

ASK THE EXPERT

Tree-Based Machine Learning Methods and Model Interpretability in SAS[®] Viya[®] – A Case Study

Christa Cody, Ph.D.

Ari Zitin





Christa Cody, Ph.D.

Data Scientist – Machine learning/Education Research

Christa Cody has a PhD in computer science with a focus on educational tech and machine learning. As a SAS data scientist, she uses data sets, feature engineering and visualizations to provide a variety of stakeholders within SAS insights into the customer's educational experience. She also works closely with SAS Education Marketing to conduct analytical experiments to solve problems and help users get the most benefit out of their educational content.



Ari Zitin

Analytics Instructor – Machine Learning, Signal Processing, and Optimization

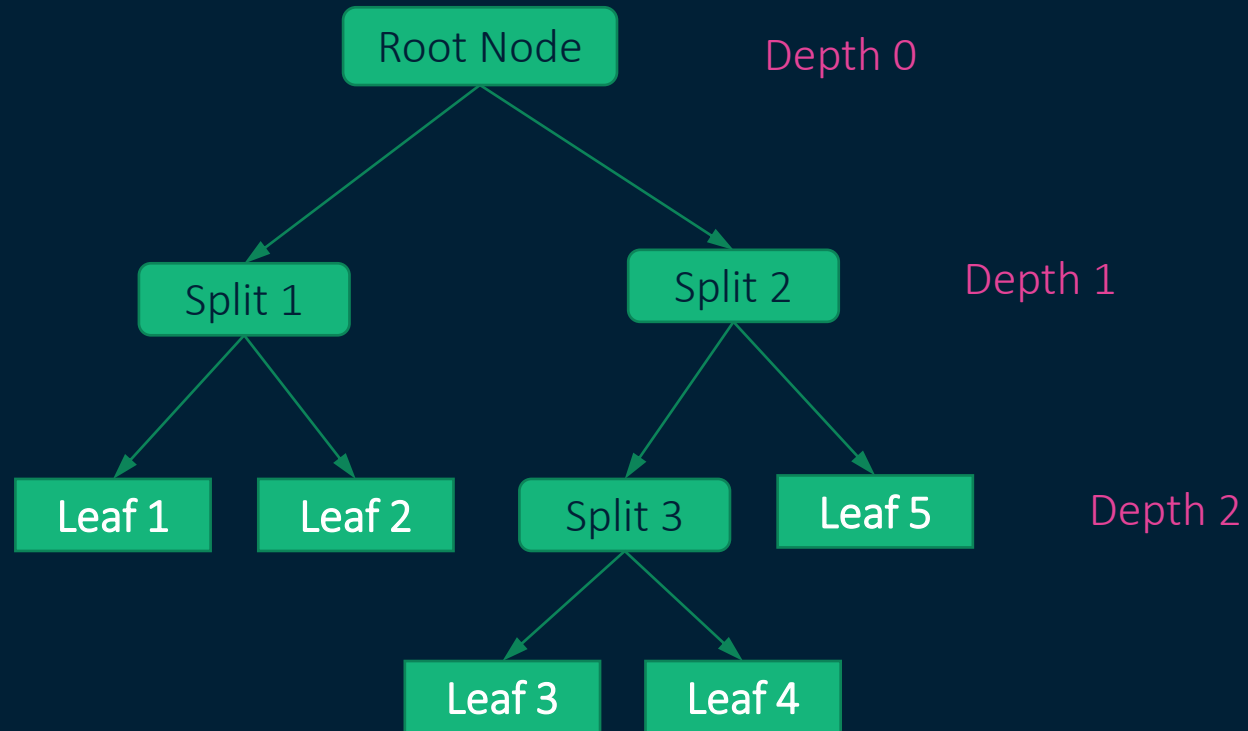
Ari Zitin holds bachelor's degrees in both physics and mathematics from UNC-Chapel Hill. His research focused on collecting and analyzing low-energy physics data to better understand the neutrino. Zitin taught introductory and advanced physics and scientific programming courses at UC-Berkeley while working on a master's in physics with a focus on nonlinear dynamics. While at SAS, he has worked to develop courses that teach how to use Python code to control SAS analytical procedures.

Agenda

- Tree-based methods + discussion
- Demo
 - Case study using banking data
 - ML model in Model Studio
 - Model interpretability plots in Model Studio

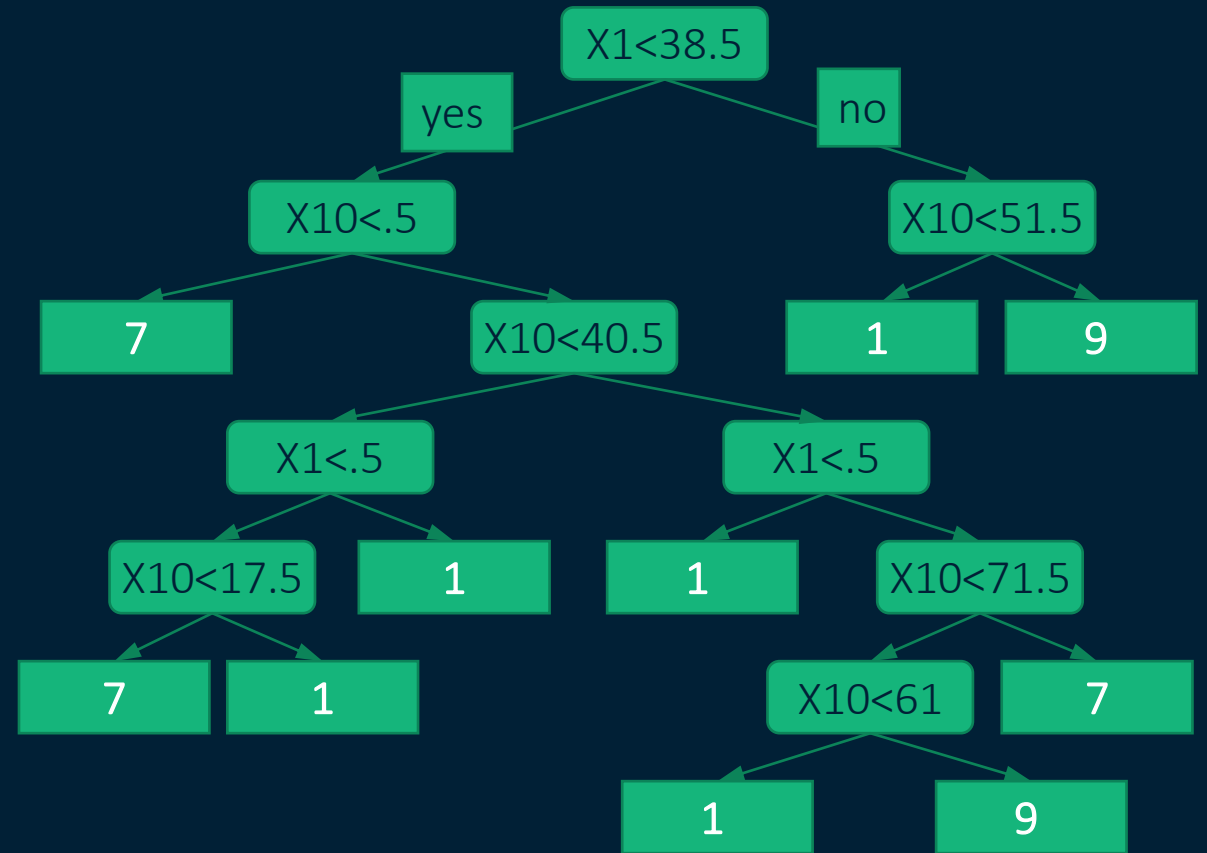
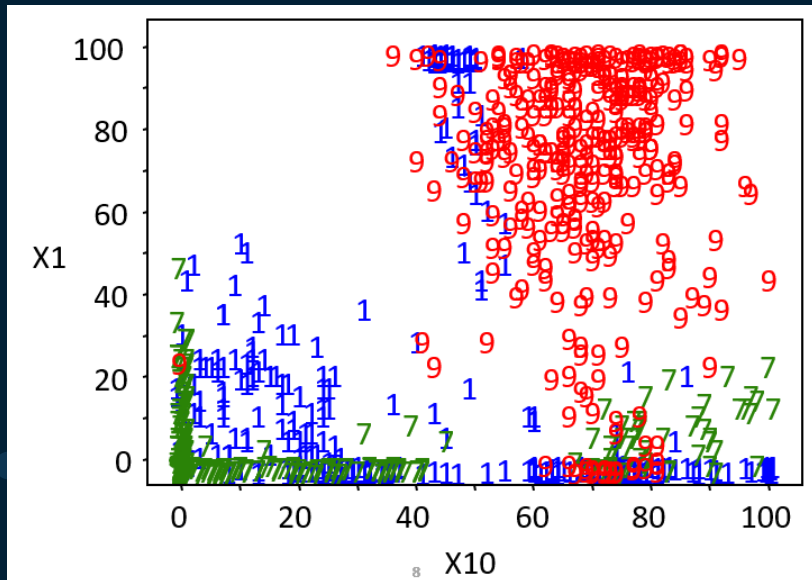
Tree-based Methods

Decision Tree



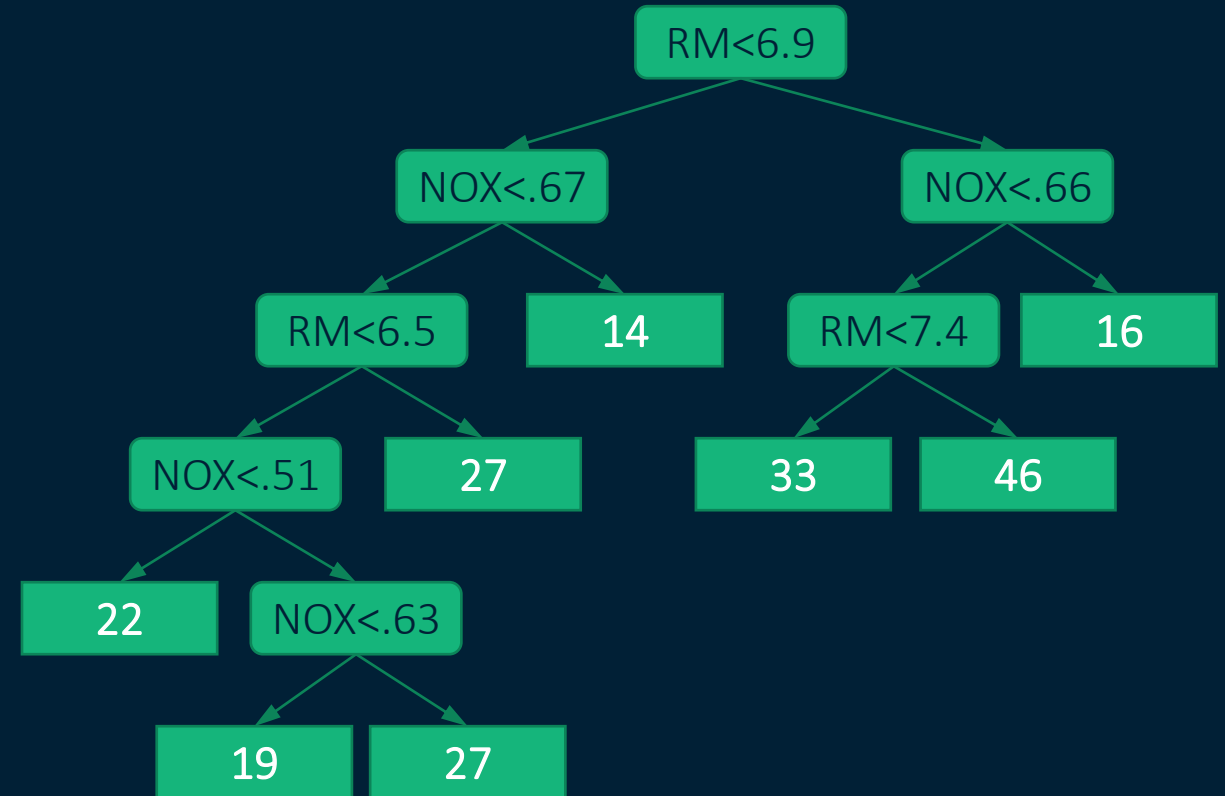
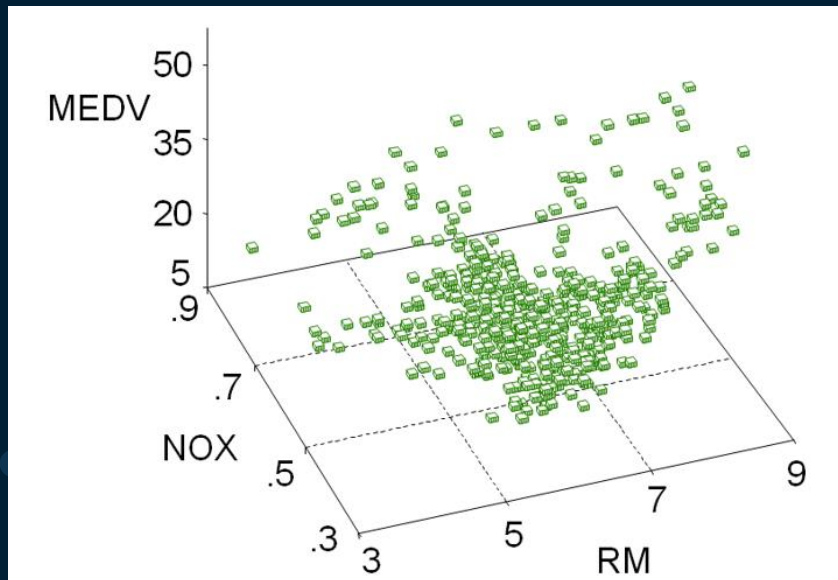
Classification Tree

- Inputs: pressure points on notepad (X1, X10)
- Target: Digits (1, 7, 9)



Regression Tree

- Inputs: nitrogen oxide gas (NOX), number of rooms (RM)
- Target: Median home value



Parameters + Hyperparameters

Tree structure

- Maximum depth
- Minimum leaf size



Parameters + Hyperparameters

Tree structure

- Maximum depth
- Minimum leaf size

Tree learning

- Grow criterion
- Pruning method



Grow

Prune

Final Model

Advantages & Disadvantages

Decision Tree

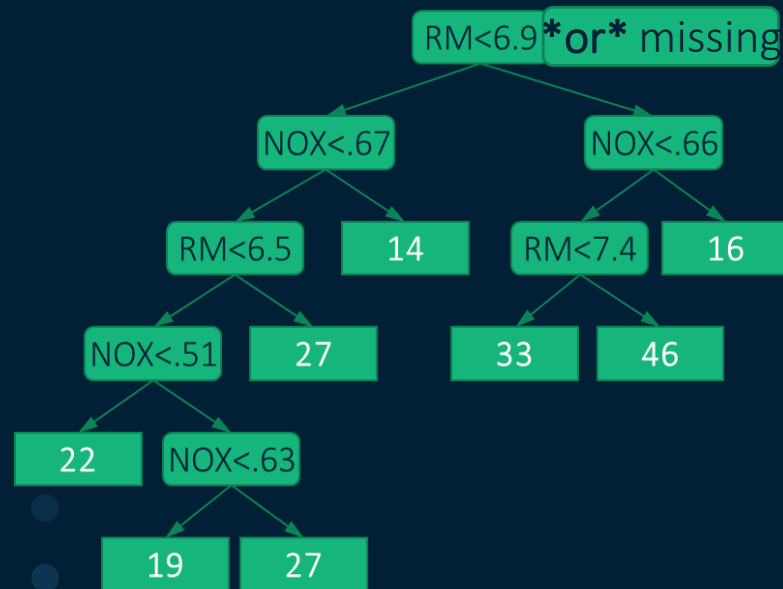
Advantages

- Interpretable
- Less data preparation

Disadvantages

- Unstable

...but this can be utilized!



Demo

Case Study: Tree-based Methods on Banking data

Decision Trees + Forests

Why worry about Model Interpretability?

#1 Need to explain predictions

*Misconception: only needed when
you need to explain predictions*



Why worry about Model Interpretability?

Understanding **why** your model makes predictions



Increased...

- Understanding of **data**
- Understanding of **model**
- **Trust**

Model Interpretability Techniques

- Global interpretability
 - Variable Importance
 - Partial Dependence Plots (PD)
- Local Interpretability
 - Individual Conditional Expectation (ICE)
 - Local Interpretable Model-Agnostic Explanations (LIME)
 - Kernel SHAP

Demo

Case Study: Model Interpretability

Explore Helpful Resources

[Ask the Expert](#)

View other user webinars that provide insights into using SAS products to make your job easier.

[FREE Training](#)

Learn from home – free for 30 days. Get software labs to practice and online support if needed.

[SAS Support Communities](#)

Ask questions, get answers and share insights with SAS users.

[SAS Analytics Explorers](#)

An exclusive platform to collaborate, learn and share your expertise. Gain access to a diverse network to advance your career. Special rewards and recognition exclusively for SAS users.

[SAS Users YouTube Channel](#)

A plethora of videos on hundreds of topics, just for SAS users.

[Newsletters](#)

Get the latest SAS news plus tips, tricks and more.

[Users Groups](#)

Meet local SAS users, network and exchange ideas – virtually.

[SAS Profile](#)

If you haven't already done so, create your SAS Profile to access free training, SAS Support Communities, technical support, software downloads, newsletters and more.



Thank you
for joining us for
this SAS webinar

