Proceedings of the 2nd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/104/20241150

A Research on Defects and Improvements to Multi-Armed Slots

Shengyu Wu

Shanxi University, Taiyuan City, Shanxi Province, 030000, China

18803424671@163.com

Abstract. In recent years, various large language models have sprung up, which are constantly being trained and updated. Most of the time, these models can provide answers more accurately. However, in some specific cases, the answers generated by training on a large corpus of non-emotion text may not be correct. This paper proposes a feasible method to reduce the sum of squares as much as possible so that the fitting function is closer to the true value. By adding the concept of emotion to deep learning, combined with the fusion model of multi-layer linear regression and neural network learning, it was found that adding emotion is more helpful to the language model through multiple trainings, so that the "arm" pulled by the AI each time gets a larger expectation than the original, which significantly improves the reliability and provides a new idea and direction for the problem in this direction.

Keywords: Deep learning, Multi-layer linear model, Emotion, Multi-armed slots.

1. Introduction

With the rise of artificial intelligence, people can always hope to produce a robot that can really think independently like a human, these AI can communicate like a real person and even put forward their own suggestions, and constantly updated algorithms are the key to ensure that these AI can be more "anthropomorphic", in recent years, chatbots such as Tongyi and Du Know have become people's main focus. Through continuous deep learning, a more fitting function is formed, so that AI can complete tasks more accurately, and even guess people's thoughts to a certain extent. At the same time, AI is also shining in many fields such as medicine and machinery. However, it is important to note that even if these AIs have a large enough database, they can still go wrong in some directions, because to a large extent, people's emotions affect people's motivation, for example, emotions can affect a patient's recovery time, and if judged only by past data, it can shorten the patient's treatment time and have unpredictable consequences. That's the case with a lot of deep learning today. Therefore, adding emotion, an important factor, to deep learning and then using multi-layer linear models to construct functions will make it more fitted, which is of far-reaching significance for updating algorithms and improving the credibility of AI.

2. Practical application of multi-armed slot machines

First, to give an example, when a person asks a question to an AI; I'm in a terrible mood right now, can you recommend today's lunch. The process is similar to a multi-armed slot machine problem, where the gambler stands in front of the slot machine and selects the lever to get the most prizes. Of course, no one gets the jackpot the first time. So, constant experimentation is essential. When the number of times is high enough, the probability of getting the jackpot each time is greater until he can finally guess exactly which lever the jackpot is hidden underneath. Similarly, we changed the decision lever to AI, and each lever represents an option. The casino owner represents a questioner when the questioner asks a simple question such as: Can you tell me what to eat today? At this time, it is the turn of the AI to make a choice. Each choice represents an option, and the AI needs to choose an arm from it and obtain the corresponding benefits, the goal is to maximize the sum of the benefits obtained in a given time, that is, the maximum probability can guess the thoughts of the casino owner at this time.

The multi-armed slot machine algorithm is a classical online learning algorithm that does not rely on prior knowledge [1], and the competitive analysis of the evaluation online algorithm has also evolved into the regret analysis of the multi-armed slot machine model [2]. The multi-armed slot machine model observes the benefits generated by the selected arm at each decision round but does not observe the gains of the other arms that could have been selected, and the returns of all arms are unknown until they are pulled. Therefore, the multi-armed slot machine model does not rely on prior knowledge but adjusts the decision-making scheme based on existing empirical knowledge. Moreover, the benefits produced by pulling the arms are assumed to satisfy the Independent and identically distributed (i. i. d), i.e., the benefits are random variables that obey the same distribution and are independent of each other.

The multi-armed slot machine model can be simplified to the following mathematical model, there are N options in a decision, that is, N arms, N = {1, ... , N} represents the set of arms, so that T = {1,2, ... , T} represents the decision round, in the T round decision-making process, the player (or decision-maker) chooses to pull a certain arm k among N arms and obtains a benefit value :

$$X_t^k \tag{1}$$

which is expected to be a random distribution of $\mu k = E[x_t^k]$, so that $\mu*$ represents the maximum possible benefit for the decision-maker

$$\mu^* = \max_{k \in N} \mu_t^k \tag{2}$$

let denote whether the arm k is pulled at time t, if it is pulled d_t^k 1, vice versa d_t^k= 0, let denote the total number of times the arm k is pulled at time t, The formula is as follows:

$$H_k(t) = \sum_{j=1}^{t} d_j^k \tag{3}$$

As mentioned above, the multi-armed slot machine can express a decision, the gambler chooses one from multiple arms to make the most profit. However, because it does not have a relevant database, the selected arm will also improve with the increasing number of times the slot machine is pulled, when its training times reach a certain number, the profit obtained is maximized, and the difference between the profit generated by the current scheme and the optimal return under the ideal situation is defined as regret. The multi-armed slot machine model aims to maximize the cumulative return, which is also equivalent to minimizing the cumulative regret, the regrets arising after the T-round decision-making process are:

$$R_t = T\mu^* - \sum_{t=1}^{T} d_t^k x_t^k \tag{4}$$

Some variants

Defects and solutions of multi-armed slotsWith the iterative evolution of algorithms and the rapid development of chip computing capabilities, various multi-armed slot machines have sprung up, including the Federation Dobby [4], the contextual multi-armed slot [5], Duling Bandit [6], and the combination multi-armed slot [7].

some flaws

Defects At a time when natural language models (NLMs) are undoubtedly the focus of attention, some AIs can only use deep learning on "surface matter" (if the AI wants to determine what a person is going to eat today, then it will judge by the salary it earns, the restaurant closest to the place where it lives), and which arm will receive the most reward. And many times, the substance on the surface is not the only influencing factor. When a person feels hungry, he changes not only in terms of gastrointestinal function, but in many ways, perhaps even in most of the functions that he has. His perception has changed (he is more likely to find food), his memory has changed (he is more likely to recall a good meal), his emotions have changed (he is more nervous, excited), and the content of his thoughts has changed (he is more inclined to think about getting food) [8]. (This is just a list of the effects of emotions on appetite, everything humans do is more or less affected by emotions, such as negative emotions after surgery, which will greatly affect the recovery time).

The surface matter is briefly described in the previous paragraph and will be explained in detail in this paragraph. The surface matter is the sum of all the material independent variables that affect the dependent variable, because this variable can be clearly observed compared to emotion: for example, when we evaluate the act of eating, then the wages earned by the individual are a superficial substance. The surface substance will change with the independent variable. For example, when a person is sick, the severity of the disease will affect the recovery time, the severity of the disease is a surface matter. However, when the dependent variable changes ,something else, the severity is not necessarily superficial

2.1. Resolution

In order to solve this problem, this paper uses data at the two levels of individuals (emotions) and "surface matter" with nested relationships. It assumes that each study emotion is nested in each "surface matter", which is highly correlated with the individuals of "surface matter", and the individuals of different "surface matter" are independent of each other. The multi-layer linear model (HLM) can effectively connect the non-independent data at the micro and macro levels to distinguish between individual and group effects. The traditional HLM requirements for the dependent variable are limited to continuous data. However, the dependent variable in this paper is dichotomous data with a value of 0 or 1, and the use of traditional HLM is no longer applicable. The Generalized Multilayer Linear Model (HGLM) is developed based on HLM and uses a binomial sampling model and a logit link function to process data with a dichotomous dependent variable. [9] The expression for the model is as follows:

$$\mathrm{Prob}(y_{ij} = 1/\beta_{ij}) = \varphi_{ij} \tag{5}$$

Level 1

$$\log[\varphi_{ij}/(1-\varphi_{ij})] = \eta_{ij} \tag{6}$$

$$\eta_{ij=\beta_{0j}+\sum_{n=1}^{k}\beta_{nj}X_{ij}} \tag{7}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \sum_{m=1}^{l}\gamma_{om}W_{mj} + \mu_{0j} \tag{8}$$

$$\beta_{nj} = \gamma_{00} + \sum_{m=1}^{l}\gamma_{om}W_{mj} + \mu_{nj} \tag{9}$$

Hybrid model

$$\eta_{ij} = \gamma_{00} + \sum_{m=1}^{l}\gamma_{0m}W_{mj} + \sum_{n=1}^{k}\gamma_{n0}X_{nij} + \sum_{m=1}^{l}\gamma_{nm}W_{mj}X_{nij} + \mu_{nj}X_{nij} + \mu_{0j} \tag{10}$$

Where: y_{ij} indicates whether the i-th individual of the j-th surface substance chooses to eat; φ_{ij} denotes the probability of the ith individual of the j-th surface substance choosing to eat; η_{ij} is a function of the i-th individual dependent variable of the j-th surface substance; X_{nij} denotes the observed value of the i-th individual independent variable of the j-th surface substance; β_{0j} and β_{nj} represent the intercept and slope of the dependent variable to the independent variable regression of the j-th surface substance, respectively; W_{mj} notes the mth independent variable of the j-th surface substance at the

DOI: 10.54254/2755-2721/104/20241150

regional level; γ_{00} and γ_{n0} represent the intercepts of β_{0j} and β_{nj}, respectively; γ_{om} and γ_{n0} represent the slope of W_{mj}, respectively; $\mu_{0j}\gamma_{00}$ and $\mu_{nj}\gamma_{00}$ are random error terms.(The individual here represents the emotion)

2.2. Analysis of influencing factors

Before using HGLM to analyze the influencing factors of endless labor, it is necessary to establish a zero model, that is, only the dependent variable is added, and whether there is a significant difference in the dependent variable is analyzed by calculating the internal correlation coefficient ICC, and then whether HGLM can be used is judged. Because human motivation is diverse and complex, this experiment only studied the simplest behavior of eating, and the results were only for the two options of wanting to eat or not wanting to eat. The results of the random-effects chi-square test were significant, different emotions have different weights on the behavior of eating, and there is a significant gap in the endless labor behavior under different emotions, as shown in the figure below.

Table 1. Data display.

eat	Fixed effect	Between-group variance	X squared
	γ_{00}	0.227	173.221

2.3. Analysis of moderating effects

When emotions are added to this model, look at the data obtained. The effect of surface matter on the probability of eating is indeed disturbed by the moderating variables. Based on the null model, the factor of emotion is added to it, and it is found that in different individuals, even if their surface material is the same, different emotions will produce different probabilities of the behavior of eating. When an individual becomes happy, the individual is more inclined to choose to eat, while when an individual becomes depressed, he begins to feel anxious, irritable, and even lose their mind. At this point, the probability of the individual eating will become very small. Even though he was close to the restaurant at this time and had enough money, he still tended to choose not to eat.

2.4. Weight analysis

Through the comparison of different individuals, it can be found that in many cases, bad emotions have a much greater impact on behavior than a good emotion. The weight of these bad emotions is usually very large, and a bad mood can often change the original trajectory of the answer.

3. Recommendations based on the results of the experiment

When people face multiple pressures, such as academics, changes in interpersonal relationships, and uncertainty about future employment, these pressures often have a profound impact on their emotions. Anxiety, stress and loneliness, as well as self-doubt are likely to ensue as you grow and learn. For example, when important nodes such as final exams and graduation theses arrive, high academic pressure can make students feel anxious and nervous. In terms of social interaction: For most people, it is likely to be difficult to adapt to an unfamiliar environment, so understanding and effectively coping with emotional problems, and providing corresponding psychological support and resources are essential for individual growth.

4. Conclusion

With the rapid development of technology, artificial intelligence, and machine learning have become increasing popular, and the problems caused by them have also attracted widespread public attention[10]. A large number of machine learning algorithms are only designed to maximize the expected value (i.e., maximize benefits) and ignore a small number of individuals, which will lead to the inability of AI to accurately determine the motivations of each individual. In addition, algorithms in real-world scenarios cannot fully understand the input data, most of the relevant data is currently inaccessible, and algorithms

DOI: 10.54254/2755-2721/104/20241150

that rely on prior information often do not fit into practical applications. Specifically, the main work of this paper is to use a multi-layer linear model to add emotion as a factor to machine learning to find out whether the fitting function generated by it is closer to the real situation and whether the results obtained are more accurate.

This article only discusses the simple act of eating, and for an individual, his behavior is not only simple: communication, labor, and learning. It can be said that an individual's life is made up of a variety of behaviors and motivations, so it is essential to make AI more human-like, build more complex models, and conduct more extensive training. At the same time, this paper only selects a few simple emotions (joy, sadness, anger, anxiety), and the influence of a person's motivation may be far more than these.It is necessary to comprehensively consider multiple influencing factors in the future to propose a more comprehensive fairness measure.

References

[1] Tobias S. Interest, prior knowledge, and learning[J]. Review of educational Research, 1994,64(1): 37-54.

[2] Bubeck S, Cesa-Bianchi N. Regret analysis of stochastic and nonstochastic multi-armed banditproblems[J]. Foundations and Trends in Machine Learning, 2012, 5(1): 1-122.

[3] Montgomery-Smith S J. Comparison of sums of independent identically distributed randomvariables[J]. arXiv preprint math/9310214, 1993

[4] Lai T L, Robbins H. Asymptotically efficient adaptive allocation rules[J]. Advances in appliedmathematics, 1985, 6(1): 4-22.

[5] Wang L, Wang C, Wang K, et al. Biucb: A contextual bandit algorithm for cold-start anddiversified recommendation[C]//2017 IEEE International Conference on Big Knowledge (ICBK).IEEE, 2017: 248-253.

[6] Yue Y, Broder J, Kleinberg R, et al. The k-armed dueling bandits problem[J]. Journal ofComputer and System Sciences, 2012, 78(5): 1538-1556.

[7] Gai Y, Krishnamachari B, Jain R. Learning multiuser channel allocations in cognitive radionetworks: A combinatorial multi-armed bandit formulation[C]//2010 IEEE Symposium on NewFrontiers in Dynamic Spectrum (DySPAN). IEEE, 2010: 1-9.

[8] Maslow, A. H. (2023). Motivation And Personality: Motivation And Personality: Unlocking Your Inner Drive and Understanding Human Behavior by AH Maslow. Prabhat Prakashan.

[9] Li, X., & Xin, T. (2006). Multilevel linear models for dichotomous data: Principles and applications. Psychological Development and Education, 4, Article numbers if available.

[10] Nie, J. (2017). Labor, income, and the differentiation of elderly care stages among rural elderly: An empirical analysis of the "no rest" phenomenon. Study and Practice, 8.