

Title of Invention: TRANSLATION QUALITY EVALUATION DEVICE AND TRANSLATION QUALITY EVALUATION METHOD

TECHNICAL FIELD

[1] The present invention relates to a translation quality evaluation device and translation quality evaluation method.

BACKGROUND

[2] Recently, as artificial intelligence technology has developed, various types of machine translation services using the artificial intelligence technology have been provided.

[3] In particular, companies that conduct Internet business based on portal search engines are actively providing machine translation services based on their existing infrastructures. Also, companies that provide machine translation services are also providing advanced machine translation services.

[4] As translation learning data is gradually being obtained, discussions about translation quality are gradually increasing. In order to increase translation quality, translation quality has to be first measured. This is because, when translation quality is explicitly displayed as a number, maximum translation quality increase may be achieved with limited resources and time.

[5] The present invention proposes technology that may provide a score necessary to evaluate the translation quality of a corpus including a pair of an original text and a translated text.

SUMMARY OF INVENTION

Technical Problem

- [6] In order to solve the problems of the conventional technology described above, the present invention provides a translation quality evaluation device and a translation quality evaluation method that provide a translation quality score of a translation corpus by inputting, to a score prediction model, respective cosine similarities calculated based on a corpus including an original text, a translated text, and a back-translated text.
- [7] However, technical objects to be achieved by the present embodiment are not limited to the technical objects described above, and there may another technical object.

Solution to Problem

- [8] In order to achieve the technical object described above, a translation quality evaluation device according to an embodiment of the present invention includes a communication module, a memory storing a translation quality evaluation program, and a processor configured to execute the translation quality evaluation program, wherein the translation quality evaluation program receives an original text written in a first language and a translated text of the original text written in a second language, generates a back-translated text written in the first language by performing machine translation on the translated text into the first language, and calculates a translation quality score by inputting, to a score prediction model, a first cosine similarity between the original text and the translated text and a second cosine similarity between the original text and the back-translated text calculated based on a corpus including the original text, the translated text, and the back-translated text.
- [9] A translation quality evaluation method using a translation quality evaluation device, according to another embodiment of the present invention, includes (a)

receiving an original text written in a first language and a translated text of the original text written in a second language, (b) generating a back-translated text written in the first language by performing machine translation on the translated text into the first language, and (c) calculating a translation quality score by inputting, to a score prediction model, a first cosine similarity between the original text and the translated text and a second cosine similarity between the original text and the back-translated text calculated based on a corpus including the original text, the translated text, and the back-translated text.

Advantageous Effects

[10] According to the above description for solving the problem of the present invention, a level of correction work required to improve of a translation corpus is presented as a translation quality score, and thus, the convenience of review work on the translation corpus may be improved.

[11] In addition, it is possible to distinguish data that needs improvement from the translation corpus, and thus, the quality of the known translation corpus may be increased, and preemptive evaluation on the translation corpus constructed in the future may be performed.

BRIEF DESCRIPTION OF THE DRAWINGS

[12] FIG. 1 is a block diagram illustrating a translation quality evaluation system according to an embodiment of the present invention.

[13] FIG. 2 is a block diagram illustrating detailed modules of a translation quality evaluation device according to an embodiment of the present invention.

[14] FIG. 3 is a diagram illustrating a translation quality evaluation device according to an embodiment of the present invention.

[15] FIG. 4 is a flowchart illustrating a translation quality evaluation method using a translation quality evaluation device, according to another embodiment of the present invention.

DETAILED DESCRIPTION OF INVENTION

[16] Hereinafter, embodiments of the present invention will be described in detail with reference to the accompanying drawings such that those skilled in the art in which the present invention belongs may easily practice the present invention. However, the present invention may be implemented in many different forms and is not limited to the embodiments described herein. In addition, in order to clearly describe the present invention, parts irrelevant to the description are omitted in the drawings, and similar reference numerals are attached to similar parts throughout the specification.

[17] When it is described that a portion is "connected" to another portion throughout the specification, this includes not only a case where the portion is "directly connected" to another portion but also a case where the portion is "electrically connected" to another portion with another component therebetween. In addition, when it is described that a portion "includes" a certain component, this means that the portion may further include another component without excluding another component unless otherwise stated.

[18] In the present invention, a "portion" includes a unit realized by hardware, a unit realized by software, and a unit realized by using both. In addition, one unit may be realized by using two or more pieces of hardware, and two or more units may be realized by using one piece of hardware. Meanwhile, a "~ portion" is not limited to software or hardware, and a "~ portion" may be configured to be included in an addressable storage medium or may be configured to

reproduce one or more processors. Therefore, in one example, "~ portion" refers to components, such as software components, object-oriented software components, class components, and task components, and includes processes, functions, properties, and procedures, subroutines, segments of program code, drivers, firmware, microcode, circuits, data, databases, data structures, tables, arrays, and variables. The functions provided within the components and "portions" may be combined into a smaller number of components and "portions" or may be further separated into additional components and "portions". Additionally, components and "portions" may be implemented to regenerate a device or one or more central processing units (CPUs).

[19] In addition, the "network" refers to a connection structure capable of exchanging information between nodes, such as a terminal and a server and includes a local area network (LAN), a wide area network (WAN), the Internet (WWW: World Wide Web), a wired and wireless data communications network, a telephone network, a wired and wireless television communication network, and so on. For example, the wireless data communication network includes third generation (3G), fourth generation (4G), fifth generation (5G), third generation partnership project (3GPP), long term evolution (LTE), world interoperability for microwave access (WIMAX), Wi-Fi, Bluetooth communication, infrared communication, ultrasonic communication, visible light communication (VLC), LiFi, and so on but are not limited thereto.

[20] A user terminal may be implemented by a computer or mobile terminal that may access a translation quality evaluation device through a network. Here, the computer may include, for example, a desktop computer or a laptop computer in which a web browser is stored, or so on, and a mobile terminal is, for example, a wireless communication device that guarantees portability and

mobility and may include all types of handheld wireless communication devices, such as various smartphones and tablet personal computers (PCs).

[21] FIG. 1 is a block diagram illustrating a translation quality evaluation system according to an embodiment of the present invention.

[22] Referring to FIG. 1, the translation quality evaluation system of the present invention includes a translation quality evaluation device 100 and a plurality of user terminals 10.

[23] The translation quality evaluation device 100 includes a communication module 110, a memory 120, a processor 130, and a database (DB) 140.

[24] The communication module 110 provides a communication interface required to provide signals transmitted and received to and from the user terminal 10 in the form of packet data in conjunction with a communication network. Here, the communication module 110 may include hardware and software required to transmit and receive signals, such as control signals or data signals, through wired or wireless connections with other network devices.

[25] The memory 120 may store a translation quality evaluation program. The translation quality evaluation program receives an original text written in a first language and a translation of the original text written in a second language, generates a back-translated text by performing machine translation on the translated text into the first language, and calculates a translation quality score by inputting, into a score prediction model, a first cosine similarity between the original text and the translated text and a second cosine similarity between the original text and the back-translated text calculated based on a corpus consisting of the original text, the translated text, and the back-translated text. Also, the memory 120 performs a function of temporarily or permanently store the data processed by the processor 130. Here, the memory 120 may include

volatile storage media or non-volatile storage media, but the scope of the present invention is not limited thereto.

[26] The memory 120 may store a separate program, such as an operating system for processing and controlling the processor 130 and may also perform a function of temporary storage of input or output data.

[27] The processor 130 executes a translation quality evaluation program stored in the memory 120 and provides a function of controlling the hardware of the translation quality evaluation device 100 according to the execution of the translation quality evaluation program. That is, the processor 130 performs hardware control functions, such as file system, memory allocation, network, basic library, timer, device control (display, media, input device, 3D, and so on), and other utilities required according to the execution of the translation quality evaluation program.

[28] Specifically, the processor 130 receives the original text written in the first language and the translated text written in the second language according to the execution of the translation quality evaluation program, generates a back-translated text written in the first language by performing machine translation on the translated text into the first language, and calculates a translation quality score by inputting the first cosine similarity between the original text and the translated text and the second cosine similarity between the original text and the back-translated text calculated based on a corpus consisting of the original text, translated text, and back-translated text into a score prediction model. Also, respective detailed steps of a translation quality evaluation process according to the execution of the translation quality evaluation program are described below with reference to FIGS. 2 and 3.

[29] This processor 130 may include all types of devices that may process data,

such as a processor. Here, a "processor" may mean, for example, a data processing device that has a physically structured circuit to perform a function expressed by code or instructions included in a program and is built into hardware. The data processing device built into hardware may include, for example, a microprocessor, a central processing unit (CPU), a processor core, a multiprocessor, an application-specific integrated circuit (ASIC), and a field programmable gate array (FPGA), or so on, but the scope of the present invention is not limited thereto.

[30] The database 140 may be a medium that organically combines data commonly required to perform a translation quality evaluation method and stores the combined data. Also, the database 140 may store learning data including a cosine similarity between an original text and a translated text and a cosine similarity between an original text and a back-translated text on a corpus consisting of respective sentence pairs of multiple languages, translated text number information on the respective sentence pairs, and translation quality scores. This database 140 may be provided as a separate component from the memory 120 or may be built into a partial area of the memory 120.

[31] Meanwhile, the translation quality evaluation device 100 may operate in a cloud computing service model, such as software as a service (SaaS), platform as a service (PaaS), or infrastructure as a service (IaaS). Also, the translation quality evaluation device 100 may be built in the form of a private cloud, public cloud, or hybrid cloud.

[32] The translation quality evaluation device 100 may perform operations for providing translation services in conjunction with a translation service providing program operating in the user terminal 10. The present invention is not limited thereto and may be used independently from the translation service provision

program.

[33] The user terminal 10 may include a communication module, a memory, and a processor, and the communication module of each terminal may perform data communication with the translation quality evaluation device 100 under control by a processor.

[34] FIG. 2 is a block diagram illustrating detailed modules of a translation quality evaluation device according to an embodiment of the present invention, and FIG. 3 is a diagram illustrating the translation quality evaluation device according to the embodiment of the present invention.

[35] Referring to FIG. 2, the processor 130 may include detailed modules that perform various functions according to the execution of a translation quality evaluation program. Here, the detailed modules may include a score prediction model 210, an embedding vector calculation unit 220, a cosine similarity calculation unit 230, and a sentence length measurement unit 240.

[36] Referring to FIG. 3, a translation quality evaluation program (hereinafter referred to as a program) receives an original text written in a first language and a translated text written in a second language and performs machine translation on the translated text into the first language to generate a back-translation written in the first language. Here, the back-translated text refers to a sentence obtained by machine-translating the translated text from a translation corpus back into an original language. Also, the back-translated text may be a machine-translated text by using another company's machine translator that is open for use by general users online.

[37] Next, the program may calculate a translation quality score by inputting the first and second cosine similarities calculated based on a corpus 20 into the score prediction model 210. Here, the corpus 20 may include a set of a sentence pair

including sentences consisting of a first language of the original text and a second language of the translated text, and a set of a sentence pair including sentences consisting of a second language of the translated text and a first language of the back-translated text.

[38] That is, the present invention provides a translation quality score of the translated text and may provide quantitative evaluation criteria for users to inspect and proofread the translated text.

[39] Next, when proofreading information on the translated text is input from the user terminal 10, the program may recalculate the translation quality score based on the corpus 20 in which the proofreading information is reflected.

[40] Specifically, the score prediction model 210 may output a translation quality score on the corpus 20 by inputting a first cosine similarity between the original text and the translated text, a second cosine similarity between the original text and the back-translated text, and translated text number information to be described below. In this case, detailed descriptions of a learning process of the score prediction model 210 are described below.

[41] Meanwhile, when predicting quality by only similarity (the second cosine similarity) of an original-back-translation sentence pair, there is a disadvantage that the quality of the back-translated text is determined by the performance of the known machine translator. In order to complement this, the present invention may improve the performance of the score prediction model 210 that determines a translation quality score by learning not only the second cosine similarity but also a similarity (the first cosine similarity) of the original-translation sentence pair as learning data of the score prediction model 210.

[42] The embedding vector calculation unit 220 may calculate embedding vectors for the original text and the translated text by using a multilingual language

model and may calculate embedding vectors for the original text and the back-translated text by using a single language model.

[43] For example, the embedding vector calculation unit 220 may calculate a sentence embedding vectors for each original text and translated text by using a multilingual language model, and calculate a sentence embedding vectors for each original text and back-translated text by using a single language model. In this case, the present invention may use a pre-learned multilingual language model and a single language model and may obtain high-quality sentence embedding with relatively few resources.

[44] The cosine similarity calculation unit 230 may calculate the first cosine similarity based on embedding vectors for the original text and the translated text, and calculate the second cosine similarity based on the embedding vectors for the original text and the back-translated text.

[45] The sentence length measurement unit 240 may calculate translated text number information including a translated text number of an original text of a corpus, a translated text number of the back-translated text, and a translated text number difference between the original text and the back-translated text. for example, the sentence length measurement unit 240 is the known sentence length measurement algorithm and may distinguish translated texts of each sentence according to preset criteria for each language including Korean and English. In this case, the translated text is each word constituting a sentence for each language and may be preset as a minimum unit of sentence components. In addition, the sentence length measurement unit 240 may calculate a translated text number difference between the original text and the back-translated text based on the translated texts classified for each sentence.

[46] The score prediction model 210 may be a model built based on training data

with matched translation quality scores for each pair of sentences, which includes first and second cosine similarities, translated text number information (a translated text number of an original text, a translated text number of a back-translated text, and a translated text number difference between the original text and the back-translated text), and feature information, as feature information on the corpus previously stored in the database 140. That is, the score prediction model 210 may be trained to output a translation quality score of a corpus by using each cosine similarity and translated text number information as an input. For example, the score prediction model 210 may apply various types of machine learning models, such as a regression analysis model and a boosting ensemble model.

[47] For example, the translation quality score may be defined as a direct assessment (DA) score obtained by evaluating the translation quality for each pair of sentences between an original text and a translated text for a predetermined corpus by a group of experts.

[48] For example, the present invention may calculate the first and second cosine similarities for sentence pairs of the original Korean text and an English translation for a predetermined corpus and may sample the same number of sentence pairs for each score range based on the calculated first and second cosine similarities. Thereafter, translation quality score data on the sampled sentence pairs may be obtained based on the DA score given by a group of three or more experts according to a quality evaluation criteria below.

[49] For example, the quality evaluation criteria may include following 10 evaluation rules that may grade DA scores for the English translation quality of each sentence pair. 1) 0-9: sentences are completely incorrectly matched, 2) 10-19: some words match but context is different, 3) 20-29: meaning is omitted or

added unnecessarily for each phrase, 4) 30-39 : there are difficulties in understanding context due to major mistranslations, omissions, and additions of terms, 5) 40-49: there are major mistranslations, omissions, and additions of terms but there is no difficulty in understanding context, 6) 50-59: some words are replaced with pronouns or so on, 7) 60-69: there are incorrect nouns or grammatical error of tense and word usage, 8) 70-79: context matches and there are no omissions or mistranslations, but nuance is slightly different, 9) 80-89: there may be expressions, which seem awkward to native speakers, such as translation trend, but overall meaning of sentence responds without omission or addition, 10) 90-100: there are no awkward expressions, omissions, duplication, or surplus expressions, and idioms and proverbs are translated into expressions appropriate for native speaker's culture. Here, detailed evaluation criteria within the same range (same score range) of the evaluation rules 1) to 10) are set to fluency. That is, the detailed evaluation criteria are to give a lower score in a case where fluency is low and translation trend is strong even within the same score range, and a unit of evaluation is set as a short sentence assuming no context.

[50] In this way, the quality evaluation criteria may be a rule set inductively by distinguishing the types and degrees of errors frequently occurring when a user actually inspects a corpus of translated texts. Also, the quality evaluation criteria may provide enough specificity to present a direction even to first-time evaluators.

[51] Therefore, when receiving the corpus 20 for an original text and a translated text from the user terminal 10, the program may calculate the first cosine similarity between the original text and the translated text, the second cosine similarity between the original text and the back-translated text, and the

translated text number information on each sentence pair. When each cosine similarity and translated text number information calculated in this case is input to the score prediction model 210, the score prediction model 210 may provide a translation quality score.

[52] In another embodiment, the program may provide a translation review step based on an inductively determined first and second threshold for the translation quality score. For example, the translation review step may be divided into three levels: very bad, bad, and good, according to the translation quality score. For example, a translated text with a score (very bad) below the first threshold may be classified as "a score that is determined to require new translation work because the translation is very different from the original text", a translated text with a score (bad) exceeding the first threshold and below the second threshold may be classified as a "score that is determined to require correction in minor parts, such as some words in the sentence", and a translated text with a score (good) above the second threshold may be classified as a translated text without error in grammar or words.

[53] Hereinafter, descriptions on components performing the same function among the components illustrated in FIGS. 1 to 3 described above are omitted.

[54] FIG. 4 is a flowchart illustrating a translation quality evaluation method using a translation quality evaluation device, according to another embodiment of the present invention.

[55] Referring to FIG. 4, the translation quality evaluation method according to the present invention includes step S110 of receiving an original text written in a first language and a translated text written in a second language, step S120 of generating a back-translated text written in the first language by performing machine translation on the translated text into the first language, and step

S130 of calculating a translation quality score by inputting, to a score prediction model, a first cosine similarity between the original text and the translated text and a second cosine similarity between the original text and the back-translated text calculated based on a corpus composed of the original text, the translated text, and the back-translated text.

[56] After step S130, the present invention may further include a step of recalculating the translation quality score based on a corpus in which correction information is reflected, when the correction information on the translated text is input from a user terminal.

[57] In step S130, embedding vectors for the original text and the translated text may be calculated by using a multilingual language model, and embedding vectors for the original text and the back-translated text may be calculated by using a single language model.

[58] In step S130, the first cosine similarity may be calculated based on the embedding vectors for the original text and the translated text, and the second cosine similarity may be calculated based on the embedding vectors for the original text and the back-translated text.

[59] In step S130, translated text number information including a translated text number of the original text of a corpus, a translated text number of the back-translated text, and a translated text number difference between the original text and the back-translated text may be calculated based on a sentence length measurement algorithm.

[60] The score prediction model 210 may include a regression model trained to output a score for evaluating a translation quality of a corpus by using the first and second cosine similarities and translated text number information as an input.

[61] The method described above may be implemented in the form of a recording medium including instructions executable by a computer, such as a program module executed by a computer. A computer readable medium may be any available medium that may be accessed by a computer and includes both volatile and nonvolatile media, removable and non-removable media. Also, the computer readable medium may include a computer storage medium. A computer storage medium includes both volatile and nonvolatile media and removable and non-removable media implemented by any method or technology for storing information, such as computer readable instructions, data structures, program modules or other data.

[62] Those skilled in the art to which the present invention belongs will understand that the present invention may be easily modified into another specific form based on the above descriptions without changing the technical idea or essential features of the present invention. Therefore, the embodiments described above should be understood as illustrative in all respects and not limiting. The scope of the present invention is indicated by the following claims, and the meaning and scope of the claims and all changes or modifications derived from the equivalent concepts should be interpreted as being included in the scope of the present invention.

CLAIMS

1. A translation quality evaluation device comprising:
a communication module;
a memory storing a translation quality evaluation program; and
a processor configured to execute the translation quality evaluation program,
wherein the translation quality evaluation program receives an original text written in a first language and a translated text of the original text written in a second language, generates a back-translated text written in the first language by performing machine translation on the translated text into the first language, and calculates a translation quality score by inputting, to a score prediction model, a first cosine similarity between the original text and the translated text and a second cosine similarity between the original text and the back-translated text calculated based on a corpus including the original text, the translated text, and the back-translated text.
2. The translation quality evaluation device of claim 1, wherein,
the translation quality evaluation program recalculates the translation quality score based on a corpus in which the proofreading information is reflected, when proofreading information on the translated text is input from a user terminal.
3. The translation quality evaluation device of claim 1, wherein
the translation quality evaluation program calculates embedding vectors for the original text and the translated text by using a multilingual language model, and calculates embedding vectors for the original text and the back-translated text by using a single language model.

4. The translation quality evaluation device of claim 3, wherein the translation quality evaluation program calculates the first cosine similarity based on the embedding vectors for the original text and the translated text, and calculates the second cosine similarity based on the embedding vectors for the original text and the back-translated text.

5. The translation quality evaluation device of claim 1, wherein the translation quality evaluation program calculates word number information including a word number of the original text on the corpus, a word number of the back-translated text, a word number difference between the original text and the back-translated text based on a sentence length measurement algorithm.

6. The translation quality evaluation device of claim 5, wherein the score prediction model includes a regression model trained to output a score for evaluating translation quality of the corpus by using the first cosine similarity, the second cosine similarity, and the word number information as an input.

7. A translation quality evaluation method using a translation quality evaluation device, the translation quality evaluation method comprising:

(a) receiving an original text written in a first language and a translated text of the original text written in a second language;

(b) generating a back-translated text written in the first language by performing machine translation on the translated text into the first language; and

(c) calculating a translation quality score by inputting, to a score prediction model, a first cosine similarity between the original text and the translated text and

a second cosine similarity between the original text and the back-translated text calculated based on a corpus including the original text, the translated text, and the back-translated text.

8. The translation quality evaluation method of claim 7, further comprising:

after the step (c), recalculating the translation quality score based on a corpus in which proofreading information is reflected, when the proofreading information on the translated text is input from a user terminal.

9. The translation quality evaluation method of claim 7, wherein,
in the step (c), calculating embedding vectors for the original text and the translated text by using a multilingual language model, and calculating embedding vectors for the original text and the back-translated text by using a single language model.

10. The translation quality evaluation method of claim 9, wherein,
in the step (c), calculating the first cosine similarity based on the embedding vectors for the original text and the translated text, and calculating the second cosine similarity based on the embedding vectors for the original text and the back-translated text.

11. The translation quality evaluation method of claim 7, wherein,
in the step (c), calculating word number information including a word number of the original text on the corpus, a word number of the back-translated text, a word number difference between the original text and the back-translated text

based on a sentence length measurement algorithm.

12. The translation quality evaluation method of claim 11, wherein the score prediction model includes a regression model trained to output a score for evaluating translation quality of the corpus by using the first cosine similarity, the second cosine similarity, and the word number information as an input.