires the development of smart technologies including sensing intelligent, control intelligence and decision-making intelligence. Please refer to [37–39] for detailed definitions of the three sub-domains (sensor, control, and decision-making intelligences). The

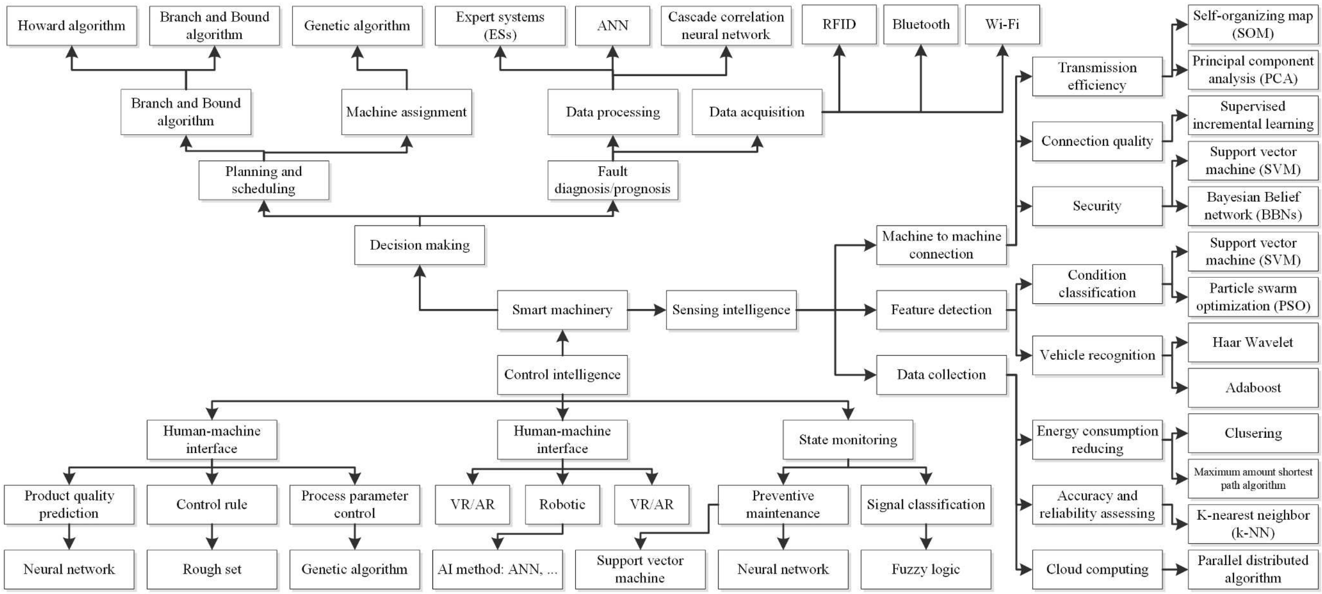

Fig. 8. Ontology of smart machinery related technologies.

brief description of the sub-domains follows:

- Sensing intelligence act as the eyes of the smart machinery. Sensors of all kinds should be able to detect the operating status of the process using measurement signals such as vibration, torque, temperature, and power.
- Control intelligence resemble the hand and feet of a smart machine tool and compensates for sensor measurement errors in perceptual intelligence. Control intelligence uses detection devices to control actions to ensure the reliability of the smart machine during operation.
- Decision making intelligence is like the brain of smart machine tool. Data processing, feature extraction, and intelligent detection, AI and other algorithms are used to improve decision making.

The ontology of smart machinery defined and the AI aspects of smart machinery are shown in Figs. 8 and 9.

4.2. Patent search strategy

Using the ontology of smart machinery and AI algorithms, the patent documents are collected from the five patent offices including Europe (EP), the World Intellectual Property Office (WIPO), China (CN), Japan (JP), and the United States (US). The search time zone is limited to five years (2013–2017). All patent data based on the ontology of smart machinery are collected from the DI platform. The smart search function of the DI platform uses search strings taken from the combination of AI techniques and smart machinery for three major sub-domains including control intelligence, intelligent decision-making, and sensor intelligence. Further sub-domains with AI techniques for more specific and detailed search strategies are shown in Table 1. The numbers of three sub-domain patent publications on control intelligence, sensor intelligence, and intelligent decision-making of smart machinery domain are 487,613 and 608, respectively (a total 1708 training patents). Table 2 shows that more specific search

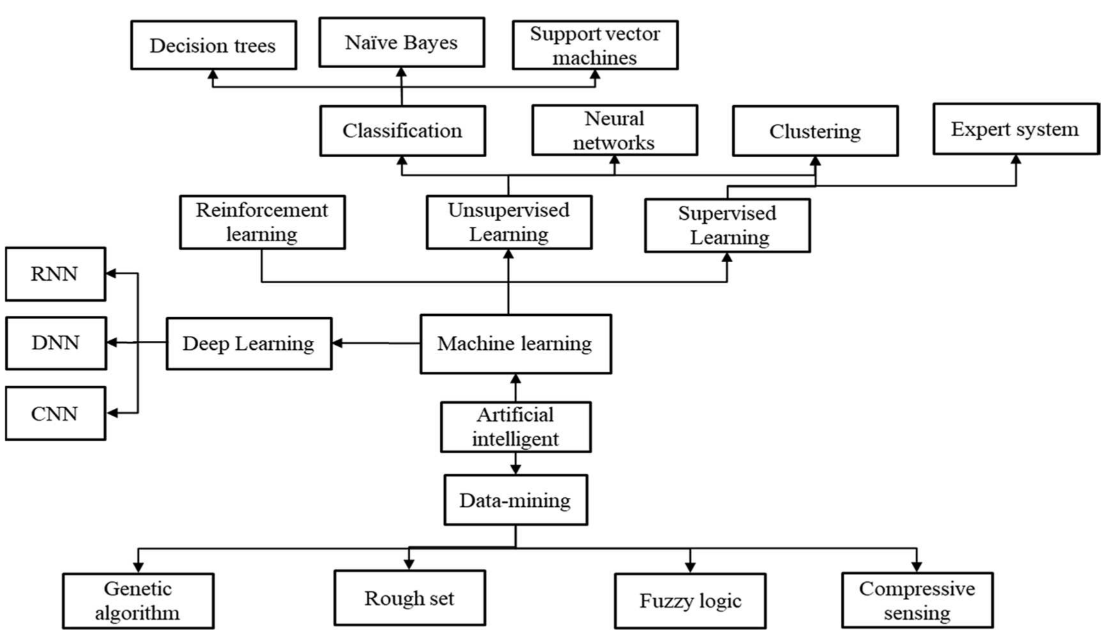

Fig. 9. Ontology of AI techniques deployed for smart machinery.

Table 1
Smart machinery multi-country search query.

Function	Value
Smart search topic	"ROBOT" "CNC" "MECH*" "SMART MACH*" "DECISION MAK*" "DEEP LEARNING*" "MACHINE* LEARN*" "ARTIFICIAL INTELLIGEN*" "CONTROL* INTELLIGENCE*" "SENS* INTELLIGENCE*"
Search time zone	From 2013 to 2017
Reign Limit	US, EP, JP, CH, WIPO

terms are needed to collect the most relevant patents from the three domains. The control intelligence domain uses search string terms like remote, and control. The sensor intelligence uses terms such as sensor and collection. The intelligent decision-making includes terms such as planning and scheduling. After the summarization system models are trained, the summary compilations of 30 relevant patents, as case study, are conducted to test the system. These related US and China patents from three sub-domains (completely different from the original training patents) are randomly selected to verify the accuracy of the summary reports. Tables A1–A3 (in Appendix A) list the 30 smart machinery patents used for summarization testing and verification.

4.3. Smart machinery patent technical summary compilation

The research collected patents for smart machinery domains of control intelligence, sensor intelligence, and intelligent decision-making from China and the US patent office. The features of filtered patents include publication number, assignee, and title. Each patent can be divided into novelty, use, and advantage so that the content of the target summary for a set patents can be quickly formed.

4.3.1. Evaluation result

Table 3 shows the patent technical summary result for the three sub-domains of smart machinery. The average precision and recall ratios are 90% and 84% respectively, with the highest precision and recall ratios being 96% and 93% for patent summaries in fault diagnosis domain. The fault diagnosis domain is more specific and, thus, consistency and similarity among patent documents are high. Subsequently, the summarization model is better trained with training document set. The extra sentence generated by the proposed system may have introduced errors between the summary sentences compared with the target summary sentences. The maximum number of error summary sentences is three, which is a small number when considering the number of inputs and outputs of the proposed system. Therefore, the proposed system is reliable to summarize information for multiple patent documents. The system automates the manual Derwent innovation system which requires patent engineers to generate summaries report of the novelty, use, and advantage of a patent.

4.3.2. Intelligent decision-making patents technical summary compilation

In this section, additional details from the two cases with the best and worst performance from the sub-domain patents including fault diagnosis, and planning and scheduling are analyzed.

4.3.2.1. Fault diagnosis (Example 1). There are five relevant patents for fault diagnosis from the intelligent decision-making domain for the patent summarization. Tables 4 and 5 show the technical summary contents for these five patents. The novelty summary indicates the fault and failure condition of a machine or robot diagnosis is based on AI algorithms to state prediction and self-learning to execute a specific action. The use summary indicates that the invention technology is applied to manufacturing and production line or shop floor control work such as material handling and assemble material inspection. The advantage summary emphasizes the invention technology that can predict the fault condition and solve the problems related to faulty machine performance. These patents improve the efficiency, accuracy, and quality of the production line and manufacturing environment. Keywords include articulated robot, abnormal state, fault detection, wafer, manufacturing process, production line, and diagnostics, which represent the key technology and purpose of the summary result.

Table 3
Performance comparisons of technical summaries for smart machinery.

Sub-domain	Ontology level	Language	Precision	Recall
Control intelligence	2	English	93.1%	90.0%
Control intelligence	2	Chinese	88.4%	82.1%
Sensor intelligence	2	English	88.8%	80.0%
Sensor intelligence	2	Chinese	88.5%	84.6%
Fault diagnosis	3	English	96.3%	92.8%
Planning and scheduling	3	English	85.7%	75.0%

According to the evaluation result, there are 28 sentences in the target summary, and the output summary has 27 sentences. There are 26 sentences covered by the output summary of the proposed system and target summary of DI platform. The precision and recall are 96% and 93%, respectively. Two essential sentences were not output by the proposed system. However, the summary result of the case yields the best performance compared to the other domains. The missing sentences indicate the purpose of the algorithm to predict the fault condition and occurred timing. The extra sentence describes the algorithm and purpose of patent US20170210009A1. The summary result of the fault diagnosis domain has the highest quality given that this domain is more specific and consistent.

4.3.2.2. Planning and scheduling (Example 2). Another example is to sue four relevant patents of the planning and scheduling domain for the summarization test. The patents cover smart robotic and machinery technology for design and manufacturing suppliers, industrial safety, production management, and improved environment solution providers. Tables 6 and 7 show the technical summary contents and system performance for these four patents.

The novelty summary indicates a focus on production planning, scheduling or order planning based using AI algorithms (genetic

Table 2
Search queries for AI-based smart machinery sub-domains.

Domain	Addition search condition
Control intelligence	(remote adj control*) or (human–machine adj interface*) or (state adj monitor*) or(artificial adj neural adj network) or (neural adj network) or (support adj vector adj machine) or (fuzzy adj logic)
Sensor intelligence	(machin* adj to adj machin*) or (feature adj detect*) or (data adj collect*) or (support adj vector adj machine) or (cluster*) or (adaboost)
Intelligent decision-making	(machin* adj diagnosis) or (machin* adj prognosis) or (planning) or (schedule*) or (ANN) or (artificial adj neural adj network) or (neural adj network) or (expert adj system)

Table 4
Technical domain summary report (of intelligent fault diagnosis patents).

Type	Content
Novelty summary	● *The apparatus has processors for determining presence or absence of a failure in an articulated robot based on an output value of each encoder and an output value of each motor.* ● *A recording unit stores the output value of each encoder and the output value of each motor, where the set of motors is stopped when one of the output values of the set of motors exceeds a threshold, and the apparatus determines presence or absence of the failure based on the output value of each encoder and an output value of each motor, which are stored after the set of motors is stopped* ● *The method involves recording movement information of joints based on an output of detection devices by a control device sb.* ● *A presence or an absence of a failure is determined in an articulated robot when detecting an abnormality in an operation of the articulated robot by the control device based on the movement information recorded in period from before detection of the abnormality until detection of the abnormality.* ● *A failure portion of the articulated robot is specified if the robot determines that the failure is present in the articulated robot in the step of determining.* ● *The method involves detecting fault in a failed arm of robotic arms.* ● *Non-failed arms of the robotic arms are placed in safe state if the fault is to be treated as the system fault.* ● *The method involves generating a defect model by synthesizing the defect template stored in a template storage unit and instruction image of wafer.* ● *The classification type of defect is set by a setting unit for calculating the feature values of the defects in the defect model.* ● *The defect of the wafer is classified from the relation between the feature value of defects and classification types of the defects, based on feature value of the defect.* ● *The method involves collecting original data of a quality characteristic during manufacturing process to simplify, calculate and organize a data necessary.* ● *The control chart is used for verifying and confirming a process quality.* ● *Judgment is made to check whether manufacturing process is monitored continuously by using the control chart if the process quality is in an abnormal state or the manufacturing process does not monitor continuously by using the control chart if the process quality does not in the abnormal state.*
Use summary	● *Diagnosis apparatus for diagnosing failure i.e. interference with an object, of an articulated robot at a production line during assembling of various products.* ● *Method for controlling robot system, involves determining presence or absence of failure in articulated robot, and specifying failure portion of robot if robot determines that failure is present in articulated robot in step of determining.* ● *Method for realizing fault reaction, fault isolation and graceful degradation in a robotic system claimed in manufacturing, construction, hazardous material handling applications.* ● *Method for classifying defects of wafer used for manufacturing semiconductor device.* ● *Multi-element quality diagnostic classifier fuzzy support vector machine manufacturing method.*
Advantage summary	● *The apparatus estimates a load state of each joint accurately and determining a possibility of damage in each joint with less error when an abnormality such as an interference with the object occurs in the articulated robot.* ● *The apparatus accurately estimates a load state of each joint and evaluates a possibility of damage with less error when movement information before detection of an abnormality cannot be used.* ● *The method enables estimating a load state of the joint accurately and determining a possibility of damage in the joint with less error when the abnormality such as an interference with an object occurs in the articulated robot.* ● *The method enables installing the robot and a peripheral device in small space in order to improve space efficiency, so that collision between the robot and the peripheral device occurs easily even with a small operational error or a setting error.* ● *The method enables utilizing a system to continue operating in a degraded state and clearing the fault and continuing use of the robotic system.* ● *The method enables providing fault handling scheme which ensures that the system is automatically placed in a safe state upon detection of the fault, and facilitates graceful degradation of the robotic system after detection of fault that is not cleared.* ● *The defects of the wafer can be identified and classified effectively.* ● *The method enables realizing accurate judging state, low algorithm complexity, quick processing time and strong applicability and reducing misjudgment probability and data offset.*

Table 5
Additional highlights (of intelligent fault diagnosis patents).

Type	Content
Extra summary sentence	Diagnosis apparatus for diagnosing interference with object, of articulated robot, has processors for determining presence or absence of failure in articulated robot based on output value of each encoder and output value of each motor. (Title)
Keywords	Processors, articulated robot, output value, encoder, motor, threshold, failure, movement, control device, abnormality, detection, detecting fault, robotic arms, safe state, image of wafer, storage, feature values, classification, calculating, quality characteristic, simply manufacturing process, control chart, process quality, verify, monitor, diagnosing failure, production line, various, controlling robot, hazardous material handling, graceful degradation, semiconductor, fuzzy support vector machine, joint accurately, load state, errors, damage, evaluates estimate, peripheral device, operation error, setting error, degraded state, handling scheme, automatically, effectively, low algorithm complexity, quick processing time, reducing misjudgment probability
Precision ratio	96% (26/27)
Recall ratio	93% (26/28)

Notes: underline words are attention words based on trained results.

algorithm) or model evaluation to process shop floor data and solve decision-making problems. The use summary indicates that the invention technology is applied to manufacturing plants, industrial systems, oil refineries, chemical plants, and power generation facilities. The advantage summary emphasizes that the technology enables better production scheduling optimization, improvement of production planning, reduced cost and waste of production cycle electricity consumption through optimization algorithm planning and control. Keywords include production schedule, genetic algorithm, quality data, production and operation planning, hierarchical layers, decision-making, which represent the key technology and purpose of the original patent content.

According to the evaluation result, there are 24 sentences in target summary, and the output summary has 21 sentences. There are 18 sentences covered by the output summary of proposed system and target summary of DI platform. The precision and recall are 86% and 75%, respectively. Thus, six essential sentences were missed by the proposed system that these sentences describe algorithmic process details. However, these attention words in the planning and scheduling domain do focus on the algorithm terms of the patents. The extra sentences represent the claim and method details of patent US20140129491A1 which is related to the method for training

Table 6
Technical summary results (of intelligent planning and scheduling patents).

Type	Content
Novelty summary	• *A buffer-based control scheme is provided to generate an optimal buffer threshold value and a production schedule.* • *The method involves providing a baseline simulation model of a manufacturing plant to obtain energy and production performance of a station.* • *A final production schedule is generated by utilizing extreme and empirical buffer utilization policies, where the scheme utilizes a genetic algorithm comprising first fitness function that includes electricity consumption minimization objective and second fitness function that includes electricity cost minimization objective.* • *A data relation model is established according to a quality influence database.* • *A comprehensive evaluation model is established according to the enterprise quality data.* • *An auditing and modeling process is performed.* • *An order generation and management system comprises hierarchical layers that are configured to solve a decision-making problem.* • *Each hierarchical layer is configured to generate solution data representing a possible solution to a sub-problem associated with the decision-making problem.* • *The method involves formulating optimization problem to determine model parameters of an industrial system to be modeled.* • *The optimization problem is solved to define an empirical model of the system.*
Use summary	• *Order generation and management system for facilitating solutions of decision-making problems.* • *Mass data depth analysis based mechanical parts production planning method, involves utilizing correlation change rule for performing visual process, and utilizing option content module for finishing production plan and operation plan.* • *Uses include but are not limited to a manufacturing plant, oil refinery, chemical plant and a power generation facility for controller-based applications, model predictive control applications, environmental management applications, production performance management applications, plant operation optimization applications and industrial scheduling system applications.*
Advantage summary	• *The method enables providing optimized production scheduling by optimizing electricity consumption and cost by using a genetic algorithm and buffer control.* • *The method enables realizing rapid and efficient optimization of production planning for a manufacturing execution system MES so as to improve level of mechanical parts production plan.* • *Enables each hierarchical layer to be configured to use the orders to simplify a search for the possible solution to the sub-problem being solved by the hierarchical layer, such as by excluding solutions inconsistent with the orders.* • *The method enables optimization modalities of the model to use test data as inputs to optimization of the model and analyzes output of the optimization to determine suitability to accurately optimize certain aspects such as cost, production outputs and power generation, of the system.* • *The method enables providing activation functions that are selected due to asymptotic properties, so that a neural network provides enhanced extrapolation behaviors and improved run time performance.*

parameters and optimization. Another extra focuses on key performance and purpose of patent EP2250552A4. The summary result of planning and scheduling domain has lower reliability given that this domain is focused on algorithm terms instead of solutions and performance methods.

4.4. Discussion

The aim of the control intelligence domain is that machine learning approaches are applied to measure model parameters to make predictions about object grasping, remote control, and command decisions. The aim of the sensor intelligence domain is to apply cameras and sensors to capture images, light sources, detect and sense objects, perform edge analysis, and notify dangerous situations. Patents in this domain are most often used for manufacturing processes or robot movement. Intelligent decision-making most frequently applies genetic algorithms or evaluation models to provide solutions such as Tayler series method to conduct production planning and scheduling optimization, diagnose machine health, and improve wafer defect analysis. Patents in this domain use algorithms to reduce cost and waste for production processes and to improve efficiency and product quality. Given the overall experimental results of the case study on smart machinery, the performance has high similarity and is reliably consistent between the output summary of the proposed system and the target summary of original patent content by DI patent engineers.

5. Conclusions

This research has proposed a novel intelligent patent summary system to provide researchers or enterprises a means to automatically generate summary reports. The summaries provide insights of the development trends of emerging industries and provides a strategic advantage to develop highly competitive products and patents as well as plan financial and marketing strategies. In the case study section, the proposed model automatically generated the summary reports related to techniques for the smart machinery patents. The experimental results show the AI-based solution for abstractive and extractive summarization not only retrieves the major sentences, but also captures the technical keywords. The summarization system has sufficient E-discovery capability to identify important information from original IP

Table 7
Additional highlights (intelligent planning and scheduling patents).

Type	Content
Extra summary sentence	• A method, comprising formulating an optimization problem to determine a plurality of model parameters of a system to be modeled solving the optimization problem to define an empirical model of the system and training the empirical model using training data, wherein the empirical model is constrained via general constraints relating to first-principles information and process knowledge of the system. (Claim) • An independent claim is also included for a method for optimizing production scheduling in a manufacturing plant. (Abstract) • The drawing shows a flowchart illustrating a method for providing optimized production scheduling by optimizing electricity consumption and cost. (Abstract)
Keyword	Production schedule, genetic algorithm, quality data, production and operation planning, hierarchical layers, decision-making, first-principles, manufacturing plant, Taylor series, MES, accurately optimize, cost production outputs, threshold value, optimal buffer, performance, simulation model, electricity consumption minimization, electricity cost minimization, quality influence database, evaluation, order generation and management, optimization problem, depth analysis, visual process, buffer control, efficient optimization, activation functions, improved run time, neural network
Precision ratio	86% (18/21)
Recall ratio	75% (18/24)

Notes: underline words are attention words based on trained results.

documents. The summary report shows that the objectivity, understandability, and readability of the automatically generated result is reliable and improves with patents in similar domains. In addition, the ROUGE indicators for precision and recall objectively demonstrate that the summary result and original IP documents have high similarity and are consistent. The proposed summarization model has high accuracy and quality and yields performance result that are nearly equivalent to the capability provided by DI patent engineers. The summary results compress the important information and yields a technical summary. The summary report identifies important IP technologies in a timely and less costly manner for companies or researchers.

This research deploys natural language processing techniques and ML models to better identify patent meanings, access important information, and, most importantly, eliminates time for manually examining large number of patent documents for knowledge extraction and summarization. The proposed system, integrating abstractive and extractive methods, is a novel methodology for the patent summarization. However, the limitation of this research is that the training document set for ML-based summarization modeling only consists of 1708 patents for smart machinery as training inputs and their corresponding summaries (manually written) as reference outputs. One might challenge that the reference outputs are subjective, although the results well match a reputable patent analytic platform. Further, the summarization system requires more training data to improve system performance and expand to other domains of interests. Nonetheless, the significant contribution of this research is that the ML-based intelligent patent summary compilation approach is presented and the case study is conducted to show the repeatability of the methodology and its promising results. For future work, the goal is to construct multiple models which combine the versatile neural network models for improved performance. New research should combine other algorithms to deal with multi-lingual training documents without reference outputs so that the system has the capability to generate summary reports from original patent documents written in different languages.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is partially supported by the Ministry of Science and Technology research projects (Grant numbers: MOST-107-2218-E-031-002-MY2, MOST-108-2221-E-007-075-MY3 and MOST-108-2410-H-009-025-MY2) in Taiwan.

Appendix A. – Some smart machinery related patents for patent summarization testing and verification.

See Tables A1–A3.

Table A1
Example **control intelligence** patents.

Language	Patent number	Title
English	US20170252922A1	Deep machine learning methods and apparatus for robotic grasping.
	US20170334066A1	Machine learning methods related to predicting motion of object in a robot's environment based on image capturing the object and based on parameter for future robot movement in the environment.
	US20160325434A1	Apparatus for remotely controlling robots and control method thereof.
	US20130297072A1	Control apparatus and method for master–slave robot, master–slave robot control programs, and integrated electronic circuits.
	US20160243701A1	Facilitating device control.
Chinese	CN107428004A	The automatic collection of object data and mark
	CN106874914A	A kind of industrial machinery arm visual spatial attention method based on depth convolutional neural networks
	CN106737661A	A kind of controlled system with self-regulation of time delay force feedback remote-controlled robot
	CN106335057A	Total-space smooth hole insertion control method applied to assembly robot and based on real-time force control
	CN106078741A	Based on a determination that the limited performance flexible mechanical arm control method of theory of learning
	CN106903690A	A kind of crane movements track recognizing method

Table A2
Example **sensor intelligence** patents.

Language	Patent number	Title
English	JP2017123147A	Locating feature position for guiding robot, pinpointing of the characteristic position for guiding a robot.
	JP2016052514A	Mobile robot, movable robot.
	WO2013033338A3	Asynchronous data stream framework.
	US9324003B2	Location of image capture device and object features in a captured image.
	WO2013185102A1	Carpet drift estimation using differential sensors or visual measurements.
Chinese	CN105856230A	ORB key frame closed-loop detection SLAM method capable of improving consistency of position and pose of robot
	CN107480597A	A kind of Obstacle Avoidance based on neural network model
	CN103731604B	Follow-up mechanism and method for tracing
	CN107111739A	The detection and tracking of article characteristics
	CN106874854A	Unmanned plane wireless vehicle tracking based on embedded platform

Table A3
Example **intelligent decision-making for smart machinery** patents.

Sub-domain	Patent number	Title
Fault diagnosis	US20170210009A1	Robot system controlling method, program, recording medium, robot system, and diagnosis apparatus.
	JP2015217468A	Method for controlling robot system, program, recording medium, and diagnostic device.
	EP3057742A1	Fault reaction, fault isolation, and graceful degradation in a robotic system.
	US8379965B2	Defect classification method, computer storage medium, and defect classification apparatus.
	CN107346122A	Improved fuzzy support vector machine, manufacturing process using multivariate quality diagnostic classifier.
Planning and scheduling	US20160179081A1	Optimized production scheduling using buffer control and genetic algorithm.
	CN107392385A	A production planning method based on quality data depth analysis.
	EP2250552A4	Apparatus and method for order generation and management to facilitate solutions of decision-making problems.
	US20140129491A1	Empirical modeling with globally enforced general constraints.

References

[1] T. Allman, Deterring e-discovery misconduct with counsel sanctions: the unintended consequences of Qualcomm v. Broadcom, Yale Law J. 118 (2009) 161–165.
[2] J. Pooley, V. Huang, Multi-national patent litigation: management of discovery and settlement issues and the role of the judiciary, Fordham Intellectual Property Media Entertain. Law J. 22 (2011) 45–68.
[3] R. Barzilay, K.R. McKeown, M. Elhadad, Information fusion in the context of multi-document summarization, ACL (1999) 550–557.
[4] A.J. Trappey, C.V. Trappey, An R&D knowledge management method for patent document summarization, Ind. Manage. Data Syst. 108 (2) (2008) 245–257.
[5] A.J. Trappey, C.V. Trappey, B.H.S. Kao, Automated patent document summarization for R&D intellectual property management, CSCWD (2006) 1–6.
[6] R. Ferreira, L. de Souza Cabral, R.D. Lins, G.P. Silva, F. Freitas, G.D.D. Cavalcanti, R. Lima, S.J. Simske, L. Favaro, Assessing sentence scoring techniques for extractive text summarization, Expert Syst. Appl. 40 (14) (2013) 5755–5764.
[7] R. Ferreira, F. Freitas, L. de Souza Cabral, R.D. Lins, R. Lima, G. Franca, L. Favaro, A context based text summarization system, DAS (2014) 66–70.
[8] A.J.C. Trappey, C.V. Trappey, An R&D knowledge management method for patent document summarization, Ind. Manage. Data Syst. 108 (2) (2008) 245–257.
[9] R. Nallapati, B. Zhou, C. Gulcehre, B. Xiang, Abstractive text summarization using sequence-to-sequence RNNs and beyond, CoNLL (2016) 280–290.
[10] D. Crevier, AI: The Tumultuous History of the Search for Artificial Intelligence, Basic Books Inc, New York, 1993.
[11] R. Poli, M. Healy, A. Kameas, Theory and Applications of Ontology: Computer Applications, Springer, Berlin, 2010.
[12] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Lves, DBpedia: a nucleus for a web of open data, in: K. Aberer, K.S. Choi, N. Noy, D. Allemang, K. Lee, L. Nixon (Eds.), The Semantic Web, Berlin, Springer, 2007, pp. 722–735.
[13] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, J. Taylor, Freebase: a collaboratively created graph database for structuring human knowledge, The 2008 ACM SIGMOD International Conference on Management of Data, 2008, pp. 1247–1250.
[14] Y. Pan, Heading toward artificial intelligence 2.0, Engineering 2 (2016) 409–413.
[15] A. Prieto, B. Prieto, E.M. Ortigosa, E. Ros, F. Pelayo, J. Ortega, I. Rojas, Neural networks: an overview of early research, current frameworks and new challenges, Neurocomputing 214 (2016) 242–268.
[16] I. Arel, D.C. Rose, T.P. Karnowski, Deep machine learning – a new frontier in artificial intelligence research, IEEE Comput. Intell. Mag. 5 (4) (2010) 13–18.
[17] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Sig. Process. 72 (2016) 303–315.
[18] P. Swietojanski, A. Ghoshal, S. Renals, Unsupervised cross-lingual knowledge transfer in DNN-based LVCSR, SLT-2012 (2012,) 246–251.
[19] R. Collobert, J. Weston, A unified architecture for natural language processing: deep neural networks with multitask learning, ICML-2008 (2008) 160–167.
[20] A. Khan, B. Baharudin, L.H. Lee, K. Khan, A review of machine learning algorithms for text-documents classification, J. Adv. Info. Technol. 1 (1) (2010) 4–20.
[21] F. Chiarello, A. Cimino, G. Fantoni, F. Dell'Orlett, Automatic users extraction from patents, World Patent Inf. 54 (2018) 28–38.
[22] D.H. Milanez, L.I.L. de Faria, R.M. do Amaral, J.A.R. Gregolin, Claim-based patent indicators: a novel approach to analyze patent content and monitor technological advances, World Patent Info. 50 (2017) 64–72.
[23] D. De Clercq, Z. Wena, Q. Song, Innovation hotspots in food waste treatment, biogas, and anaerobic digestion technology: a natural language processing approach, Sci. Total Environ. 673 (2019) 402–413.
[24] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, NIPS-2013 (2013) 3111–3119.
[25] M. Whitehead, D.K.N. Johnson, A tool for visualizing and exploring relationships among cancer-related patents, AAAI-2017 (2017) 235–238.
[26] A. Nenkova, K. McKeown, A survey of text summarization techniques, in: C.C. Aggarwal, C. Zhai (Eds.), Mining Text Data, Springer Science and Business Media, New York, 2012, pp. 43–76.
[27] Z. Cao, F. Wei, L. Dong, S. Li, M. Zhou, Ranking with recursive neural networks and its application to multi-document summarization, AAAI-2015 (2015) 2153–2159.
[28] Y.J. Kumar, O.S. Goh, H. Basiron, N.H. Choon, P.C. Suppiah, A review on automatic text summarization approaches, J. Comput. Sci. 12 (4) (2016) 178–190.
[29] A. Severyn, A. Moschitti, Learning to rank short text pairs with convolutional deep neural networks, ACM SIGIR-215 (2015) 373–382.
[30] C. Yao, J. Shen, G. Chen, Automatic document summarization via deep neural networks, ISCID-2015 (2015) 291–296.
[31] A.J. Trappey, C.V. Trappey, C.Y. Wu, Automatic patent document summarization for collaborative knowledge systems and services, J. Syst. Sci. Syst. Eng. 18 (1) (2009) 71–94.
[32] S. Brügmann, N. Bouayad-Agha, A. Burga, S. Carrascosa, A. Ciaramella, M. Ciaramella, J. Codina-Filba, E. Escorsa, A. Judea, S. Mille, A. Müller, H. Saggion, P. Ziering, H. Schütze, L. Wanner, Towards content-oriented patent document processing: intelligent patent analysis and summarization, World Patent Inf. 40 (2015) 30–42.
[33] K. Girthana, S. Swamynathan, Semantic query-based patent summarization system (SQPSS), International Conference on Intelligent Information Technologies, 2018, pp. 169–179.
[34] Clarivate Analytics, Derwent World Patents Index (White-sheet) – The world's most trusted source of curated patent intelligence, (2019) [Online]. Available: https://clarivate.com/products/derwent-world-patents-index/ (Accessed: 3 June, 2019).
[35] C.Y. Lin, ROUGE: A Package for Automatic Evaluation of Summaries, ACL, 2004.
[36] Y.S. Lin, J.Y. Jiang, S.J. Lee, A similarity measure for text classification and clustering, IEEE Trans. Knowl. Data Eng. 26 (7) (2014) 1575–1590.
[37] S. Atluru, S.H. Huang, J.P. Snyder, A smart machine supervisory system framework, Int. J. Adv. Manuf. Technol. 58 (5–8) (2012) 563–572.
[38] Q. Dai, X. Wang, The principle and application of intelligent machine tools, MMECEB-2015 (2015) 322–325.
[39] H. Cao, X. Zhang, X. Chen, The concept and progress of intelligent spindles: a review, Int. J. Mach. Tools Manuf. 112 (2017) 21–52.