

DESCRIPTION

EVALUATION SYSTEM AND EVALUATION PROGRAM

5

Technical Field

[0001] The present invention relates to an evaluation system for evaluating a business plan using a plurality of evaluation criteria, and an evaluation program causing a computer to execute such evaluation.

10

Background Art

[0002] A business planning assistance system according to Patent Document 1 stores a plurality of pieces of regional information and case information about a business planned for each region, obtains an index for a problem to be solved by a business plan, extracts, as recommended case information, case information with a common index from the pieces of case information, and generates and outputs information based on a business evaluation item indicating a business result out of the recommended case information. Further, the business planning assistance system generates, based on the recommended case information, an evaluation model trained with training data with one business evaluation item as a response variable and another business evaluation item as an explanatory variable, and outputs a value of a response variable obtained by inputting a value of a business evaluation item of the business plan to the evaluation model as an explanatory variable.

25

[0003] A language communication ability evaluation system according to Patent Document 2 includes an evaluation target obtaining section that receives a user input including a question and an evaluation target, an evaluation criterion obtaining section that obtains an evaluation criterion for a language ability evaluation test, a prompt providing section that provides a generative AI with a prompt including the user input, and instructions and descriptions for the generative AI, a response obtaining section that obtains a response to the prompt from the generative AI, and an evaluation display section that displays an evaluation of the evaluation target included in the response. The instructions constituting the prompt include a role assigning instruction to assign a role as an evaluator for the language ability evaluation test to the generative AI, an evaluation creation instruction to evaluate the

35

evaluation target based on the evaluation criterion, and a confirmation instruction to confirm that the evaluation according to the evaluation creation instruction is based on the evaluation criterion. Further, the descriptions constituting the prompt include a description of the evaluation criterion and a description of an evaluation method based on the evaluation criterion.

Prior Art Documents

Patent Documents

[0004] Patent Document 1: JP 2024-011875A
Patent Document 2: JP 7521860B

Disclosure of the Invention

Problem to be Solved by the Invention

[0005] In the business planning assistance system according to Patent Document 1, a learning model needs to be constructed that is trained with training data with one business evaluation item as a response variable and another business evaluation item as an explanatory variable. When there are many business evaluation items, a significant burden is imposed on the creation of the training data.

[0006] In the language communication ability evaluation system according to Patent Document 2, evaluation information regarding an evaluation target is obtained from the generative AI by giving a prompt, and a preparatory work burden for the evaluation is therefore small. In particular, in the evaluation of a language communication ability as in an English certification test, evaluation criteria are limited to those for the language communication ability and are well known. Thus, the evaluation is relatively simple. However, if the evaluation target is a business plan, items constituting the business plan and items of evaluation criteria are complicated, and it is difficult to apply an evaluation technique such as that described in Patent Document 2 directly to business plan evaluation.

[0007] In view of the foregoing circumstances, the present invention aims to provide an evaluation system capable of accurately evaluating a business plan using a plurality of evaluation criteria with a small burden.

Means for Solving Problem

[0008] An evaluation system according to the present invention is a system configured to evaluate a business plan with use of a plurality of evaluation criteria, and includes: a plan input section configured to input the business plan; a criterion input section configured to input the evaluation criteria; an evaluation level calculation unit configured to calculate an evaluation level for each of the evaluation criteria with respect to the business plan, with use of the input business plan and the evaluation criteria; and an evaluation unit configured to perform a final evaluation of the business plan based on the evaluation level.

[0009] With this configuration, a plurality of evaluation criteria to be used to evaluate a business plan are prepared. A business plan to be evaluated is evaluated for each evaluation criterion, and the evaluation level for the evaluation criterion is calculated. Even if the business plan covers a lot of content, the burden of the evaluation work is small if the evaluation is performed according to one evaluation criterion. Furthermore, the final evaluation of the business plan is determined based on all evaluation levels calculated for the respective evaluation criteria. Thus, the business plan is checked from all perspectives, and the final evaluation thereof is accurate.

[0010] To check whether the business plan satisfies an evaluation criterion, it is preferable to check the similarity between sentences included in the business plan and sentences included in the evaluation criterion. The similarity between sentences can be calculated by converting the sentences into vector representations and performing vector calculation, such as dot product calculation. Using a technique such as word embeddings for vector representations of sentences is convenient for subsequent evaluation processing (natural language processing) as well. Therefore, the present invention proposes that the evaluation level calculation unit is configured to generate plan vector data by converting a sentence included in the business plan into a vector representation, generate criterion vector data by converting a sentence included in the evaluation criteria into a vector representation, and calculate the evaluation level based on a similarity between the business plan and each of the evaluation criteria calculated with use of the plan vector data and the criterion vector data.

[0011] If a classification method such as the K-means method is used to calculate the similarity between the business plan and each evaluation criterion (evaluation criterion satisfaction degree) using the plan vector data

and the criterion vector data, a quantitative examination (such as a Euclidean distance from a representative point of a cluster) can be given with respect to a calculated cluster (equivalent to an evaluation level), which is convenient. For this reason, the present invention proposes that the evaluation level is

5 represented by a plurality of clusters separated in a vector space including the plan vector data and the criterion vector data, and the evaluation level is calculated through computation of distribution among the clusters. The evaluation level of the business plan can be calculated from the distribution among the clusters. For example, it is possible to classify the vector data into

10 a plurality of clusters (evaluation levels) using the K-means method, and classify the vector data into each cluster (each evaluation level).

[0012] The evaluation level for a business plan using the plan vector data and the criterion vector data as described above can be relatively easily calculated with high accuracy by creating a proper prompt and using various types of

15 generative AIs. For this reason, the present invention proposes that the evaluation level calculation unit includes a generative AI, and the business plan and the evaluation criteria are provided to the generative AI as a prompt.

[0013] One example of business plans that have recently been the subject of much debate due to the difficulty in evaluation is a business plan to generate

20 carbon credits. Carbon credits refer to a system that mainly allows companies to buy and sell greenhouse gas emission reductions. Companies can quantify the amount of greenhouse gas reduction or absorption resulting from their environmental activities and trade the emission rights certified as carbon credits with other companies. Therefore, a carbon credit-related business

25 plan needs to be evaluated from multiple perspectives and involves many evaluation criterion items, and this evaluation is important. Further, the evaluation criteria may include those formulated by each country as well as criteria and know-how unique to each company. For this reason, in the present invention, the business plan is a business plan for earning a carbon

30 credit, and the evaluation criteria include: a general evaluation criterion created according to a carbon credit certification standard; and an independent evaluation criterion created independently. With this, a business plan to generate carbon credits can be accurately evaluated with a small burden.

[0014] Carbon credits are traded between companies or brokers, and the

35 transaction value of carbon credits to be produced is important for the business plan to produce carbon credits. Therefore, an expected increase in the

transaction value of carbon credits to be produced may also be added to the evaluation criteria. For this reason, the present invention also proposes that the final evaluation of the business plan for earning a carbon credit is affected by a transaction value of the carbon credit.

5 [0015] The present invention also encompasses in its scope of the right an evaluation program causing a computer to execute processing for evaluating a business plan with use of a plurality of evaluation criteria. This evaluation program causes the computer to execute: plan input processing for inputting the business plan; criterion input processing for inputting the evaluation
10 criteria; evaluation level calculation processing for calculating an evaluation level for each of the evaluation criteria with respect to the business plan, with use of the input business plan and the evaluation criteria; and evaluation processing for performing a final evaluation of the business plan based on the evaluation level. The processing performed through the execution of the
15 above program exhibits various effects described regarding the above evaluation system.

[0016] Further features, operations, and effects of the present invention will be apparent from the following description of the present invention with reference to the drawings.

20

Brief Description of the Drawings

[0017] FIG. 1 is a schematic diagram showing a schematic configuration of an evaluation system.

FIG. 2 is a flow diagram illustrating a processing flow when a business
25 plan is quantitatively evaluated.

FIG. 3 is a document diagram showing an example of a prompt.

FIG. 4 is an explanatory diagram illustrating classification of evaluation according to the K-means method.

FIG. 5 is a table of evaluation levels for respective evaluation criteria
30 with respect to a plurality of business plans.

FIG. 6 is a table of evaluation levels with respect to a plurality of business plans.

Best Mode for Carrying out the Invention

35 [0018] First, a schematic configuration of an evaluation system according to the present invention will be described with reference to FIG. 1. This

evaluation system evaluates a business plan using a plurality of evaluation criteria. The system includes a plan input section 11, a criterion input section 12, an evaluation level calculation unit 20, and an evaluation unit 30.

5 [0019] The plan input section 11 converts an input business plan into data processible by the evaluation level calculation unit 20, and provides the converted data to the evaluation level calculation unit 20. The criterion input section 12 converts evaluation criteria used in evaluating the business plan into data processible by the evaluation level calculation unit 20, and provides the converted data to the evaluation level calculation unit 20. The evaluation
10 criteria include several to hundreds of items, and these items may be grouped together in one or more documents.

[0020] The evaluation level calculation unit 20 is a core component of the evaluation system that is substantially constructed using a computer unit, and calculates an evaluation level for each evaluation criterion with respect to a
15 business plan using the provided business plan and evaluation criteria. Thus, the evaluation level calculation unit 20 includes, as its functional sections, an instruction generation section 21, a vector representation module 22, and a computing section 23. Note that at least some part of processing performed by each component constituting the evaluation system is substantially
20 implemented by a computer executing a computer program.

[0021] The instruction generation section 21 (prompt generation section 21) functions as an interface for receiving instructions from a user and generating an instruction script for the computing section 23 of this evaluation system. The instruction script for the computing section 23 that is generated by the
25 instruction generation section 21 includes the content of the user instructions, the content of the business plan, and the content of the evaluation criteria. Examples of the content of the user instructions include: "The document to be input is a report on a carbon credit project"; "For each criterion, find evidence for a question of whether this project complies with the criteria"; and "Output
30 the obtained evidence in strict compliance with the output format."

[0022] The vector representation module 22 generates plan vector data by converting sentences (including words and phrases) constituting the business plan included in the instruction script into vector representations, and similarly generates criterion vector data by converting sentences constituting
35 the evaluation criteria included in the instruction script into vector representations. For example, sentences are converted into vector

representations using the word embedding technique, which is used in Bidirectional Encoder Representations from Transformers (BERT) and the like.

The generated plan vector data and criterion vector data are provided to the computing section 23 together with the content of the user instructions.

5 [0023] The computing section 23 calculates an evaluation level of the input business plan by calculating a similarity indicating a satisfaction degree of each evaluation criterion in the business plan with reference to the user instructions, using the plan vector data and the criterion vector data. The similarity can be calculated with, for example, a technique of obtaining a
10 similarity between sentences of the plan vector data and sentences of the criterion vector data by vector operation (e.g., dot product operation). The evaluation level is calculated for each of a plurality of criterion items included in the evaluation criteria. Here, the evaluation level can be regarded as an aggregate of evaluation levels per criterion item.

15 [0024] The evaluation unit 30 conducts a final evaluation of the business plan to be evaluated, based on the evaluation level as the computation result of the computing section 23, i.e., the output of the evaluation level calculation unit 20, and outputs the final evaluation of the business plan. Note that the evaluation unit 30 may be incorporated into the evaluation level calculation
20 unit 20 as one of the functional sections of the evaluation level calculation unit 20.

[0025] The vector representation module 22 and the evaluation level calculation unit 20 may be integrated as a generative AI 25. In this case, the instruction generation section 21 functions as the prompt generation section
25 21. Further, the functions of the evaluation unit 30 can also be incorporated into the generative AI 25.

Alternatively, the evaluation unit 30 may be constituted by an independent machine learning model independent of the generative AI 25.

[0026] Next, one embodiment of the evaluation system according to the
30 present invention will be described. In this embodiment, the evaluation system evaluates a business plan to obtain carbon credits, based on carbon credit evaluation criteria. The evaluation level calculation unit 20 of the evaluation system includes a generative AI 25 that outputs an evaluation level in response to the provision of a prompt generated by the prompt generation
35 section 21. Note that the carbon credit evaluation criteria include general evaluation criteria created according to carbon credit certification criteria, and

independent evaluation criteria created independently.

[0027] Carbon credits are emissions rights that companies and individuals can earn by funding projects that reduce greenhouse gas emissions, and the value of the emissions rights corresponds to the amount of greenhouse gas reduction. Such greenhouse gas emission projects (hereinafter referred to simply as “projects”) require large amounts of capital investment. It is, therefore, important to carefully examine a project proposal in advance to ensure that the carbon credits generated by the project have an appropriate value. Project proposals may include a project proposal related to reducing greenhouse gas emissions from rice paddies, a project proposal related to afforestation and deforestation control, a project proposal related to biochar, and the like. This type of project proposal is the business plan in the present invention. Although the examples shown in FIGS. 3 and 6 deal with a project proposal related to biochar as a business plan, it is needless to say that the business plan in the present invention is not limited to a project plan related to biochar.

[0028] FIG. 2 shows a processing flow when the evaluation system quantitatively evaluates a business plan, which is a project proposal related to biochar, and a forestry business plan. The evaluation system can sequentially evaluate various business plans according to the same evaluation criteria. The following description deals with the business plan that is a project proposal related to biochar.

[0029] A user creates a business plan related to a biochar project that is to be evaluated. At the same time, the user creates carbon credit evaluation criteria (hereinafter referred to simply as evaluation criteria) for evaluating this business plan. The created business plan and evaluation criteria are converted into text data and provided to the evaluation system.

[0030] The evaluation criteria include more than 100 items, and examples of these include the following.

- (1) Feedstock must be purely biogenic waste biomass and not purpose grown.
- (2) Feedstock must have been otherwise left to decay or combusted for purposes other than energy production.
- (3) It is expressed in dry tons of biochar. Care must be taken to exclude any moisture, as including moisture would lead to an overestimation of the carbon sequestered.
- (4) The biomass used must not contain any paint residues, solvents, or other

potentially toxic impurities.

(5) Feedstock must not have been imported from other countries.

[0031] A prompt can be generated by the evaluation system using the provided business plan and evaluation criteria, or may be generated by the user using a user interface of the prompt generation section 21 and provided to the evaluation system. In this embodiment, the generative AI 25 uses a large-scale language model (LLM), and a prompt suitable for the specifications of this LLM is generated. FIG. 3 shows an example of such a prompt. In the example in FIG. 3, the prompt generation section 21 generates a prompt based on instructions (content of user instructions), the business plan, and the evaluation criteria that are provided by the user via the Internet or the like. Here, the prompt is written in English to increase the output accuracy of the generative AI 25. Of course, the prompt may be written in Japanese.

[0032] When the data volume of the input business plan is large, the name of a document file designated to be read from an appropriate memory is written in the prompt. Further, when the input evaluation criteria includes a large number of items, prompts including a predetermined number of items of the evaluation criteria are sequentially provided to the generative AI 25, and a plurality of output results output from the generative AI 25 are integrated.

[0033]

The degree to which the business plan satisfies the evaluation criteria is determined based on the similarity between the business plan and each evaluation criterion item (the degree to which the business plan satisfies the evaluation criteria). The evaluation level as the result of this determination, which is the degree to which the business plan satisfies each evaluation criterion, is set to any of the following six evaluation levels: “Yes”, “Likely Yes”, “Likely No”, “No”, “Irrelevant”, and “Not mentioned”.

[0034] To perform the determination based on the similarity, the evaluation system uses proximity (similarity: distance) between data in a vector space of vector data (plan vector data) of sentences constituting the business plan and vector data (evaluation criterion vector data) of sentences constituting the evaluation criteria. That is to say, each data is classified into a cluster that is one of the evaluation levels: “Yes”, “Likely Yes”, “Likely No”, “No”, “Irrelevant”, and “Not Mentioned”. In other words, the classification of each vector data into a specific cluster (specific evaluation level) is calculated based on the distribution of vector groups among the clusters in the vector space. The K-

means method is used for this classification.

[0035] Classification according to the K-means method in this embodiment is as shown in FIG. 4, in which six classification domains (equivalent to evaluation levels) corresponding to “Yes”, “Likely Yes”, “Likely No”, “No”, “Irrelevant”, and “Not Mentioned” are set in a vector space in which data groups including the plan vector data and the criterion vector data disperse, the data groups are classified into the respective classification domains, a distance (Euclidean distance: L2 distance) between a centroid of a data group included in each classification domain and a centroid of each classification domain is calculated, and this distance provides quantitative support for calculating the evaluation level.

More specifically, the classification is performed according to the following procedure, for example:

- (1) a vector data group is generated by converting a description of each piece of evidence into a vector representation;
- (2) the vector data are clustered into classification domains;
- (3) a centroid vector of the cluster of each classification domain is calculated; and
- (4) for the description of each piece of evidence, the classification domain of a cluster with the lowest Euclidean distance among the centroid vectors calculated in (3) is regarded as the classification result according to the K-means method.

[0036] Thus, a prompt includes, as output instructions, any of the six evaluation levels, a description of the evidence serving as the basis for the evaluation, and the like.

[0037] The evaluation unit 30 evaluates the value (quality) of the business plan to be evaluated from the evaluation level calculated for each of 100 or more evaluation criteria, and generates evaluation data (evaluation report) including the value of carbon credits of the business plan.

[0038] FIG. 5 is a table output as a reference example illustrating six evaluation levels for a large number of evaluation criteria with respect to a plurality of business plans. Note that the data groups subjected to the K-means method in this reference example are vector data that are output by the generative AI and obtained by converting sentences (evidence) extracted as the support for the evaluation calculation into vector representations. Examples of the sentences extracted as the basis for the evaluation calculation include

the following.

- (1) A certain company produces charcoal in batches. Each kiln is filled with firewood, the combustion chamber is sealed, and pyrolysis is performed in the kiln. In the cooling stage, heat is radiated from the walls and the top of the kiln, and the kiln is cooled as it is. After cooling, the combustion chamber is opened, and carbides are processed for transport.
 - (2) There is no mention of the moisture content of biochar for preventing dust and explosion during transport.
 - (3) This document does not provide details of the moisture content of biochar for preventing dust and explosion during transport.
 - (4) The moisture content was confirmed to be 20% to 30% by measuring the moisture content in individual large bags.
 - (5) This document does not contain information regarding the transport of biochar over 200 kilometers or the accompanying emissions of greenhouse gases.
- [0039] In the table shown in FIG. 5, each item in the vertical column indicates a document number that specifies a business plan, and each item in the horizontal row indicates an evaluation criterion item number that specifies an evaluation criterion item. For example, each cell in the second column indicates the evaluation level for a corresponding evaluation criterion item with respect to the business plan of the document number 2. The cell (2, 2) in the second row and second column show “Likely Yes” as the evaluation level for the evaluation criterion item 2 with respect to the business plan of the document number 2.
- [0040] Based on evaluation level data such as those shown in FIG. 5, the evaluation unit 30 can estimate the transaction value of carbon credits in each business plan and generate an overall evaluation report indicating the order of the value.
- [0041] In a table output shown as a reference example in FIG. 6, the vertical column is the same as that in FIG. 5. A first cell (Document ID) in the horizontal row indicates a document number specifying a business plan, a second cell (labeled class) in the horizontal row indicates the evaluation level of the business plan, a third cell (evidence) in the horizontal row indicates sentences extracted as the support for the evaluation calculation, a fourth cell (matched) in the horizontal row indicates the match (TRUE) or mismatch (FALSE) between the evaluation by the generative AI and the evaluation

according to the K-means method, a fifth cell (K-means classification) in the horizontal row indicates the evaluation level according to the K-means method, and a sixth cell (L2 distance) in the horizontal row indicates a distance value calculated according to the K-means method. The fifth cell in the horizontal

5 row indicates the evaluation level that ranks first according to the K-means method, and the sixth cell in the horizontal row indicates the distance value as the support for the first rank. The seventh and subsequent cells in the horizontal row sequentially indicate the evaluation levels and the distance values for the second and subsequent ranks according to the K-means method.

10 The final evaluation of the business plan is the evaluation level at the first rank. In FIG. 6, the final evaluation of the business plan denoted as the document number 1 is “Yes”. Here, the evaluation levels and the distance values at the second and subsequent ranks are also indicated, enabling quantitative examination of the final evaluation.

15 [0042] Other Embodiments

(1) Although the evaluation system in the above embodiment is integrally configured, the components thereof may be distributed via a network to configure the evaluation system as a distributed system. For example, a configuration may be used in which the prompt generation section 21 of the

20 evaluation level calculation unit 20 is constructed on a user terminal, while the generative AI 25 of the evaluation level calculation unit 20 and the evaluation unit 30 are constructed on a remote cloud computing service, and final evaluation results are transmitted to the user terminal. Further, the generative AI 25 may be constructed on another server or user terminal.

25 [0043] (2) Although the evaluation unit 30 can be integrated with the generative AI 25, the evaluation unit 30 may also be constructed as a machine learning model that outputs a final result indicating the order of the value of a business plan using various data outputs (e.g., table data such as those shown in FIGS. 5 and 6) from the generative AI 25 as input data.

30 [0044] (3) The above embodiment mainly takes a business plan related to a biochar project as an embodiment of a business plan. However, the present invention is not limited thereto, and the evaluation system according to the present invention can also be applied to various business plans such as a plan for a business related to energy saving, a plan for a food self-sufficiency

35 business, with similar advantages.

[0045] It should be noted that the configurations disclosed in the above

embodiments (including other embodiments; the same applies below) can be applied in combination with configurations disclosed in other embodiments, as long as no contradiction arises. Further, the embodiments disclosed herein are illustrative, and the embodiments of the invention are not limited thereto
5 and can be modified as appropriate without departing from the object of the invention.

Industrial Applicability

[0046] The present invention can be applied to an evaluation system and an
10 evaluation program for evaluating a business plan using a plurality of evaluation criteria.

Description of Reference Signs

[0047] 2: Evaluation criterion item
15 11: Plan input section
12: Criterion input section
20: Evaluation level calculation unit
21: Instruction generation section (prompt generation section)
22: Vector representation module
20 23: Computing section
25: Generative AI
30: Evaluation unit

CLAIMS

1. An evaluation system configured to evaluate a business plan with use of a plurality of evaluation criteria, the evaluation system comprising:
 - 5 a plan input section configured to input the business plan;
 - a criterion input section configured to input the evaluation criteria;
 - an evaluation level calculation unit configured to calculate an evaluation level for each of the evaluation criteria with respect to the business plan, with use of the input business plan and the evaluation criteria; and
 - 10 an evaluation unit configured to perform a final evaluation of the business plan based on the evaluation level.
2. The evaluation system according to claim 1,
 - 15 wherein the evaluation level calculation unit is configured to generate plan vector data by converting a sentence included in the business plan into a vector representation, generate criterion vector data by converting a sentence included in the evaluation criteria into a vector representation, and calculate the evaluation level based on a similarity between the business plan and each of the evaluation criteria calculated with use of the plan vector data and the
 - 20 criterion vector data.
3. The evaluation system according to claim 2,
 - 25 wherein the evaluation level is represented by a plurality of clusters separated in a vector space including the plan vector data and the criterion vector data, and the evaluation level is calculated through computation of distribution among the clusters.
4. The evaluation system according to any one of claims 1 to 3,
 - 30 wherein the evaluation level calculation unit includes a generative AI, and the business plan and the evaluation criteria are provided to the generative AI as a prompt.
5. The evaluation system according to claim 4,
 - 35 wherein the business plan is a business plan for earning a carbon credit, and the evaluation criteria include: a general evaluation criterion created according to a carbon credit certification standard; and an independent

evaluation criterion created independently.

6. The evaluation system according to claim 5,
wherein the final evaluation of the business plan for earning a carbon
5 credit is affected by a transaction value of the carbon credit.

7. An evaluation program causing a computer to execute processing for
evaluating a business plan with use of a plurality of evaluation criteria, the
evaluation program causing the program to execute:
10 plan input processing for inputting the business plan;
criterion input processing for inputting the evaluation criteria;
evaluation level calculation processing for calculating an evaluation
level for each of the evaluation criteria with respect to the business plan, with
use of the input business plan and the evaluation criteria; and
15 evaluation processing for performing a final evaluation of the business
plan based on the evaluation level.