Reading 3 : AI ML Challenges

Name: Pushpak Vijay Katkhede **Date: 01/30/2023**

---

1. *Imagine that you (personally) have $10 to gamble in a horse race, and you have a model that predicts which horse will win the race. What minimum level of accuracy (from 0 to 100%) would **you** require before you would follow the model's predictions, and why?*

Ans: If I was to gamble the amount, I would follow the model prediction if its accuracy is at least above 50 percentage. The reason is that if I follow my instinct the probability of success is ½ given that the race is fair. So, accuracy of model above 50 percentage will have more chance of winning in the race or else I would prefer to follow my instincts.

2. *If you are currently working on a research task, (a) What part of Figure 1 have you found most time-consuming so far? (b) What "impact" metric(s) would be appropriate for your task? If you are not currently working on a research task, imagine developing a voice recognition system for computer login (if the speaker is recognized, the computer logs them in; otherwise, they must type a password.) (a) What part of Figure 1 do you expect would take the most time to develop this system? (b) What "impact" metric(s) would be appropriate for this task?*

Ans : For the voice recognition system, if we are not using any benchmark dataset the most tedious task would be collect the dataset and then preparing the dataset for further processing. As the part of the login process the authentication would be important to recognize the voice accurately, hence, accuracy would be the most important impact metric. Also, as time is also an important concern as they authentication should not be very time consuming. At the time of authentication, the noise in the input voice could also be considered, which needs to be addressed as well.

3. *What is one **question** you have after finishing the reading? (What wasn't clear? What needs more investigation?)*

Ans: I agree that there is high disparity between the actual metrics under consideration for assessing ML solutions and their actual impact on the domain for which the solution has been worked out. So, my question basically is there any pipeline or approach that you have in mind to deal with this? I believe an initial education introduced in standard operating procedures for any research project would help suffice to solve this problem.

4. ***Reflect on your work for the course over the past week.*** *What did you do that was effective and increased your knowledge? What could you do or change to increase what you gain from this course? Is there anything about this course you are anxious / worried about? (There are no wrong answers here; this is your chance to maximize what you get out of the course and to let me know about any concerns.)*

Ans: I have looked over multiple datasets for the project in the past week and performed preliminary analysis on them. This exposed me to diverse types of data from varied domains. I would be more focused on the dataset that I choose in this week and would be implementing the learned concepts in lectures on the fly over that dataset. I'm currently not worried about anything in this course.