NYR CONFERENCE - MAY 2024

ENGAGING STAKEHOLDERS THROUGH DATA SCIENCE

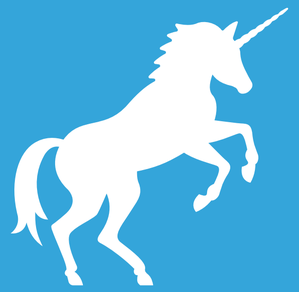
NOT YOUR COLLEGE STATS COURSE

MEGAN ROBERTSON

OUTLINE

1. Why do we care?
2. Challenges
3. How to Engage Stakeholders

WHY DO WE CARE?



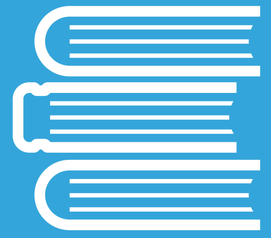
Data science is not magic

Develop stakeholder relationships

CHALLENGES

Stakeholder fear or apprehension

Where does their unique knowledge fit into the analysis?



Information overload

How can you adjust the material based on the audience?

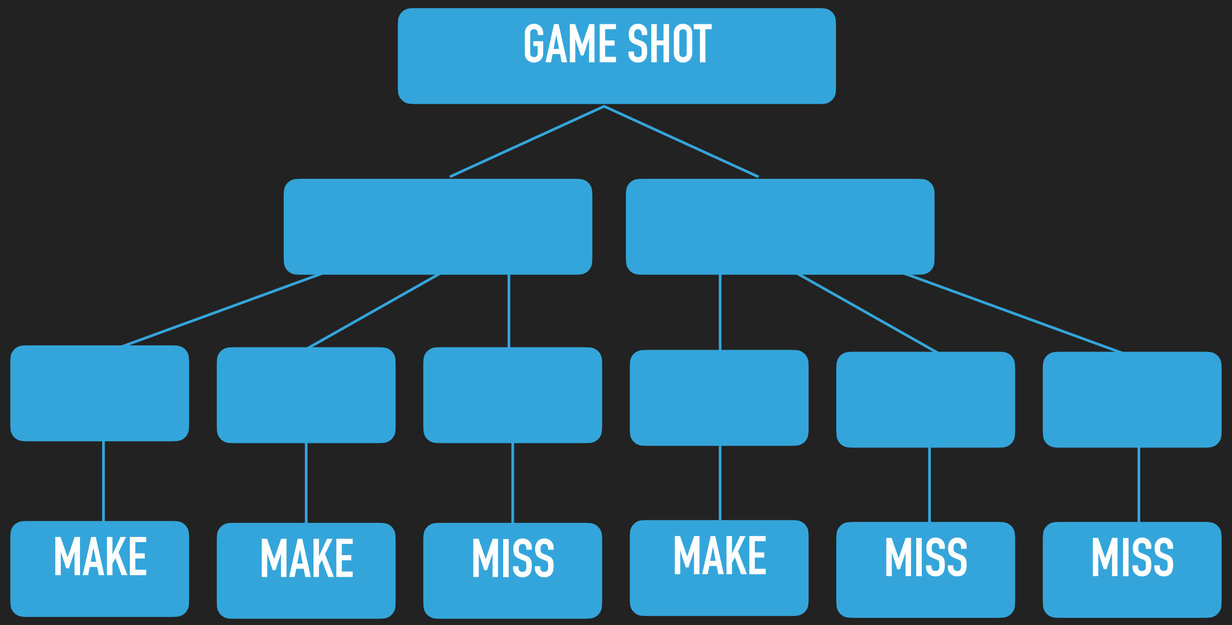
HOW DO WE DO THIS ?

DOS AND DONTS

- Do relate to known concepts
- Do use visual aids

- Don’t use overly complex language
- Don’t be condescending

OUR USE CASE



Shooter velocity
Distance to closest defender
Distance to closest teammate
Time on shot clock
Angle with defender
Court location
Distance from basket

HOW DO WE EXPLAIN THE MODEL?

1. Identify the key components
   - What are the main ideas of this analysis?
   - When do you apply this technique?
   - How would I label this model with keywords?

HOW DO WE EXPLAIN THE MODEL?

1. Identify the key components
   - What are the main ideas of this analysis?
     - *Combine multiple decision trees, learns rules from data*
   - When do you apply this technique?
     - *Predictive modeling*
   - How would I label this model with keywords?
     - *Prediction, ensemble, bagging, decision tree, variance, random sample*

HOW DO WE EXPLAIN THE MODEL?

2. Share limitations

   - What assumptions were made?

   - What risks are there from incorrectly applying this analysis?

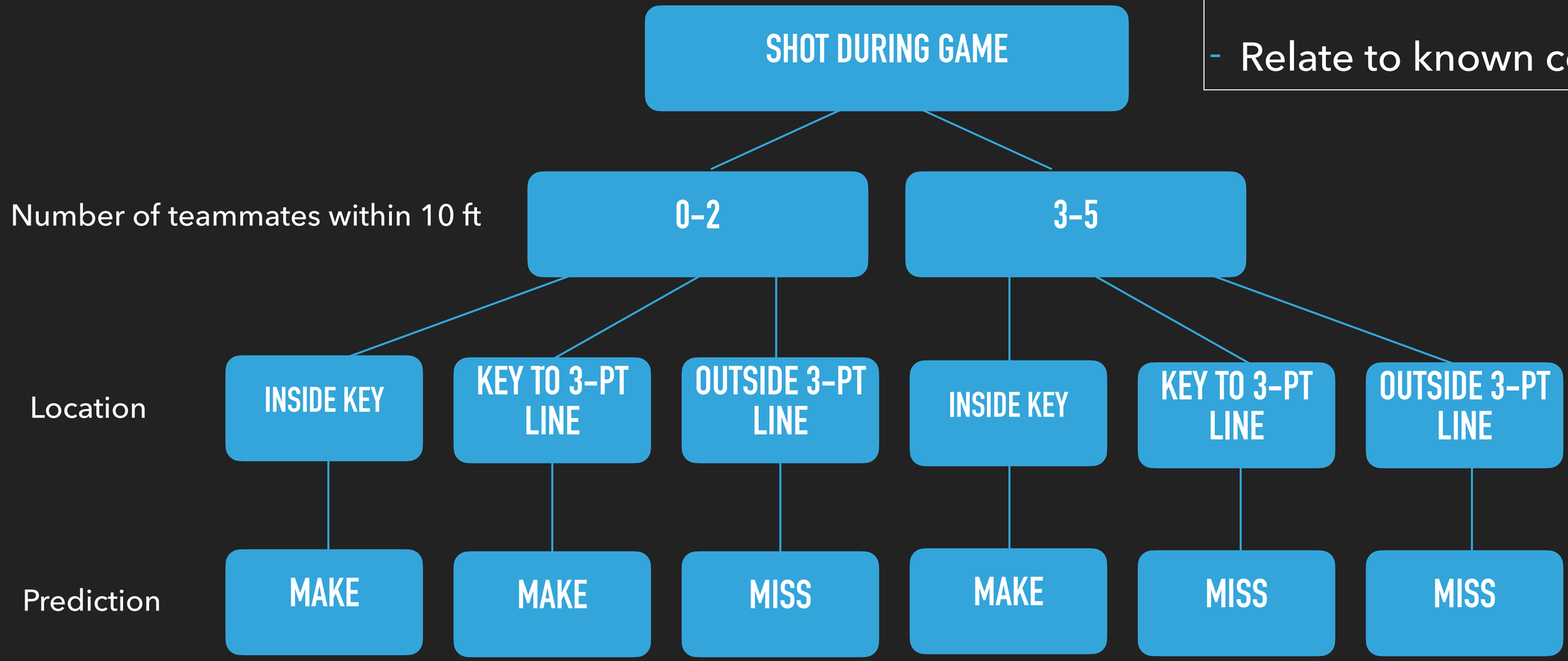
HOW DO WE EXPLAIN THE MODEL?

2. Share limitations

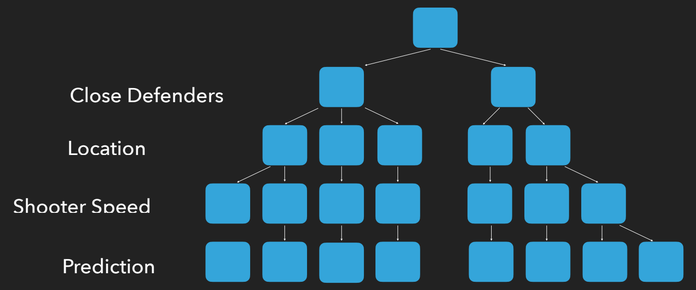
- What assumptions were made?
  - *Analysis is only as good as our data, applies to our team*
  - *Representative sampling*
- What risks are there from incorrectly applying this analysis?
  - *Might not create optimized practice/game strategies*

TREE MODEL

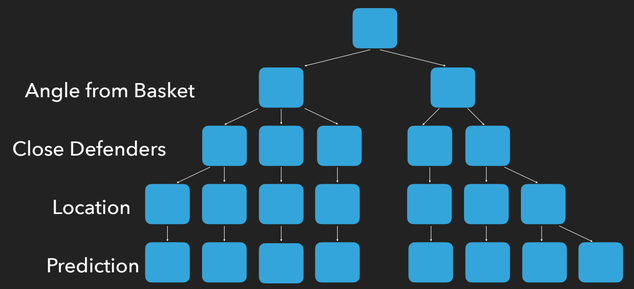
- Simple visualization
- Relate to known concepts



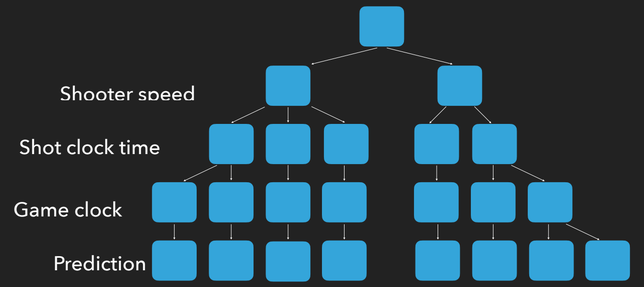
COMBINING TREES



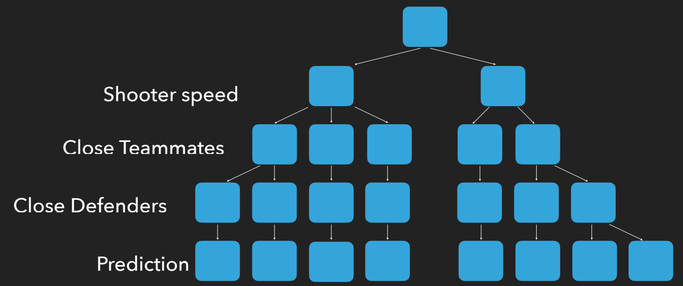
Make



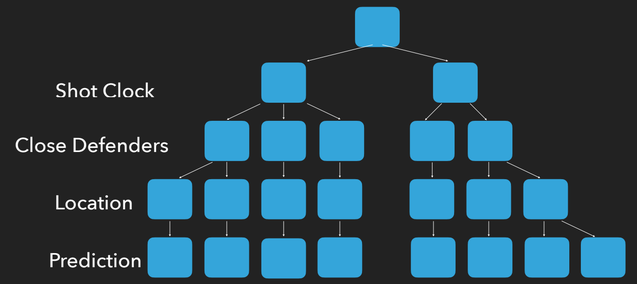
Make



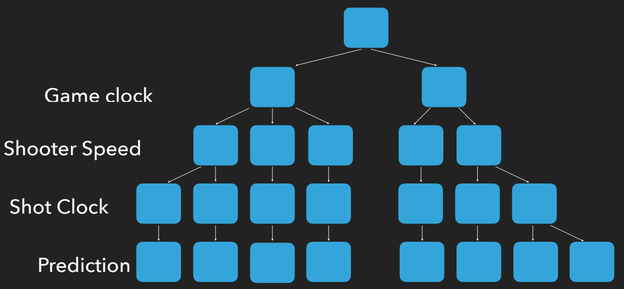
Miss



Make



Miss



Make

Majority of trees say *Make* so we predict *Make*

CONCLUDING REMARKS

- Data science analysis should not be a black box and can be accessible
- Presenting results is a balance between providing appropriate details for the audience and not overwhelming them
- Communication is an often overlooked skill when hiring and training data scientists

MEGAN ROBERTSON
SENIOR DATA SCIENCE PROFESSIONAL

- megan.alice.robertson@gmail.com
- megrobertson.weebly.com