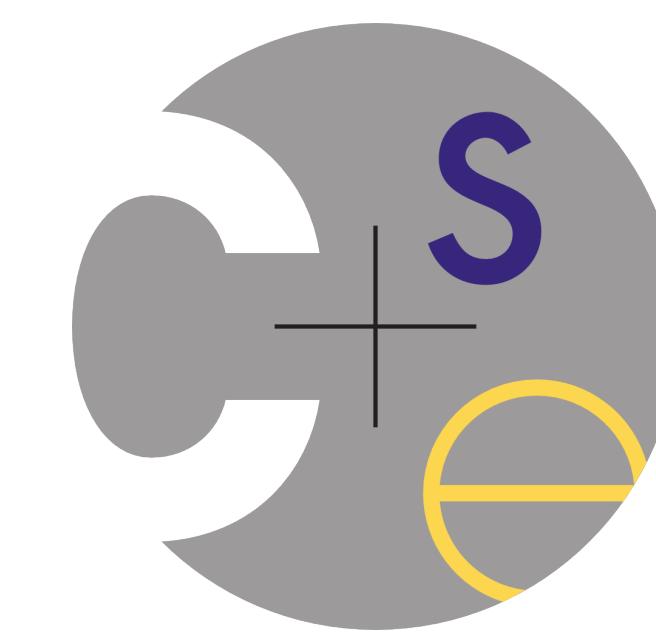




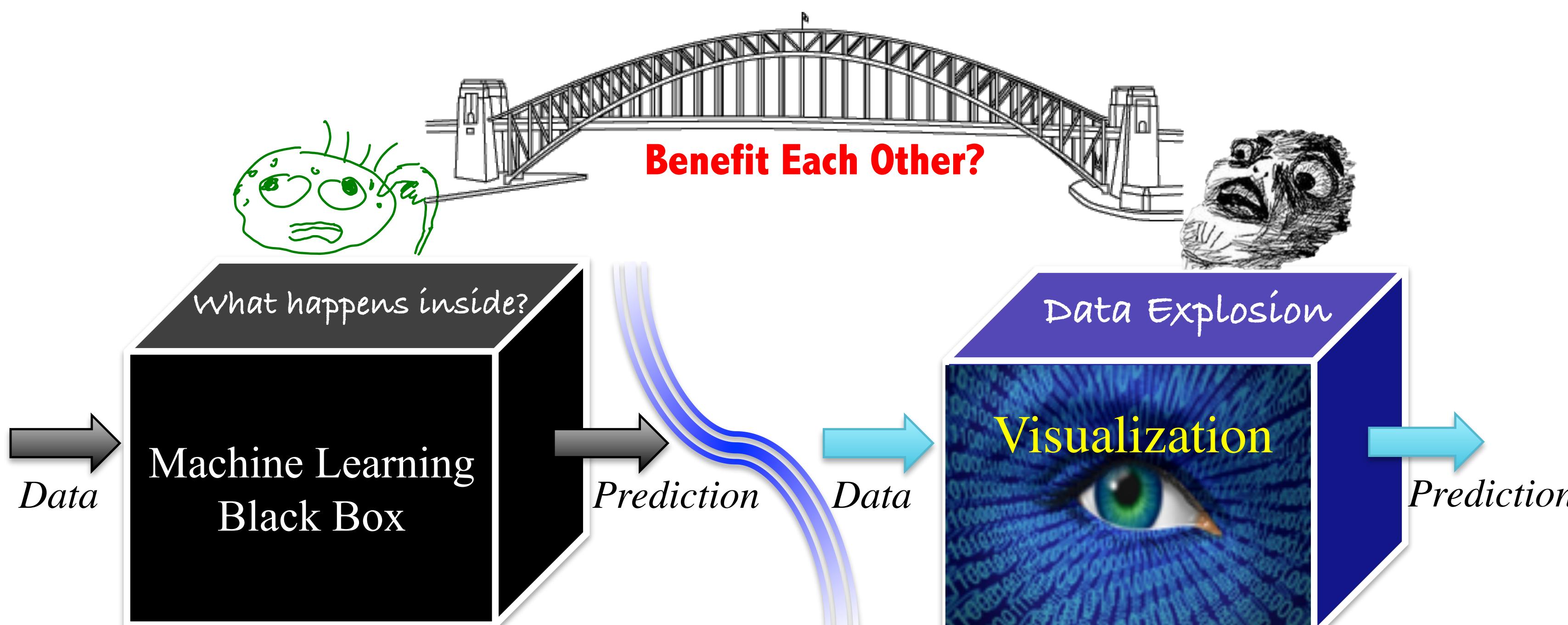
TRANSPARENT BOOSTING TREE: AN INTERACTIVE MACHINE LEARNING FRAMEWORK



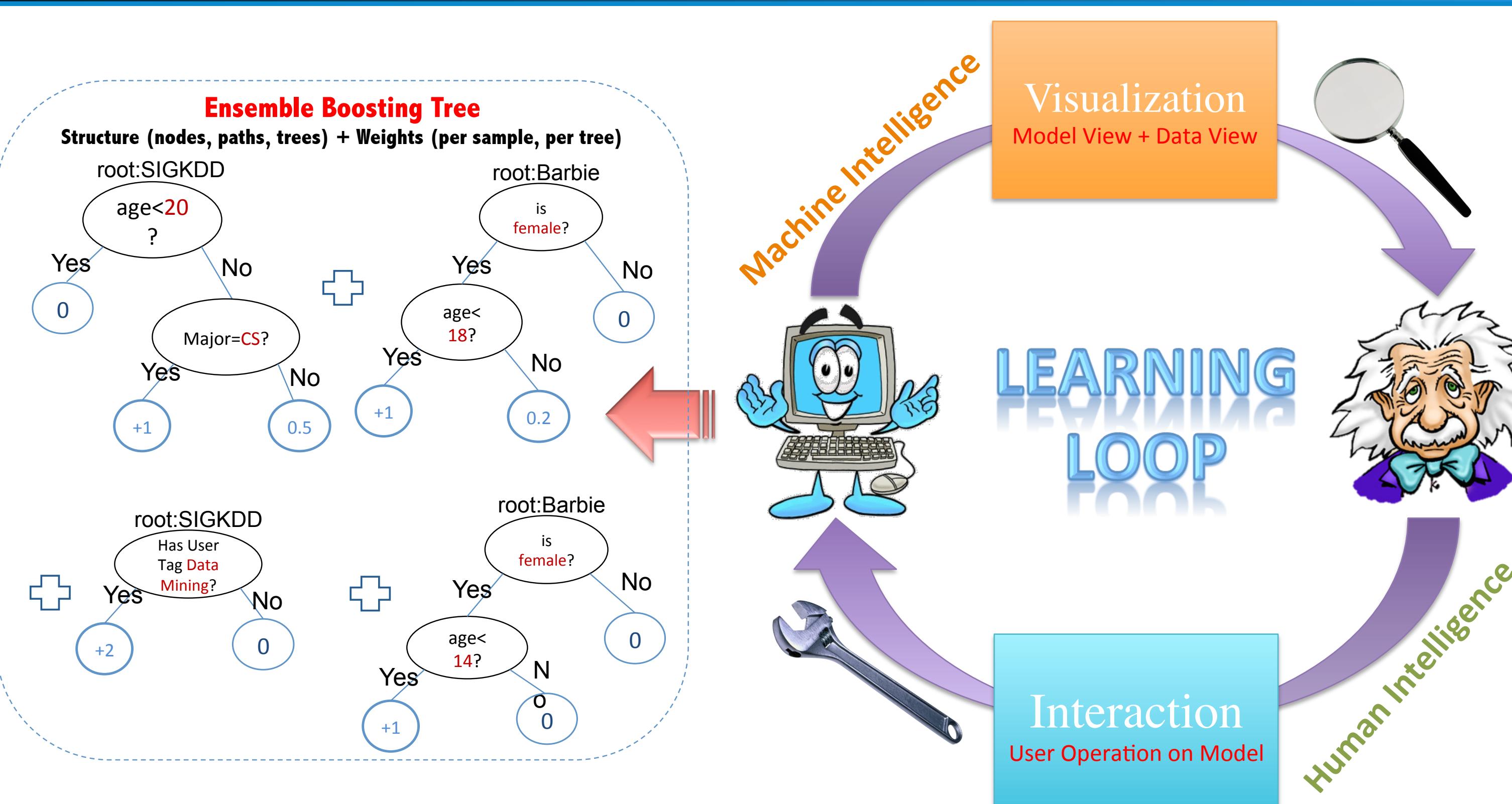
Tianyi Zhou, Tianqi Chen, Luheng He, Computer Science & Engineering, University of Washington

PROBLEM: MACHINE LEARNING MEET VISUALIZATION

- Machine Learning
 - (a) A fancy black box using machine Intelligence, and statistics.
 - (b) Popular, but few people know why (**not transparent**).
 - (c) Good at fitting data, but difficult/slow to fit domain knowledge.
- Visualization
 - (a) Explore data by using human intelligence, and user feedback.
 - (b) Effective on clean, low-dim, small data; hard on noisy, high-dim, big data.
 - (c) Good at guiding people to useful judgement, not good at processing numbers.



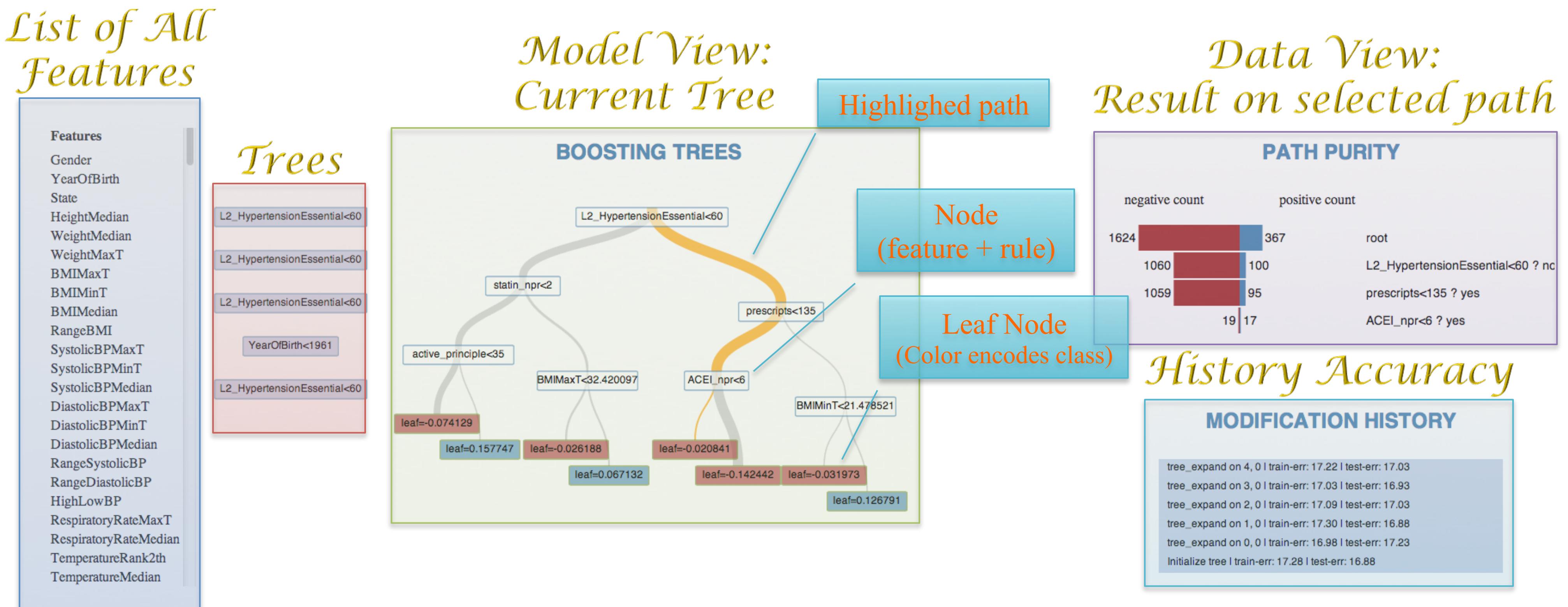
IDEA: CLOSE A LOOP BETWEEN ML AND VISUALIZATION



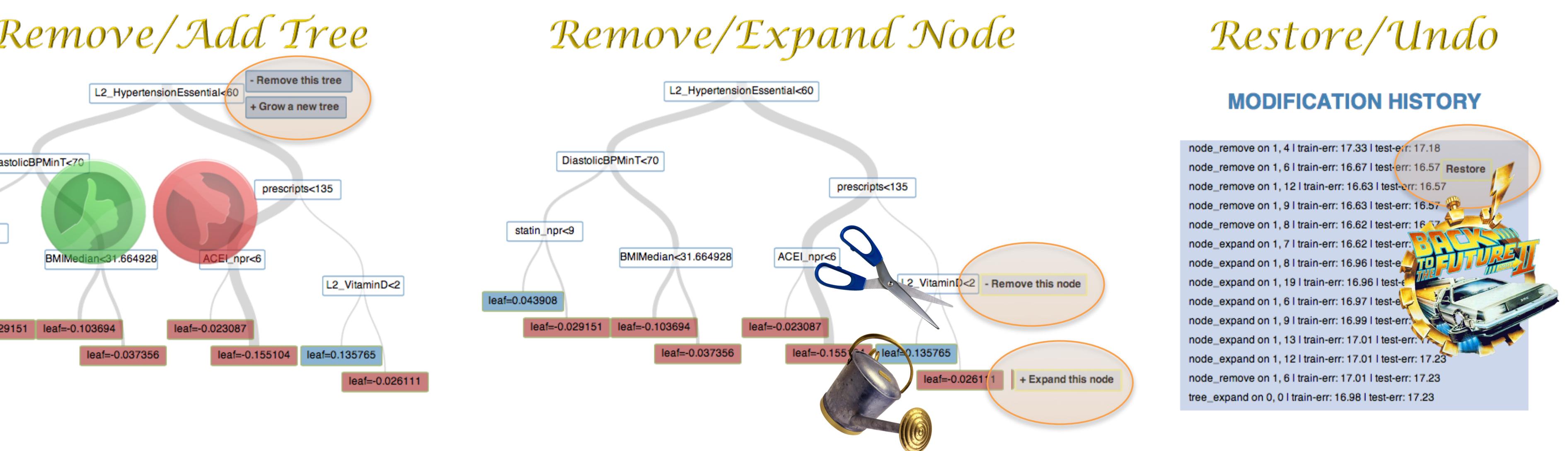
- Visualize machine learning process, improve the interpretability of ML model.
- Let machine learning and visualization improve each other in a mutually rewarding way.

TRANSPARENT BOOSTING TREE

- User Interface and Encoding in Visualization:



- User Operations and Interaction with ML model:



- Based on User feedback and operations, learning algorithm implemented in xgboost will:
 - Remove/Add tree: remove the current tree/learn a new different tree.
 - Remove/Expand node: update all the influenced weights of trees and samples.
 - Restore: back to the model obtained at any previous step.
 - User can choose to give/not give feedback on any parts of model at any step of learning process, and xgboost will return an ensemble boosting tree obeying user's setting.

CONCLUSION

- Visualize the learning process and mechanism of ensemble boosting tree; Make machine learning more transparent and interpretable, and thus more convincing.
- Involve domain knowledge and user feedback into ML (reduce generalization error); Enable visualization to handle big data (auto-choose what to show to users).
- Will allow user to select features, and extend this transparent learning idea to other ML techniques.