International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 12 (2017) pp. 3542-3545

A Probability Learning Model for Constructing Artificial Minds

Sunghae Jun

Department of Statistics,

Cheongju University, Chungbuk, 28503, Korea.

[1]ORCID: 0000-0003-1961-0055

Abstract

Artificial intelligence (AI) has received a lot of attention from diverse fields. Many researchers have studied on AI theories and developed the necessary technologies for AI applications. The main goal of AI research is to find making systems always do the optimal things. But such AI systems are different in general human behaviors. Though people act optimally at all times, they conduct rationally. We need to AI systems thinking and acting like human being in some cases such as care assistance. In this paper, we propose a probability learning model to decide not optimally but rationally. To show how our model can be applied to real problem, we carry out a case study using data set from University of California, Irvine machine learning repository.

Keywords: Probability learning, Artificial intelligence, Artificial mind, Multiple linear regression, Probability distribution

INTRODUCTION

Intelligence is a mental capability to solve diverse problems [1-4]. We can make computer to do right decision as much as we have optimal decision with intelligence. In addition, we have improved our intelligence by learning. Learning is to acquire the knowledge through experience or study [5]. Machine learning is to make computers learn from data (experience). We have had immense results related to machine learning researches [6-7]. We also have developed artificial intelligence (AI) using machine learning. AI is making computers intelligent [8]. Most of AI researches were focused on optimal decisions. We want to make the AI do the right thing. But we need to the computers with mind like humans. Franklin (2001) introduced the artificial mind (AM) [9], and he showed the AM from animal minds to AI with minds. In the AI systems for health care such as personal support services for seniors, we need the AM system supporting seniors. Because most humans including seniors want to the computers with mind or thinking similar to themselves. In this paper, we propose a methodology to construct AM using probability distribution and learning. Ross (2012) defined statistics as a learning art from data [10]. Statistical learning theory and Bayesian learning are two approaches to learning in statistics [11-12]. So we construct AM system using statistical approach based on probability learning. The remainder of this paper consists of as follows. Next section introduces machine learning and artificial mind. We propose a methodology of AM construction using probability learning methods in the following section. To verify the performance of our proposed, we carry out an experiment in the section of experimental result. Lastly, we present our conclusions and future works in last section.

MACHINE LEARNING AND ARTIFICIAL MIND

A learning process is an estimation method between input and output in a system [13]. Machine learning is a learning process of computer using observed data to make computer carrying out optimal decision [12]. Machine learning is also learning from experience (data) to make computer to do right behaviors. Statistics is another tool for learning from data. Ross (2012) defined statistics as follow; "Statistics is the art of learning from data" [10]. Statistics and machine learning all try to carry out optimal decision in given problem. This is representative AI approach and uses only one fixed value estimated from analytical models for optimal decision. In comparison, humans do not act on a single value for optimal decision. Human makes various decisions about a given problem. However, this does not deviate significantly from the optimal value. In this paper, we propose probabilistic learning of human behavior. This process is based on an AM approach to optimal decision. Unlike artificial intelligence, the artificial mind is being studied relatively recently in computer science, statistics, or applied mathematics [9, 14-15]. We study on probability distribution and learning for the AM approach.

CONSTRUCTING ARTIFICIAL MIND USING PROBABILITY LEARNING MODEL

We propose a probability learning model to construct an AM. In general, the AI performs the optimal decision by machine learning including statistics under observed data. In addition, the process provides only one predicted value for finding optimal decision. Every time we should use a fixed value to carry out optimal decision in AI approach. But human do not make a decision by one fixed value in given problem.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 12 (2017) pp. 3542-3545

Sometimes, human's decision may change in the same situation. But, the decision does not deviate significantly from the optimum. In this paper, we study on this behavior of human as an AM approach. Figure 1 shows to compare general AM and our AM approaches.

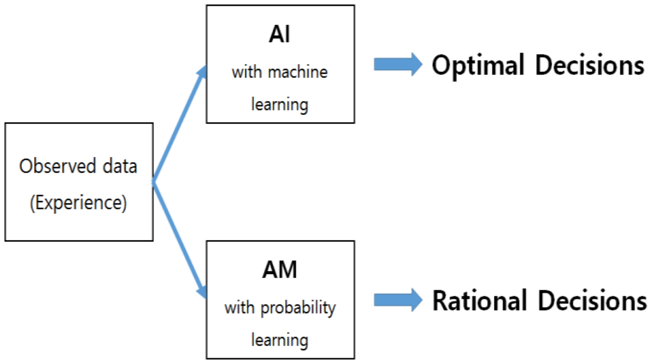

Figure 1. Comparison between artificial intelligence and artificial mind

Under observed data, the AI is based on machine learning algorithms for optimal decisions. The observed data in human behavior can be replaced by experience. In this paper, we introduce AM based on probability learning for rational decisions. The optimal decision means to find only one value to solve given problem in our study. But the rational decision of our research provides various values to solve the problem. In this paper, we consider a regression problem to compare AI and AM. In multiple linear regression model, we fit the regression parameters $\beta_0, \beta_1, \dots, \beta_k$ using observed data $(y, x_1, x_2, \dots, x_k)$ as follow [16].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

The error term ε follows the normal distribution with mean 0 and variance σ^2. In general, we estimate the parameters by least squares point estimation, and compute the predictive value of response y in given new x variables. This is representative approach to AI with machine learning for optimal decision. In this approach, the prediction of y ($\hat{y}$) comes with the following equation.

$$\hat{y}_{AI} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k$$

Where b_i is point estimation of β_i. Every time we get the same result for the predictive value of y. In comparison, we propose another approach to predict y. We do not always get the same results for the predictive value of y. In this paper, we propose to predict y based on AM concept. All regression parameters have their probability distribution with mean and variance in our research. We use normal distribution for the probability distribution of each parameter as follow.

$$\beta_i \sim \mathrm{N}\left(b_i, SE_{b_i}^2\right), \qquad \mathrm{i} = 0,1,2, \dots, \mathrm{k}$$

Where SE_{b_i} represents the standard error of b_i. We generate each random number from this distribution according to $\mathrm{i} = 0,1,2, \dots, \mathrm{k}$, and use the numbers as regression parameters to compute the predicted value of y. So we compute the predicted y by $\tilde{b}_i$ generated from the distribution in the following.

$$\tilde{b}_0\text{: random number from } \mathrm{N}\left(b_0, SE_{b_0}^2\right)$$
$$\tilde{b}_1\text{: random number from } \mathrm{N}\left(b_1, SE_{b_1}^2\right)$$
$$\vdots$$
$$\tilde{b}_k\text{: random number from } \mathrm{N}\left(b_k, SE_{b_k}^2\right)$$

Using the extracted (k+1) parameters, we get the prediction of y ($\hat{y}_{AM_j}$) as follow.

$$\hat{y}_{AM_j} = \tilde{b}_{0_j} + \tilde{b}_{1_j} x_1 + \tilde{b}_{2_j} x_2 + \cdots + \tilde{b}_{k_j} x_k, j = 1,2, \dots$$

This value is different every time, but it does not deviate significantly from the optimal value. The biggest difference between $\hat{y}_{AI}$ and $\hat{y}_{AM_j}$ is the possibility of change every time. That is, $\hat{y}_{AI}$ is always fixed, but $\hat{y}_{AM_j}$ is changed by the generated random numbers for the parameters. The change near optimal value is similar to human mind. Therefore, we construct the artificial mind with probability distribution similar to human.

EXPERIMENTAL RESULTS

To illustrate how our research could be applied to practical problem, we carried out an experiment using popular machine learning data set from UCI machine learning repository [17]. We selected the IRIS data set consisting of one categorical variable (Species) and four continuous variables (Sepal.Length(SL), Sepal.Width(SW), Petal.Length(PL), Petal.Width(PW)). From this data set, we considered the following multiple regression model.

$$SL = \beta_0 + \beta_1 SW + \beta_2 PL + \beta_3 PW + \epsilon$$

While ϵ is error term, and it follows normal distribution with mean=0 and variance=σ^2. Using total 150 instances, we fitted regression model as follow.

$$SL = 1.8560 + 0.6508SW + 0.7091PL - 0.5565PW$$

When the values of SW, PL, and PW are 3.1, 3.8, and 1.2 respectively, we can predict the value of SL for 5.9003 (1.8560+0.6508*3.1+0.7091*3.8-0.5565*1.2). This is traditional AI approach to draw optimal decision. In this paper, we deal with this problem in another way of AM with probability learning. From the result of multiple linear regression, we got the following probability distributions for regression parameters.

$$\tilde{b}_0 \sim N(1.8560, 0.2508^2)$$
$$\tilde{b}_1 \sim N(0.6508, 0.0667^2)$$
$$\tilde{b}_2 \sim N(0.7091, 0.0567^2)$$
$$\tilde{b}_3 \sim N(-0.5565, 0.1276^2)$$

Using the probability distributions, we computed the predicted values of SL as follow.

$$\widehat{SL}_{AM} = \tilde{b}_0 + \tilde{b}_1 SW + \tilde{b}_2 PL + \tilde{b}_3 PW$$

In addition, we predicted the values of SL when the values of SW, PL, and PW are 3.1, 3.8, and 1.2 respectively. Table 1 shows ten values for prediction of SL ($\widehat{SL}_{AM}$) using generated parameters from their probability distributions.

Table 1. Parameter generations and predictions of y for rational decisions

No.	$\tilde{b}_0$	$\tilde{b}_1$	$\tilde{b}_2$	$\tilde{b}_3$	$\widehat{SL}_{AM}$
1	1.8018	0.6806	0.7307	-0.8174	5.7075
2	1.7189	0.6122	0.6571	-0.4819	5.5356
3	1.8488	0.6258	0.6912	-0.5383	5.7689
4	2.1651	0.6781	0.7316	-0.4950	6.4533
5	1.6320	0.7253	0.6663	-0.4272	5.8997
6	1.8784	0.7331	0.8034	-0.4529	6.6605
7	1.6601	0.6295	0.6507	-0.3599	5.6524
8	1.9387	0.6007	0.7502	-0.5847	5.9501
9	1.4042	0.5836	0.7123	-0.6514	5.1385
10	1.6137	0.6757	0.7511	-0.6094	5.8315

In the AI approach, the predicted value was 5.9003. Ten values of $\widehat{SL}_{AM}$ are all different each other, but they are located near 5.9003. We knew that this result of AM reflects the human behaviors. Therefore, we verified the validity of our proposed AM approach.

CONCLUSIONS

Considering human decision making procedures, we distinguished AI and AM according to the method of decision. Both AI and AM make decisions, but AI depends on one fixed value and AM can make decisions through multiple values from the proposed approach of probability distribution. Human behavior and thinking are not always fixed. In our research, every time data (experience) is collected, the model for the decision of AM is evolved. To illustrate how this study could be applied to solve a problem, we made an experiment of multiple linear regression analysis. The regression result based on AI provided only one value for response variable. In comparison, the result of AM based regression constructed the probability distributions of regression parameters. So, we could predict different response every time. But the predicted values were located near optimal value. This research will contribute to the development of artificial intelligence systems that need to be equipped with human minds such as healthcare robots. In this paper, the proposed approach is the beginning to make a computer with a mind. In our future works, we will study on more advanced probability models to make diverse AMs continuously.

REFERENCES

[1] J. R. Flynn, “What is intelligence?: Beyond the Flynn effect”, Cambridge University Press, 2007.

[2] L. S. Gottfredson, “Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography”, Intelligence, vol. 24, no. 1, pp. 13-23, 1997.

[3] U. Neisser, G. Boodoo, T. J. Bouchard Jr, A. W. Boykin, N. Brody, S. J. Ceci and S. Urbina, “Intelligence: Knowns and unknowns”, American psychologist, vol. 51, no. 2, pp. 77, 1996.

[4] R. J. Sternberg, “The theory of successful intelligence”, Review of General psychology, vol. 3, no. 4, pp. 292, 1999.

[5] M. E. Bouton, “Learning and behavior: A contemporary synthesis”, Sinauer, Sunderland, MA, 2007.

[6] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects”, Science, vol. 349, no. 6245, pp. 255-260, 2015.

[7] T. Mitchell, “Machine Learning”, McGraw-Hill, Boston, 1997.

[8] S. Russell and P. Norvig, “Artificial Intelligence: A Modern Approach, Third Edition”, Pearson, Essex, UK, 2014.

[9] S. Franklin, S. Wolpert, S. R. McKay and W. Christian, “Artificial minds”, Computers in Physics, vol. 11, no. 3, pp. 258-259, 1997.

[10] S. M. Ross, “Introduction to Probability and Statistics for Engineers and Scientists, Fourth Edition”, Elsevier, Seoul, Korea, 2012.

[11] R. M. Neal, “Bayesian learning for neural networks”, Springer Science & Business Media, 2012.

[12] V. Vapnik, “Statistical learning theory”, Wiley, New

York, 1998.

[13] V. Cherkassky and F. M. Mulier, “Learning from data: concepts, theory, and methods”, John Wiley & Sons, New York, 2007.

[14] S. Franklin, “Artificial Minds”, MIT Press, Cambridge, MA, 2001.

[15] R. Kurzweil, “How to create a mind: The secret of human thought revealed”, Penguin, 2013.

[16] B. L. Bowerman, R. T. O’Connell and A. B. Koehler, “Forecasting, Time Series, and Regression, An Applied Approach”, Brooks/Cole, Independence, KY, 2005.

[17] UCI ML Repository, The UC Irvine Machine Learning Repository, http://archive.ics.uci.edu/ml, 2017.