Study on the Learning Algorithms of Artificial Intelligence

Qinghua Sun¹, a and Fengxiang Yin², b
¹Dept. of Telecommunication Engineering, Shijiazhuang Post and Telecommunications Vocational and Technical College, 050031, Hebei Shijiazhuang, P.R.China
²Dept. of Electronic and Information, Shanghai Electronic Information Vocational and Technical College, 201411, Shanghai, P.R.China

¹18931368679@189.cn, b461165392@qq.com

Abstract. This article delves into artificial intelligence algorithms. From a functional and formal point of view, the artificial intelligence algorithms are divided into four categories which named: statistical based algorithms, tree-based algorithms, neural network-based algorithms, and comprehensive algorithms. On this basis, the common algorithm model and its principle in the field of artificial intelligence are analyzed in this paper. In order to further study the semantic logic in artificial intelligence field, based on the tools of semantic vector space model, this paper abstract and analyze the process of understanding “Knowledge Set” on the empirical level and "Control Knowledge Set" on the thinking level. This paper lays a theoretical foundation for the research of artificial intelligence and robots.

1. AI and Cognitive Neuroscience
Cognitive neuroscience generally studies human psychological activities from the following three levels:

- The nervous system mechanism in which psychological activities trigger related behaviors, many scholars have studied the changes in the nervous system during the process of human cognition or behavior, including the corresponding changes in the nervous system during the process of human thinking, perception, learning, language, and execution;
- Study the principles and rules of consciousness and psychological activities. Existing research shows that knowledge is the consciousness of various things. It is the layer exchange of human sensory incoming information through the human perception system, and the sample with a certain meaning and value is obtained. Eventually, the awareness of the inherent meaning of things is generated by the brain which will generate awareness of the intrinsic meaning of the characteristics and uses of things;
- Theoretical model research on the construction of psychological mechanisms of cognitive behavioral based on brain neuroscience. The related research is mainly based on the achievements of neuroscience and the construction of models of cognitive behavioral psychological mechanisms, such as neural networks, in-depth learning and other algorithm models.

The research of cognitive neuroscience has changed the inherent concept of traditional psychology. Through the experimental research of cognitive psychology science, we have a more objective...
research method on human consciousness process and unconscious process, and accumulated a large amount of data for the study of artificial intelligence cognitive process.

2. Study on the Learning Mode of AI
At present, the learning mode of artificial intelligence is a hot topic in the field of data analysis. Many studies apply the theories such as philosophy and cognitive psychology, cognitive neurology, and artificial intelligence to analyze philosophy process of understanding and human consciousness activities from the perspectives of physical symbol systems, knowledge engineering, and artificial intelligence.

Many products use several algorithms for machine learning. The following figure summarizes the common artificial intelligence learning algorithms.

![Figure 1. Common Artificial Intelligence Learning Algorithm.](image)

Here we introduce the learning logic of artificial intelligence from two aspects.

In the learning mode of artificial intelligence, researchers will choose different modeling methods according to different data types. In the field of artificial intelligence, there are several common learning methods.

2.1. Supervised Learning Mode
In the supervised learning mode, we enter "training data" and mark each group of training data, such as the identification of "spam" in the anti-spam system. The corresponding learning process is established through supervised learning, the forecast results are compared with the actual results of "training data", and the prediction model is gradually adjusted until the forecast results reach the expected accuracy rate. Common types of supervised learning include classification approach and regression predictions. Common algorithms include regression models and reverse neural networks.

2.2. Non-supervised Learning Mode
In the non-supervised learning mode, the data does not need to be specifically marked, and the learning model can independently infer the internal structure of the data. Common learning methods include association rule learning and cluster analysis. Common algorithms include the Apriori algorithm and the k-Means algorithm.

2.3. Semi-supervised Learning Mode
In the semi-supervised learning mode, only part of the input data is marked. At first, this learning model needs to learn the internal structure of the data so that the reasonable organization data can be used for prediction and analysis. Specific algorithms are usually extensions of the supervised learning algorithms, such as graph theory inference algorithms or Laplace support vector machines. At first, the semi-supervised learning algorithm model with the unidentified data, and then predict the identified data.
2.4. **Intensive Learning**

In the enhanced learning mode, unlike the supervision model, the input data is only used as the basis for checking whether the model is correct or not. In this way, the input data is directly fed back to the model, and the model is adjusted accordingly. Dynamic systems and robot control are such learning methods. Typical algorithms include Q-Learning and time difference learning.

At present, the common learning method used by enterprises is supervised and non-supervised learning. Due to the large number of non-marked data and a small amount of identifiable data in the field of image recognition, the focus of this field is on semi-supervised learning. And in artificial intelligence and robots, the enhanced learning mode has widely applications.

2.5. **Artificial intelligence algorithm classification**

Now there are many artificial intelligence algorithms, but from the point of view of function and formal, artificial intelligence algorithms can be divided into four categories: statistical based algorithms, tree-based algorithms, neural network-based algorithms, and comprehensive algorithms.

3. **Semantic vector space model analysis**

With the development of the Internet, the information retrieval model based on keyword matching is widely used, but with the maturity of artificial intelligence technology, semantic retrieval has become a hot research topic. In semantic retrieval, the vector space model proposed by Salton is relatively famous. This kind of retrieval system can abstract query conditions and query documents into two vectors in vector space according to keywords. By comparing the Cosine angles between the two vectors to determine the Relevance of the query and the document, we can receive the document that satisfies the conditions.

This method adopts the method of multiple keywords matching, which can reflect the connotation of the document more than a single keyword match, but this method is based on the independence of keywords, and can not reflect the situation of the term meaning or meaning multiple words. Therefore, in many cases, the central meaning of the document can not be truly reflected. By establishing the mathematical basis of the semantic vector space model, this paper designs the mechanism of article interpretation based on the semantic vector space model, and deeply studies the semantic compatibility calculation method, so as to establish the mathematical foundation for effectively understanding the empirical knowledge and Control Knowledge of intelligence in the article.

3.1. **Semantic vector space model (SVSM, Semantic Vector Space Model)**

We use the semantic vector space model SVSM to abstract the processing of document content into vector operations in vector space. After any document is represented as a related space vector, the match between documents can be measured by calculating the degree of compatibility between vectors. The commonly used matching measure in general text processing is cosine distance.

- Text: refers to any text (article) or fragment in text.
- Key term (Key): The content feature of the text is usually represented by the basic language units (words, words, phrases, etc.) which it contains. These basic language units are called key items of the text. Therefore, Any text can be represented by Key List as K (k₁, k₂,..., kₙ) of which key items.
- Key Weight: For text T (k₁, k₂,..., kₙ) containing N key items, key items are usually given the corresponding weight to indicate their importance in text T.
- Semantic vector space model (SVSM): Given a text T, because the key items in the text can be repeated and ordered. In order to simplify the analysis, the sequential relationship is not considered for the time being. At this time, we regard the key item as an N-dimensional coordinate, and any document is the value corresponding to the n-dimensional coordinate. Therefore, any document T can be seen as an N-dimensional vector.
3.2. Mathematical Abstract of Semantic Vector Spatial Model

(1) Model definition

Definition 1: Assuming that there are M document sets \( T = \{t_1, t_2, \ldots, t_m\} \) (M is the total number of documents), for any document \( t_i \in \{t_1, t_2, \ldots, t_m\} \), the keyword \( k_j \in \{k_1, k_2, \ldots, k_n\} \) (N is the number of keywords), \( \exists K(i,0) \in \{k_1, k_2, \ldots, k_n\} \), such that:

\[
K(i,0) = \max \{k_f\} = \max \left\{ \frac{n_j}{\max \{n_j\}} \right\}
\]  

(1)

It is called the main keyword \( K(i,0) \) of the document \( t_i \). Among them, \( i \in \{1,2,\ldots, m\} \), \( j \in \{1,2,\ldots, n\} \), \( n_j \) is for the number of times the keyword \( k_j \) appears in the document \( t_i \), \( \max \{n_j\} \) is the number of times the keyword appears most frequently in the document set \( T \).

Definition 2: Assuming that \( K(i,0) \) is the main keyword of the document \( t_i \in \{t_1, t_2, \ldots, t_m\} \), \( n(i,0) \) is the frequency that \( K(i,0) \) appears in the document \( t_i \), \( k_j \in \{k_1, k_2, \ldots, k_m\} \), \( k_j \neq k(i,0) \), and \( n_{ij} \) is the frequency that the keyword \( k_j \) appears in the document \( t_i \), so that:

\[
\Delta I(i, j) = \Delta(i, j) \times n(i,0)
\]  

(2)

It is called the semantic increment \( \Delta I(i,j) \) that is the increment between keyword \( k_j \) and the main keyword \( K(i,0) \) in the document \( t_i \). Among them, \( \Delta(i,j) \) is the semantic compatibility between the keyword and the main keyword \( K(i,0) \) of the document \( t_i \).

Definition 3: Given a document set \( T = \{t_1, t_2, \ldots, t_m\} \) and a keyword set \( K = \{k_1, k_2, \ldots, k_n\} \), the semantic vector space model of a single document \( t_i \in \{t_1, t_2, \ldots, t_m\} \) can be described as:

\[
t_i = \{k_{1i}, w_{1i}, \Delta(1,i)\}, \ldots, (k_{ni}, w_{ni}, \Delta(n,i))\} \quad \text{Among them,} \quad i \in \{1,2,\ldots, m\} , \quad j \in \{1,2,\ldots, n\} , \quad w_{ij} \text{ is called the weighted value of keywords } k_j \text{ in the document } t_i , \quad \Delta(i,j) \text{ is the semantic compatibility between keywords } k_j \text{ and the main keywords } K(i,0) \text{ of the document } t_i .
\]

The entire document set \( T \) can be expressed as the following weight matrix:

\[
T = \begin{bmatrix}
[w_{11},\Delta(1,1)] & [w_{12},\Delta(1,2)] & \cdots & [w_{1n},\Delta(1,n)] \\
[w_{21},\Delta(2,1)] & [w_{22},\Delta(2,2)] & \cdots & [w_{2n},\Delta(2,n)] \\
\vdots & \vdots & \ddots & \vdots \\
[w_{m1},\Delta(m,1)] & [w_{m2},\Delta(m,2)] & \cdots & [w_{mn},\Delta(m,n)]
\end{bmatrix}
\]

In the text vector space model, we generally firstly extract the basic language units (words, words, phrases, and phrases) of the text which form key feature items. According to the importance in the text, weights \( w_j \) can be given. The calculation of weights \( w_j \) generally has the following methods:

- Boolean value: The desirable value of \( w_j \) is 1 or 0, which indicates whether the feature appears in the text.
- Word frequency weight value: \( w_j \) expressed by the frequency of features appearing in the document.
Word frequency / file frequency weight: There are two formulas, one considering the amount of text information and the other not considering.

We can use different formulas to calculate word frequency/file frequency weights. Generally speaking, word frequency refers to the value which the number of times a word appears in a document divided by the total number of words in the document. If the total number of words in a document is 1000 and the word "communication" appears five times, the word frequency of "communication" in the document is 5/1000 = 0.005. File frequency refers to how many files in the total files set have appeared the word "communication", the file frequency is indicated the value which the number divide by the total number of files contained in the files set. For example, if the word "communication" appears in 100 documents and the total number of documents is 1,000,000, the reverse file frequency is log (1,000,000/100) = 4. The final word frequency/file frequency weight score is 0.005*4 = 0.02.

Definition 4: Assuming that \( K(i,0) \) is the main key word of the document \( t_i \), \( n_j \) is for the number of times the keyword \( k_j \) appears in the document \( t_i \), for any key word \( k_j \in (k_1, k_2, ..., k_n) \), the formula (3) is the frequency formula of keyword \( k_j \):

\[
\Delta k f_j = \frac{\Delta I(i,j) + n_j}{\text{max} n_j + \Delta I(i,j)}
\]

Among them, \( \Delta I(i,j) \) is called semantic increment, weighting formula (4) is the formula for calculating the \( \Delta KF\text{-ITF} \) weight of keywords \( k_j \).

\[
\Delta k f_j \times \text{itf}_j = \frac{\Delta I(i,j) + n_j}{\text{max} n_j + \Delta I(i,j)} \times \log_2 \left( \frac{m}{n_j} + 1 \right)
\]

(2) Norms of Semantic Vector Space

If we abstract the text into a semantic vector space model, we define a norm with the concept of "length", which can assign non-zero positive lengths to all vectors in the vector space.

If we define the Euclidean norm in two-dimensional Euclidean geometric space \( R \), then in this vector space, the element can be drawn as a directed line segment with arrows starting from the origin, and the length of the line segment of each vector is the Euclidean norm of the vector. Then, the model should satisfy the following characteristics.

Definition 5: Let \( R \) be a linear space over the field of real numbers, functional \( \| \cdot \| : X \to R \) satisfies the following conditions:

- Positive definiteness: \( \| x \| \geq 0 \), and \( \| x \| = 0 \iff x = 0 \)
- Positive homogeneity: \( \| cx \| = c \| x \| \)
- Subadditivity: \( \| x + y \| \leq \| x \| + \| y \| \)

So, we call above \( \| \cdot \| \) is a norm of \( \| x \| \).

If norm is defined in linear space, it is called normed linear space.

(3) Definition of Semantic Matching Degree

Based on the above model, we can describe the matching degree of two documents as the correlation degree between two vectors (Degree of Relevance). When the text is represented as a semantic vector space model, we can measure the semantic matching degree by the similarity (correlation) between vectors. Several distance norms are discussed below to represent the similarity between vectors.
Definition 6: For document set \( T = \{t_1, t_2, \ldots, t_m\} \) and keyword set \( K = \{k_1, k_2, \ldots, k_n\} \), the semantic vector space of a single document is \( \vec{t}_i = \{(k_1, \Delta(k_i, i)), \ldots, (k_j, \Delta(k_j, i)), \ldots, (k_n, \Delta(k_n, i))\} \):

Among them, \( i \in \{1, 2, \ldots, m\} \), \( j \in \{1, 2, \ldots, n\} \), \( \Delta(i, j) \) is the semantic matching degree of the keywords \( k_i \) and the main keywords \( (i, 0) \) of the document \( t_i \).

We can use different functions to define the semantic matching degree of the above space, but the most common measurement method is cosine.

Definition 7: For any two document vectors \( \vec{t}_i \) and \( \vec{t}_j \), the correlation formula is:

\[
R(t_i, t_j) = \cos(\vec{t}_i, \vec{t}_j) = \frac{\vec{t}_i \cdot \vec{t}_j}{|\vec{t}_i| \times |\vec{t}_j|} = \frac{\sum_{k=1}^{n} w_{ik} \times w_{jk}}{\sqrt{\sum_{k=1}^{n} w_{ik}^2 \times \sum_{k=1}^{n} w_{jk}^2}}
\] (5)

(4) Completeness of Semantic Vector Space

Definition 8: Suppose that norm \( d(i, j) = ||t_i - t_j|| \) is used to define the semantic matching degree of any two semantic vectors \( t_i \) and \( t_j \) in the semantic vector space, which called norm semantic matching degree.

Corollary 1: For document set \( T = \{t_1, t_2, \ldots, t_m\} \) and keyword set \( K = \{k_1, k_2, \ldots, k_n\} \), the semantic vector space of a single document is \( \vec{t}_i = \{(k_1, d(k_1, i)), \ldots, (k_j, d(k_j, i)), \ldots, (k_n, d(k_n, i))\} \):

Among them, \( i \in \{1, 2, \ldots, m\} \), \( j \in \{1, 2, \ldots, n\} \), \( d(i, j) \) is the semantic matching degree of two semantic vectors \( t_i \) and \( t_j \). The conditions for the above semantic vector space to be Banach space are as follows: The conditions for the above semantic vector space to be Banach space are as follows:

The normed linear space of the semantic vector space is complete as the metric space (which is naturally induced by its norm).

The normed linear space of the semantic vector space is complete as the metric space by its norm \( d(i, j) = ||t_i - t_j|| \).

Proof:

1) Because \( d(i, j) = ||t_i - t_j|| \) is a norm of the semantic vector space, the above-mentioned semantic vector space is a normed linear space.

2) Conditions show that the metric space normed by \( d(i, j) = ||t_i - t_j|| \) is complete in the normed linear space of the semantic vector space, that is, any Cauchy sequence is convergent in it.

Therefore, the semantic vector space satisfying the above conditions is Banach space.

It can be proved that norms in finite dimensional space are equivalent.

During the development of AI technology, people have many ways to understand human thought and experience knowledge. Using the semantic vector space model, this paper analyses the process of AI technology’s understanding of "knowledge" at the level of experience and thinking, and establishes corresponding mathematical models, which lays a theoretical foundation for further study of AI technology.

Acknowledgments
The supported project Number is JRS-2019-1143 (Research Project of Human Resources Social Security in Hebei Province).
References
[1] Gazzaniga MS. The Cognitive neuro science [M]. Cambridge: MITpress, 1995.
[2] Brechler S J. The newest age of ductionnism [J]. Monitoron Psychology, 2006, 37(8).