Received: 26 July 2020 | Accepted: 4 December 2020

DOI: 10.1002/agj2.20622

Agronomy Journal

AGRICULTURAL AND BIOLOGICAL SCIENCES: PLANT, SOIL, ANIMAL AND ENVIRONMENT

Intelligent pig-raising knowledge question-answering system based on neural network schemes

Hsu-Yang Kung[1] | Ren-Wu Yu[1] | Chi-Hua Chen[2,3] | Chan-Wei Tsai[4] | Chia-Yu Lin[5]

[1] Department of Management Information Systems, National Pingtung Univ. of Science and Technology, Pingtung, Taiwan

[2] Business College, Southwest Univ., Chongqing, China

[3] Key Laboratory of Intelligent Metro of University in Fujian, Fuzhou University, Fuzhou, China

[4] Department of Business Administration, National Pingtung Univ. of Science and Technology, Pingtung, Taiwan

[5] Department of Computer Science and Engineering, Yuan Ze Univ., Taoyuan, Taiwan

Correspondence
Chi-Hua Chen, Business College, Southwest University, Chongqing, China.
Email: chihua0826@gmail.com

Associate Editor: Mohamed Nabil Abd-El-Mageed Omar

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 61906043, 11501114, 11901100, 61877010, 61902313, 62002063, 61902072, 11901100; Fundamental Research Funds for the Central Universities of China, Grant/Award Number: SWU2009107; Chongqing Science and Technology Bureau, Grant/Award Number: cstc2020ngzx0010; Funds of Education Department of Fujian Province, Grant/Award Number: JAT190026; Ministry of Science and Technology, Taiwan, Grant/Award Numbers: MOST 108-2637-E-020-00, MOST 109-2637-E-020-004; Fuzhou University, Grant/Award Numbers: 0330/50016703, 0330/50009113, 510649/XRC-18049, 510650/XRC-18050, 510730/XRC-18075, 510872/GXRC-20016, 510930/XRC-20060; Fujian Natural Science Funds, Grant/Award Numbers: 2019J01243, 2020J01498, 2020J05111, 2020J05112

Abstract

The most important parts of Taiwan's agriculture are animal husbandry and the pig industry, of which the output value reached NT$75.6 billion in 2017. Taiwan has a high technical level of pig raising. However, practical pig-raising skills rely mainly on the inheritance of mentoring experience. The livestock and pig breeding industry in Taiwan has no relevant pig breeding knowledge management information system and no intelligent knowledge question-answering system. Therefore, this study designs and implements an intelligent knowledge question-and-answer system for pig farming. To identify intelligent questions and answers for raising livestock pigs, this study addresses the following issues: (a) to determine the semantic meaning of a sentence, the system needs to accurately interpret the meaning of a question and to identify the expression of a knowledge entity. Therefore, this study applies the lattice long-short-term memory (LSTM) and structured perceptron methods to parse sentences accurately and correctly perform word segmentation. A set of stop words for pig raising in Taiwan is initially established to ensure accurate sentence parsing; (b) to understand the intent of each sentence, the bidirectional gated recurrent init (bi-GRU) method is adopted to realize the knowledge extraction of the livestock in the question and complete the intent detection and slot filling. The bi-GRU extracts the correct livestock knowledge and classifies it into suitable topics; (c) the wide range of knowledge sources in questions often leads to unrelated vocabulary, repeated words, and structural loss in potential answers. To establish an effective knowledge search graph, the method infers implicit knowledge from the existing knowledge and conducts related knowledge retrieval based on the inference results. The knowledge data model is

Abbreviations: Bi-GRU, bidirectional gated recurrent unit; Bi-LTSM, bidirectional long-short-term memory; GRU, gated recurrent unit; LSTM, long-short-term memory; NLP, natural language processing; RDF, resource description framework; RDFS, resource description framework schemas; SNN, Siamese neural network; SP, structured perceptron; SPARQL, simple protocol and RDF query language.

defined in the open standard resource description framework Knowledge filtering and knowledge reasoning strategies are presented to produce candidate answers to the livestock problem; (d) to obtain the final answer, most related works search established knowledge graphs, often producing multiple ambiguous candidate answers. The Siamese neural network (SNN) is adopted to obtain accurate answers. An SNN compares candidate answers to a livestock problem as two input values, previously trained by the bidirectional LSTM neural network. The final answer is determined from the cosine similarity value, which represents the highest relevance to the question. Finally, this study implements the intelligent pig-raising knowledge question-answering system based on the proposed methodology. The evaluation results reveal that the proposed system is accurate and practicable.

1 | INTRODUCTION

An investigation by the Executive Committee of the Executive Yuan concluded that the market for Taiwanese pigs has been in good condition for 106 years, according to market transaction prices. The average input/output ratio was 1.48, and the livestock industry in Taiwan is a well-developed and important industry affecting the national economy and the people's livelihood. The current animal husbandry industry is heavily promoting production technology improvement, variety improvement, and disease prevention and control programs to reduce the production costs of livestock products and improve the quality of animal products. Taiwan's livestock breeding experience and knowledge have been handed down by experts and passed down by paper, pencil records, and oral traditions. Many experiences and knowledge cannot be fully recorded or learned and have no clear standard basis, and animal husbandry practitioners are not willing to share their successful experiences. The knowledge of the entire livestock market is closed and has no clear introduction method, increasing the threshold for young people to enter livestock farming. Network technology and communication technology have flourished in recent years; the Internet is becoming an important part of people's daily lives, resulting in a growth in demand. Therefore, the Internet can be used to find knowledge on subjects such as managing sow delivery, most suitable temperature and humidity for growing piglets, and using smart question-and-answer technology to obtain relevant animal husbandry knowledge.

Young people are slowly starting to work in animal husbandry, but it is not an easy task for those who are new to the industry. Animal husbandry requires a lot of research, knowledge, and experience, which is time-consuming and difficult for young animal farmers. Pig breeding requires years of experience and knowledge accumulation. Knowledge of Taiwanese pigs has been passed down from generation to generation. Households are not all the same, and no standard knowledge set is available, leading to demand for complete feeding management practices and question-and-answer intelligent education training. The continuous development of computer science and technology has led to advances in knowledge engineering, in which knowledge is accumulated to create a knowledge question-answering system that can understand the description and semantic meaning of each problem and can answer each question clearly and correctly. With the advancement of science and technology, many people in the livestock industry share information on the Internet and accumulate a large amount of animal husbandry data. However, extracting information from these materials and applying knowledge management, sharing, and application are key to promoting the intelligentization of animal husbandry in Taiwan.

In view of this development, users hope to use the network information to quickly find the most needed answer information, rather than searching through the information provided on a web page. Users are increasingly relying on search engines to provide information. Therefore, researchers are increasingly focusing on how to extract knowledge from the problem and effectively implement knowledge representation, form a knowledge network, and obtain efficient and accurate information-indexing technology for search engines. Producing intelligent animal husbandry questions and answers involves solving many natural language processing (NLP) technical issues, including: (a) correctly determining the livestock participle, so that a computer can understand statements about livestock; (b) understanding livestock issues and extracting relevant knowledge category; (c) using knowledge retrieval for answer selection; (d) selecting the correct answer among candidate answers to give to the user in the animal husbandry knowledge question-and-answer system.

Core Ideas

- Determining the livestock participle correctly
- Understanding livestock issues, and extracting relevant knowledge category
- Using knowledge retrieval for answer selection
- Selecting the correct answer among candidate answers

2 | LITERATURE REVIEWS

2.1 | Natural language processing

Natural language processing is a branch of artificial intelligence and linguistics that transforms complex natural languages into forms that are easy for computers to process and understand (Otter et al., 2020). The rules of NLP were originally set manually, but now the machine learns by itself. In the late 1980s, NLP introduced machine-learning algorithms (Bates, 1995), instead of writing all the transformation rules in a programming language. The algorithmic model allows a computer to learn from the training materials, looking for information including specific patterns and trends. The main aim of NLP is to enable computers to understand human language, to handle words and languages properly, and finally to allow computers to understand the natural language of words and to convert unstructured data into natural language.

2.1.1 | Chinese word segmentation

Word segmentation is the process of splitting a continuous piece of text into a sequence of words according to semantics (Zhang et al., 2018). In English sentences, words are separated by natural delimiters such as spaces. Chinese text is divided into paragraphs by punctuation, but words have no formal demarcation line, making Chinese more complicated than English. In the developed world, users frequently search for information on the Internet. Much of this information comprises unstructured data. To search for articles in Chinese on the Internet, the Chinese words need to be classified. Therefore, Chinese words need to be separated by spaces, as in English. Word segmentation (shown in Figure 1) is the most basic method in natural language.

2.1.2 | Named entity recognition

Named entity recognition can be performed on words by segmenting the non-structural sentences of the livestock problem action (Li et al., 2020). A so-called named entity may refer to a wide range of things in Chinese, such as names of people, places, and institutions, as shown in Figure 2. The categories and rules of the entities are named to complete data classification according to different industry needs.

2.1.3 | Lattice long-short-term memory

Lattice long-short-term memory (LSTM) is a new model method for Chinese-named entity recognition that was proposed in 2018 (Zhang & Yang, 2018) and is superior to character-based and word-based methods (Lample et al., 2016). Lattice LSTM can simultaneously encode character-level sequence information and word messages corresponding to the sequence for automatic model retrieval, thus enriching the semantic expression and minimizing the influence of word segmentation error. The shortcoming of character-based named entity recognition is that it cannot fully express information contained in words and word order, which is important for understanding. This study adopts the lattice LSTM structure to automatically control the information flow from the beginning to the end of the sentence, as shown in Figure 3. The gating cell dynamically transmits information from different paths to each character. After deep-learning framework training on the data, lattice LSTM finds more useful words based on the semantics, leading to better named entity recognition performance.

2.1.4 | Structured perceptron

Structured perceptron (SP) was presented by Collins (Collins, 2002) to solve the problem of sequence labeling in word segmentation. This study encountered many popular words, such as “森 77, 9487, Awesome,” that are not in the animal husbandry dictionary. To handle these words, word segmentation is performed using the SP, which records the cumulative value of the feature weights in each word and thus derives the perceptron of the final model. The advantage of this model is that ‘noise points’ in the data do not affect the overall word segmentation process. For instance, if a document has 1,000 data points, and one of them is a noise point, then the model is

FIGURE 1 Chinese word segmentation based on pig raising

FIGURE 2 Named entity recognition based on pig raising

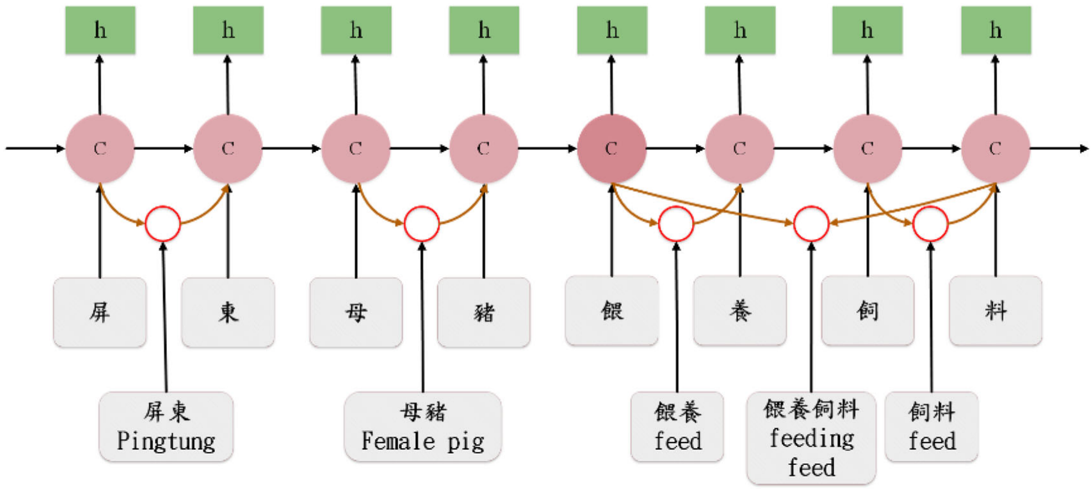

FIGURE 3 Lattice long-short-term memory (LSTM) based on pig raising

trained by the remaining 999 data points. If the prediction is inaccurate, then the SP can accumulate and average the 1,000-model weight vectors to improve the accuracy and practicability of the overall word segmentation model.

2.2 | Knowledge extraction

A document is converted by word embedding in NLP to train the computer to understand the words. This study introduces the deep learning method that will be introduced into the knowledge extraction and strengthens the training mode to increase the learning space. This section introduces the word vector and gate thatform the controlled loop gated recurrent unit (GRU).

2.2.1 | Word embedding

Resources on the Internet are becoming increasingly abundant. If the problem of understanding natural language needs to be converted into machine learning, then the first step is to turn the language into mathematics, for example using Word vectors. Word vectors represent words as vectorsthat are understood by the computer and can be adopted for subsequent applications (Xing et al., 2018).

2.2.2 | GRU and bidirectional GRU (bi-GRU)

Gated recurrent unit and recurrent neural networks are the most different from the neuron. Two control switches (i.e., gates) are added, namely the reset gate and update gate, and the gate structure adopts a sigmoid as the activation function. A fully connected neural network layer outputs a value between 0 and 1, depending on how much data is currently being input and able to pass this structure. The reset gate determines how the new input information is combined with the previous input, and the update gate defines the amount of the previous memory stored to the current time, as shown in Figure 4. The gated loop unit neural networks in the proposed model eliminate the gradient disappearance problem that often occurs in recurrent neural networks, and make it faster than LSTM with better performance and output (Fukushima, 1980).

The main aim of the bi-GRU neural network is to increase the accuracy of the classification and to obtain a complete sentence context. For many NLP tasks, good input is required to train a model well. Forward and backward processing is performed to obtain context information, which provides a vocabulary to create a neural network to generate forward and reverse output data to be passed to the output layer, thus providing the output of each location with context information. The bi-GRU neural network can also be adopted to obtain the context information, which trains the network to identify complete words, as shown in Figure 5 (Zhang et al., 2018).

2.3 | Knowledge retrieval

When a user queries the animal husbandry system, the search query performs classification after the processing is completed. The system queries the knowledge base to produce and return a set of potential answers to the question. This section describes knowledge filtering, knowledge representation, and association searching.

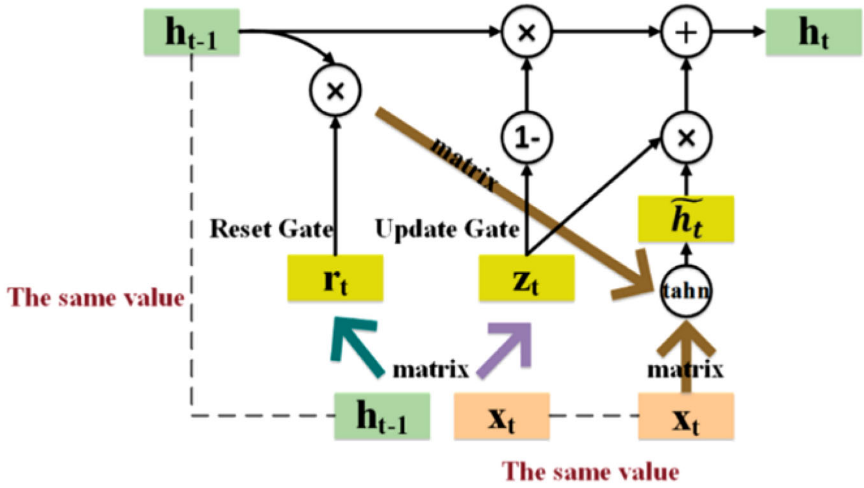

FIGURE 4 Gated recurrent unit

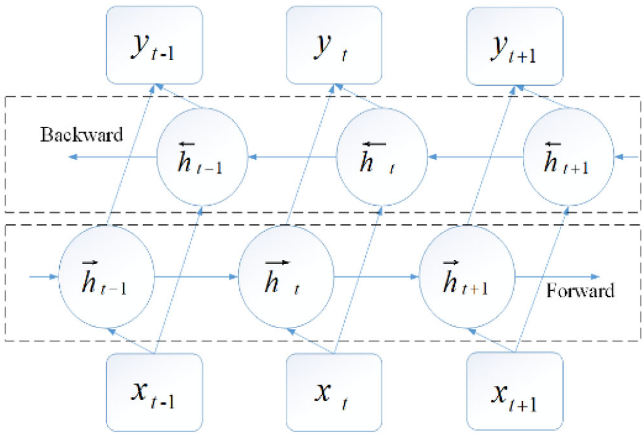

FIGURE 5 Bidirectional gated recurrent unit

2.3.1 | Knowledge filtering

After the knowledge classification is established, key classification words are selected as search keywords for livestock knowledge. A high correlation means that relevant answers are easy to find. In the case of livestock problems, relevant knowledge in areas unrelated to livestock should be considered in order to improve the search accuracy. After the re-classification is completed, the knowledge is filtered again, the animal husbandry dictionary is matched to find such a classification, and the non-herbal related classification words are excluded.

2.3.2 | Knowledge representation

Resource description framework (RDF) is a general description language presented by the World Wide Web Consortium that is written in XML to describe Internet resources and represent correlations between them (Liu et al., 2019). The control vocabulary in an RDF can be customized for a specific subject, such as animal husbandry. RDF schemas (RDFS) build on RDF by providing specifications for the use of RDF categories and attributes, thus making descriptions more meaningful than in a standard RDF. The relationship between classes and sub–classes defines the attributes owned by the category. The qualified attributes point to a range of values, allowing RDFS to perform inference and searching. Therefore, users are able to establish a hierarchical relationship between concepts and attributes based on the inferences.

```
Prefix pro : <http://properties.mysite/>
Select ?site ?owner-name Where {
?site pro : owner-name ?owner-name }
```

FIGURE 6 Simple protocol and RDF query language (SPARQL) syntax query

2.3.3 | Associated search

A structured knowledge map needs to be queried by using simple protocol and RDF query language (SPARQL), which is a query language used to obtain information from RDF graphics, and RDFS (Peng et al., 2018). Figure 6 shows the basic query syntax of RDFS.

2.4 | Knowledge computing

The path of different knowledge is adopted to calculate the category of the question according to the most similar answer to the calculation of the candidate answers generated by the knowledge retrieval and the livestock problem then identify the knowledge base corresponding to the animal husbandry knowledge, and finally answer the animal husbandry question. In the knowledge base question-and-answer system for

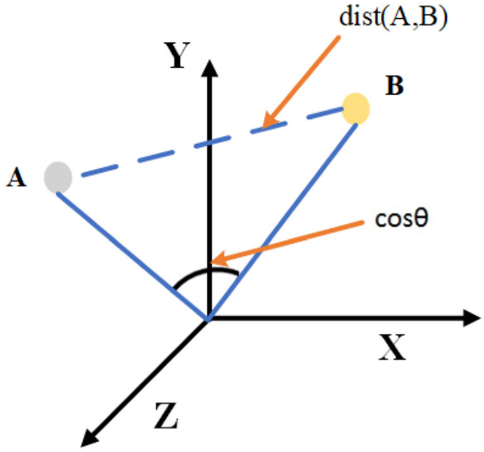

FIGURE 7 Cosine similarity

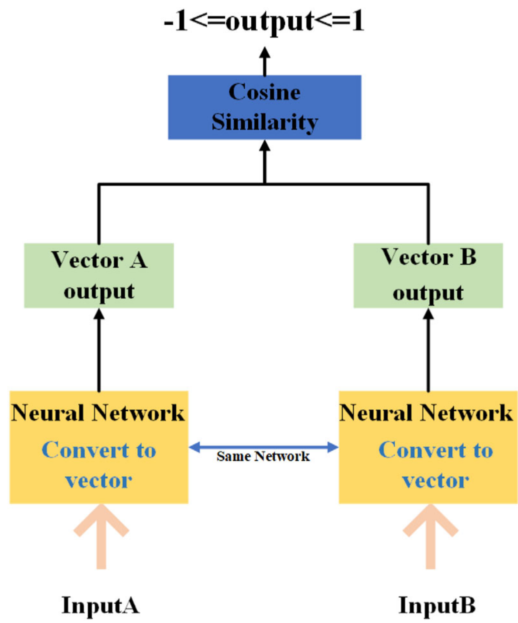

FIGURE 8 Siamese neural network

livestock, the knowledge of the livestock sector provides decision-making data and knowledge to animal husbandry practitioners. Additionally, the system provides advice on decision-making and a smart search method that allows users to search in natural language.

2.4.1 | Cosine similarity

Cosine similarity is a commonly adopted similarity calculation method in knowledge calculation. It can be adopted to calculate the similarity between files, between words, and between query strings (Eghbali & Tahvildari, 2019). The principle of cosine similarity calculation is to measure the similarity between two items by measuring the cosine of the angle between the two vectors, as shown in Equation (1). The cosine of a 0° angle is 1, while the cosine of any other angle is no greater than 1, and the minimum cosine is −1. Thus, the cosine of the angle between two vectors determines whether the two vectors point in similar directions. The cosine similarity is generally used in the positive space, giving it a value between 0 and 1, as shown in Figure 7.

$$\text{similarity} = \cos(\theta) = \frac{A \bullet B}{\|A\|\,\|B\|} = \frac{\sum_{i=1}^{n}\left(a_i b_i\right)}{\sqrt{\sum_{i=1}^{n}\left(a_i\right)^2}\sqrt{\sum_{i=1}^{n}\left(b_i\right)^2}} \tag{1}$$

2.4.2 | Siamese neural network

In 1993, Jane Bromley (Bromley et al., 1993) presented the SNN for signature verification on American checks. Signature verification is a process of determining whether the signature on a check is consistent with the bank reservation signature. The SNN is a measure of similarity used to assess the similarity of two input samples (Bromley et al., 1993). Figure 8 shows the network framework, which has two subnets of the same structure and sharing weights. The network receives two inputs separately and converts them into vectors and then calculates the distance between the two output vectors by means of the cosine similarity distance metric. The SNN learns a similarity measure from the data and compares the learned metrics to match the samples of the new unknown category.

3 | MATERIALS AND METHODS

The intelligent pig-raising knowledge question system based on bi-GRU and SNN schemes is based on the bi-GRU and SNN methods and applies deep learning to the livestock production sector. A smart question-answering system can be built on the basis of the livestock knowledge map.

First, the livestock problem is processed by NLP. The livestock problem is input into the deep learning training mode to segment each sentence accurately and thus detect the intent of the livestock problem. Intent detection and slot filling are adopted to develop an effective knowledge extraction method, find the best extraction process, and classify the problem into the correct topic for subsequent knowledge retrieval.

Due to a wide range of problems, the text often contains non-livestock-related vocabulary, repeated words, and structural defects, so knowledge filtering must be performed before knowledge retrieval to ensure the correctness and integrity of knowledge classification. Knowledge classification is represented using triples. Knowledge reasoning is adopted to obtain the implicit knowledge existing in the data. The knowledge-reasoning results are then searched for the

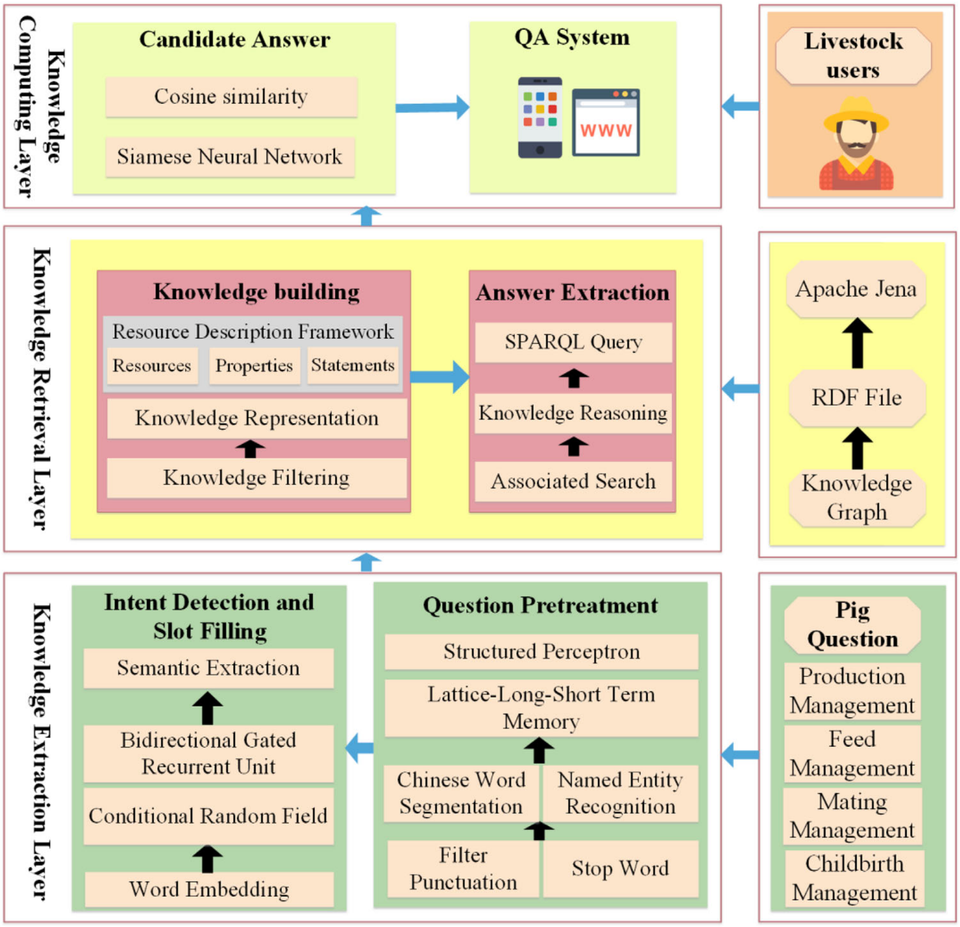

FIGURE 9 Pig-raising knowledge question-answering system

candidate answer by SPARQL, and the candidate answers are finally calculated. Knowledge computation is performed to find the most relevant relevance optimization between the question and the answer. The question-and-answer system can query the knowledge base and automatically find the corresponding answers to questions about production management, delivery management, and breeding management, thereby achieving an increase in production value and ease of operation.

According to the above system function overview, the system studied herein can be divided into three levels: data processing, data search, and data presentation. This system thus has three layers, as shown in Figure 9 below: (a) knowledge extraction layer; (b) knowledge retrieval layer; and (c) knowledge computing layer.

3.1 | Knowledge extraction layer

The Internet has many animal husbandry knowledge bases, including "Taiwan animal breeding original knowledge base," "pig production medical platform," and "administrative hospital livestock production laboratory." However, they are not widely used by animal husbandry workers because the knowledge is scattered and difficult to use, the format is not uniform, and the query process is quite inconvenient. Therefore, the proposed system has a unified research–knowledge extraction layer to input animal husbandry problems. The knowledge extraction layer adopts Chinese word segmentation to process livestock information and extract the scope of the problem. The two-way gated loop unit neural network then learns and classifies the knowledge, according to the probability of semantic occurrence, as birth knowledge, breeding knowledge, breeding knowledge, or breastfeeding knowledge. The query basis for knowledge retrieval is shown in Figure 10 below and described in the following sub-sections.

3.1.1 | Filter punctuation and stop word

Before performing the Chinese word segmentation, the question needs to be processed into a computer-recognizable pattern to obtain the answer. However, if the preprocessing of the question is not filtered, then an abnormality will occur in the knowledge retrieval stage, making the information input by the user difficult to understand and leading to an incorrect

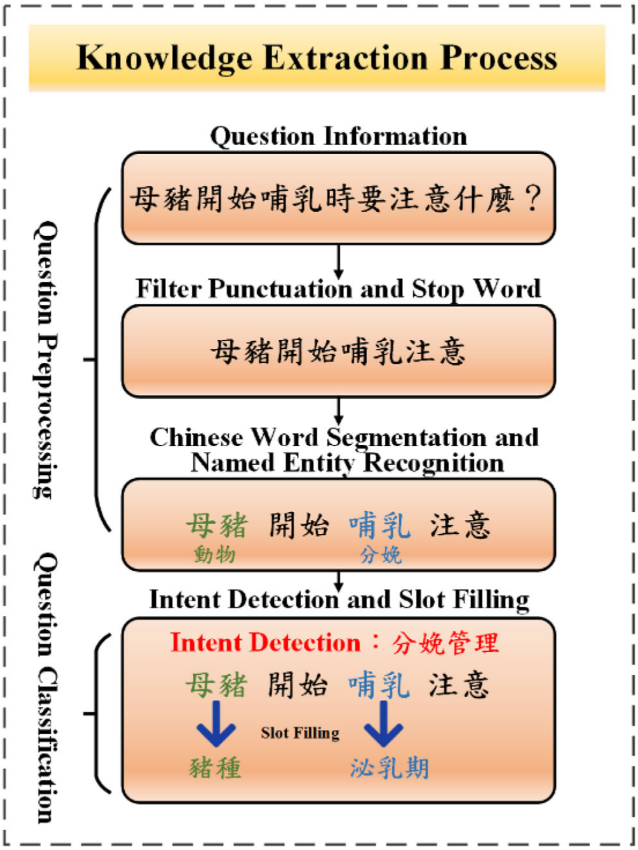

FIGURE 10 Knowledge extraction process

answer. When the computer recognizes Chinese sentences, the punctuation does not affect the computer's interpretation of the semantic mechanism but does affect the result of the word segmentation. Therefore, so that the sentence is segmented accurately, regular expressions are used to filter punctuation in the question.

A stop word is a word that repeatedly appears in a question and has no practical meaning in itself. Stop words in Chinese include most prepositions and auxiliary words. The removal of stop words is often performed in data preprocessing as a prestep using a stop word list, which can be created based on the actual dataset. Stop words are function words contained in human language. These function words are extremely common. Function words, such as "的," "是," "也," and "很," have no practical meaning compared with other words.To improve search efficiency and condense semantics, the system automatically filters out some unwanted words or words before processing natural language data. Therefore, this study establishes a stop word list for livestock issues. For instance, the Chinese words "的," "是," and "什麼," can be filtered out from the sentence "母豬開始哺乳的注意事項是什麼?" by the system for helping narrow the search scope to improve the efficiency and performance of the search engine.

TABLE 1 Begin intermediate other end and single (BIOES) labeling instructions

Label	Full name
B	Begin
I	Intermediate
E	End
S	Single
O	Other

3.1.2 | Chinese word segmentation

The word is the smallest meaningful and freely available language unit. Any language processing system must first distinguish the words in the data for further processing, such as machine translation, language analysis, language understanding, and information extraction. Therefore, the work of Chinese word segmentation has become an indispensable technology for natural language processing. This study adopts the Chinese sentence tree database and the word dictionary tree in the animal language to process the word segmentation of the question, as shown in Figure 11.

3.1.3 | Lattice LSTM Chinese word segmentation

Since Chinese dictionaries cannot list all Chinese words, especially when dealing with documents in different fields, domain-specific special vocabulary or proper nouns often cause word segmentation systems to produce incorrect segmentation due to insufficient reference vocabulary. In order to solve this problem, this study presents a novel word segmentation method, lattice LSTM, to strengthen the collection and segmentation of vocabulary and integrate potential vocabulary information into the segmentation training mode to improve segmentation and search results. The entity recognition effect improves the named entity recognition performance. Each time point input is a character, using BIOES annotation, as shown in Table 1, and find the continuous word of relevance from the vocabulary of the animal husbandry dictionary. The element is calculated in detail as shown in Figure 12.

(I) In the LSTM structure, each cell contains an input gate, a forgetting gate, an output gate, and a sigmoid function of 0 to 1. The sigmoid function is calculated based on the current input and the output of the previous cell. The superscript c represents a character-based model, and superscript w stands for word-based model.

(II) In Equations (2), (3), and (4), parameters i, f, and o denote the input, forget, and output gates, respectively.

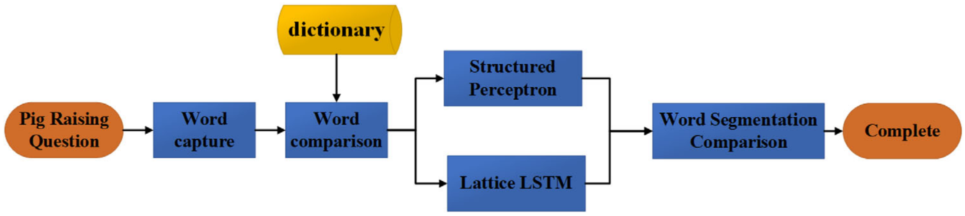

FIGURE 11 Word segmentation process

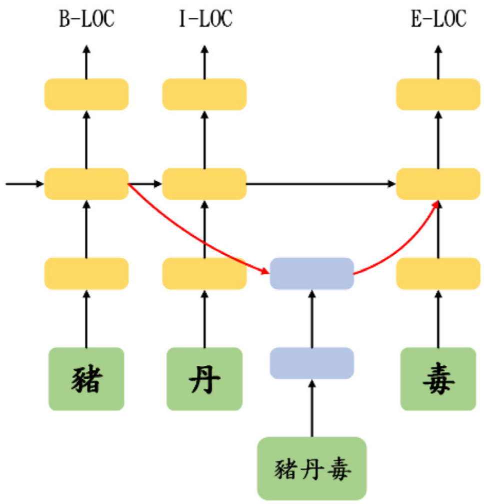

FIGURE 12 Begin intermediate other end and single (BIOES)

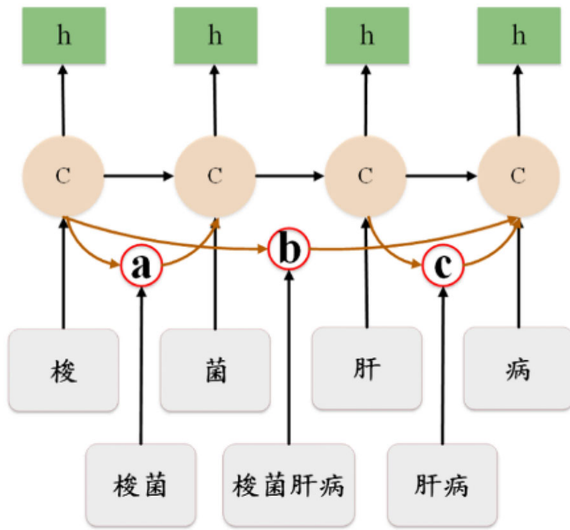

FIGURE 13 Pig-raising: An example of output

Parameter c denotes the input vector through linear transformation, x denotes the input vector, σ denotes the sigmoid function, W denotes the word, and b denotes the bias value.

$$\begin{bmatrix} i_j^c \\ o_j^c \\ f_j^c \\ \tilde{c}_j^c \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \sigma \end{bmatrix} \left(\mathrm{W}^{c\mathrm{T}} \begin{bmatrix} x_j^c \\ h_{j-1}^c \end{bmatrix} + b^c \right) \tag{2}$$

T represents the matrix transpose. b is a matrix.

$$c_j^c = f_j^c \odot c_{j-1}^c + i_j^c \odot \tilde{c}_j^c \tag{3}$$

$$h_j^c = o_j^c \odot \tanh(c_j^c) \tag{4}$$

(III) For instance, in Figure 13, the words in the sentence that end with “病” are: “梭菌肝病” and “肝病.” Therefore, the current character cell must consider these two words in addition to the word “sick.” From the picture, the two red circles, cell a and cell c, are connected to the arrow, which represents the information of these two words “梭菌” and “肝病.”

(IV) The red cells of each vocabulary are independent of each other, where x and h in the matrix denote the output of the word vector and the first word cell, respectively, as in Equations (5) and (6):

$$\begin{bmatrix} i_{b,e}^w \\ f_{b,e}^w \\ \tilde{c}_{b,e}^w \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \tanh \end{bmatrix} \left(\mathrm{W}^{w\mathrm{T}} \begin{bmatrix} x_{b,e}^w \\ h_b^c \end{bmatrix} + b^w \right) \tag{5}$$

$$c_{b,e}^w = f_{b,e}^w \odot c_b^c + i_{b,e}^w \odot \tilde{c}_{b,e}^w \tag{6}$$

(V) These vocabulary messages do not all fit into the current term cell, so trade-offs are made. Therefore, the additional gating unit is adopted to calculate the vocabulary message weight based on the current word and vocabulary information, as shown in Equation (7). The x and c in the matrix denote the vector of the current word and the cell state of the current vocabulary, respectively.

$$i_{b,e}^c = \sigma \left(\mathrm{W}^{l\mathrm{T}} \begin{bmatrix} x_e^c \\ c_{b,e}^w \end{bmatrix} + b^l \right) \tag{7}$$

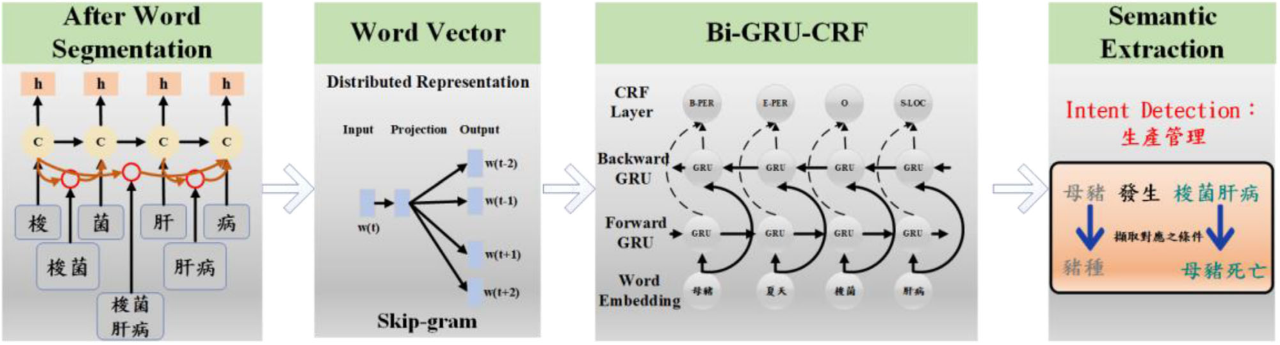

FIGURE 14 Intelligent question-answering system

(VI) Finally, all candidate terms are calculated to be linked to the cell and the weight corresponding to the input.

3.1.4 | Structured perceptron

The SP is adopted to find words that lattice LSTM does not find in the animal husbandry dictionary. Invented words are added to the database to discriminate and filter out meaningless words, such as "森77," "9478," and "好棒棒." The concept of global learning in SP is adopted for structural prediction. The global learning construction model adopts the maximum entropy model (Cho et al., 2014) with a probability P(Y|X), where X denotes the input sequence x_1^n and y denotes the label sequence y_1^n. Equation (8) describes the SP model. The score function is based on the principle of maximum entropy, and the target result is the Y sequence corresponding to the maximum value of the score function, as in Equation (8):

$$\arg\max S(Y,X)\,,\,\text{where}S\,(Y,X) = \sum\nolimits_s \alpha_s \Phi_s\,(Y,X)\,,\,s \in Y \tag{8}$$

Equation (9) shows the calculation of output t, where a_1 to a_n denote the respective component values of the n-dimensional input vector, w_1 to w_n denote the weight values of the inputs connected to the perceptron, b denotes the bias value, and f denotes the activation function.

$$t = f\left(\sum_{i=1}^{n} w_i a_i + b_i\right) \tag{9}$$

3.1.5 | Intent detection and slot filling

This part analyzes the meaning of livestock knowledge based on the knowledge of animal husbandry to form a livestock wisdom question-and-answer system. Breeding knowledge of pigs is classified into production management, feed management, breeding management, and childbirth management to find the correct answer based on these classifications. Figure 14 shows the intelligent question-answering system. First, based on the knowledge map structure in the animal husbandry knowledge, the input problem is analyzed, the related entity relationship is searched from the animal husbandry problem, and the candidate answers are identified from these relationship chains. Degree sorting is then performed to select the correct answer from these candidate answers. That is, each candidate answer is classified into the correct topic in the animal domain knowledge, and the correct search answer is found from the knowledge map.

3.1.6 | Word vector

After the preexisting participle step in this study, for the intent classification, the part of speech must first be converted into distributed representation to the question vector, and the distributed representation is to learn the semantic meaning in an unsupervised way. The distributed representation model can extract more hidden semantic information of sentences than the commonly used methods for calculating semantic similarity based on the direct use of word features or the logical representation of generated sentences. Thus, the model not only solves the problem of data sparseness based on lexical representation; it is also more convenient than traditional methods for extracting the semantic relevance between sentences or words. This method expresses words in a low-dimensional real vector, and each dimension represents a feature of a word. This representation can reflect the complete word message characteristics of the word. For instance, "豬瘟" is expressed as [0.673, 0.726, …, 0.439]. The advantage of this representation is that similar words have smaller distances. The grammatical and semantic similarities of words can be judged from the distance between them, reflecting the correlation between words. This study adopts the skip-gram model for distributed representation. This model was presented by Tomas Mikolov (Mikolov et al., 2013), with the livestock problem as input

TABLE 2 Parameter explanation

Parameter	Explanation
r_t	Reset gate vector
z_t	Update gate vector
x_t	Input vector
h_t	Output vector
W, U	Parameter matrices
σ	Logistic sigmoid

and the words in a livestock question as the input word w_t. The prediction context is $w_{(t-1)}$, $w_{(t-2)}$, $w_{(t+1)}$, $w_{(t+2)}$, , and the output passes through a SoftMax regression classifier. The result is a matrix of probability distributions. The values in the vector represent the probability of occurrence of words other than input words, thus improving the accuracy of the model. The method aims to obtain a distributed representation of the words with lower computational complexity.

3.1.7 | Bidirectional-GRU

According to the Natural Language Computing Group paper, the GRU is superior to the recurrent neural network when using the neural model. The GRU is also better than the LSTM according to Faster (Mikolov et al., 2010). In this study, animal husbandry professional data were accumulated for training an input model, and the word segmentation results were converted into word vector information as input to neural networks. After the forward and reverse two GRU deep-learning models, the word vector message learning is temporarily stored in the hidden layer h to maximize the accuracy of the context. The output is combined with the conditional random field (CRF), and the GRU is adopted to extract the temporary knowledge. The problem classification score corresponding to each value is obtained according to the same weighting data. Finally, the classification topic is selected through the CRF, and the overall score of the comprehensive sentence is divided into its category. The parameters in bi-GRU–CRF are defined and explained in Table 2, and the operational structure diagram is shown in Figure 15.

(I) GRU mainly has two gated values: reset gate and update gate. The value is between 0 and 1 through the logistic sigmoid function, as shown in Equations (10) and (11):

$$\text{Reset Gate}: r_t = \sigma\left(W_r x_t + U_r h_{t-1}\right) \tag{10}$$

$$\text{Update Gate}: z_t = \sigma\left(W_z x_t + U_z h_{t-1}\right) \tag{11}$$

(II) The candidate hidden layer at time t is then calculated as $\tilde{h}$, as shown in Equation (12), where the current time new data is recorded and where r_t denotes the amount of memory that needs to be retained. If $r_t = 0$, then only the current $\tilde{h}$ is required.

$$\tilde{h}_t = \tanh\left(W_h x_t + U_h r_t h_{t-1}\right) \tag{12}$$

(III) The final control needs to forget data from the previous moment h^{t-1} and add data from the current time $\tilde{h}$ to obtain h_t, the hidden layer information of the last output, as shown in the following Equation (13):

$$h_t = z_t \tilde{h}_t + \left(1 - z_t\right) h_{t-1} \tag{13}$$

(IV) If the Reset Gate is close to 0, then the previous message is discarded, so that information related to the non-class can be excluded. The update gate controls how much hidden layer status information is required to control the

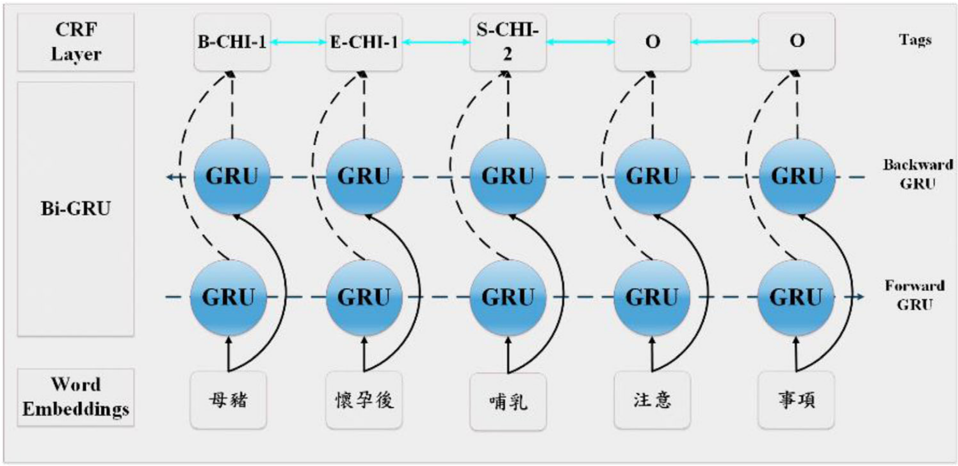

FIGURE 15 Bidirectional gated recurrent unit CRF (bi-GRU-CRF)

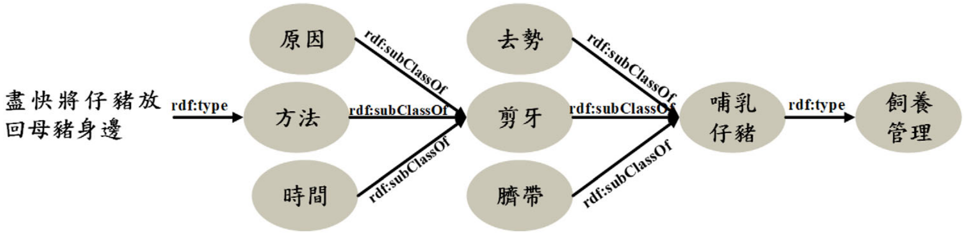

FIGURE 16 Knowledge retrieval

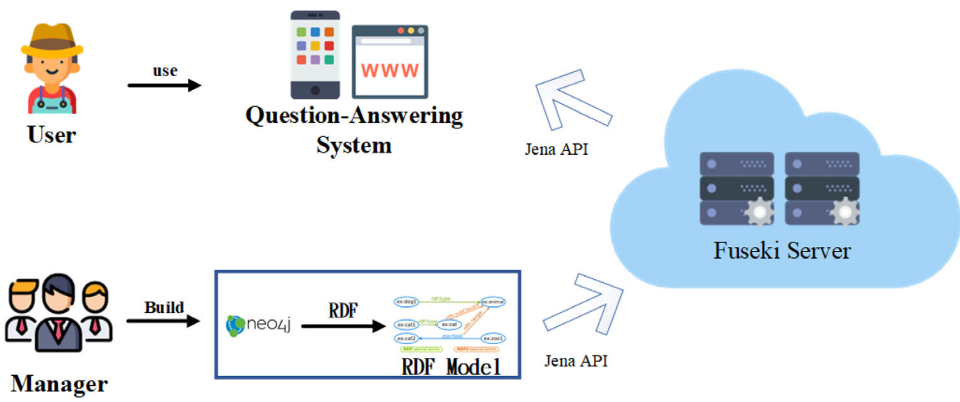

FIGURE 17 Server construction and query flow chart

current hidden layer output h_t. The value of 1 is equivalent to retaining the previous hidden layer state information to the current time, and can learn long distance dependence, so that the context can be clearer.

3.2 | Knowledge retrieval layer

Knowledge retrieval means searching for knowledge maps from the intent classification to find the answers to the corresponding conditions. The knowledge must be filtered to exclude non-related livestock classification words before the answer is selected. The classification attribute of the problem is defined and processed by the RDF. Metadata comprises three data modules, which are resources, properties, and statements, and are presented in the form of the graph-based data dodel triples. The SPARQL map search is performed in parallel to produce candidate answers.

3.2.1 | Knowledge representation

This study obtains entity, relationship, and attribute information from unstructured and semi-structured data after the classification of animal husbandry data. The result is but a series of discrete knowledge sets based on four themes (i.e., birth management, feeding management, production management, and breeding management). Knowledge is represented by RDF triples. This triad model is adopted to search the livestock knowledge map, as shown in Figure 16.

3.2.2 | Knowledge reasoning

Knowledge reasoning ability is an important feature of human intelligence. Some words contain implicit information. For example, knowing the father of a wife infers the father-in-law. Thus, hidden knowledge can be discovered from existing knowledge. Therefore, reasoning needs the support of relevant rules. Knowledge inference of the animal husbandry industry through the knowledge base can infer complete knowledge, reasoning “肉豬” from “公豬” + “育肥,” and inferring “種豬” from “豬” + “配種.” These automated reasoning processes solve the past time-consuming and labor-intensive construction but cannot find relevant meanings from the past database. This study adds TransE model space association to rule–reasoning to optimize the map.

3.2.3 | Rule-based reasoning

The knowledge map is represented by RDF using resources, properties, and statements. However, word definitions have

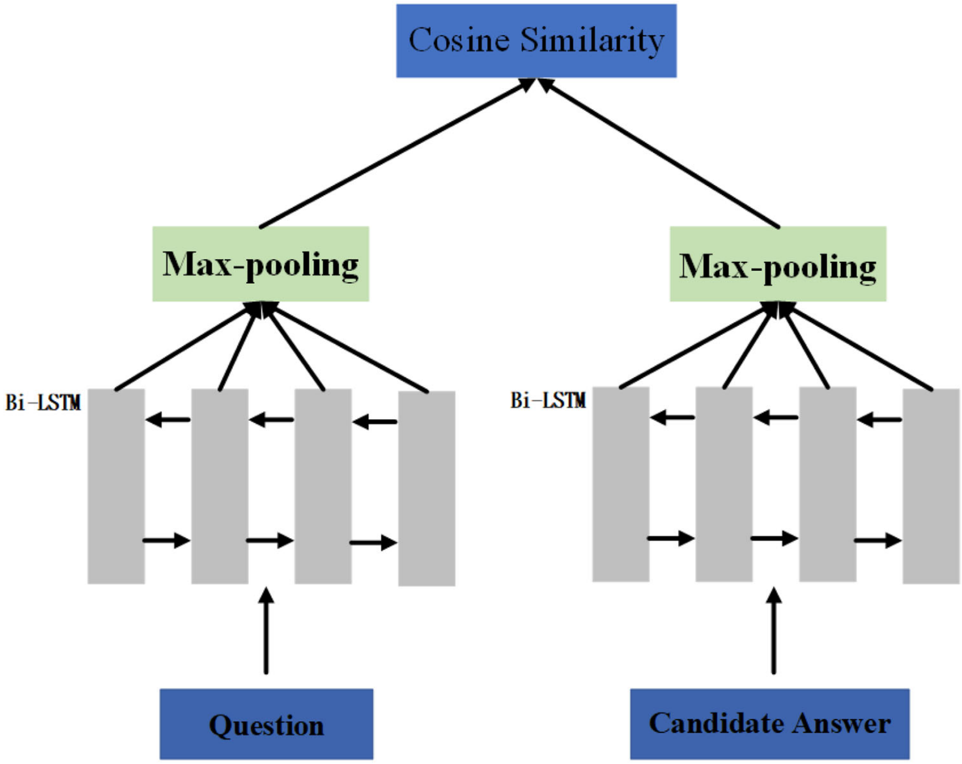

FIGURE 18 Knowledge computing layer

many dependencies. To solve problems in the field of livestock production, this thesis conducts statistical-translating embedding to strengthen the rule reasoning and express the entities and relationships in a sentence as low-dimensional vectors. The relationship between the sentences and the triples is expressed as follows: the subject is the head (h), the property is the relation (r), and object is the tail (t). These have multiple associations in general semantics; that is, the vector of the head vector plus relation is equal to the tail vector. Therefore, TransE transforms an entity into a vector and defines a distance function [$d(h + r, t)$] as a rule-based reasoning learning method. During the learning process, the value of the entity-combined relation vector is constantly adjusted and updated. In the implementation, TransE adopts the maximum spacing method to calculate the distance between $h + r$ and t. S denotes a training set, containing h, l, and t, which belong to the set of connections, and the margin value is set. In the training process of the model, the objective function is as shown in Equation (14):

$$L = \sum_{(h,l,t)\in S} \sum_{(h',l,t')\in S} \left[r+d\,(h+l,t) - d\,\left(h'+l,t'\right)\right] \tag{14}$$

When the objective function dimension is positive, the objective function value is greater than 0, and the associated value h changes, thus obtaining the vector representation of the entity and the relationship. This study performed correlation adjustment in the part of speech between sentences to find the correlations between words and thus obtain the inferences of words.

3.2.4 | Server query

This study adopted the Fuseki Server of Apache Jena's official website as the storage server for the livestock knowledge map. The data were accessed in the form of RDF triples, so they could be input into the SPARQL query service server on Apache Jena. The flowchart of the establishment and query of the livestock knowledge map of this system was first constructed by Neo4j and stored in the Fuseki Server in RDF three-dimensional form. The user can query the knowledge ontology through the system using the Jena API interface, as shown in Figure 17.

3.3 | Knowledge computing layer

An animal husbandry question is answered from the candidate answers generated by the livestock knowledge map. The

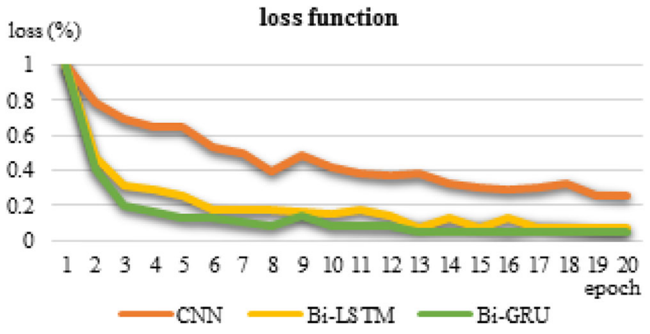

FIGURE 19 Loss function

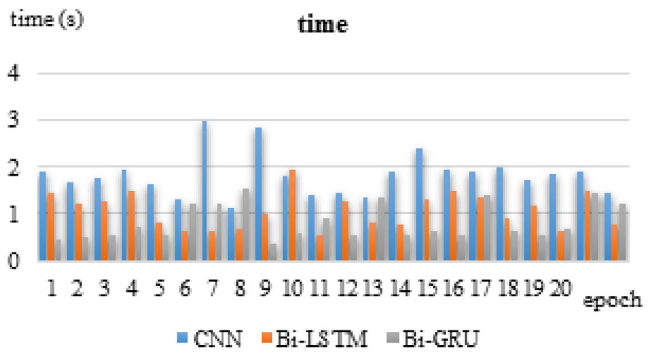

FIGURE 20 Running time cost

knowledge path is adopted to categorize the question. The knowledge base corresponds to the animal husbandry knowledge. The model provides a question and answer on the knowledge base of livestock, knowledge to the livestock industry and advice on decision–making, and a smart search method that allows users to search in natural language. In this part of the study, cosine similarity and SNN are adopted to select the candidate answers and find the answer solution that best fits the problem, as shown in Figure 18. The layer is described in detail as follows.

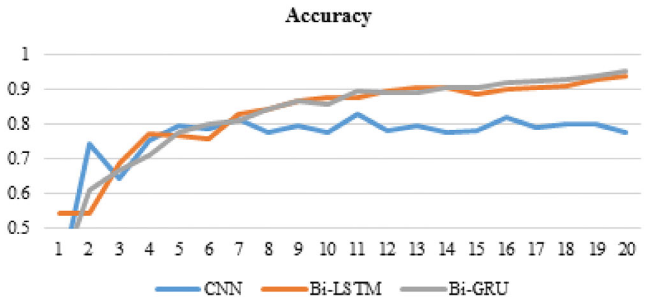

FIGURE 21 Accuracy

3.3.1 | SNN and cosine similarity

According to Paul Neculoiu, any question has only one answer in a question-and-answer system, but it may be expressed in a variety of ways, requiring additional semantic-related information to be captured (Neculoiu et al., 2016). This study adopted the SNN to compare the candidate answers with the livestock problem as two input values. The two inputs were trained by the bidirectional LSTM (bi-LSTM) neural network, and their results were compared by cosine similarity to find the answer that best fit the question. Collecting a large number of questions and answers in the livestock sector and matching the answers to the questions often yielded many answers for each question. The search questions in the multiple options are similar to the answers, so the candidate answers were converted to vector values using word2vec. Questions and answers were then trained by bi-LSTM, using max–pooling to find the most probable answer. The candidate answer path was then calculated via the cosine similarity method to finally find the best answer to the question. The inputs were the questions and candidate answers, which were then processed to generate semantic representations of questions and candidate answers and to the SoftMax layer. Internally, bi-LSTM was adopted to generate state outputs for each node, and max-pooling operations were adopted to generate distributed semantic representations of questions and candidate answers. Max-pooling means taking the maximum value for each dimension of all outputs, generating a distributed semantic representation of the problem and a candidate answer, and randomly suppressing from output some nodes during the training process.

4 | RESULTS AND DISCUSSIONS

4.1 | Data processing

The accuracy, recall, and F-measure of the lattice LSTM, SP, and Jieba Chinese word segmentation were compared by running them on the animal husbandry article. Accuracy, recall, and F-measure were selected to evaluate the proposal method. Experimental results indicate that the proposed lattice LSTM SP word segmentation has a 3.9% higher F-measure than Jieba Chinese word segmentation, which has an F-measure of 91.4%.

4.2 | Knowledge extraction

Knowledge extraction using bi-LSTM, with CRF as the decoding layer, can learn deeper features from the original data to understand the semantic meanings between words and thus train their feature relationships.

FIGURE 22 Knowledge server implementation

TABLE 3 Experimental comparison

Method	Correct rate
Convolutional neural network (CNN)	78.58
Bidirectional long-short-term memory (Bi-LSTM)	86.63
Bidirectional gated recurrent unit (Bi-GRU)	87.79

The main purpose of this model is to find the minimum value after the learning answer is subtracted from the standard answer. Deep learning is adopted to train the network to find the best answer, namely that which is closest to the standard answer, to the loss function. The lowest point, which means that the model has the closest correct answer, will let other neurons learn. The accuracy rate rises with rising number of iterations, as shown in Equation (15):

$$Loss\left(y, y'\right) = \sum_{i=1}^{n} q\left(y, y'\right), \text{where}$$

$$q\left(y, y'\right) = \left\{ \frac{a\left(y' - y\right), \text{if } y' > y}{b\left(y - y'\right), \text{if } y' \leq y} \right\} \tag{15}$$

This study compares the convergence of convolutional neural network (CNN), bi-LSTM, and bi-GRU. As shown in Figure 19, CNN has an erratic learning curve and is thus prone to learning errors. Both bi-LSTM and bi-GRU have good smooth learning curves. However, bi-GRU has a lower loss function than bi-LSTM, so bi-GRU has a smooth curve and lower error and running time cost, as shown in Figure 20.

The final answers from CNN, bi-LSTM, and bi-GRU neural networks were compared for accuracy. Figure 21 shows the accuracy of the CNN model, which was not as high as that of bi-LSTM or bi-GRU. Bidirectional LSTM had a similar accuracy to bi-GRU but a faster convergence speed and smoother curve. Thus, bi-GRU had high performance.

The knowledge extraction method for deep learning was performed on CNN, bi-LSTM, and bi-GRU to verify their performance. Table 3 shows the accuracy rate of each model. Bidirectional GRU had 10.49% higher accuracy than CNN and 1.36% higher than bi-LSTM.

4.3 | Knowledge retrieval

The knowledge map was built by the Neo4j graphics database and stored in the portable document format (FDF) on the Fuseki Server. The user can query the knowledge ontology using the Jena API interface, as shown in Figure 22.

4.4 | Knowledge computing

Training and test data simulation were performed in this model based on one-on-one question and answer training through the neural network. The topic classification judgment was first screened out. The spatial similarity was then judged through the word vector, and select 1,000 original data were used for training. Finally, 50 candidate data were selected for model verification. The candidate answer difference samples were designed separately and verified by three samples. Table 4 shows the accuracy rates of prediction for each test set, which were 100% for test set 1, 92% for test set 2, and 78% for test set 3.

4.5 | Pig-raising knowledge question-answering system

According to the above-mentioned model combination, a pig-raising knowledge question-and-answer system was established to interpret meanings of questions in the field of animal husbandry and give the user correct answers. The proposed system, as shown in Figure 23, provides questions and answers for the animal husbandry knowledge base, decision-making data and knowledge to livestock farmers, advice on decision-making, and smart search methods that allow users to search in natural language.

TABLE 4 Experimental comparison

Test 1: Large-sample difference			Test 2: Medium-sample difference			Test 3: Small-sample difference		
	Result	Rate		Result	Rate		Result	Rate
correct	50	100%	correct	46	92%	correct	39	78%
error	0	0%	error	4	8%	error	11	22%
total	50	100%	total	50	100%	total	50	100%

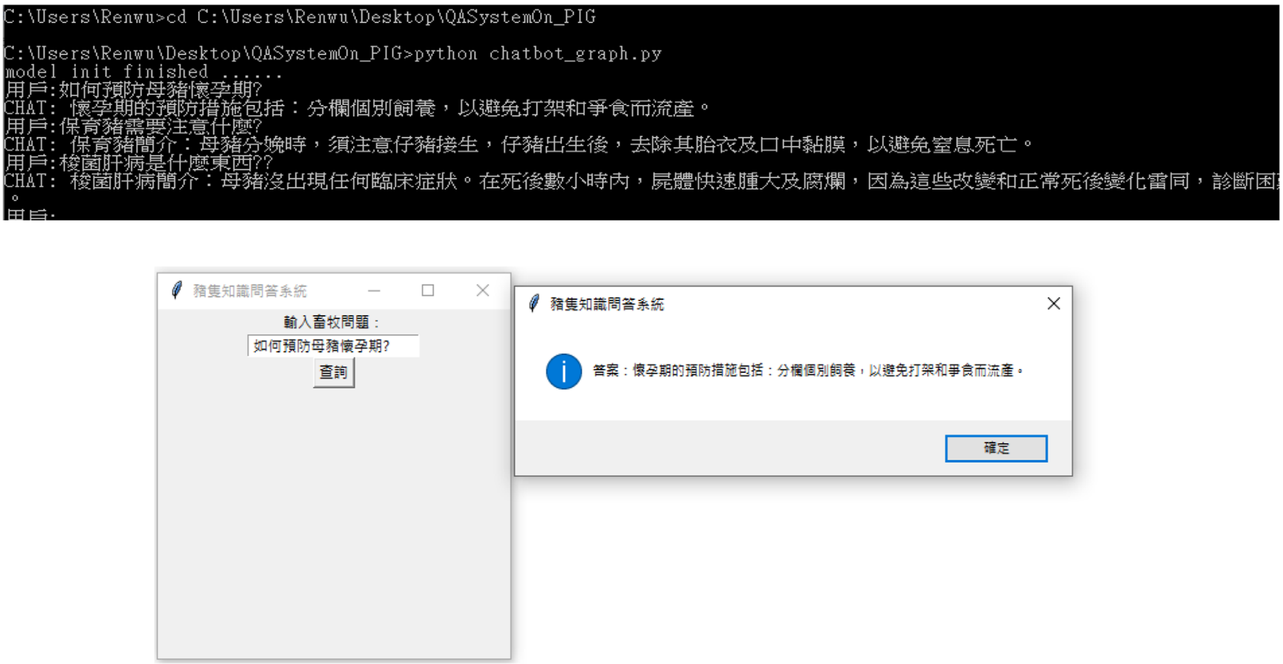

FIGURE 23 Pig-raising knowledge question-answering system

5 | CONCLUSIONS

This study presents a set of questions and answers relating to knowledge of pig–rearing based on bi-GRU and SNN methods, combined with the LTSM deep-learning method, in the following steps. (1) Accurate animal segmentation decision: This study introduces a deep learning lattice LSTM training mode, which allows for precise cutting of statements, and thus achieves accurate word segmentation. (2) Implementation accurate knowledge classification: The bi-GRU method is adopted to link the livestock problem to the intent detection and slot filling, develop effective knowledge extraction methods, find the best extraction process, and categorize them into the correct topics. (3) Knowledge retrieval mechanism design: After excluding the knowledge classification non-livestock related topics, the livestock knowledge is converted into the image data model triplet, and TransE is added to find the correlation between words. The data search is carried out using SPARQL to find candidate answers. (4) Accurate answer decision: Each candidate answer and question is compared with the SNN to find the best answer.

Through system implementation and testing of the effectiveness, this study builds the first deep-questioning and animal husbandry wisdom question-and-answer system in Taiwan, which can provide reference for Taiwanese pig farmers to improve the quality and efficiency of pig breeding.

Because the knowledge map is small, this system cannot fully answer all the questions in the knowledge base. Future work will be to expand the map data to include more livestock fields and increase the classification of various topics. Experts in the field of animal husbandry can help to expand the system to form a more complete and credible knowledge question-answering system.

ACKNOWLEDGMENTS

The work was partially supported by the Ministry of Science and Technology of Taiwan under grant number MOST 109-2637-E-020-004. This work was also supported by the National Natural Science Foundation of China (Nos. 61906043, 61902313, 61902072, 62002063, 61877010, 11501114, and 11901100), Southwest University (No. SWU2009107), Natural Science Foundation of Fujian

Province (Nos. 2020J05112, 2020J05111, 2020J01498, and 2019J01243), Funds of Education Department of Fujian Province (No. JAT190026), and Fuzhou University (Nos. 0330/50016703, 0330/50009113, 510930/GXRC- 20060, 510872/ GXRC-20016, 510930/XRC-20060, 510730/XRC-18075, 510809/GXRC −19037, 510649/XRC-18049, and 510650/XRC-18050).

ORCID

Chi-Hua Chen https://orcid.org/0000-0001-7668-7425

REFERENCES

Bates, M. (1995). Models of natural language understanding. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(22), 9977–9982. https://doi.org/10.1073/pnas.92.22.9977

Bromley, J., Guyon, I., LeCun, Y., Säckinge, E., & Shah, R. (1993). Signature verification using a "Siamese" time delay neural network. In J. D. Cowan, G. Tesauro, and J. Alspector (Eds.). *Proceedings of the 6th International Conference on Neural Information Processing Systems* (pp. 737–744). San Francisco, CA, USA: Association for Computing Machinery.

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In A. Moschitti, B. Pang, and W. Daelemans (Eds.). *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (pp. 1724–1734). Doha, Qatar: Association for Computational Linguistics.

Collins, M. (2002). Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. In K. Knight, A. Nenkova, and O. Rambow (Eds.). *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* (pp. 1–8). Prague, Czech Republic: Association for Computational Linguistics.

Eghbali, S., & Tahvildari, L. (2019). Fast cosine similarity search in binary space with angular multi-index hashing. *IEEE Transactions on Knowledge and Data Engineering*, *31*(2): 329–342. https://doi.org/10.1109/TKDE.2018.2828095

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, *36*, 193–202. https://doi.org/10.1007/BF00344251

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016). Neural architectures for named entity recognition. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 260–270). San Diego, CA, USA: Association for Computational Linguistics.

Li, J., Sun, A., Han, J., & Li, C. (2020). A survey on deep learning for named entity recognition. *IEEE Transactions on Knowledge and Data Engineering*, early access. https://doi.org/10.1109/TKDE.2020.2981314

Liu, J., Yang, M., Zhang, L., & Zhou, W. (2019). An effective biomedical data migration tool from resource description framework to JSON. *Database*, *2019*, baz088. https://doi.org/10.1093/database/baz088

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space: In P. Blunsom, K. Cho, S. Cohen, E. Grefenstette, K. M. Hermann, L. Rimell, J. Weston, and S. Wen-tau Yih (Eds.). *International Conference on Learning Representations*. Scottsdale, AZ: International Conference on Representation Learning.

Mikolov, T., Karafiat, M., Burget, L., Cernocky, J. H., & Khudanpur, S. (2010). Recurrent neural network based language model: *INTERSPEECH 2010, the 11th Annual Conference of the International Speech Communication Association*. Makuhari, Japan.

Neculoiu, P., Versteegh, M., & Rotaru, M. (2016). Learning text similarity with Siamese recurrent networks. In Gurevych and Y. Miyao (Eds.). *Proceedings of the 1st Workshop on Representation Learning for NLP* (pp. 148–157). Berlin, Germany: Association for Computational Linguistics.

Otter, D. W., Medina, J. R., & Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, early access. https://doi.org/10.1109/TNNLS.2020.2979670

Peng, P., Ge, Q., Zou, L., Özsu, M. T., Xu, Z., & Zhao, D. (2018). Optimizing multi-query evaluation in federated RDF systems. *IEEE Transactions on Knowledge and Data Engineering*, early access. https://doi.org/10.1109/TKDE.2019.2947050

Xing, W., Yuan, X., Li, L., Hu, L., & Peng, J. (2018). Phenotype extraction based on word embedding to sentence embedding cascaded approach. *IEEE Transactions on NanoBioscience*, *17*(3): 172–180. https://doi.org/10.1109/TNB.2018.2838137

Zhang, D., Tian, L., Hong, M., Han, F., Ren, Y., & Chen, Y. (2018). Combining convolution neural network and bidirectional gated recurrent unit for sentence semantic classification. *IEEE Access*, *6*, 73750–73759. https://doi.org/10.1109/ACCESS.2018.2882878

Zhang, M., Yu, N., & Fu, G. (2018). A simple and effective neural model for joint word segmentation and POS tagging. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *26*(9): 1528–1538. https://doi.org/10.1109/TASLP.2018.2830117

Zhang, Y., & Yang, J. (2018). Chinese NER using lattice LSTM. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (pp. 1554–1564). Melbourne, Australia: Association for Computational Linguistics.

How to cite this article: Kung H-Y, Yu R-W, Chen C-H, Tsai C-W, Lin C-Y. Intelligent pig-raising knowledge question-answering system based on neural network schemes. *Agronomy Journal*. 2021;1–17. https://doi.org/10.1002/agj2.20622