Source:

#ret

0.1 | **Prompt:**

For each of the scenarios below, answer the following questions. You do not have to explain your answers

1. What type of machine learning problem (regression, classification, clustering) do you think this is?
2. If this is a supervised problem, what will you use as your targets (aka labels) and how will you get
3. What processing do you need to do to your input data? (How will you handle non-numerical inputs? Do
4. What type(s) of model(s) would you try? Remember to start with the simplest thing that might work! TI
5. What validation metric(s) would you use to decide how well you're doing?
6. What ethical challenges do the data collection, creation, and/or use of this model create? If you fee

Scenarios:

1. You are playing fantasy football and want to predict how many points each player will score next seas
2. You have customer reviews, each one of which has a rating from 1 (worst) to 10 (best) and some text.
3. You have data from a movie streaming service that consists of lists of movies that each user has wat
4. You want to predict whether a random stranger owns a cat, a dog, or neither, based on things that the
5. You want a model to predict the number of deer that will be born in a breeding season. You have a lai
        number of fawns born
        the genus and species
        number of does sighted during the mating season
        vegetation quality during the mating season ("low", "average", or "high")

0.2 | **Scenarios:**

- Football

    1. Note: Since we only have one season of point values, and hence cannot see cross season
       change in point values, the old season players will be used as training data.
    2. Regression
    3. Label: Point value
    4. One Hot Encoding, 0-1 normalization
    5. Linear Regression or Neural Networks
    6. RMSE
    7. None

- Customer Reviews

    1. Classification
    2. Positive, Negative, Neutral
    3. Some form of text processing – BOW, TFIDF, word vectors, ect.
    4. Out of the models we have learned, Naive Bayes.
    5. F score
    6. Could misrepresent reviews and allow for automation

- Movie Recommendations

    1. Classification
    2. Semi-supervised.

3. One Hot Encoding and BOW or TFIDF
4. Random forest? NN?
5. Click rate or watch time, depending on goal.
6. Ethics crumble in the face of capitalism. We gotta get our clients the right recommendations!

- Facebook Pet

  1. Classification
  2. Supervised: Dog, Cat, Neither
  3. WPIE type system.
  4. NN
  5. F score

  6. Collection of likes is invasive. Categorizing of people is also problematic, as is trying to determine private info.

- Deer Born

  1. Regression
  2. Supervised: Number of deer
  3. One Hot
  4. Linear regression, NN
  5. RMSE
  6. None

0.3 | Comment Response

- 2b. How do you convert from the 1-10 rating scale to positive/negative/neutral?

  - You would have to pick ranges that define what is positive/negative/neutral

- 2c. Given that the reviews vary greatly in their length, is one of these preferred over the other?

  - BOW doesn't work well when length varies. TFIDF, however, does.

- 2e. Why not accuracy, precision, or recall?

  - Those would all work as well. I just decided to list one possible validation metric as opposed to all of them.

- 3a. This could work, but I think you will find this problem more straightforward as a different kind of problem.

  - I guess this problem could be done as a clustering problem, representing each movie as a location in a multidimensional space then placing users in the space and clustering.

- 3b. Semi-supervised usually refers to models where we have some labels, and we generate additional labels. Where do our labels come from? How do we generate the additional ones?

  - The original labels would most likely be generated. As time progresses, we get labels back from the users that are interacting with our model's recommendations.

- 3c. Be more specific: which features would you use which techniques on? For example, you could use bag of words or OHE on the names of the stars, but you'd get pretty different results depending on which one you picked.

  - OHE: name of stars, genre. BOW or TFID: title.

- 3e. A user not watching a recommendation is not necessarily good signal on whether it's a good recommendation (i.e. I might really like a movie, and just not have time to watch it right now). Conversely, a user watching a movie is not necessarily signal that it was a good recommendation (maybe I watched it because you recommended it, but I hated it).

  - This depends on the goal of the recommendation. If you get a good recommendation and don't watch it, was it really a good recommendation? If the goal is simply to increase watch time, and thus increase the number of ads viewed, then yes watching or not watching is a good metric. Think click-bait. Of course, this has to be weighed against the long term effects of the recommendation.

- 3f. I know you are kidding (I hope so, at least!), but the whole point of us learning about ethics is precisely so that they do not crumble in the face of capitalism!

  - I think maybe we should shift the goal to rebuilding the already crumbled ethics...

- 4b. Where do these labels come from?

  - From the friends you asked.

- 4c/d. Why WPIE/NN? Is there a simpler approach we might try first?

  - BOW, TFIDF, or word vectors would work, but Facebook's method itself would most likely work better. As for the type of model, Decision Trees and Random Forests would work.

- 4e. Why not accuracy, precision, or recall?

  - Again, just decided to list one validation metric. Accuracy, precision, and recall would also work.

[[**KBdataProsResponses**]]