

Using Visual Analytics to Interpret Predictive Machine Learning Models

Josua Krause, Adam Perer, **Enrico Bertini**



NYU

**TANDON SCHOOL
OF ENGINEERING**

NYU InfoVis Lab

We study and develop interactive visual interfaces to help people **understand discover** and **communicate** information out of complex, noisy, and often large data sets.



NYU

TANDON SCHOOL
OF ENGINEERING

Co-Authors

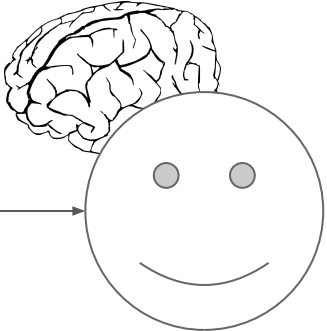
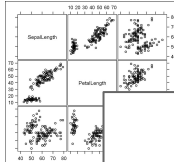
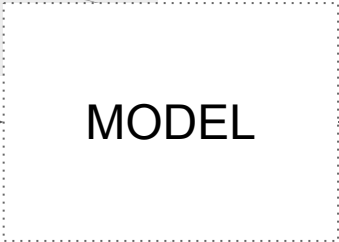
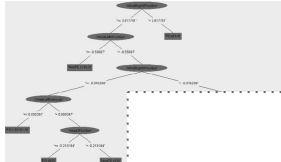
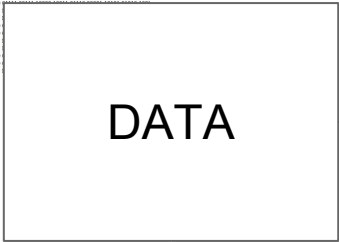
Josua Krause, NYU



Adam Perer, IBM



111 00001 00001 11111 0111 10000 10011 0111 00000 10001 10000 11111 00000 10101 0111 00111 00000 10101
111 00111 10000 10011 01110 00001 10101 01101 10000 01111 00000 10101
111 01110 00001 10101 01010 00001 00000 11111 00000 10101 0111 00111 00111 10001
111 00001 10000 01111 00000 01010 01111 0111 10000 10011 01110 00001 01110 01110 0000
111 00000 01010 01111 00000 10011 0111 10000 10011 0111 00001 10101 0111 00000 10101
111 00111 10000 10011 01110 00001 10101 01101 10000 01111 00000 01111 00000 01111
111 01010 10000 01111 00000 01010 01111 0111 10000 10011 01110 00001 01110 00000
111 01110 00001 10101 01101 00000 11111 00000 10101 0111 10000 10011 01110 00001 01110
111 00111 10000 01111 01100 00001 10101 01101 00000 01111 00000 01111 00000 01110
111 00111 10000 01111 01100 00001 10101 01101 00000 01111 00000 01111 00000 01110
111 00000 01010 01111 00000 10101 0111 10000 10011 01110 00001 01110 00000
111 00111 10000
111 01110 00001
111 01110 00001
111 01110 00001
111 01110 00001
111 00000 01010
111 00111 10000

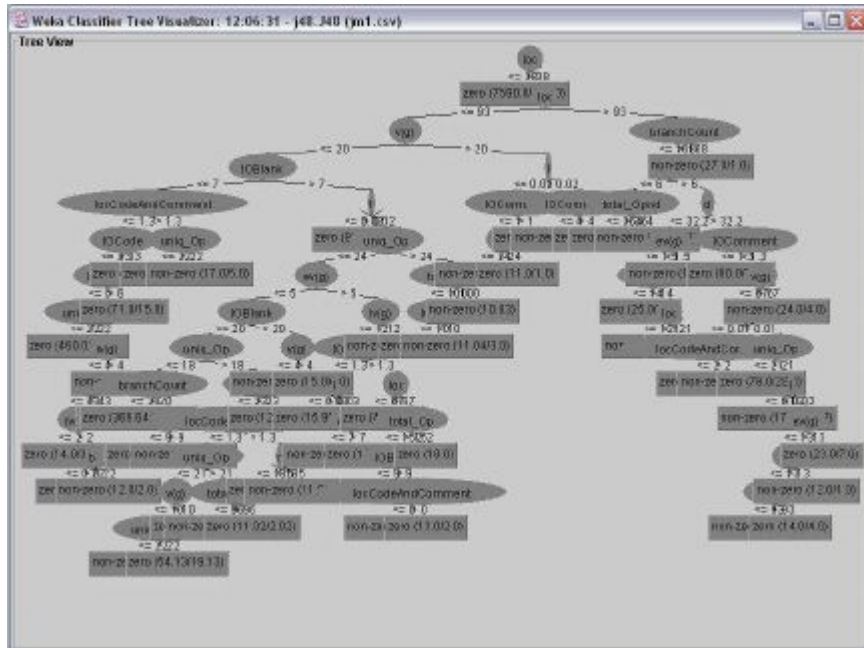


What is the role of Visual Analytics in ML model interpretation? How can we use it to enable/improve interpretation?

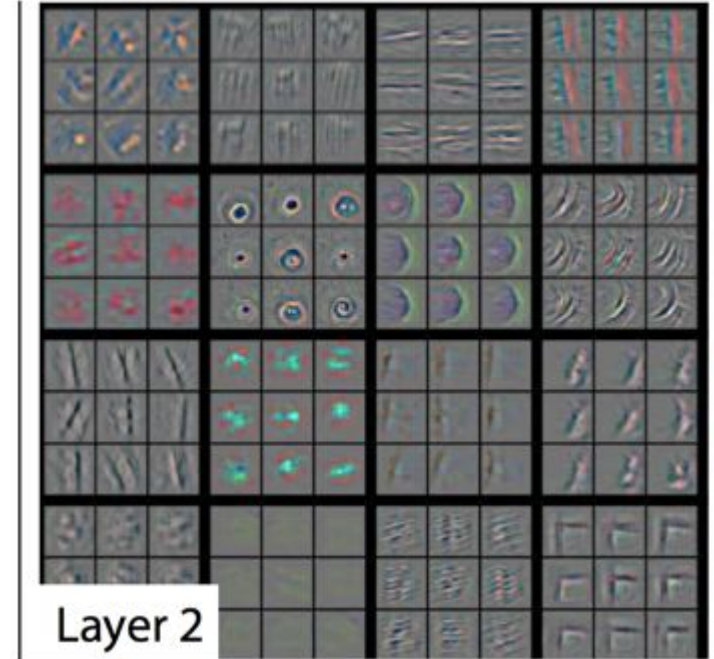
A few thoughts ...

Does it make sense to think of ML techniques as being intrinsically more or less interpretable?

Decision Tree



CNN



Vs.

Model Structure \neq Model Representation

Transparency of models
Interpretability of representation

Two modalities for model visualization:

White-Box: Visualizing model structure.

Black-Box: Visualizing model behavior.

Three modalities to inspect *model behavior*:

1. Item(s) to outcome.
2. Single feature to outcome.
3. Multiple features to outcome.

(With training, or hold-out, or simulated data.)

Prospector



Model: Random Forest (0/10)

Features: 50

Patients: 11 / 3624

ACC: 0.74656

AUC: 0.84810



Select

Inspect

Partial Dependence

Predictive Risk Color Key:



Patient: 3530 Truth: 1 Original: 0.42753 Current: 0.42753 ↻

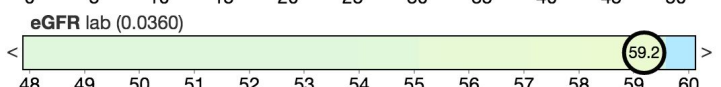
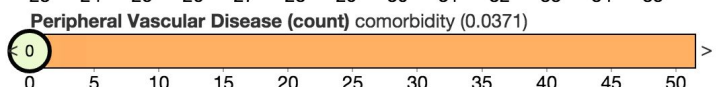
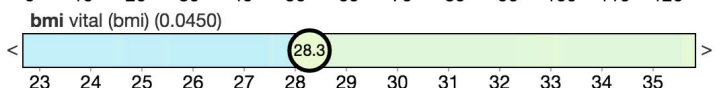
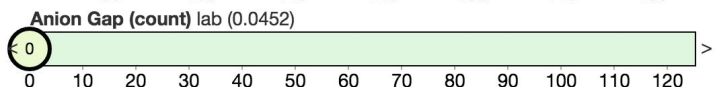
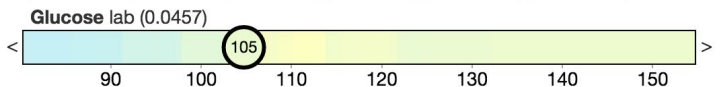
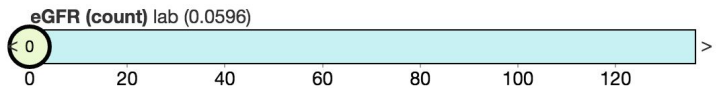
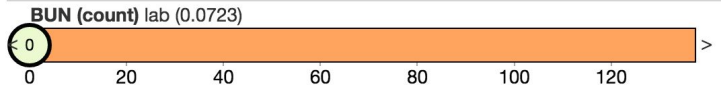
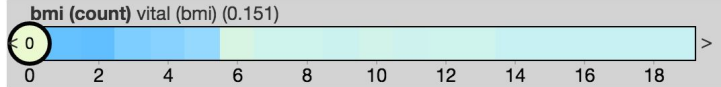
Show Neighbors

Sort by: Weight

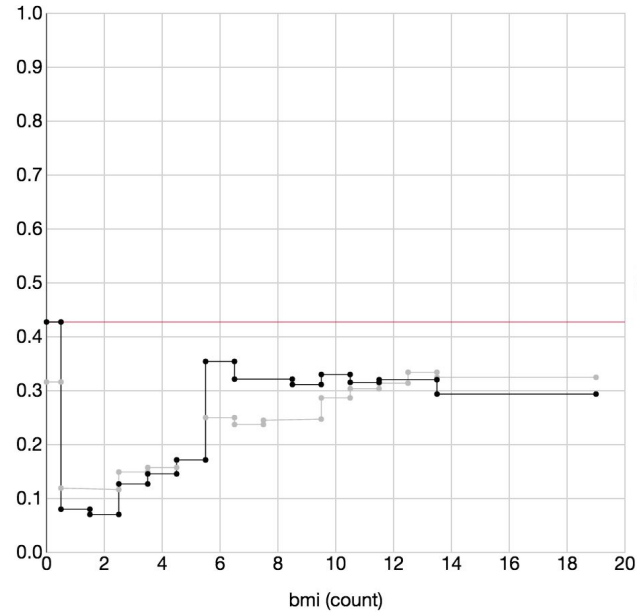
Relevance

Inc. Risk

Dec. Risk



vital (bmi) - r: 0.15119



Risk

bmi (count)

■ Current State
■ All
■ Original Score

Class Signatures



Thanks!

NOTE: Very happy to discuss and learn from you!
And very eager to establish collaborations.

Enrico Bertini

enrico.bertini@nyu.edu

Twitter: @FILWD