

Workshop Session

Mentor Name: Siviwe

Mentor - Welcome/Housekeeping

- ☐ The use of disrespectful language is prohibited during the session. This is a supportive, learning environment, please engage accordingly. (FBV: **Mutual Respect.**)
- □ No question is daft or silly ask away!
- For all non-academic questions/support, please submit a query: www.hyperiondev.com/support
- Report a safeguarding incident: hyperiondev.com/safeguardreporting
- ☐ We would love your feedback on workshop session, please complete the form that is sent after the session.

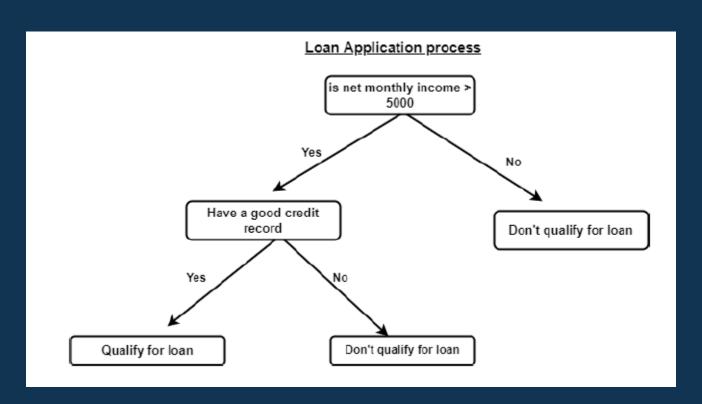
Session Objectives

Today we will revise decision trees.

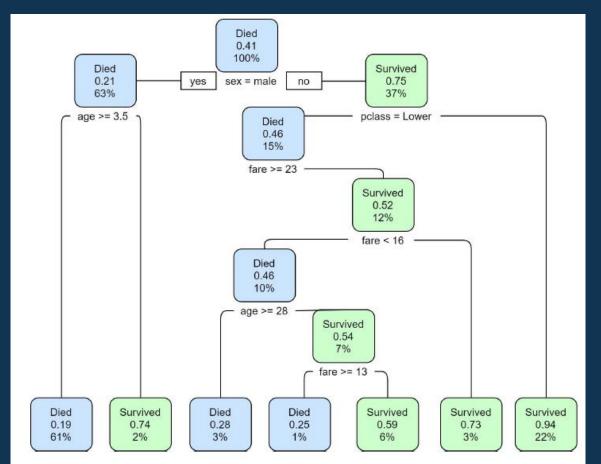
What are decision trees?

Decision trees work by formulating simple rules that partition data into ever-smaller regions. Each partitioning is like a fork in the road, where a decision must be made. The decision is made based on rules which are derived from previous experiences.

Decision Making

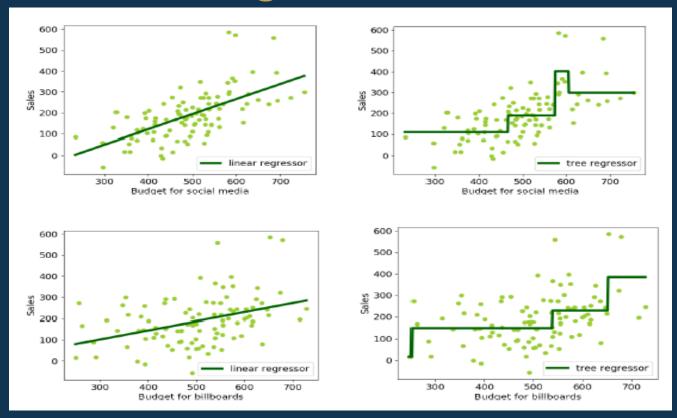


Classification Trees

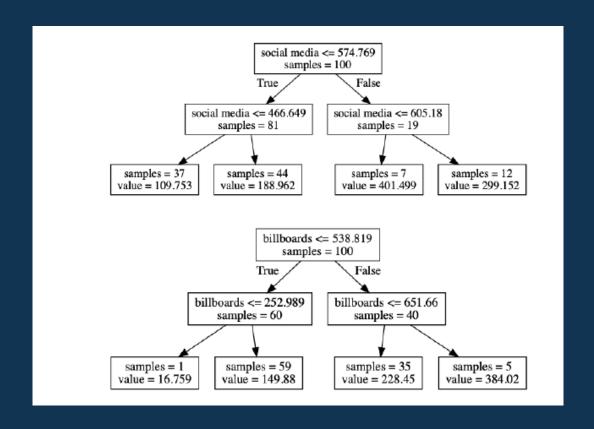


Hyperiondev

Regression Trees



Interpreting Decision Trees



Overfitting \ Underfitting



Fixing Overfitting and Underfitting

- Pruning
- Development set

Ensemble Methods

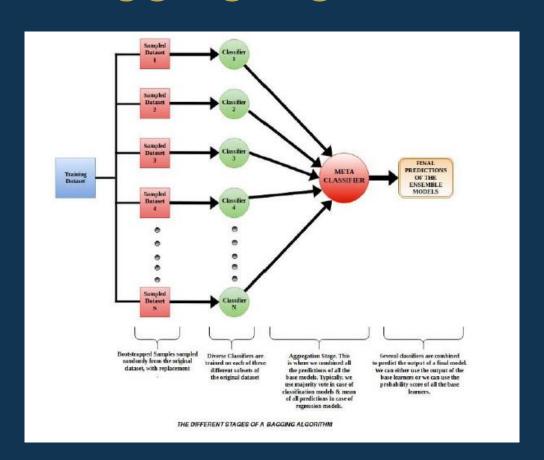
Why the need for ensemble methods? Types of ensemble methods:

- Reduce variance
- Improve accuracy
- Capture complex patterns

Bootstrapping

```
def subsample(dataset, ratio=0.7):
    sample = list()
    n_sample = round(len(dataset) * ratio)
    while len(sample) < n_sample:
        index = randrange(len(dataset))
        sample.append(dataset[index])
    return sample</pre>
```

Bagging Algorithm



Random Forests

- Decorrelating tree classifiers
- Randomizing the splitting variable

Boosting

Boosting attempts to incorporate error(residuals)
of earlier models into the data provided to later
models.



Hyperiondev

Thank you for joining us

- . Take regular breaks
- 2. Stay hydrated
- 3. Avoid prolonged screen time
- 4. Practice good posture
- 5. Get regular exercise