

## Week 4 Application Assignment

Quiz, 5 questions

✓ **Congratulations! You passed!**

Next Item



1 / 1  
point

1.  
Let's once again consider the customer reward program dataset. For your convenience, here the original data set.

data-download-wk4.xlsx

In this exercise, we use machine learning methods including trees and neural networks. Set all random seeds to 12345 in XLMiner; this should be the default value if you have not changed them. Also use default options in XLMiner unless directed otherwise.

**1.** Build a full classification tree model with Reward column as the target variable and the following set of predictor variables: Salerank, X2013USSales, X2013WorldSales, ProfitMargin, NumStores, and IndustryType. Which feature has the highest feature importance?

[Hint: Use Classify in the Data Mining ribbon since Reward is a 0-1 variable.]

- ☐ X2013USSales
- ☐ IndustryType
- ☐ ProfitMargin
- ☒ NumStores

**Correct**

Bravo

- ☐ X2013WorldSales
- ☐ Salerank



0 / 1  
point

2.

Split the dataset into training and validation sets using a 60:40 split. Use the same columns as in the previous question.

## Week 4 Application Assignment

What is the precision for on the validation data for bagged tree (round the answer to 4 decimal places, i.e. x.xxxx)?

Enter answer here

### Incorrect Response

The answer you gave is not a number.



1 / 1  
point

3.  
What is the precision on the validation data for boosted tree (round the answer to 4 decimal places, i.e. x.xxxx)?

0.5517

### Correct Response

Bravo!



1 / 1  
point

4.  
Build a neural network model with RewardSize column as the target variable and the following set of predictor variables: Salerank, X2013USSales, X2013WorldSales, NumStores, and ProfitMargin. Note that RewardSize variable is only relevant for retailers that use reward programs. For your convenience, here is the data file with values for RewardSize, which should be used for model building.

data-file-NN-question.xlsx

Split the dataset into training and validation set using a 60:40 split. Also remember to scale the data. Report RMSE on the validation data for the following models: (i) bagged neural net, (ii) boosted neural net. [Hint: Use Predict in the Data Mining ribbon since RewardSize is a continuous variable.]

Report RMSE on the validation data for the bagged neural net (round the answer to 4 decimal places, i.e. x.xxxx)?

8.9210

### Correct Response

Bravo!

---

## Week 4 Application Assignment

Quiz, 5 questions / 1



point

5.

Report RMSE on the validation data for the boosted neural net (round the answer to 4 decimal places, **i.e. x.xxxx**):

18.1487

**Correct Response**

Bravo!

---

