Week 6

Evaluating a Learning Algorithm

Deciding What to Try Next

Source: Deciding What to Try Next

How to improve performance of a learning algorithm?

1. Get more training examples: For example, going to different houses and setting up more surveys to get more data on how much each house is sold for.

NOTE: A lot of people spend a lot of time collecting more and more data. It is a misconception that more the data collected means better is the learning algorithm. But this is not entirely true.

2. Try a smaller set of features: If the number of initial features are m where x_1, x_2, x_3,...x_m are the features then choosing a smaller subset of features after careful analysis can give n features where x_1, x_2, x_3,...x_n and m > n. This prevents overfitting.

3. Try a bigger set of features: If the number of initial features are m where x_1, x_2, x_3,...x_m are the features then finding more features after careful analysis can give n features where x_1, x_2, x_3,...x_n and m < n. This prevents underfitting.

NOTE: The idea of increasing the number of features can scale up a project which can then cause a lot of time to be spent to collect data for each and every feature irrespective of the knowledge of whether the feature is going to be useful or not. Imagine collecting information like area of land, pieces of land, phone number and so on which might not turn out to be related.

4. Adding polynomial features: For example, adding a feature, x_1 to be x_1^2 and x_2 feature to be x_2^3 and so on.

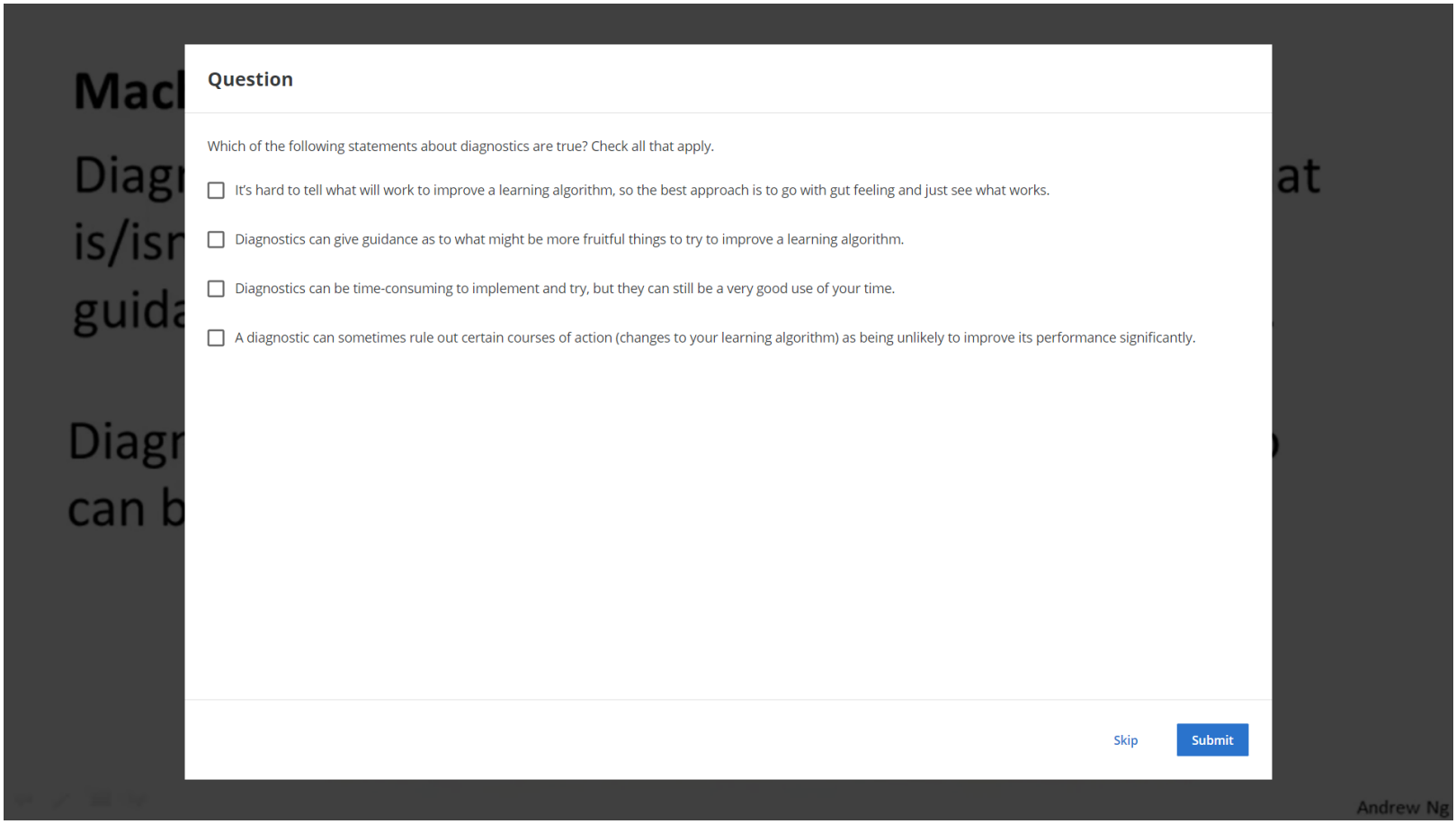
5. Changing the regularization parameter, λ: If the regularization value is decreased or increased, the performance can change.

Most of these options can scale up and take a lot of time to improve the performance of the learning algorithm. Sometimes one of the mentioned options is chosen randomly and this mostly results in unfortunate results.

What is machine learning diagnostics?

Machine learning diagnostics is to rule out the options and find the best possible options to improve the performance of the algorithm.

NOTE: The machine learning diagnostic tests can take a lot of time to implement but can eventually save a lot of time from irrelevant steps to improve the performance of the machine learning algorithm.



Question

Which of the following statements about diagnostics are true? Check all that apply.

- [ ] It's hard to tell what will work to improve a learning algorithm, so the best approach is to go with gut feeling and just see what works.
- [ ] Diagnostics can give guidance as to what might be more fruitful things to try to improve a learning algorithm.
- [ ] Diagnostics can be time-consuming to implement and try, but they can still be a very good use of your time.
- [ ] A diagnostic can sometimes rule out certain courses of action (changes to your learning algorithm) as being unlikely to improve its performance significantly.

Skip Submit

Question

Which of the following statements about diagnostics are true? Check all that apply.

- [ ] It's hard to tell what will work to improve a learning algorithm, so the best approach is to go with gut feeling and just see what works.
- [x] Diagnostics can give guidance as to what might be more fruitful things to try to improve a learning algorithm.

✓ **Correct**

- [x] Diagnostics can be time-consuming to implement and try, but they can still be a very good use of your time.

✓ **Correct**

- [x] A diagnostic can sometimes rule out certain courses of action (changes to your learning algorithm) as being unlikely to improve its performance significantly.

✓ **Correct**

Skip Continue

Andrew Ng