(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
03 December 2020 (03.12.2020)

WIPO | PCT

(10) **International Publication Number**
WO 2020/240572 A1

(51) **International Patent Classification:** *G06Q 50/00* (2012.01)

(21) **International Application Number:** PCT/IN2019/050422

(22) **International Filing Date:** 29 May 2019 (29.05.2019)

(25) **Filing Language:** English

(26) **Publication Language:** English

(71) **Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)** [SE/SE]; SE-164 83 Stockholm (SE).

(72) **Inventor; and**
(71) **Applicant** *(for SC only)*: **MOHAN, Saravanan** [IN/IN]; 39/1 Akbar Street, Triplicane, Chennai 600005 (IN).

(72) **Inventor: SATHEESH KUMAR, Perepu**; F3, Megapearl Shri, Mahalakshmi Street, Gandhi road, Velachery, Chennai 600042 (IN).

(74) **Agent: SINGH, Manisha**; LexOrbis, 709-710 Tolstoy House, 15-17 Tolstoy Marg, New Delhi 110001 (IN).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(54) **Title:** METHOD FOR TRAINING A DISCRIMINATOR

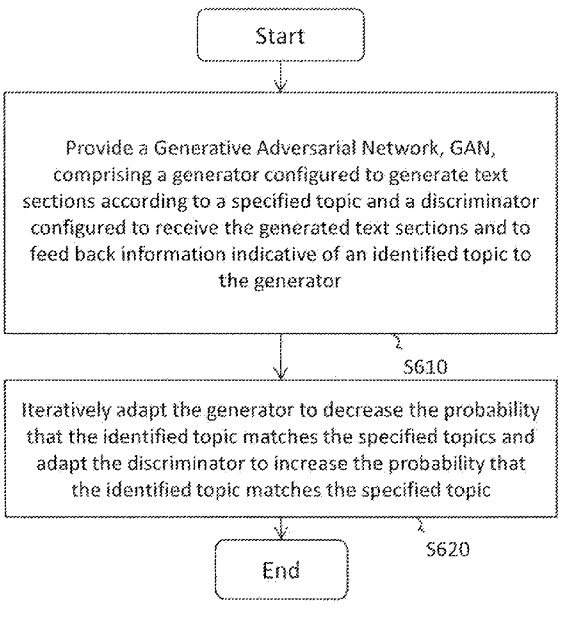

Fig. 6A

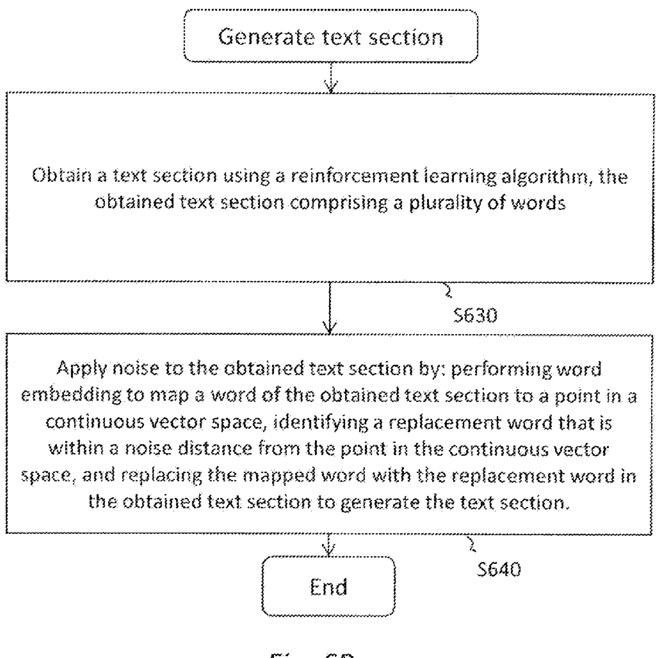

Fig. 6B

(57) **Abstract:** A computer-implemented method for training a discriminator to identify a topic to which a received text section is related. The method comprises providing a generative adversarial network, GAN, comprising a generator configured to generate text sections according to a specified topic, and the discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator. The method further comprises iteratively adapting the generator to decrease the probability that the identified topic matches the specified topic and adapting the discriminator to increase the probability that the identified topic matches the specified topic. The generator is configured to generate a text section by obtaining a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words and applying noise to the obtained text section. Noise is applied to the obtained text section by performing word embedding to map a word of the obtained text section to a

WO 2020/240572 A1

[Continued on next page]

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:
— *with international search report (Art. 21(3))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.

METHOD FOR TRAINING A DISCRIMINATOR

TECHNICAL FIELD

The present invention relates to field service operations for system troubleshooting and maintenance. In particular, the present invention relates to guide systems for guiding a field service operative in troubleshooting.

Additionally, the present invention relates to machine learning and natural language processing.

BACKGROUND

Systems are commonly distributed with system elements at sites separated by significant geographical distances. As a result, maintenance of such systems can require technicians to go “into the field” in order to perform in-person maintenance at any of the sites, in response to a trouble ticket which indicates one or more problems affecting the system. In other words, the technicians act as field service operatives.

In such systems, the range of possible problems is large, and addressing such problems requires a range of different skills. On the other hand, maintenance is provided using a limited number of available technicians, each having different or overlapping skill-sets, and with an expected maintenance speed (e.g. a Service Level Agreement, SLA). As a result, it can be necessary to assign a technician to respond to a trouble ticket even though they do not have expertise in one or more skills required for addressing the problems of the trouble ticket.

In order to enable the technician to respond to a trouble ticket, the technician is provided with access to a document database. The technician must search through the database based on the trouble ticket and based on their on-site observations, in order to find documents relevant for responding to the problems affecting the system. Alternatively, experts relevant to the problems can assist the technician remotely, if such experts are available.

The document database can be provided in the form of a searchable document database which can be accessed using a technician's terminal (such as a smartphone, tablet, laptop, etc.), for example via an application (app) installed on the terminal. However, the limited size of a screen on a terminal can make it harder to identify relevant information.
In view of this problem, it is desirable to provide a summary of the document database which is appropriate for the trouble ticket. Such a summary can be produced manually by a person who understands the technical content of the database. However, such human summarization techniques are too complex to be replicated in a computer.

Instead, machine learning techniques in natural language processing applications have been used for classifying and summarizing text sections. However, one type of machine learning technique which has problems when used for text processing is Generative Adversarial Network, GAN, models. GAN models have been used for some types of text processing with limited success, but have not been used for generating new text or extractive summaries and would not produce optimal results in these applications.

GAN models have the ability to learn robust, reusable feature representations from small sets of unlabelled data, and have been applied in other machine learning tasks. This is achieved by training deep generative models that can learn to capture the complex distributions of real-world data. GAN models have been used for semi-supervised learning based on an imbalanced learning dataset. Additionally, GAN models have been used in computer vision, where training is formalized as a game in which a discriminator model is trained to discriminate between sample images from a training set and images which are not from the training set, and a generative model is simultaneously trained to generate images which fool the discriminator model. When trained in this way, the generator can be trained to generate realistic artificial images, and the discriminator can be trained to a high degree of accuracy.

SUMMARY

In a first aspect, a computer-implemented method is provided for training a discriminator to identify a topic to which a received text section is related. The method comprises providing a generative adversarial network, GAN, comprising a generator configured to generate text sections according to a specified topic, and the discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator. The method further comprises iteratively adapting the generator to decrease the probability that the identified topic matches the specified topic and adapting the discriminator to increase the probability that the identified topic matches the specified topic. The generator is configured to generate a text section by obtaining a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words and applying noise to the obtained text section. Noise is applied to the obtained text section by performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.

In a second aspect, an apparatus is provided for training a discriminator to identify a topic to which a received text section is related. The apparatus comprises a generative adversarial network, GAN, comprising a generator configured to generate text sections according to a specified topic, and the discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator. The GAN is configured to iteratively adapt the generator to decrease the probability that the identified topic matches the specified topic and adapting the discriminator to increase the probability that the identified topic matches the specified topic. The generator is configured to generate a text section by obtaining a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words and applying noise to the obtained text section. Noise is applied to the obtained text section by performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.

In a third aspect, an apparatus is provided comprising a processor, a memory and an instruction store, the instruction store storing computer program instructions which, when executed by the processor, cause the processor to perform a method for training a discriminator to identify a topic to which a received text section is related. The method comprises providing a generative adversarial network, GAN, comprising a generator configured to generate text sections according to a specified topic, and the discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator. The method further comprises iteratively adapting the generator to decrease the probability that the identified topic matches the specified topic and adapting the discriminator to increase the probability that the identified topic matches the specified topic. The generator is configured to generate a text section by obtaining a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words and applying noise to the obtained text section. Noise is applied to the obtained text section by performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.

In a fourth aspect, a computer-readable storage medium is provided, having stored thereon computer program instructions which, when executed by a processor, cause the processor to perform a method for training a discriminator to identify a topic to which a received text section is related. The method comprises providing a generative adversarial network, GAN, comprising a generator configured to generate text sections according to a specified topic, and the discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator. The method further comprises iteratively adapting the generator to decrease the probability that the identified topic matches the specified topic and adapting the discriminator to increase the probability that the identified topic matches the specified topic. The generator is configured to generate a text section by obtaining a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words and applying noise to the obtained text section. Noise is applied to the obtained text section by performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention will now be described in detail, by way of example only, with reference to the accompanying figures, in which:

Fig. 1 is a schematic illustration of a GAN model for computer vision.

Fig. 2 is a schematic illustration of a GAN model for natural language processing.

Fig. 3 is a schematic illustration of functional components of the discriminator training apparatus of a first embodiment.

Fig. 4 is a schematic illustration of a Q-learning process.

Fig. 5 is a schematic illustration of a programmable processing apparatus in which a discriminator training apparatus of the first embodiment may be implemented.

Figs. 6A and 6B are flow diagrams showing processing operations performed by the discriminator training apparatus of the first embodiment.

Fig. 7 is a schematic illustration of a discriminator training apparatus of a second embodiment.

Fig. 8 is a schematic illustration of a term/text section matrix for a set of text sections.

Fig. 9 is a flow diagram showing processing operations performed by the discriminator training apparatus of the second embodiment.

Fig. 10 is a schematic illustration of a summary generation apparatus of a third embodiment.

Fig. 11 is a flow diagram showing processing operations performed by the summary generation apparatus of the third embodiment.

Fig. 12 is a schematic illustration of a maintained system in which a technician assistance apparatus of a fourth embodiment may be used.

Fig. 13 is a schematic illustration of the technician assistance apparatus of the fourth embodiment.

Fig. 14 is a flow diagram showing processing operations performed by the technician assistance apparatus of the fourth embodiment.

DETAILED DESCRIPTION

As used herein, system means any arrangement of industrial technical devices where devices are located at a plurality of geographically distant sites. For example, such systems may be connected systems like communication infrastructure networks and fog computing networks, or may be unconnected systems such as collections of independent local area networks, or more generally may be any industrial system. The term “geographically distant” has a meaning dependent upon the system, but generally means a distance associated with a technician travelling time that is substantial when compared to an expected time for solving a problem at a site.

As used herein, a document database comprises one or more manuals and/or documentation. The database may also be updated by technicians, for example by providing documentation or summaries of their experiences solving problems. The database provides information about main tasks, key pain points, process knowledge and relevant tools.

As used herein, technician means any skilled person or team of skilled people with at least one technical skill relevant for maintaining a system. For example, technicians may be people working in Serviceability, Product Management or PDU system departments.

Fig. 1 is a schematic illustration of a Generative Adversarial Network, GAN, 100 for computer vision. The GAN comprises a generator 110 and a discriminator 120.

The generator 110 is a model which receives random noise 130 as an input and outputs new instances of a data type being modelled, in this case an image 140. The image may have any format such as pixel data or vector graphic data. The generator 110 is implemented using a neural network has a form that corresponds to the inverse of a Convolutional Neural Network, CNN. In other words, while a convolutional classifier takes inputs and down-samples to produce a probability that the input is from a training set, the generator 110 takes an input of random noise 130, for example in the form of a vector, and up-samples to produce a result. The generator is trained to produce images with required characteristics, such as images which provide a good impersonation of a human face.

The discriminator 120 is a model which receives data instance of the same type as generated by the generator 110 (i.e. an image) as an input and outputs a classification 160. The discriminator 120 is implemented using a neural network such as a Convolutional Neural Network, CNN or a Recursive Neural Network, RNN, which down-samples an input using techniques such as maxpooling. In this example, the discriminator 120 uses a Long Short-Term Memory, LSTM, model to perform classification. The input image may be an image from a training set 150 (also herein called a real image), or may be an image 140 which has not come from the training set (also herein called a fake image). The discriminator 120 does not know whether the image is from the training set 150 or not, and a mixture of “real” and “fake” images are fed to the discriminator 120. Correspondingly, the classification 160 is a classification of authenticity indicating that the input image is either “real” or “fake”. This classification may be binary, or may be on a continuous scale, for example a value between zero and one, according to a degree of confidence that the input is “real” or “fake”. The training set 150 is a large set of images, such that it is not possible for the discriminator 120 to identify all images in the set individually, and the discriminator is trained to instead discriminate between images having properties correlated with characteristics of the training set and images having different properties.

In the GAN 100, the discriminator 120 is configured to receive images from either the generator 110 or the training set 150, and both of the generator 110 and the discriminator 120 are provided with feedback information for determining whether or not the discriminator 120 correctly classifies the images. The generator 110 and discriminator 120 are trained such that they compete. In other words, the generator 110 is trained to produce images that are likely to trick the discriminator 120 into producing the wrong classification, and the discriminator 120 is trained classify images with increasing accuracy. Both of the generator 110 and discriminator 120 are trained to optimize a different and opposing objective function, or loss function, in a zero-sum game. This approach is similar to employing an actor-critic model of reinforcement learning. As the discriminator 120 is modified and changes its behaviour, the behaviour of the generator 110 also changes, and vice versa, and their losses push against each other.

An overall effect of this feedback is to improve the accuracy of classification by the discriminator 120, by training the discriminator 120 to handle increasingly convincing images generated by the generator 110 which are used to attempt to fool the discriminator.

In more detail, in the GAN 100, there is a double feedback loop for training.

Training of the generator model 110 is performed by the GAN 100 using the generator 110 and the random noise 130 to produce randomly varying images 140, generating feedback indicating how well each image matches the required characteristics, and modifying the generator model on the basis of the feedback. The feedback may take the form of a probability that the discriminator 120 correctly classifies an input image, or may be another reward function.

Training of the discriminator model 120 is performed by the GAN 100 using the discriminator model 120 to classify images, generating feedback indicating whether or not the classification 160 is correct, and modifying the discriminator model on the basis of the feedback in order to increase the probability that the classification 160 correctly indicates whether or not the input image is from the training set 150. While the discriminator model 120 is being trained, the GAN 100 knows whether or not the input image comes from the training set 150 or is a “fake” image 140 from the generator 110, which enables the GAN 100 to determine a ground truth of whether or not the classification 160 is correct, and thereby provide feedback to the generator 110 and discriminator 120.

The GAN 100 can be formally described as a min-max adversarial game between a generative model G and a discriminative model D. The generative model G (which may be a neural network) is trained to map samples s from a noise distribution to a data space of samples G(s) of the data type (images in this case). The discriminative model D (which may be a neural network) takes a data sample t as input and outputs a scalar value D(t) representing the probability that sample t came from a training distribution x and is not a sample G(s) of the generative model. D is trained to maximise the probability of assigning the correct label to the input t, while G is trained to maximally confuse D, i.e. to minimize the probability of D assigning the correct label to the input t. D is provided with continuously varying samples t, so that the gradient of D(t) with respect to t can be used to update and improve the parameters of the discriminative model D.

The training goal can be summarized formally as:

$$\min_G \max_D \left\{E\left[p(x)\log\left(D(x)\right)\right] + E\left[p(s)\log\left(1 - D\left(G(s)\right)\right)\right]\right\}$$

wherein E is an expectation, energy or entropy value, depending upon to the probability of each value x and s in the training and random distributions respectively, and the logarithm of a function of the discriminator output D(x) and (1 - D(G(s))) respectively.

Furthermore, the GAN 100 can train the discriminator 120 using an energy-type reward function to increase the energy of “real” inputs and reduce the energy of “fake” inputs by maximising the following function:

$$f_D(t,s) = D(t) + \max(0, m - D\left(G(s)\right))$$

wherein, as above, t is an input to the discriminator that may be taken from the training distribution x or may be a generated sample G(s), and m is a weighting margin which may be varied during training in order to change the balance between positive training to reward “real” examples and negative training to reject “fake” examples.

On the other hand, the GAN 100 can train the generator 110 using an energy-type reward function by maximizing the following function:

$$f_G(s) = D(G(s))$$

Training and optimization of the generator 110 and discriminator 120 can be unstable, and therefore it can be necessary to carefully design the architecture of the generator 110 and discriminator 120 so that a balance is maintained and training converges. Using neural network architectures reduces complexity and assists in ensuring that the training converges. Additionally, it is necessary to prevent mode collapse wherein the generator learns to fool the discriminator with only a small range of different images but a high success rate, leading to an inadequate discriminator that cannot discriminate well across the range of the training set 150.

A GAN 100 as shown in Fig. 1 can be adapted from a use in computer vision to a use in Natural Language Processing, NLP.

Fig. 2 is a schematic illustration of a GAN model for natural language processing. In Fig. 2, a generator 210 generates text sections 240, for example in the form of a document matrix that replaces the generated image of Fig. 1, and a discriminator 220 produces classifications 260 indicating whether received text sections are from a training set 250 (i.e. “real” text sections) or not. The goal of the generator 210 is to generate passable text sections (for example, document extracts), that is to generate text sections which imitate natural language and which the discriminator cannot classify correctly. On the other hand, the goal of the discriminator 220 is to correctly discriminate between text sections from the training set 250 and text sections 240 from the generator 210.

However, adapting a GAN for natural language processing has a problem when adapting random noise 130 to give an input 230 for the generator 210. In particular, natural language text is discrete in the sense that not all combinations of linguistic characters are valid words. Accordingly, random noise 130 cannot be applied in order to vary text sections in the way that images are varied in computer vision applications. Furthermore, feedback from the discriminator cannot be used to control application of the random noise. These problems are addressed in embodiments of the invention.

Fig. 3 is a schematic illustration of functional components of the discriminator training apparatus 300 of a first embodiment which implements the described functions of Fig. 2, and which is similar to the machine vision application of Fig. 1 except where described otherwise.

In the first embodiment, the discriminator training apparatus 300 comprises a GAN 310, in which a generator 320 and a discriminator 330 are trained. The generator 320 and discriminator 330 can be similar to the generator 110 and discriminator 120, except as described in the following.

In particular, the generator 320 comprises a reinforcement learning algorithm 340 and a word embedding noise adder 350. The reinforcement learning algorithm 340 is trained to generate one or more text sections comprising a plurality of words. The word embedding noise adder 350 is configured to then apply noise to the one or more text sections using word embedding.

The reinforcement learning algorithm 340 generates a sequence of words, that is, a text section. In this embodiment, the reinforcement learning algorithm uses a Q-learning process. However, other reinforcement learning algorithms may be used.

Fig. 4 is a schematic illustration of a Q-learning process, which is a model-less reinforcement learning technique.

In the Q-learning process, the generation of a text section is modelled as a series of states 410. The states here represent partial text sections, which may be stored as word embeddings i.e. as vectors in a continuous space. For example, in a first state the text section is “Repairing”, in a second state the text section is “Repairing power”, in a third state the text is “Repairing network”.

Each transition between two states is defined as an action 420. The actions represent which word to choose next. More specifically, each action 420 is the addition of one or more words to a partial text section. For example, in the above example, there is a first action to go from “Repairing” to “Repairing power”, and a second action to go from “Repairing” to “Repairing network”.

The states 410 and actions 420 are arranged in a network 430 which defines their connections, and associated value functions. That is, for a given current state and a given action, a reward value is defined initially based on predetermined information and updated based on feedback from the discriminator 330. More specifically, if the feedback from the discriminator 330 indicates that classification accuracy is low for a text section generated using a certain series of states, the rewards for states and actions which lead to a similar text section are increased, so as to promote generation of text sections which the discriminator 330 cannot classify accurately. On the other hand, if the feedback indicates that classification accuracy is high for a text section generated using the certain series of states, the rewards for states and actions which lead to a similar text section are decreased. The set of reward values of the network is also called a “policy”.

The Q-value 440 is an expected total reward associated with a series of actions under a given policy, and is evaluated as follows at time t, state s and next action a:

$$Q^{\pi}(s,a) = E[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ \dots \mid s,a]$$

wherein π is the policy, E is the expected total reward, r is the reward for a given action taken at a specified time, and γ is a discount factor which weights the Q-value towards actions which occur closer to the current time t. In this expression, each time step corresponds to the addition of one or more words to the text section. Accordingly, the Bellman equation to be solved in Q-learning is expressed as:

$$Q^{\pi}(s,a) = E_{s',a'}[r \ + \ \gamma Q^{\pi}(s',a') \mid s,a]$$

By increasingly optimizing the Q-learning, it becomes possible to add each word of a text section based on a maximum expected reward for the corresponding state and action based on feedback for previously generated text sections, in order to generate a text section which is more likely to successfully fool the discriminator 330.

Referring back to Fig. 3, next the word embedding noise adder 350 is used to add noise to the text section obtained from the reinforcement learning algorithm. In the case of images, noise can be added by adding noise values to pixel values or to values defining a vector graphic. However, this is not possible with text sections that are made up of discrete words. Instead, the word embedding noise adder 350 uses word embedding, which is a technique wherein a set of discrete words are mapped to a continuous multidimensional, i.e. vector, space, wherein each word corresponds to a point in the space (although not all points in the space correspond to a word). In such a vector space, a difference between two words is represented by a displacement which has an associated length, and a corresponding distance between the words can be defined. As a result, it is possible to apply noise to an initial word by defining a noise distance representing the magnitude of the noise, identifying a set of words which differ from the initial word by less than the noise distance in the vector space, and randomly selecting a word from the identified set using a uniform distribution, to select a final word that replaces the initial word. The scale of the noise distance is set to be large enough that an acceptable level of randomness can be achieved, for example by setting the scale to be large enough that the set of words is expected to include thousands of words. Additionally, the scale of the effect of the noise adding can be controlled by the way the word embedding is defined. For example, in many word embeddings, words at nearby points in the vector space have similar semantic meanings, and in such embeddings randomness can be applied to change individual words without changing the overall semantics of the text section.

Alternatively, noise may be added by identifying a set of words which differ from the initial word by a distance that falls within a range corresponding to the noise distance. For example, the final word may be selected as the word whose distance from the initial word is closest to the noise distance.

The noise distance, that is the magnitude of the applied noise, is dependent upon feedback received for a previously generated text section. For a first generated text section, a predetermined magnitude of noise, or no noise, is applied. Thereafter, the magnitude of noise applied in a given text section has an inverse relationship with a degree to which the identified topic matches the specified topic for a preceding generated text section. More specifically, when the identified topic by the discriminator 330 matches the specified topic for the generator 320, this means that the generator 320 has not succeeded in fooling the discriminator, and a significant change should be applied using noise in order to generate a text section which is likely to have a significantly different meaning and to increase the probability of fooling the discriminator 330. On the other hand, when the identified topic by the discriminator 330 does not match the specified topic for the generator 320, this means that the generator 320 has succeeded and a smaller change should be applied using noise, in order to further explore similar inputs to the discriminator 330. If the classification output by topic classifier 360 is a continuous degree of confidence value, the noise distance may be controlled continuously as an inverse function of the classification output.

By comparison to the output of the reinforcement learning algorithm 340, the word embedding noise adder 350 allows the generator 320 to reach a higher level of precision when trained to generate natural language text sections, by providing a mechanism for adding noise.

Furthermore, in this embodiment, the precision of the generator 320 is further increased by training the reinforcement learning algorithm 340 using Q-learning in addition to training in the GAN 310.

More specifically the generator 320 uses a Q-learning model that defines the probability of generating a given sequence of words, that is a text section, and the discriminator of the GAN labels the sequence of words as corresponding or not corresponding to a specified topic. The generator in this case is reward based and it can change the topic of the text sections it creates so that they do or do not correspond to the specified topic.

The discriminator 330 comprises a topic classifier 360 which is trained to discriminate whether a received text section corresponds to a required topic, using a Long Short-Term Memory, LSTM, model to perform classification.

The GAN 310 is configured to provide feedback to the topic classifier 360, the reinforcement learning algorithm 340 and the word embedding noise adder 350 for training, wherein the feedback indicates how successfully the topic classifier 360 classifies each text section that it receives.

Fig. 5 is a schematic illustration of a general kind of programmable processing apparatus 500 in which a discriminator training apparatus of an embodiment may be implemented. The processing apparatus 500 may, for example, be a server, a terminal device, or even virtual hardware of a node in a cloud server system.

The programmable processing apparatus 500 comprises an input/output (I/O) section 510 that functions as an interface module of the apparatus. This may be used, for example, for communicating with a remote terminal in a client-server relationship.

The programmable processing apparatus further comprises a processor 520, a memory 530 and an instruction store 540 for storing computer-readable instructions which, when executed by the processor 520, cause the processor 520 to perform the processing operations hereinafter described. The memory 530 includes volatile memory and non-volatile memory. The instruction store 540 may comprise a RAM or similar type of memory, and the computer-readable instructions can be input thereto from a computer program product, such as a computer-readable storage medium 550 such as a CD-ROM, etc. or a computer-readable signal 560 carrying the computer-readable instructions.

In the present embodiment, the combination of the hardware components shown in Fig. 5, comprising the processor 520, the memory 530 and the instruction store 540, is configured to implement the functionality of the discriminator training apparatus 300 shown in Fig. 3. Similarly, a programmable signal processing apparatus of the kind shown in Fig. 5 may be used to implement the functional components of the discriminator training apparatus 700 shown in Fig. 7, the summary generation apparatus 1000 shown in Fig. 10, and the technician assistance apparatus 1300 shown in Fig. 13.

Figs. 6A and 6B are flow diagrams showing the main processing operations performed by the discriminator training apparatus 300 of the first embodiment. According to this method it is possible to improve the training of a discriminator 330 which has already been trained at a lower standard. For example, this method may be used to improve a crude topic classifier 360 which has not previously benefitted from training according to the invention.

At step S610 in Fig. 6A, a Generative Adversarial Network, GAN 310 is provided, the GAN having a generator 320 configured to generate text sections according to a specified topic, and a discriminator 330 configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator 320. The discriminator 330 preferably comprises a Long Short-Term Memory, LSTM, classifier for identifying whether or not received text sections match the specified topic.

At step S620, the generator 320 and the discriminator 330 are trained. More specifically, the generator 320 repeatedly generates text sections for the discriminator 330 and the discriminator 330 classifies text sections which may be from the generator 320 or may be real text sections in the specified topic, thereby identifying a topic of the text section. The GAN 310 oversees this process and generates feedback indicating whether the classification output from the discriminator 330 is correct for given text sections. More specifically, the GAN 310 iteratively adapts the generator 320 to decrease the probability that the identified topic matches the specified topic, and adapts the discriminator 330 to increase the probability that the identified topic matches the specified topic.

During the iterative process of step S620, the steps of Fig. 6B are performed each time a text section is generated by generator 320.

At step S630, a reinforcement learning algorithm 340 of the generator 320 is used to obtain an initial text section comprising a plurality of words.

The reinforcement learning algorithm 340 is based on Q-learning and is trained following generation and discrimination of each text section, in order to become more likely to generate a text section that the topic classifier 360 of the discriminator 330 will classify incorrectly.

At step S640 a word embedding noise adder 350 applies noise to the text section obtained in step S630.

Noise is applied by mapping an initial word of the obtained text section from a discrete space of natural language words to a continuous vector space of word embedding. As described above, in the vector space, words are arranged at points and a distance can be defined between any two words based on the mapping. Additionally a sphere can be defined around a word with a given radius.

In order to apply the noise, a sphere is defined around the initial word of the obtained text section, the sphere having a radius of a noise distance that inversely corresponds to a degree to which the discriminator 330 correctly identified the topic of a previous text section. A set of words located within the sphere are identified, and one of the set of words is randomly selected as a final word to replace the initial word.

Then, the text section, in which the initial word has been replaced by the final word, is provided as an output from the generator 320.

When used in the above method, the mechanism of added randomness based on word embedding has the effect of improving the accuracy of the discriminator 330 after being trained in the GAN 310. Additionally, this mechanism has the effect of reducing the chance of mode collapse in which a generator 320 focusses only on a very small range of text sections and, as a result, the discriminator 330 is not trained to be widely applicable.

Fig. 7 is a schematic illustration of functional modules and data stores of a discriminator training apparatus 700 of a second embodiment. The second embodiment is based on the discriminator training apparatus 300 of the first embodiment, and can be similarly implemented except as described below.

Referring to Fig. 7, the discriminator training apparatus 700 comprises a document database 710, a topic modeller 720, a discriminator 730, a long short-term memory training controller 740, a generator 750, a reinforcement learning algorithm training controller 760 and a generative adversarial network, GAN, training controller 770.

The discriminator 730, the generator 750 and the GAN training controller 770 correspond respectively to the discriminator 330, the generator 320 and the GAN 310 respectively of Fig. 3, and the GAN training process is the same as in the first embodiment, and accordingly will not be described again here.

The document database 710 is a data store comprising a plurality of text sections each related to a topic of a set of possible topics. The text sections may be stored in separate data files, or may be individual phrases, sentences, paragraphs etc. within data files. Certain text sections may additionally relate to multiple topics.

The topic modeller 720 is a functional program module implemented by a processer and instruction store, configured to perform topic modelling in order to identify a set of topics which are present in the document database 710.

The topic modelling is performed using a Latent Dirichlet Allocation, LDA, method. The topic modeller 720 produces a topic model comprising an N-by-K matrix, where N is the number of text sections in the document database 710 and K is the number of topics identified by the LDA method to be present in the document database 710.

The long short-term memory, LSTM, training controller 740 is a functional program module implemented by a processor and instruction store, configured to construct the discriminator 730 initially and to train the discriminator 730 under the control of the GAN training controller 770. The training of the discriminator 730 under the control of the GAN training controller 770 is as described above with respect to the first embodiment, and the constructing the discriminator 730 initially is described in the following.

In this embodiment, a copy of the text sections in the document database 710 is pre-processed. All stop words, numbers and punctuation are removed. Then, the remaining words of the text sections are stemmed and lemmatized to ensure uniformity. Subsequently, a term/text section matrix is constructed.

Fig. 8 is a schematic illustration of a term/text section matrix for a set of text sections. In this example, Fig. 8 is a subset of the matrix, which would normally have a large number of rows and columns corresponding to the number of different text sections in the document database 710 and the number of different terms appearing in the text sections.

In Fig. 8, each value 830 of the matrix identifies a number of times a term 810 occurs in a text section, wherein each text section corresponds to a row (or column) of the matrix identified by a text section ID 820.

In Fig. 8, it may be noted that some spelling mistakes are included such as the term “nodesfor”. Such spelling mistakes would almost inevitably occur in a document database 710. However, such spelling mistakes could be eliminated as a further pre-processing step.

After generating the term/text section matrix, a term frequency/inverse section frequency, tf-isf, matrix is generated. The ts-isf matrix is constructed to include only a limited number of the most repeated words, for example the top hundred words, in order to remove noise.

Returning to Fig. 7, finally, the LSTM training controller 740 constructs the discriminator 730. In this embodiment, the discriminator 730 is a neural network of LSTM architecture with two hidden layers, trained using the term/text section matrix to produce information indicative of an identified topic as the output. The discriminator 730 is stored in the discriminator training apparatus 700 as shown in Fig. 7. In a test case, a network was trained for 20 epochs and achieved a training accuracy of 92%. However, such training alone does not provide a robust enough discriminator for many applications.

The reinforcement learning algorithm training controller 760 is a functional program module implemented by a processor and instruction store, configured to construct the reinforcement learning algorithm 340 of generator 750 and train the reinforcement learning algorithm 340 under the control of the GAN training controller 770. The training of the reinforcement learning algorithm 340 under the control of the GAN training controller 770 is as described above with respect to the first embodiment, and the reinforcement learning algorithm 340 is constructed initially with states as next words of the text section and rewards that are initially predetermined or random and that are updated based on outputs from the discriminator 730 on the basis of previous generated text sections.

Fig. 9 is a flow diagram showing the main processing operations performed by the discriminator training apparatus of the second embodiment.

At step S910, the topic modeller 720 identifies a set of possible topics by performing topic modelling on the document database 710, wherein the document database comprises a plurality of text sections each related to a topic of the set of possible topics.

At step S920, the long short-term memory, LSTM, training controller 740 trains the discriminator 730 based on a long short-term memory model with a term frequency-inverse section frequency matrix as input and topics as output, using text sections from the document database 710.

At step S930, the reinforcement learning algorithm training controller 760 constructs a reinforcement learning model (reinforcement learning algorithm 340) with states of partial text sections, actions of adding a next word to a text section, and rewards based on corresponding outputs from the discriminator, the reinforced learning algorithm being used to initialise the generator 750.

At step S940, the generative adversarial network, GAN, training controller 770 further trains the discriminator 730 and generator 750 in a generative adversarial network, GAN, as described above for the first embodiment.

Above are described two embodiments of a discriminator training apparatus for training a discriminator. Below are described two embodiments in which such a trained discriminator may be used to generate a summary of a document database.

Fig. 10 is a schematic illustration of functional modules and data stores of a summary generation apparatus 1000 of a third embodiment.

Referring to Fig. 10, the summary generation apparatus 1000 comprises a topic information obtainer 1010, a discriminator 1020, a document database 1030, and a summary generation controller 1040.

The topic information obtainer 1010 is a functional program module implemented by a processor and instruction store, configured to obtain topic information indicating a topic to be included in a summary of the document database 1030. The topic information may be obtained by various means depending on the content of the database. For example, the topic information may be defined manually by a user's request for a summary. Alternatively, the topic information may be calculated based on sensor readings, such as maintenance alarm sensors, environmental condition sensors etc. which may obtain relevant information about a context in which the summary is to be provided. Additionally, the topic information may be calculated using historical data regarding previously generated summaries. The topic information may only specify one topic, or may specify multiple topics.

The discriminator 1020 is a discriminator that has been trained as described above in the first or second embodiment, or otherwise in accordance with the invention. As described above, each discriminator 1020 is a binary classifier trained for discriminating one topic, and therefore multiple discriminators are required if the topic information can indicate any of multiple possible topics.

The document database 1030 is a data store comprising a plurality of text sections each related to a topic of a set of possible topics. The text sections may be stored in separate data files, or may be individual phrases, sentences, paragraphs etc. within data files. Certain text sections may additionally relate to multiple topics.

The summary generation controller 1040 is a functional program module implemented by a processor and instruction store, configured to use the discriminator 1020 (or a discriminator 1020 if there are multiple discriminators) to identify a first text section in the document database 1030 that is related to a topic indicated in the obtained topic information, and to generate a summary of the document database 1030 that includes the first text section. The first text section is identified by passing text sections from the document database 1030 to the discriminator 1020 trained for the identified topic, until a text section is discriminated as being related to the identified topic. Additionally, multiple text sections from the document database may be discriminated as being related to the identified topic, and the multiple text sections may be included in the summary.

Furthermore, in some cases, the summary must also include information which is not related to a topic indicated in the topic information. For example, there may be a “general” topic which is always included in the summary regardless of the topic information. In such cases, the summary generation controller 1040 may use a second discriminator to identify a second text section of the document database 1030 that is not related to a topic indicated by the topic information, and generate a summary of the document database 1030 which includes the first section related to the indicated topic, and the second section not related to the indicated topic.

Additionally, text sections included in the summary may be summarized using extract-based summarization techniques. In particular, it is advantageous to provide a summary of the document database 1030 in which the first text section is provided in full, and a second text section not related to the topic indicated in the topic information is only included as an extract-based summary. This means that the summary provides more detail on the topic which has been indicated in the topic information as being of particular interest, while still identifying the existence of other another topic using an extract-based summary.

More generally, the summary generation controller 1040 may apply one or more discriminators 1020 to all text sections in the document database 1030, in order to determine which text sections of the document database 1030 should be included in the summary.

The summary generation apparatus may itself additionally comprise a discriminator training apparatus for training the discriminator(s) 1020 as described above. It is advantageous for a summary document to be easy to read, understand and generate. Additionally, it is advantageous for the summary document to be easily summarized. Additionally, in the context of generating summaries, it is advantageous to train a discriminator to classify text sections into topics (which may be domain-dependent topics) with minimal information. Accordingly, a reward function for the GAN model used to train the discriminator 1020 is defined to reward these properties.

Fig. 11 is a flow diagram showing the main processing operations performed by the summary generation apparatus of the third embodiment.

Referring to Fig. 11, at step S1110, the topic information obtainer 1010 obtains topic information indicating a topic to be included in a summary of the document database 1030, wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics.

Then, at step S1120, the summary generation controller 1040 uses a trained discriminator to identify a first text section of the document database 1030 that is related to the obtained topic information.

Then, at step S1130, the summary generation controller 1040 generates a summary of the document database 1030 wherein the summary includes the first text section.

Fig. 12 is a schematic illustration of a maintained system in which a technician assistance apparatus of a fourth embodiment may be used.

More specifically, referring to Fig. 12, there is illustrated a particularly practical technical scenario in which the above-described document summarization is useful.

In Fig. 12, four sites 1220-1, 1220-2, 1220-3 and 1220-4 form a system which is to be managed. As mentioned above, the sites may be connected or unconnected. The sites may also be geographically distant from one another, such that technicians cannot attend all sites while using their time efficiently, due to significant travel times in the travel network 1230 between sites.

The sites may, for example, be locations of various industrial devices.

In order to oversee and maintain the sites 1220, technicians 1240-1, 1240-2 may be sent to each site periodically or in response to sensors which detect the need for a technician. For example, each site may have sensors to detect equipment failures.

More specifically, each technician 1240 can receive a work order indicating the need to attend a site 1220. This work order may, for example, be received via an application on a terminal 1250 provided to the technician, and may provide details of problems at the site or work to be done at the site. The work order may contain details of triggered alarms, that is, information indicating a problem at the site. Ideally, the work order contains correct and accurate information, although this is outside the scope of the described invention herein. The work order may also include information related to solving the problem, such as information regarding already-attempted strategies for solving the problem.

The technician uses their terminal 1250 to communicate with a technician assistance apparatus 1210 in order to obtain a summary of a document database of technical documentation, wherein each of a plurality of text sections of the document database are related to a topic of a problem type which may occur at the site. It is particularly advantageous for the summary to be provided as an issue-based (e.g. trouble ticket-based) solution document with appropriate details and an appropriate level of summarization to solve a problem at the site 1220 without requiring the technician to, in advance, have detailed specialized knowledge of the problem and without requiring the technician to spend large amounts of time searching a document database manually. In order to achieve this, the technician assistance apparatus of the fourth embodiment generates a summary of the document database as a scope-enabled regenerative solution document.

Fig. 13 is a schematic illustration of functional modules and data stores of the technician assistance apparatus of the fourth embodiment.

The technician assistance apparatus 1300 comprises a topic information obtainer 1310, a site history storage 1320, a technician proficiency storage 1330, a discriminator 1340, a document database 1350 and a summary generation controller 1360.

The topic information obtainer 1310, discriminator 1340, document database 1350 and summary generation controller 1360 may be similar to the corresponding features of the third embodiment as described above, except where otherwise described below, and similar details are not repeated for brevity.

The topic information obtainer 1310 in the fourth embodiment calculates topic information on the basis of a request for a summary document.

For example, referring to Fig. 12, the technician assistance apparatus 1210 may receive a request for assistance from a technician 1240-1 who is attending a site 1220-1, via their terminal 1250-1. The request includes a technician ID which identifies the technician 1240-1, and a site ID which identifies the site 1220-1 currently being attended by the technician.

The topic information obtainer 1310 then applies rules to determine topic information based on site history information associated with the site ID in the site history storage 1320 and/or technician proficiency information associated with the technician ID in the technician proficiency storage 1330.

More specifically, the site history storage 1320 stores, for each site 1220 in the managed system, information regarding problems which have previously occurred in at the site. This information may include a frequency of each type of problem, a last date on which the problem occurred, a burden to the system associated with fixing the problem, a burden to the system associated with failing to fix the problem (some problems only have isolated impact whereas others may impact a whole site or system), and so on. The site history storage 1320 may also store information regarding summary documents which have been previously generated for a site, known characteristics of a site (such as what equipment, parts, tools etc. are present at the site). All of this information may be used to determine which problems are most likely to occur at the site 1220 and which potential problems should be prioritised by the technician 1240.

Similarly, the technician proficiency storage 1330 stores, for each technician 1240 in the managed system, information regarding problems which they have previously solved or which they are likely to be able to solve given their specialization.

In particular, this information may be linked to the document database 1350 and may indicate particular text sections with which the technician already has proficiency.

The information associated with each technician is obtained by constructing a Latent Dirichlet Allocation model for each text section with which they are familiar, in order to obtain more general information about which topics the technician is likely to be familiar with. For example, if the technician is known to be familiar with M text sections out of N text sections in the document database 1350, and it has been established that the document database 1350 has text sections corresponding to K topics, an M by K LDA model can be constructed to represent the proficiency (or proficiencies) of the technician.

The rules used by the topic information obtainer 1310 may be static predetermined rules or may be determined by an additional model which undergoes machine learning based on the successes and failures of the technician assistance apparatus. For example, the topic information obtainer 1310 can use rules to prioritise topics associated with relatively important possible problems and to prioritise topics with which the technician has lower proficiency.

In particular, the topic information obtainer 1310 may advantageously determine topic information which comprises information indicating a topic associated with a type of problem which has occurred previously at the site.

Additionally or alternatively, the topic information obtainer 1310 may advantageously determine topic information which comprises information indicating a topic associated with type of problem with which the technician does not have experience and therefore is not proficient in. This determination is particularly useful in a case where the technicians 1240-1 and 1240-2 have different proficiencies (i.e. different levels of experience), because, due to the geographical distances and travel times between sites, it is not always possible to assign the most proficient technician to each site for dealing with corresponding reported problems, and therefore assistance to the technician is more useful if the technician assistance apparatus can identify topics in the document database with which the technician may require more assistance.

This intelligent selection of topics to be included in the summary leads to a summary which is more likely to be successful in assisting the technician in solving a problem, and more likely to reduce the time taken by the technician to solve the problem.

The discriminator 1340 is a discriminator that has been trained as described above in the first or second embodiment, or otherwise in accordance with the invention, to efficiently and correctly identify topics associated with text sections in the document database 1350. As described above, each discriminator 1020 is a binary classifier trained for discriminating one topic, and therefore multiple discriminators are required if the topic information can indicate any of multiple possible topics.

As described above for the third embodiment, the technician assistance apparatus 1300 may comprise discriminator training apparatus in accordance with the first or second embodiment. In this case, discriminators are trained in advance for various topics (which may correspond to different potential problems in the system or different proficiencies of technicians) and are stored ready for use as required. The discriminators 1340 can be updated when there is any change to the document database 1350, for example by addition of new text sections to the database or removal of text sections from the database.

With respect to computation time, it is noteworthy that according to the invention a GAN model is only used for training a robust discriminator 1340, and is not required at the time of generating a summary. Instead, only the discriminator(s) are required at the time of summarizing a document database 1350, in order to identify text sections to include in the summary. At that stage, the discriminator(s) are already trained and are able to efficiently and accurately classify the text sections without requiring processing or time resources that are comparable to the training stage. Accordingly, the invention is useful even in contexts where GAN training would be too slow or computationally expensive for assisting a technician in real-time.

As mentioned above, the document database 1350 is a database of technical documentation, wherein each of a plurality of text sections of the document database are related to a topic of a problem type which may occur at the site. The database otherwise works similarly to the third embodiment.

The summary generation controller 1360 also works similarly to the third embodiment, and uses the discriminator(s) 1340 to identify one or more text sections of the document database 1350 to be included in a generated summary. Ideally, the summary generation controller 1360 should be configured to selectively include whole text sections and/or extractively-summarized text sections, in order to generate a summary which is simple for the technician to use and understand.

As an additional optional feature of the technician assistance apparatus 1300, the apparatus may, after providing a summary for assisting a technician, obtain feedback from the technician regarding whether problem was solved and/or whether the summary was helpful in solving the problem. If the summary was not effective in assisting the technician, the technician assistance apparatus 1300 will regenerate a further solution document. At the same time, the technician assistance apparatus may provide feedback to the topic information obtainer 1310 in order to improve its rules for determining topic information. Additionally, the summary itself may be stored for future use if it was successful in assisting the technician in solving the problem. The stored summary may be reviewed by an expert once it is known to be successful, in order to finalize the summary as a solution document to be retained. Then, if a similar problem arises again, the stored summary can be provided as an initial attempt to assist in solving the similar problem with minimal resources.

There will now be described an example that can assist in understanding the utility of the described technician assistance apparatus.

Referring to Fig. 12, assume that there is a problem at site 1220-1 and, referring to Fig. 13, assume that the document database 1350 contains ten text sections which each describe troubleshooting techniques for a respective component of the system at site 1220-1.

In a comparative example, the technician only has access to the whole document database, and is not provided with a summary. In that case, the technician has to read through up to ten text sections and perform corresponding troubleshooting until the problem solved. For example, the technician may troubleshoot components associated with each text section until a broken or incorrectly operating component is identified and repaired. If the last text section is the one related to the current problem, then the technician must read ten text sections and troubleshoot ten components, where the work on nine out of ten components does not solve the problem.

On the other hand, when using a technician assistance apparatus 1300 according to the invention, the site history storage 1320 indicates a frequency of one or more problems which have previously occurred. For example, the site history storage 1320 may indicate that, in the past two times a same or similar problem was reported at site 1220-1, the problem was solved by troubleshooting a particular component. Accordingly, the summary is generated to indicate that troubleshooting for the particular component should be prioritized over troubleshooting for another component. This will decrease the average time for solving the problem.

Furthermore, in the case where the information associated with a technician ID in the technician proficiency storage 1330 corresponding to a technician 1240 attending site 1220-1 indicates that they are proficient with the problem that has occurred in the past two times, but they are less proficient with another problem which has previously occurred at site 1220-1, this time associated with a second component found at the site, the summary is now generated to provide a summary of information regarding the particular component that has been the cause of the problem the past two times, and to provide more detailed information regarding the other problem.

Fig. 14 is a flow diagram showing the main processing operations performed by the technician assistance apparatus of the fourth embodiment.

Referring to Fig. 14, at step S1410, the technician assistance apparatus 1300 receives a request for a summary document, the request including a technician ID and a site ID. This request is received from the technician's terminal 1250.

At step S1420, the topic information obtainer 1310 determines topic information indicating a topic to be included in a summary of a document database 1350, based on site history information corresponding to the site ID stored in the site history storage 1320 and a proficiency corresponding to a technician ID of the technician in the technician proficiency storage 1330.

At step S1430, the summary generation controller 1360 uses a discriminator 1340 to identify a first text section of the document database 1350 that is related to the topic information determined by the topic information obtainer 1310 in step S1420.

At step S1440, the summary generation controller 1360 generates the summary of the document database 1350, where the summary includes the first text section identified in step S1430.

Then, at step S1450, the technician assistance apparatus 1300 transmits the generated summary to the terminal 1250 from which the request was received, to assist the technician associated with the technician ID in solving a problem at the site associated with the site ID.

Modifications

In the above description, the discriminator and generator are implemented using neural networks such as Convolutional Neural Networks, CNN, or Recursive Neural Networks, RNN. However, the discriminator and generator may each be any type of machine learning model capable of receiving the above-described input(s), producing the above-described output(s) and learning from the above-described feedback(s).

In the above description, noise is added using word embedding and randomly selecting a final word within a noise distance of an initial word that is to be replaced. In some embodiments, the random selection may be guided by an additional reinforcement learning, RL, model in order to ensure that the final word is semantically similar to the initial word, so that the text section remains, as a whole, a natural language text section (or, conversely, in order to ensure that the final word is not semantically similar to the initial word, when the previous text section generated by the generator was correctly discriminated).

In the above description of Fig. 3, a reinforcement learning algorithm 340 generates a text section and then a word embedding noise adder 350 adds noise to the text section. However, in other embodiments, the word embedding noise adder 350 may add noise to a partial text section while it is still being generated in the reinforcement learning algorithm. In other words, the word embedding noise adder 350 may be used to randomly vary an intermediate state in the above-described Q-learning model such that the corresponding state and Q-values of the partially-generated text section are altered in between multiple actions of adding words according to the Q-learning model.

In the above description, the discriminator is a binary classifier that identifies whether or not a text section matches one topic or not, or identifies a probability of the one topic being associated with the text section. Accordingly, where identification of multiple topics is required, multiple discriminators are used. However, in other embodiments, the discriminator may be configured to produce a binary output or probability output with respect to multiple topics. In such other embodiments, the discriminator may be trained in a Generative Adversarial Network with one or more generators that are capable of generating text sections for any of a plurality of topics.

In the above description, it is only mentioned that the document database includes text sections. However, the document database may also include images such as photographs or graphs. Where such images are associated with text sections, for example by being located adjacent to a text section or by being mentioned in a text section, the above-described techniques can be used to generate a summary which also includes images associated with the text sections used in the summary.

In the above description, the summary is generated from a document database. However, the trained discriminator(s) may be used to summarize other information. For example, the trained discriminator(s) may be used to summarize a work order regarding a problem at a site. More specifically, work orders often include copy/pasting of, for example, alarm data or error data and history not relevant to or related to the current problem at the site. As a result, it is useful to be able to summarize either or both of the work order, which provides specific information about a current problem at a site, and the document database, which provides general information about known problems and solutions.

In the above description, a discriminator training apparatus which trains a discriminator, and a summary generation apparatus and a technician assistance apparatus each of which use a discriminator, are described as separate embodiments. However, these embodiments may be combined to provide a summary generation apparatus comprising a discriminator training apparatus, or a technician assistance apparatus comprising a discriminator training apparatus. The corresponding described methods may also be combined into single methods in which the same discriminator is trained and used.

In the above description, a technician assistance apparatus is used to assist a technician in solving a problem at a site. However, more generally, the invention can be used in any domain where technicians manage a system and handle various trouble shooting problems. For example, a technician working in a cloud-based service to assist customers remotely in solving their problems can use the described technician assistance apparatus to generate a regenerative solution for the problems and to increase the likelihood of successfully helping the customer to solve the problems.

Claims:

1. A computer-implemented method for training a discriminator (330) to identify a topic to which a received text section is related, the method comprising:

providing (S610) a generative adversarial network, GAN, (310) comprising:

a generator (320) configured to generate text sections according to a specified topic, and

the discriminator (330) configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator; and

iteratively adapting (S620) the generator (320) to decrease the probability that the identified topic matches the specified topic and adapting the discriminator (330) to increase the probability that the identified topic matches the specified topic,

wherein the generator (320) is configured to generate a text section by:

obtaining (S630) a text section using a reinforcement learning algorithm (340), the obtained text section comprising a plurality of words; and

applying (S640) noise to the obtained text section by:

performing word embedding to map a word of the obtained text section to a point in a continuous vector space,

identifying a replacement word that is within a noise distance from the point in the continuous vector space, and

replacing the mapped word with the replacement word in the obtained text section to generate the text section.

2. A computer-implemented method according to claim 1, wherein a magnitude of the noise applied (S640) to the obtained text section for a given generated text section has an inverse relationship with a degree to which the identified topic matches the specified topic for a preceding generated text section.

3. A computer-implemented method according to claim 1 or claim 2, wherein the reinforcement learning algorithm (340) is a Q-learning algorithm.

4. A computer-implemented method according to any of claims 1 to 3, wherein the discriminator (330) comprises a Long Short-Term Memory, LSTM, classifier.

5. A computer-implemented method for generating a summary of a document database (1030), wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics, the method comprising:

obtaining (S1110) topic information indicating a topic to be included in the summary;

using (S1120) a discriminator (1020) that has been trained in accordance with any of claims 1 to 4 to identify a first text section of the document database that is related to the topic to be included in the summary; and

generating (S1130) the summary of the document database to include the first text section.

6. A computer-implemented method according to claim 5, further comprising training (S940) the discriminator (730, 1020) in accordance with any of claims 1 to 4.

7. A computer-implemented method according to claim 6, wherein training the discriminator (730, 1020) further comprises identifying (S910) the set of possible topics by performing topic modelling on the document database (710, 1030).

8. A computer-implemented method according to any of claims 5 to 7, the method further comprising:

using a second discriminator to identify a second text section of the document database that is not related to a topic indicated by the topic information to be included in the summary; and

generating the summary of the document database to include the first text section and a summary of the second text section.

9. A computer-implemented method for assisting a technician (1240-1) in solving a problem at a site (1220-1), the method comprising providing (S1450) a summary of a document database (1350) which has been generated (S1440) in accordance with any of claims 5 to 8, wherein the document database comprises technical documentation and the plurality of text sections are each related to a topic of a problem type which may occur at the site.

10. A computer-implemented method according to claim 9, further comprising generating (S1440) the summary in accordance with any of claims 5 to 8.

11. A computer-implemented method according to claim 9 or claim 10, wherein the topic information comprises information indicating a topic of a problem type which has occurred previously at the site (1220-1).

12. A computer-implemented method according to any of claims 9 to 11, wherein the topic information comprises information indicating a topic of a problem type with which the technician (1240-1) does not have experience.

13. A computer-implemented method according to claim 12, wherein the site (1220-1) is one of a plurality of geographically distant sites (1220-1, 1220-2, 1220-3, 1220-4) of a system maintained by a plurality of different technicians (1240-1, 1240-2) with different experience in different problem types.

14. An apparatus (300) for training a discriminator (330) to identify a topic to which a received text section is related, the apparatus comprising:

a generative adversarial network, GAN, (310) comprising:

a generator (320) configured to generate text sections according to a specified topic, and

the discriminator (330) configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator,

wherein the GAN is configured to iteratively adapt (S620) the generator (320) to decrease the probability that the identified topic matches the specified topic and adapting the discriminator (330) to increase the probability that the identified topic matches the specified topic, and

wherein the generator (320) is configured to generate a text section by:

obtaining (S630) a text section using a reinforcement learning algorithm (340), the obtained text section comprising a plurality of words; and

applying (S640) noise to the obtained text section by:

performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and

replacing the mapped word with the replacement word in the obtained text section to generate the text section.

15. An apparatus according to claim 14, wherein a magnitude of the noise applied (S640) by the generator (320) to the obtained text section for a given generated text section has an inverse relationship with a degree to which the identified topic matches the specified topic for a preceding generated text section.

16. An apparatus according to claim 14 or claim 15, wherein the reinforcement learning algorithm (340) is a Q-learning algorithm.

17. An apparatus according to any of claims 14 to 16, wherein the discriminator (330) comprises a Long Short-Term Memory, LSTM, classifier.

18. An apparatus (1000) for generating a summary of a document database (1030), wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics, the apparatus comprising:

a topic information obtainer (1010) configured to obtain (S1110) topic information indicating a topic to be included in the summary;

a discriminator (1020) that has been trained using apparatus according to any of claims 14 to 17 to identify a first text section of the document database that is related to the topic to be included in the summary; and

a summary generation controller (1040) configured to generate (S1130) the summary of the document database to include the first text section.

19. An apparatus (1000) according to claim 18, further comprising a discriminator training apparatus (300) according to any of claims 14 to 17.

20. An apparatus (1000) according to claim 19, wherein training the discriminator (730, 1020) further comprises identifying (S910) the set of possible topics by performing topic modelling on the document database (710, 1030).

21. An apparatus (1000) according to any of claims 18 to 20, the apparatus further comprising:

a second discriminator configured to identify a second text section of the document database that is not related to a topic indicated by the topic information to be included in the summary,

wherein the summary generation controller (1040) is configured to generate the summary of the document database to include the first text section and a summary of the second text section.

22. An apparatus (1210, 1300) for assisting a technician (1240-1) in solving a problem at a site (1220-1), the apparatus being configured to provide (S1450) a summary of a document database (1350) which has been generated (S1440) using apparatus according to any of claims 18 to 21, wherein the document database comprises technical documentation and the plurality of text sections are each related to a topic of a problem type which may occur at the site.

23. An apparatus (1300) according to claim 22, further comprising apparatus (1310, 1340, 1350, 1360) according to any of claims 18 to 22 configured to generate (S1440) the summary.

24. An apparatus (1300) according to claim 22 or claim 23, further comprising a site history storage (1320) storing information indicating a topic of a problem type which has occurred previously at the site (1220-1).

25. An apparatus (1300) according to any of claims 22 to 24, further comprising a technician proficiency storage (1330) storing information indicating a topic of a problem type with which the technician (1240-1) does not have experience.

26. An apparatus (1300) according to claim 25, wherein the site (1220-1) is one of a plurality of geographically distant sites (1220-1, 1220-2, 1220-3, 1220-4) of a system maintained by a plurality of different technicians (1240-1, 1240-2) with different experience in different problem types.

27. An apparatus (1300) according to any of claims 22 to 26, further comprising an input/output section (510) for communicating with a terminal (1250-1) of the technician (1240-1).

28. An apparatus according to any of claims 14 to 27, wherein the apparatus is a cloud node.

29. An apparatus (500) comprising a processor (520), a memory (530) and an instruction store (540), the instruction store storing computer program instructions which, when executed by the processor, cause the processor to perform a method for training a discriminator (330) to identify a topic to which a received text section is related, the method comprising:

providing (S610) a generative adversarial network, GAN, (310) comprising:

a generator (320) configured to generate text sections according to a specified topic, and

the discriminator (330) configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator; and

iteratively adapting (S620) the generator (320) to decrease the probability that the identified topic matches the specified topic and adapting the discriminator (330) to increase the probability that the identified topic matches the specified topic,

wherein the generator (320) is configured to generate a text section by:

obtaining (S630) a text section using a reinforcement learning algorithm (340), the obtained text section comprising a plurality of words; and

applying (S640) noise to the obtained text section by:

performing word embedding to map a word of the obtained text section to a point in a continuous vector space,

identifying a replacement word that is within a noise distance from the point in the continuous vector space, and

replacing the mapped word with the replacement word in the obtained text section to generate the text section.

30. An apparatus (500) according to claim 29, wherein a magnitude of the noise applied (S640) to the obtained text section for a given generated text section has an inverse relationship with a degree to which the identified topic matches the specified topic for a preceding generated text section.

31. An apparatus (500) according to claim 29 or claim 30, wherein the reinforcement learning algorithm (340) is a Q-learning algorithm.

32. An apparatus (500) according to any of claims 29 to 31, wherein the discriminator (330) comprises a Long Short-Term Memory, LSTM, classifier.

33. An apparatus (500) comprising a processor (520), a memory (530) and an instruction store (540), the instruction store storing computer program instructions which, when executed by the processor, cause the processor to perform a method for generating a summary of a document database (1030), wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics, the method comprising:

obtaining (S1110) topic information indicating a topic to be included in the summary;

using (S1120) a discriminator (1020) that has been trained by an apparatus according to any of claims 29 to 32 to identify a first text section of the document database that is related to the topic to be included in the summary; and

generating (S1130) the summary of the document database to include the first text section.

34. An apparatus (500) according to claim 33, the computer program instructions further causing the processor to perform training (S940) the discriminator (730, 1020) in accordance with any of claims 29 to 32.

35. An apparatus (500) according to claim 34, wherein training the discriminator (730, 1020) further comprises identifying (S910) the set of possible topics by performing topic modelling on the document database (710, 1030).

36. An apparatus (500) according to any of claims 33 to 35, the computer program instructions further causing the processor to perform:

using a second discriminator to identify a second text section of the document database that is not related to a topic indicated by the topic information to be included in the summary; and

generating the summary of the document database to include the first text section and a summary of the second text section.

37. An apparatus (500) comprising a processor (520), a memory (530) and an instruction store (540), the instruction store storing computer program instructions which, when executed by the processor, cause the processor to perform a method for assisting a technician (1240-1) in solving a problem at a site (1220-1), the method comprising providing (S1450) a summary of a document database (1350) which has been generated (S1440) by an apparatus according to any of claims 33 to 36, wherein the document database comprises technical documentation and the plurality of text sections are each related to a topic of a problem type which may occur at the site.

38. An apparatus (500) according to claim 37, the computer program instructions further causing the processor to perform generating (S1440) the summary in accordance with any of claims 33 to 36.

39. An apparatus (500) according to claim 37 or claim 38, wherein the topic information comprises information indicating a topic of a problem type which has occurred previously at the site (1220-1).

40. An apparatus (500) according to any of claims 37 to 39, wherein the topic information comprises information indicating a topic of a problem type with which the technician (1240-1) does not have experience.

41. An apparatus (500) according to claim 40, wherein the site (1220-1) is one of a plurality of geographically distant sites (1220-1, 1220-2, 1220-3, 1220-4) of a system maintained by a plurality of different technicians (1240-1, 1240-2) with different experience in different problem types.

42. An apparatus (1300) according to any of claims 37 to 41, further comprising an input/output section (510) for communicating with a terminal (1250-1) of the technician (1240-1).

43. An apparatus according to any of claims 29 to 42, wherein the apparatus is a cloud node.

44. A computer-readable storage medium (550), having stored thereon computer program instructions which, when executed by a processor, cause the processor to perform a method for training a discriminator (330) to identify a topic to which a received text section is related, the method comprising:

providing (S610) a generative adversarial network, GAN, (310) comprising:

a generator (320) configured to generate text sections according to a specified topic, and

the discriminator (330) configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator; and

iteratively adapting (S620) the generator (320) to decrease the probability that the identified topic matches the specified topic and adapting the discriminator (330) to increase the probability that the identified topic matches the specified topic,

wherein the generator (320) is configured to generate a text section by:

obtaining (S630) a text section using a reinforcement learning algorithm (340), the obtained text section comprising a plurality of words; and

applying (S640) noise to the obtained text section by:

performing word embedding to map a word of the obtained text section to a point in a continuous vector space,

identifying a replacement word that is within a noise distance from the point in the continuous vector space, and

replacing the mapped word with the replacement word in the obtained text section to generate the text section.

45. A computer-readable storage medium (550) according to claim 44, wherein a magnitude of the noise applied (S640) to the obtained text section for a given generated

text section has an inverse relationship with a degree to which the identified topic matches the specified topic for a preceding generated text section.

46. A computer-readable storage medium (550) according to claim 44 or claim 45, wherein the reinforcement learning algorithm (340) is a Q-learning algorithm.

47. A computer-readable storage medium (550) according to any of claims 44 to 46, wherein the discriminator (330) comprises a Long Short-Term Memory, LSTM, classifier.

48. A computer-readable storage medium (550), having stored thereon computer program instructions which, when executed by a processor, cause the processor to perform a method for generating a summary of a document database (1030), wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics, the method comprising:

obtaining (S1110) topic information indicating a topic to be included in the summary;

using (S1120) a discriminator (1020) that has been trained by an apparatus according to any of claims 44 to 47 to identify a first text section of the document database that is related to the topic to be included in the summary; and

generating (S1130) the summary of the document database to include the first text section.

49. A computer-readable storage medium (550) according to claim 48, the computer program instructions further causing the processor to perform training (S940) the discriminator (730, 1020) in accordance with any of claims 44 to 47.

50. A computer-readable storage medium (550) according to claim 49, wherein training the discriminator (730, 1020) further comprises identifying (S910) the set of possible topics by performing topic modelling on the document database (710, 1030).

51. A computer-readable storage medium (550) according to any of claims 48 to 50, the computer program instructions further causing the processor to perform:

using a second discriminator to identify a second text section of the document database that is not related to a topic indicated by the topic information to be included in the summary; and

generating the summary of the document database to include the first text section and a summary of the second text section.

52. A computer-readable storage medium (550), having stored thereon computer program instructions which, when executed by a processor, cause the processor to perform a method for assisting a technician (1240-1) in solving a problem at a site (1220-1), the method comprising providing (S1450) a summary of a document database (1350) which has been generated (S1440) by an apparatus according to any of claims 48 to 51, wherein the document database comprises technical documentation and the plurality of text sections are each related to a topic of a problem type which may occur at the site.

53. A computer-readable storage medium (550) according to claim 52, the computer program instructions further causing the processor to perform generating (S1440) the summary in accordance with any of claims 33 to 36.

54. A computer-readable storage medium (550) according to claim 52 or claim 53, wherein the topic information comprises information indicating a topic of a problem type which has occurred previously at the site (1220-1).

55. A computer-readable storage medium (550) according to any of claims 52 to 54, wherein the topic information comprises information indicating a topic of a problem type with which the technician (1240-1) does not have experience.

56. A computer-readable storage medium (550) according to claim 55, wherein the site (1220-1) is one of a plurality of geographically distant sites (1220-1, 1220-2, 1220-3, 1220-4) of a system maintained by a plurality of different technicians (1240-1, 1240-2) with different experience in different problem types.

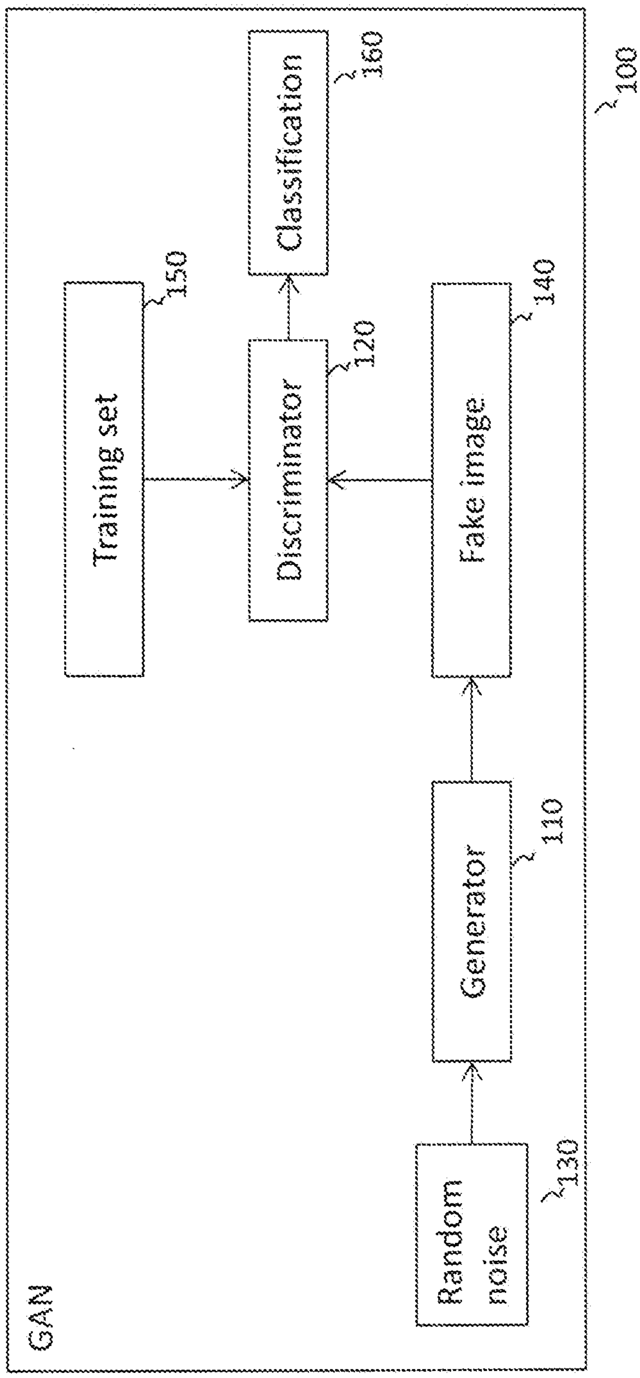
GAN
Random noise
130
Generator
110
Fake image
140
Training set
150
Discriminator
120
Classification
160
100

Fig. 1

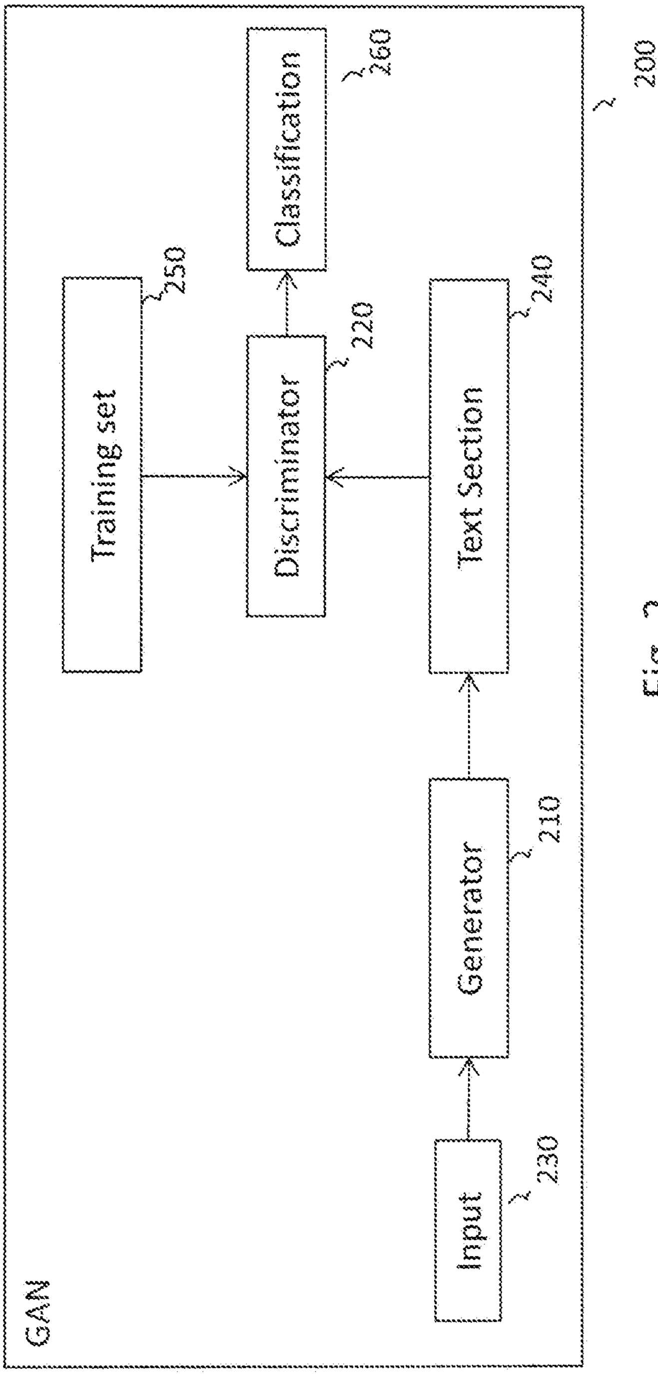
GAN
Input
230
Generator
210
Text Section
240
Training set
250
Discriminator
220
Classification
260
200

Fig. 2

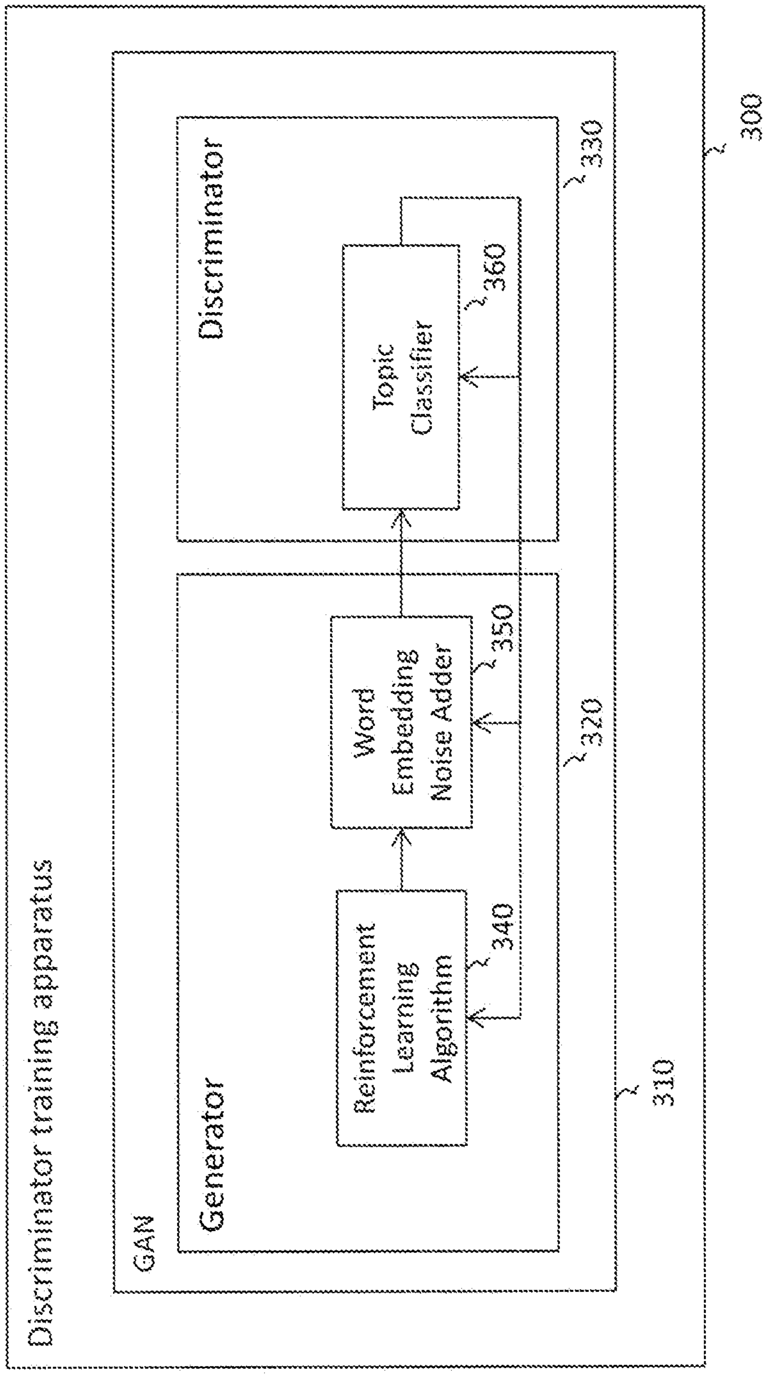
Discriminator training apparatus
GAN
Generator
Discriminator
Reinforcement Learning Algorithm
Word Embedding Noise Adder
Topic Classifier
300
310
320
330
340
350
360

Fig. 3

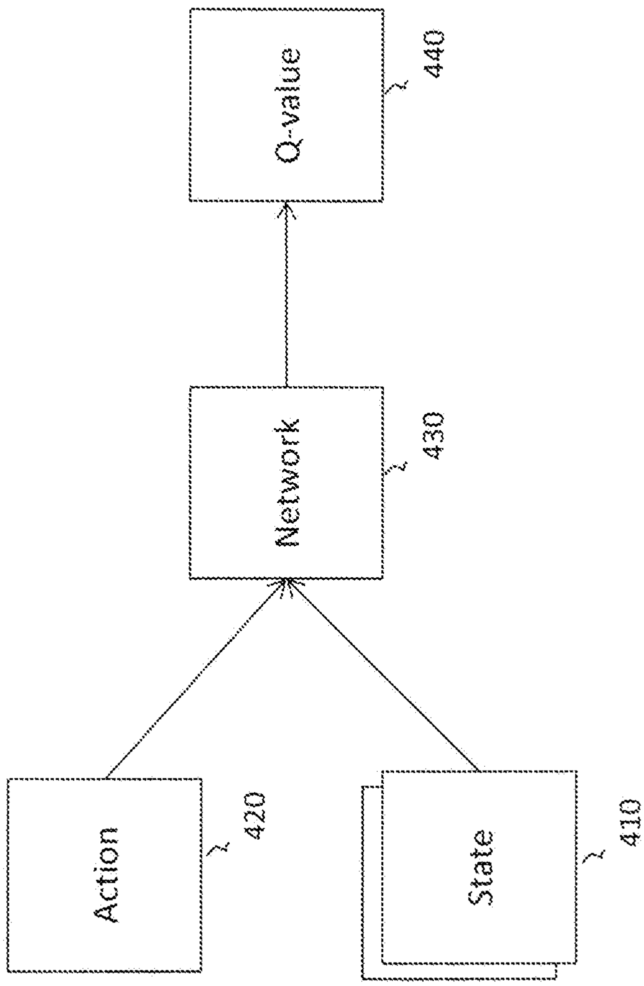

Fig. 4

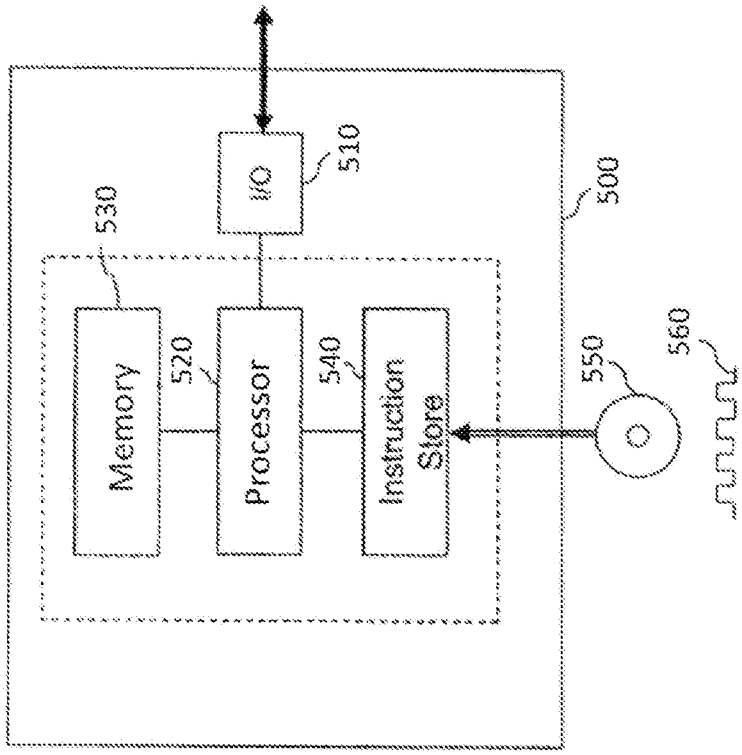
I/O
510
530
Memory
520
Processor
540
Instruction Store
500
550
560

Fig. 5

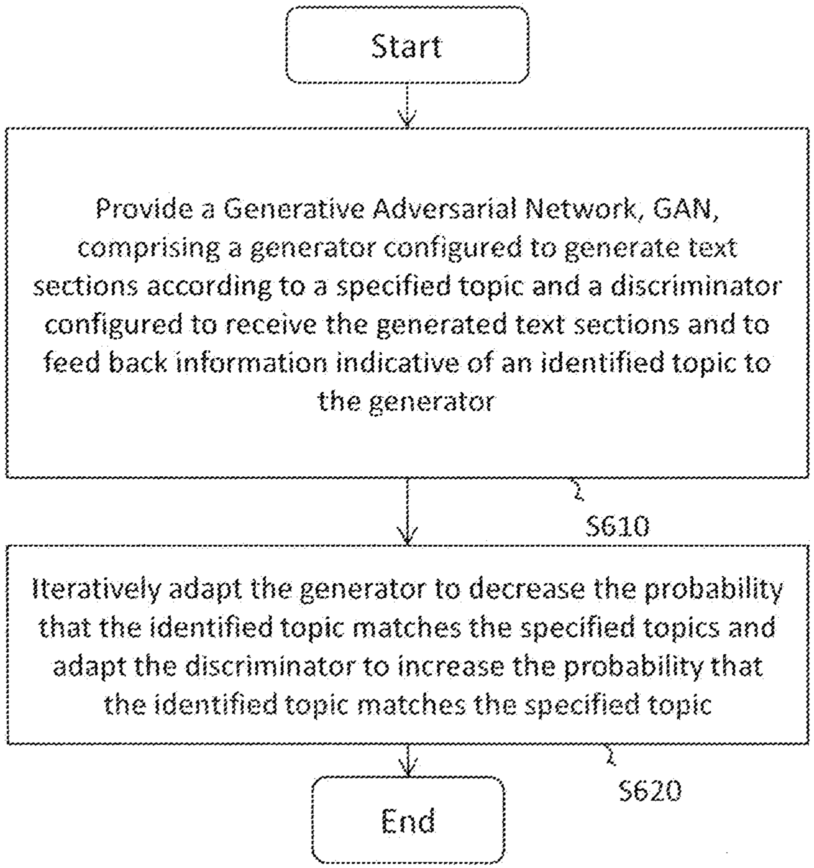
Start
Provide a Generative Adversarial Network, GAN, comprising a generator configured to generate text sections according to a specified topic and a discriminator configured to receive the generated text sections and to feed back information indicative of an identified topic to the generator
S610
Iteratively adapt the generator to decrease the probability that the identified topic matches the specified topics and adapt the discriminator to increase the probability that the identified topic matches the specified topic
S620
End

Fig. 6A

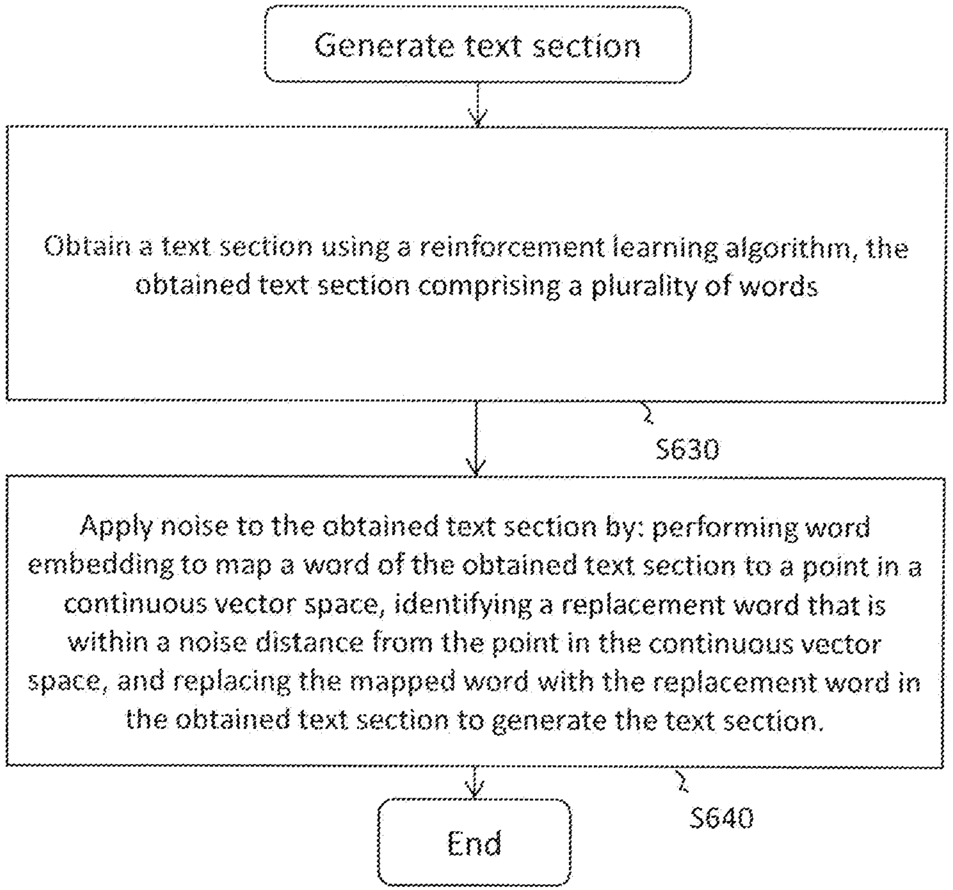
Generate text section
Obtain a text section using a reinforcement learning algorithm, the obtained text section comprising a plurality of words
S630
Apply noise to the obtained text section by: performing word embedding to map a word of the obtained text section to a point in a continuous vector space, identifying a replacement word that is within a noise distance from the point in the continuous vector space, and replacing the mapped word with the replacement word in the obtained text section to generate the text section.
S640
End

Fig. 6B

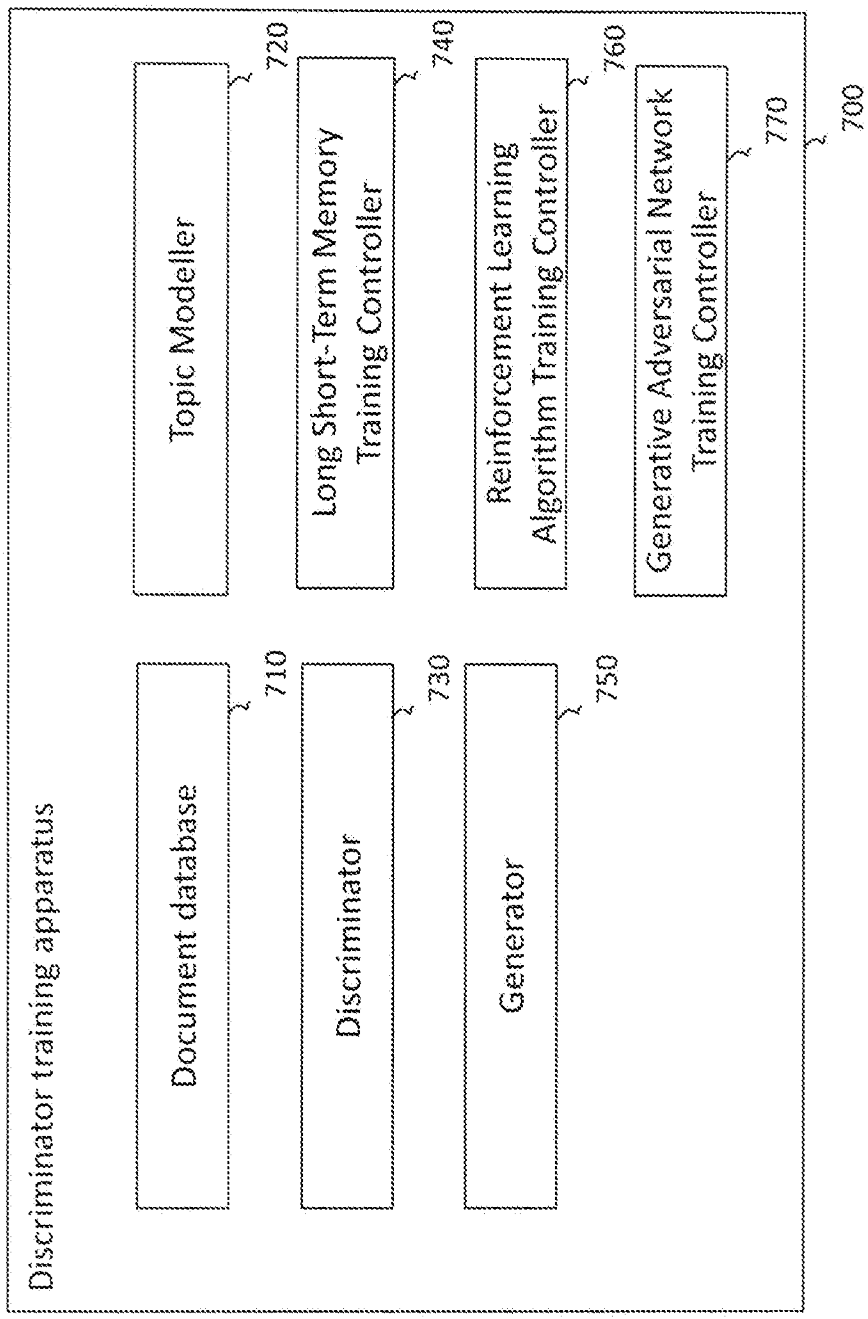
Discriminator training apparatus
Document database
710
Discriminator
730
Generator
750
Topic Modeller
720
Long Short-Term Memory Training Controller
740
Reinforcement Learning Algorithm Training Controller
760
Generative Adversarial Network Training Controller
770
700

Fig. 7

810

Text section ID	Term						
	Next	No	Node	Nodesfor	Non	Nonlinear	Normal
0	0	1	2	1	0	0	0
1	0	0	15	0	2	0	0
2	0	0	8	1	0	0	0
3	0	0	1	0	0	0	0
4	5	4	69	0	3	1	1
5	2	0	8	0	1	0	0

820

830

800

Fig. 8

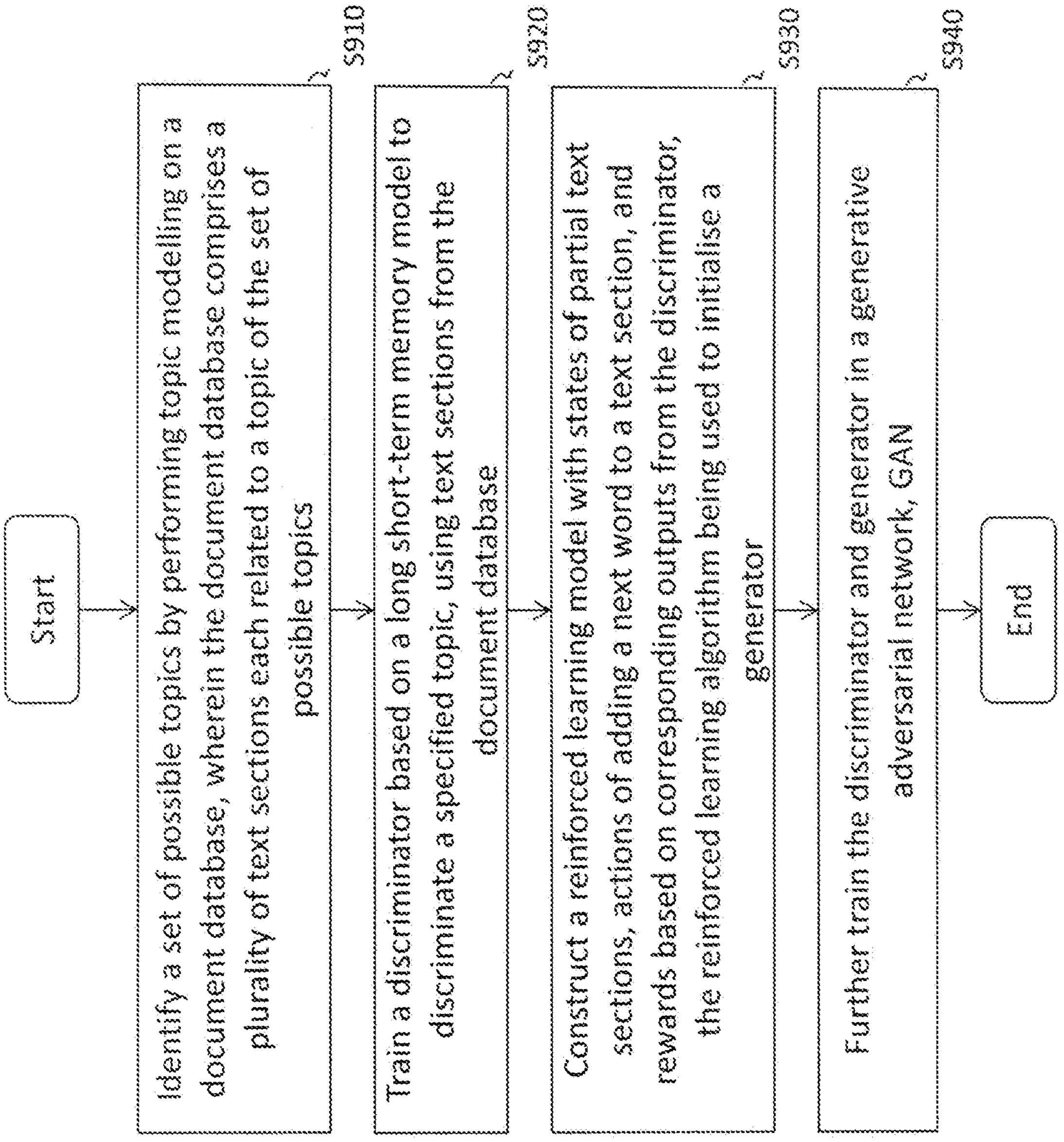
Start
Identify a set of possible topics by performing topic modelling on a document database, wherein the document database comprises a plurality of text sections each related to a topic of the set of possible topics
S910
Train a discriminator based on a long short-term memory model to discriminate a specified topic, using text sections from the document database
S920
Construct a reinforced learning model with states of partial text sections, actions of adding a next word to a text section, and rewards based on corresponding outputs from the discriminator, the reinforced learning algorithm being used to initialise a generator
S930
Further train the discriminator and generator in a generative adversarial network, GAN
S940
End

Fig. 9

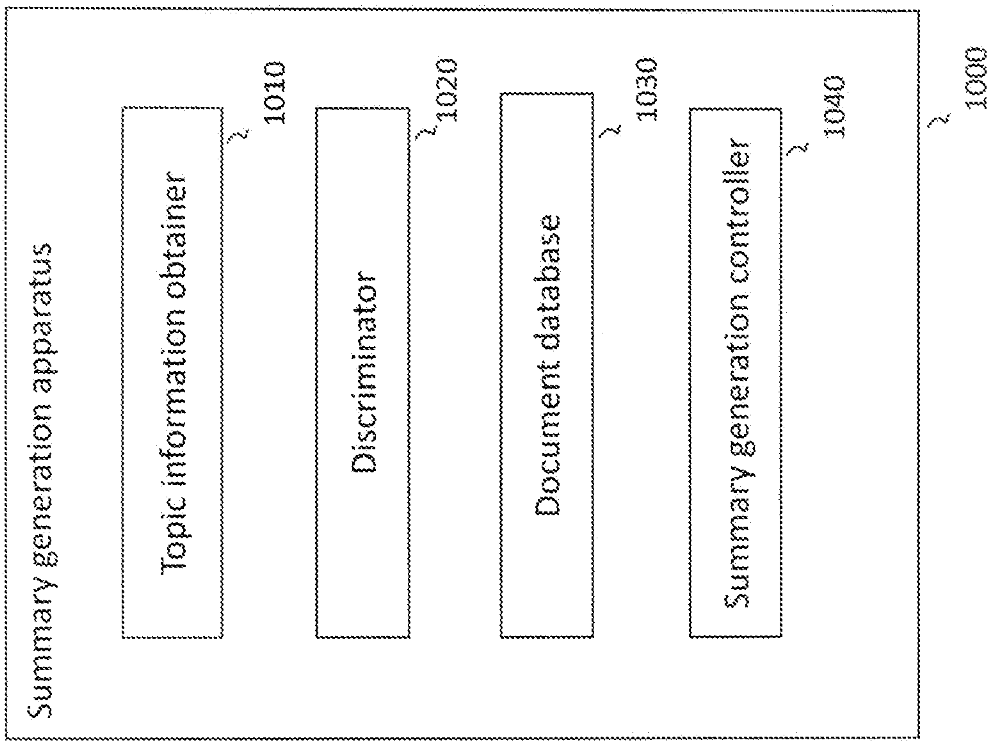
Summary generation apparatus
Topic information obtainer
1010
Discriminator
1020
Document database
1030
Summary generation controller
1040
1000

Fig. 10

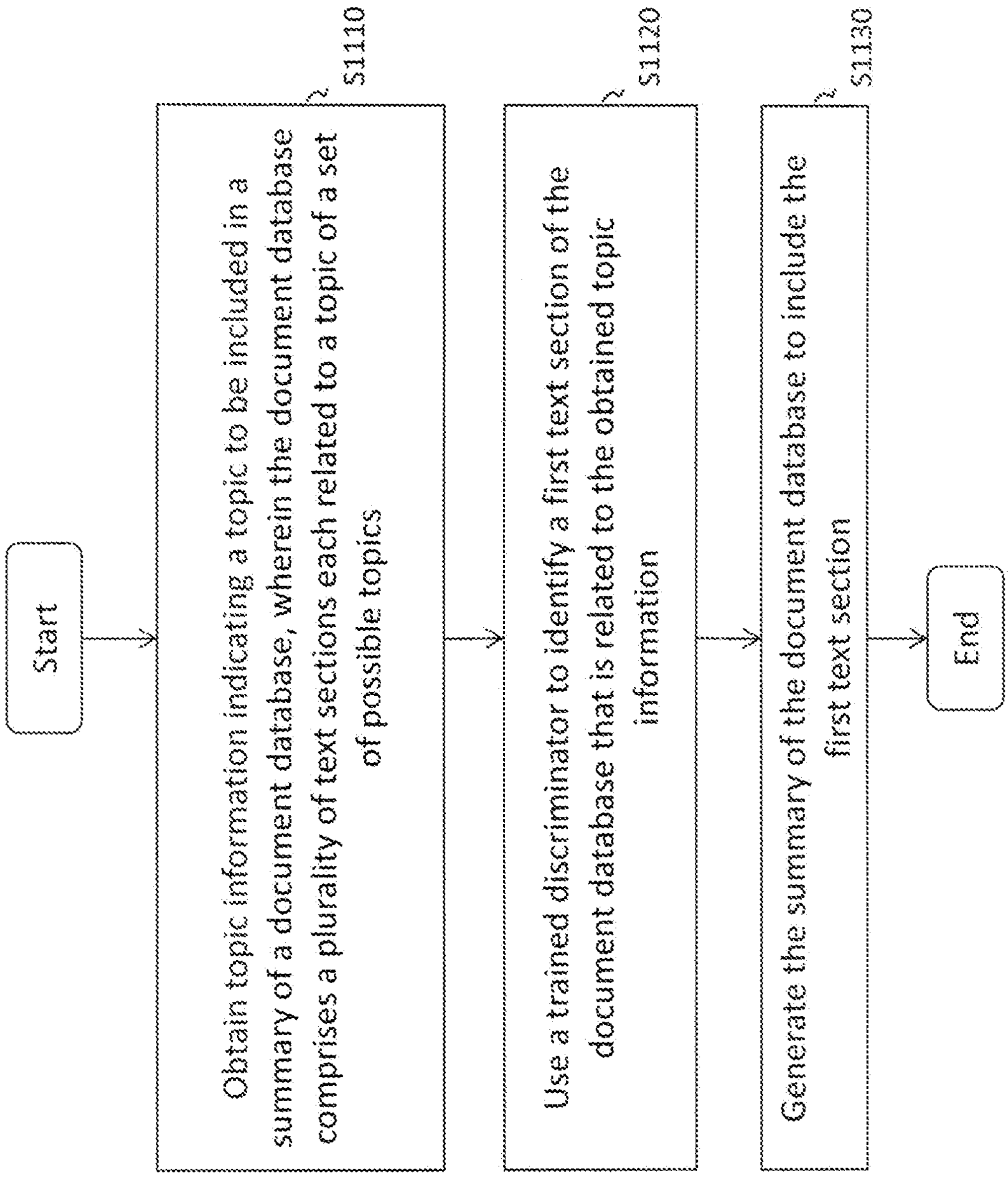
Start
Obtain topic information indicating a topic to be included in a summary of a document database, wherein the document database comprises a plurality of text sections each related to a topic of a set of possible topics
S1110
Use a trained discriminator to identify a first text section of the document database that is related to the obtained topic information
S1120
Generate the summary of the document database to include the first text section
S1130
End

Fig. 11

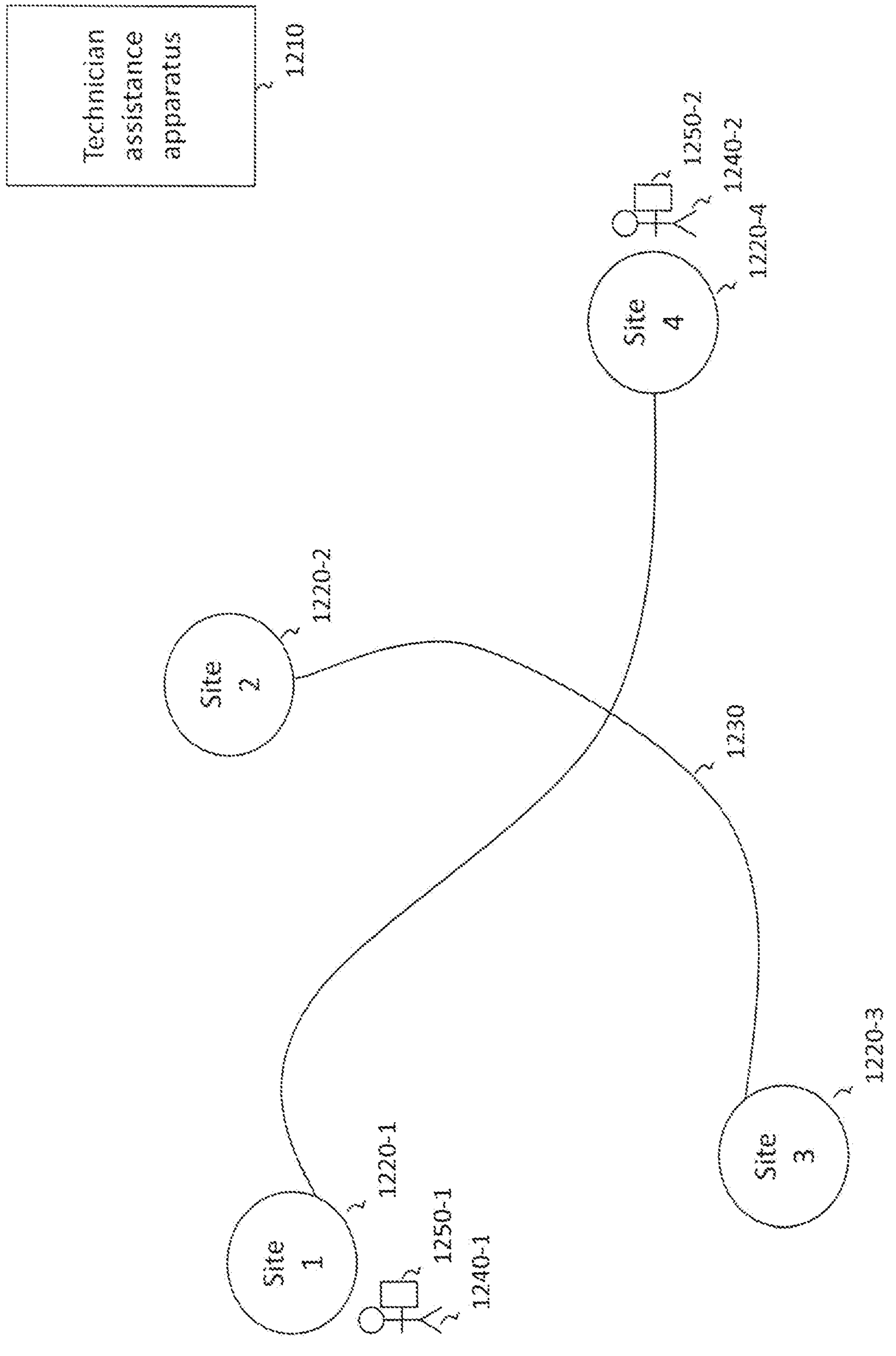
Technician assistance apparatus
1210
1250-2
1240-2
Site 4
1220-4
Site 2
1220-2
1230
Site 3
1220-3
Site 1
1220-1
1250-1
1240-1

Fig. 12

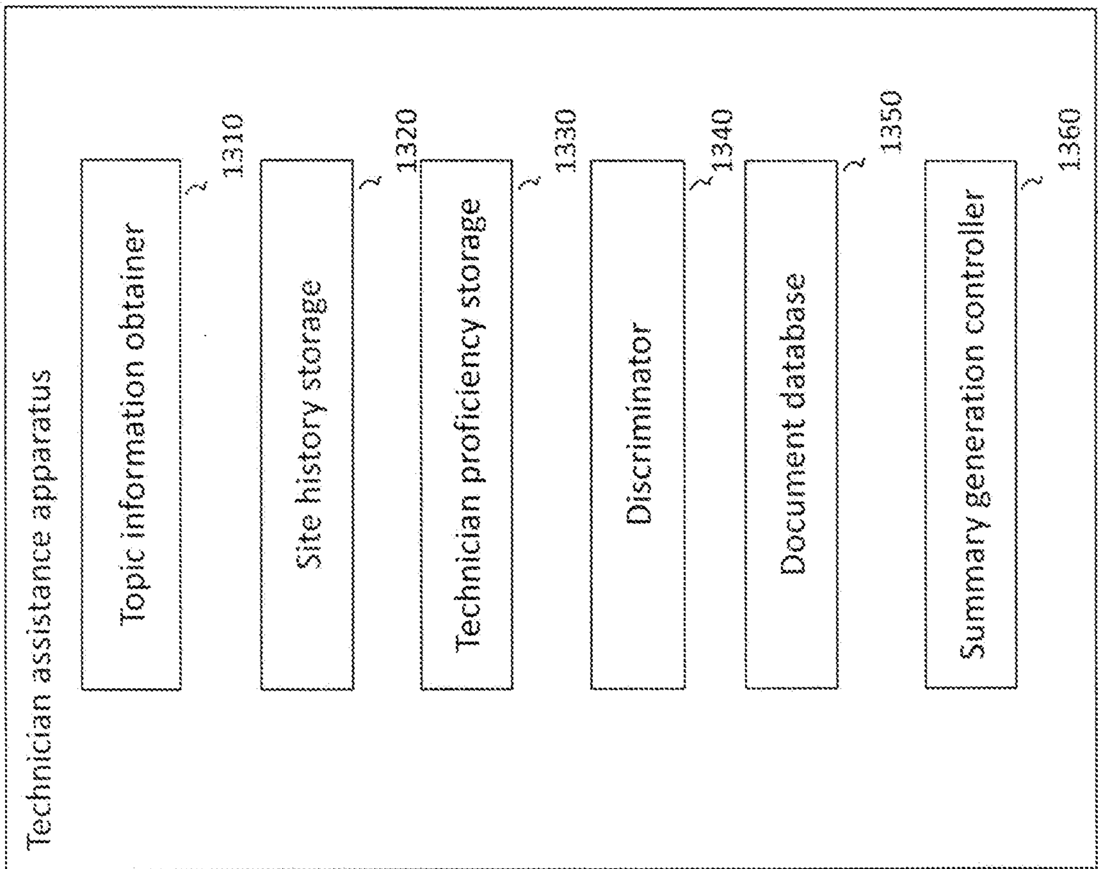
Technician assistance apparatus
Topic information obtainer
1310
Site history storage
1320
Technician proficiency storage
1330
Discriminator
1340
Document database
1350
Summary generation controller
1360
1300

Fig. 13

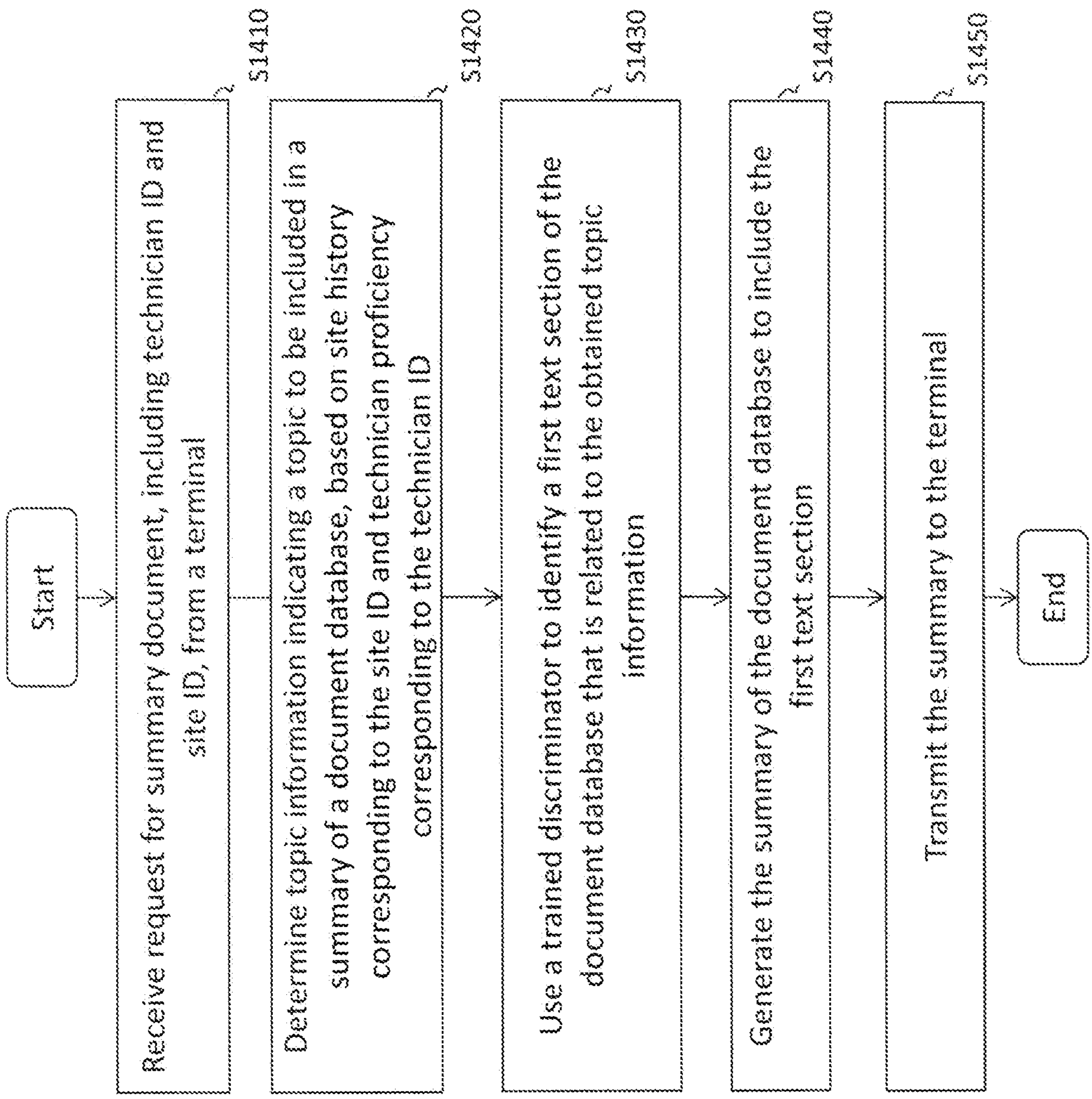
Start
Receive request for summary document, including technician ID and site ID, from a terminal
S1410
Determine topic information indicating a topic to be included in a summary of a document database, based on site history corresponding to the site ID and technician proficiency corresponding to the technician ID
S1420
Use a trained discriminator to identify a first text section of the document database that is related to the obtained topic information
S1430
Generate the summary of the document database to include the first text section
S1440
Transmit the summary to the terminal
S1450
End

Fig. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IN2019/050422

A. CLASSIFICATION OF SUBJECT MATTER
G06Q50/00 Version=2019.01

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q,G10L,G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

TotalPatent One, IPO Internal Database
Keywords: Discriminator,generative adversarial network,reinforcement

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US10152970B1 (CAPITAL ONE SERVICES LLC) 11 December 2018 (11-12-2018) (Abstract; column 4,lines 10-23; column 6,lines 62-65; column 8, lines 10-55;	1-8, 14-21, 29-51
Y	column 10, line 61 - column 14, line 27; column 16, lines 30-36)	9-13, 22-28, 52-56
Y	US20030134599A1 (PANGRAC & ASSOCIATES DEVELOPMENT INC) 17 July 2003 (17-07-2003) Whole Document	9-13,22-28, 52-56

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16-09-2019	16-09-2019

Name and mailing address of the ISA/	Authorized officer
Indian Patent Office Plot No.32, Sector 14,Dwarka,New Delhi-110075 Facsimile No.	Lal Ratnakar Telephone No. +91-1125300200

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IN2019/050422

Citation	Pub.Date	Family	Pub.Date
WO 2018227462 A1	20-12-2018	CN 109690526 A	26-04-2019