Machine Learning
3. Teaching ML

Ekpe Okorafor ekpe@okorafor.com

Advice on Teaching Machine Learning

- Assumptions
    - Had to determine all the variables to give tips but the following are some prerequisites
    - Students have a background in Python or R. But preferably Python
    - Machine learning and AI are built on mathematical principles like Calculus, Linear Algebra, Probability, Statistics, and Optimization; and many would-be AI practitioners find this daunting.
    - If there's one universal course for Machine Learning, it must be Andrew Ng's course.
- Tips & Considerations
    - Check if the assumptions hold true
    - If it is a varied class with diverse skills, prior knowledge and backgrounds, use your judgement
    - Adjust to the required length of the course – 1 day, 2 days, or 1-week etc
    - Successful practical sessions are much more valuable than full on lectures!

Focus on the lab exercises

- A successful course is one where the students go through the exercises and finish
- This means, ensure you have a plan to un-block the students who are stuck. Keeping them moving on is key
  - If possible, have them pair up. The interactions help
  - They can also succeed by applying divide-and-conquer to complete the exercises
- Minimize the lecture, instead provide exercises that illustrate the ML concepts
- Give enough time for the “building recommender system” exercises
  - If you can’t do both collaborative filtering approaches, don’t sweat it. Do one!
  - User-based or content-based – to produce recommendations

Key things to focus on

1. Difference between supervised and unsupervised learning
2. Collaborative filtering-based Recommendation
3. The concept of Similarity in making recommendations
4. An appropriate dataset like MovieLens
5. Exploring the dataset is crucial
6. Have them create plots or visualizations of the data set – very engaging!
7. The different similarity algorithms
8. Validating and Interpreting the recommendation

How best should they do the exercises?

- The question around should they use R or Python
  - I suggest Python but we have done this course with R
  - Python is more widely used in ML and big data environments
  - R is popular in academia and research
- Should they write their code themselves or use packages
  - Depends on how much time, expected performance and students' programming background
  - The first section requires the students to write their own code (or complete the code)
  - Packages are the introduced later. At this time, they should understand what they are doing

Final Thoughts

1. I am sure you have a good sense of what ML is by now. You can discuss the basic concept. Recommender system is simply a way for us to apply those concepts on a dataset

2. There are so many cool AI/ML tools to leverage. You can also take a free coursera course to brush up if you don't feel confident. But really, you can do it.

3. The key is to have them successfully do simple incremental exercises that result in some recommendation!

Data Schools