Cpts_570: Machine Learning

Homework IV

12/3/2019

Instructor: Janardhan Rao (Jana) Doppa

Student: Yu-Chieh Wang 11641327

1. Analytical Part

1. Finite-Horizon MDPs.

a. Two game consoles are very popular in the market.

The first machine which use state-action reward function have a pair of scissors and a prize dropped on a rope. Players must use scissors to cut the rope to win prizes. The machine is designed to have only one button, so players can only press this button to control the distance the scissors move forward(Actions depends on how long a person press the button). When the button is pressed, the scissors in the machine will move forward slowly. When the button is released, the scissors stop moving and cut forward. If the rope is cut successfully, the prize will fall down to the exit(State 1). Conversely, the player fails the challenge(State 0).

The other which use state reward function machine is a slot machine. Players can know if they have won or not(State 1 and 0) by pulling the lever(Action). If a machine has only one action in each state and the reward of any action is the same, we can ignore the actions and simplify its function as a state reward function.

b. We set the initial value to 0 because there will be no action at first.

$V^0(s) = R(s), \forall s \Rightarrow V^0(s) = 0, \forall s$

$$V^k(s) = R(s) + \underbrace{max}_{a} \sum_{s'} T(s,a,s')V^{k-1}(s')$$

$$\Rightarrow V^k(s) = \underbrace{max}_{a} F(s,a) + \sum T(s,a,s')V^{k-1}(s')$$

$$\pi^*(s,k) = arg\underbrace{max}_{a}\sum_{s'} T(s,a,s')V^{k-1}(s')$$

$$\Rightarrow \pi^*(s,k) = arg\underbrace{max}_{a}\sum_{s'} T(s,a,s')V^{k-1}s'$$

c. **New MDP**

We modify the transition function and reward of MDP to get the new MDP M^{new} .

$T^{new}(s,a,q_{s,a}) = 1$

$T^{new}(q_{s,a}, a^{new}, s^{new}) = T(s,a,s^{new})$

$R^{new}(s) = 0$

$R^{new}(q_{s,a}) = R(s,a)$

2. **K-th Order MDPs.**

For any k-order MDM M, we build a new equivalent MDP M' :

The state space, action and reward function for M' are $S' = S^k$, $A' = A$ and

$R'((s, s_1, s_2 \dots s_{k-1})) = R(s)$.

Because the reward in the new MDP depends on the current state, we can write the

transition function for M' as following:

$$T'((s, s_1, s_2 \dots s_{k-1}, a, \overrightarrow{s})) = \begin{cases} P_r(s' | a, s, s_1, s_2 \dots s_{k-1}), if\ \overrightarrow{s} = s, s_1, s_2 \dots s_{k-2}\,. \\ 0, otherwise\,. \end{cases}$$

Then, we get the policy for M' as following:

$\pi'((s, s_1, s_2 \dots s_{k-1})) = pi((s, s_1, s_2 \dots s_{k-1}))$.

3. **Bellman Optimality Equation**

1. Reward function R(s)

$$V^*(s) = R(s) + \beta \underbrace{max}_{a} \sum_{s'} T(s,a,s')V^*(s')$$

2. State-Action Reward function F(s, a)

$$V^*(s) = \underbrace{max}_{a} R(s,a) + \beta \sum_{s'} T(s,a,s')V^*(s')$$

3. State-Action-State Reward function R(s, a, s')

$$V^*(s) = \underbrace{max}_{a} \sum_{s'} T(s,a,s')R(s,a,s') + \beta V^*(s')$$

4. Single Policy for MDP that takes action a in both states.

a. Linear Equations

$$V_0 = R(s_0) + \beta V_1 = \beta V_1$$

$$V_1 = R(s_1) + \beta V_1 = 1 + \beta V_1$$

When $\beta = 1$, we get $V_0 = V_1$ and $V_1 = 1 + V_1$

The results show that this equation has no solution which means the value of β does not work.

b. Linear Equations with discount factor

When $\beta = 0.9$, we get $V_0 = 0.9V_1$ and $V_1 = 1 + 0.9V_1$

Then, we can solve the equation and get $V_0 = 9$ and $V_1 = 10$.

5. **Ensemble Methods in Machine Learning.**

The article points out that ensembles are often much more accurate than the individual classifiers that make them up. To make ensemble algorithm successful, diverse classifiers are necessary because the classification results are making by voting of classifiers. However, if any classifier’s accuracy is lower than 0.5, it will reduce the accuracy of ensembles. In addition, the reason why ensemble method is better than individual one is because it solve statistical, computational, and representational problems which are related to overfitting. To sum up, this method the probability of overfitting.

6. **Ten simple rules for responsible big data**

With the increasing development of big data and the dependence of human beings on big data, people are paying more and more attention to moral issues. Therefore, this article provides ten simple rules to solve complex ethical issues.

First of all, It is important to acknowledge that data are people and can do harm. Since the data comes from people's lives, it may reveal people's privacy or vices, and ultimately lead to harm to people. Second, people should recognize that privacy is more than a binary value. Researchers must obtain their information with the consent of the people, otherwise they will violate their privacy. Third, guard against the reidentification of your data. We should sort out the collected information. Only the necessary information is left to prevent others from using the data to trace back to the individual to undermine personal privacy. Fourth, practice ethical data sharing. In order to maximize human interests, researchers should selflessly share the results obtained in research so that human science and technology continue to progress. Fifth, consider the strengths and limitations of your data; big does not automatically mean better. Although a large amount of data is beneficial for our research and observation, big data does not mean blindly gathering all the information together. In the process of collecting data, we need to consider the role of information in different situations and distinguish them. Sixth, Debate the tough, ethical choices. Scientists debate the pros and cons of new technology through debate. Ultimately, make ethical decisions. Seventh, develop a code of conduct for your organization, research community, or industry, which means to develop a code of ethics to limit everyone's behavior. Eighth, Continuing the previous point, scientists should design data and systems to audit people's ethical behavior. Next, Responsible researchers should break away from traditional thinking. In addition to researching and analyzing data, we must also understand what kind of impact these data and analysis results will have on society. Finally, as times and technology advance, so do people's thinking. Some old codes of conduct may not apply in the future, so researchers should know under what circumstances to break the rules and welcome the new society.

7. **Hidden Technical Debt in Machine Learning Systems.**

With the development of machine learning technology, people find that it takes less time and money to develop a machine learning system, but it is very expensive and time-consuming to maintain the system, and a technical debt for machine learning is born. This article mentions several concepts related to technical debt. The first concept is that strong boundaries help create maintainable code, but there are three ways to erode the boundaries of machine learning systems: angles, correction cascades, and undeclared consumers. Second, reliance on debt complicates the code of the system. In machine learning, data dependence can also accumulate technical debt than code, and it is more difficult to detect than code dependence. Third, feedback loops mean that the increasing amount of data in the system over time has resulted in analytical technical debt. If the computer cannot update the training model frequently, the prediction results will no longer be accurate. Fourth, Machine learning has a feedback loop feature. Researchers must update the training model frequently to ensure the accuracy of the calculation results. Fifth, researchers should avoid refactoring the following two anti-patterns of machine learning: glue code, pipeline jungles, dead experimental code paths, abstraction debt, and common smells. Sixth, researchers should understand the configuration system for machine learning to avoid potential errors. Seventh, because the world is always changing, the results of machine learning should not be limited to past data, and training results should be changed as new data is added. Next, In addition to testing data and code debt, engineers also care about reproducible debt and process management debt. Finally, engineers not only need to belittle the debt, but also have to measure and pay off the debt to make the system sustainable.

8. **A rubric for ML production readiness and technical debt reduction.**

According to this article, the authors provide 28 tested methods to improve the debt of machine learning. These 28 methods include 7 basic structures, 7 models, and 7 training levels and 7 rules for machine learning. First of all, Seven models describe training, unit testing, pipeline architecture, model quality, model adaptation, Canary process testing, and fast and safe pre-recovery versions of machine learning. Second, The seven rules include the dependence of the data on each other, the unchanged data, the unchanged training results, the updated training model, the stability of the model, the training speed of the model, and the time to wait for results.

9. Bias in Data Analysis

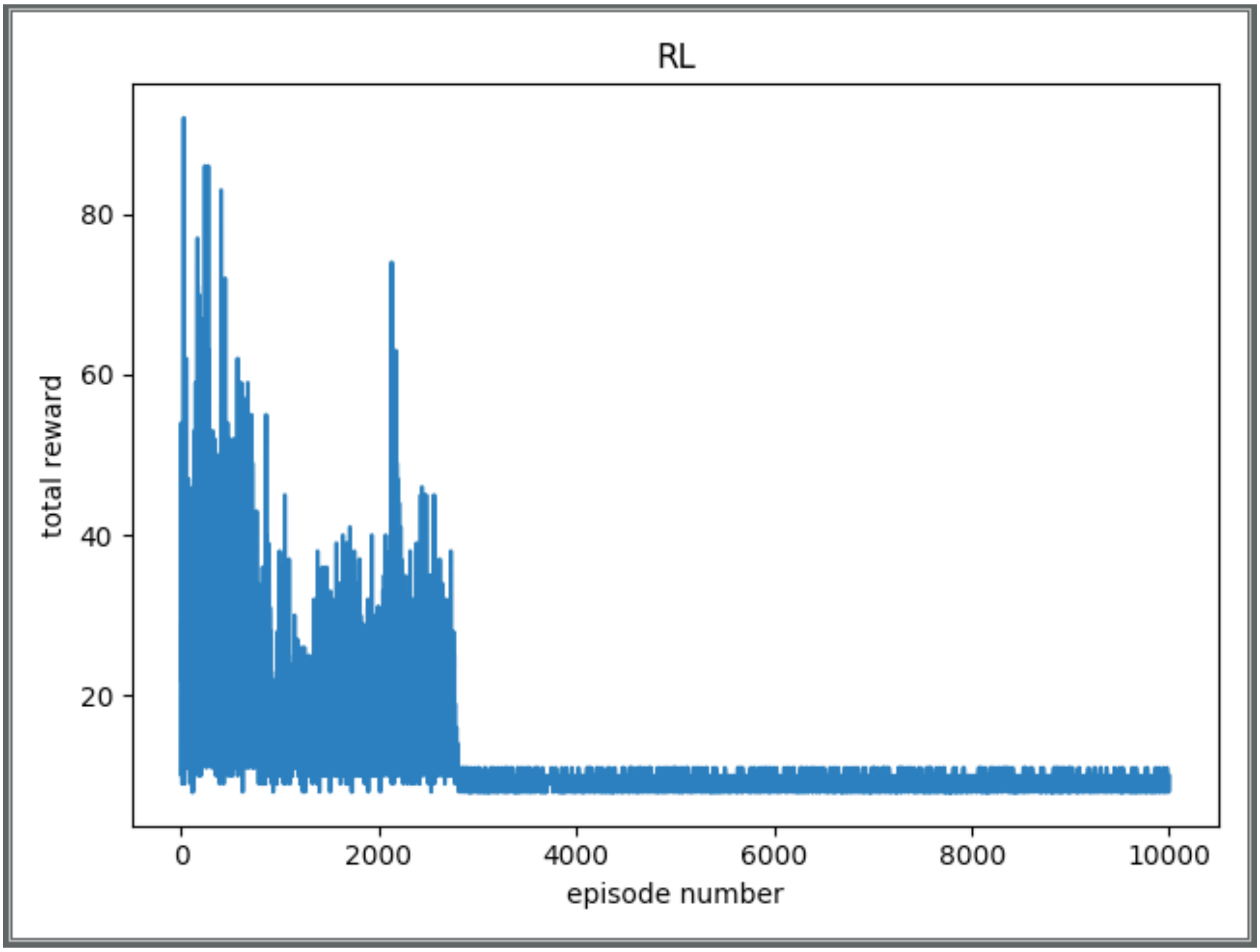
From this video we learned that some people think that machine learning will be biased in classification. The speaker explained what prejudice is. Prejudice actually comes from people for labeling data, so prejudice does not come from machines but humans. When tagging materials, people take into account the culture, religion, and politics of the time. For example, whether a person has committed a crime and the years of the sentence are determined by the judge as to what happened and the law at the time. However, our law has been amended, and the same crimes may have been punished differently ten years ago and years later. But the law is also modified by humans, so the real prejudice comes from humans. So, when we think that machine learning is biased, should we reflect that bias comes from ourselves? Machine learning is learned from past experience, so when the predicted result is biased, does it mean that people think about things differently than in the past? In conclusion, people's lives and society are constantly changing. Past concepts and ideas may not apply to the future, so people should not rely too much on machine learning.

10. **Societal impacts of AI.**

    The video begins by explaining how AlphaGo learns from past gaming experiences and ultimately defeats humans. In addition to this example, the speaker also provided five other examples of artificial intelligence: virtual payment methods in China, autonomous trucks, artificial intelligence replacing human work, analysis of people's lives to provide advertising, and China's surveillance system. From these examples we can find out how advanced the technology of artificial intelligence and how close it is to people's lives. However, these technology industries also bring convenience and inconvenience. For example, when we enjoy an ad tailored to us, we also feel that our privacy has been violated because everything we like is recorded. Another example is that when people find ways to overcome their labor problems, they also face unemployment.

2. Programming Part

1. It seems that running the iteration 1000 times is perfect. Running over 2000 times has overfitting problem.



2. Group discussion with classmate HungWei Lee.