

Assignment_1_MAT3373

John Bute

2024-09-19

Question1

1. This is supervised learning as we have a database of similar data on prior customers, including whether they have been accepted a loan or not, thus we have an idea of what the correct output values should be.
2. This is unsupervised learning, as we don't know have data that tells us that Customer X will buy book B with A. Instead, we just have a list of prior transactions, and analyzing them together to find a pattern to hopefully make a recommendation (with clustering or arm).
3. This is supervised learning, as we know the characteristics of other packets with a known threat status, thus we have a labelled output.
4. This is unsupervised learning, as we are looking for a correlation between certain customers in order to create segments from them.
5. This is supervised learning, as we have labelled data telling us firms that are bankrupt and non-bankrupt firms. Thus, we expect a certain correct output when testing our supervised learning model.
6. Supervised learning as we give our model certain inputs, which are paired with expected outputs (repair time).
7. This is an example of a classification model where we are scanning zipcodes and sorting them in the correct bin. Therefore we have a preset destination for each zipcode and therefore an expected outcome for each time we sort a zipcode.
8. This is supervised learning, as we have data regarding customers and if they left or not. Thus, we can train a model to score a likelihood of if the customer will leave or not based on the historical data that we have of them.
9. We are in an unsupervised learning case, as we are simply detecting anomalies within the patterns inside the audit data, then identifying those data points which represent cases of unnecessary testing and overbilling.
10. This is an example of association rules mining, which is an unsupervised learning model based on finding causal relations between products that are bought together. (We are trying to find a pattern in the data)
11. This is supervised learning, as you have historical medical data, which is labelled data, and a predictive modeling through decision trees, which means that we will need test data (which we have) in order to determine the accuracy of our model, therefore we have expected outputs for certain inputs
12. This is supervised learning, as we have attributes of students who do not return to school, and we can train a model to determine the probability of students with some of these attributes of leaving school, thus leaving us with labelled data and therefore supervised learning.

13. Unsupervised learning, as we do not have a labeled outcome, we merely just profile customers and group them together based on their CDRs.
14. We have access to historical data of machines that failed, which is labelled data. Therefore we have supervised learning.
15. This is supervised learning, as in order to build a model that distinguishes legitimate tweets from fake ones, we need to have access to a labelled dataset with tweets and whether they are fake or not.