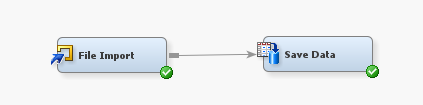
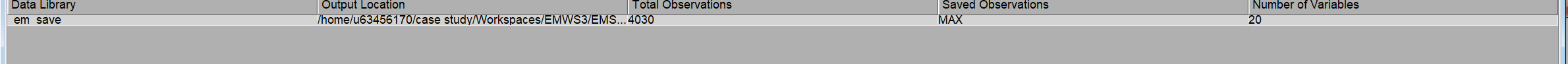
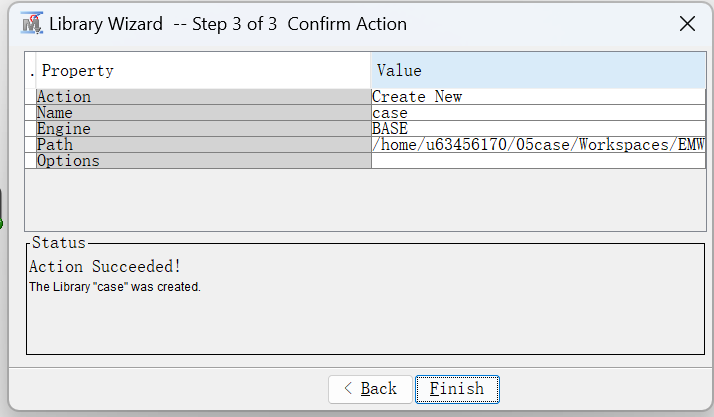
Create a new diagram

The import of the file and save data:

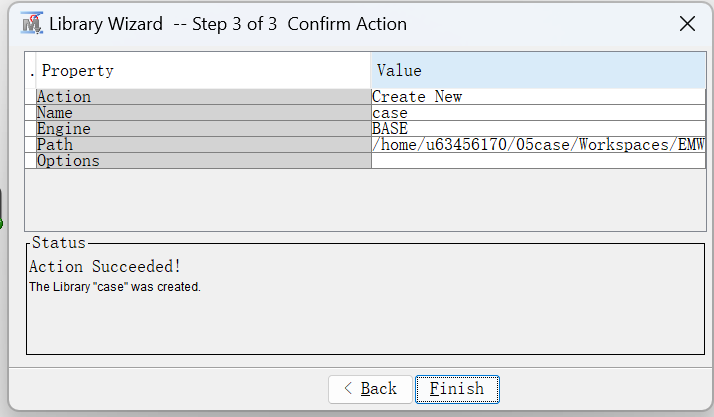


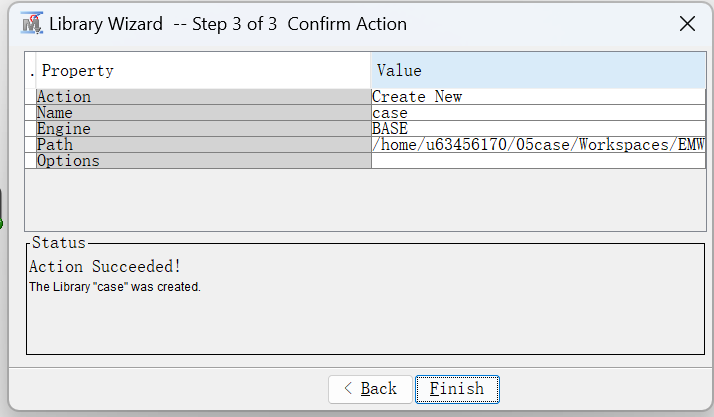


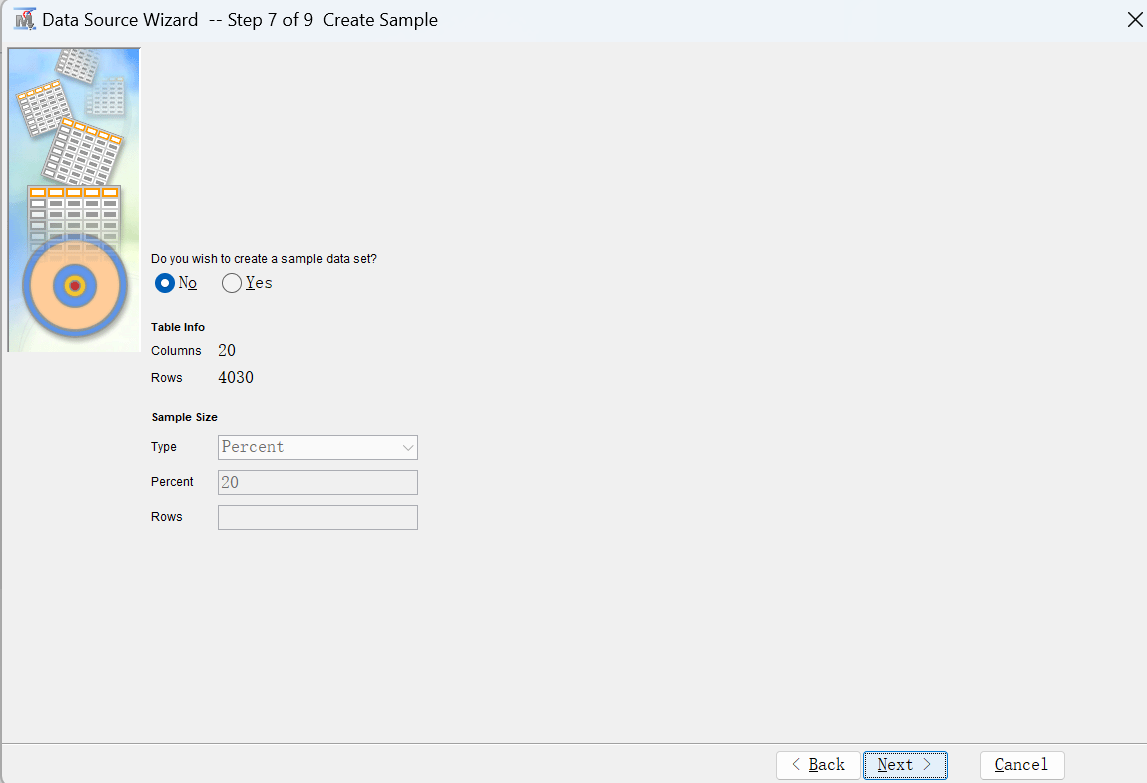
Build the library:



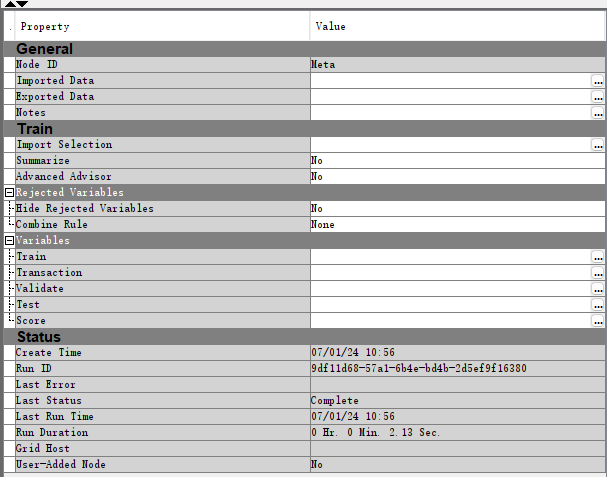
Change the level:







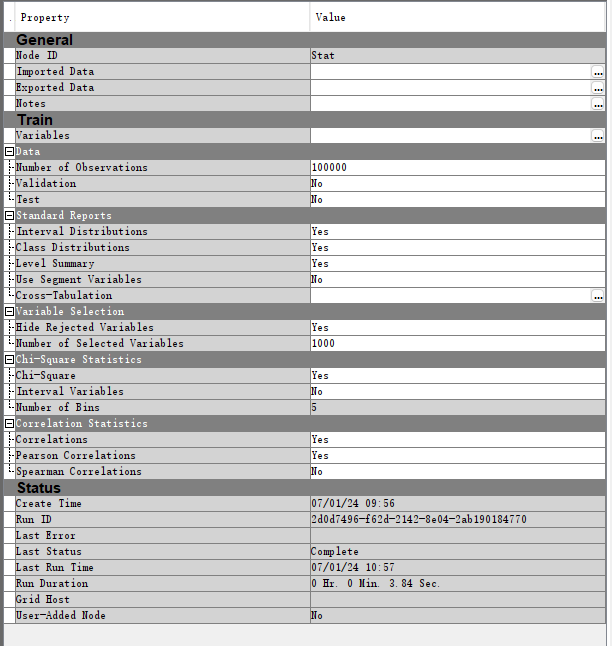
Run metadata:

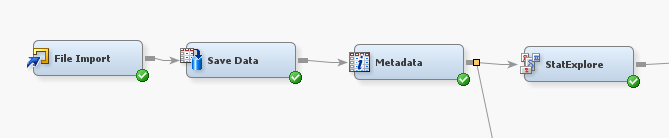


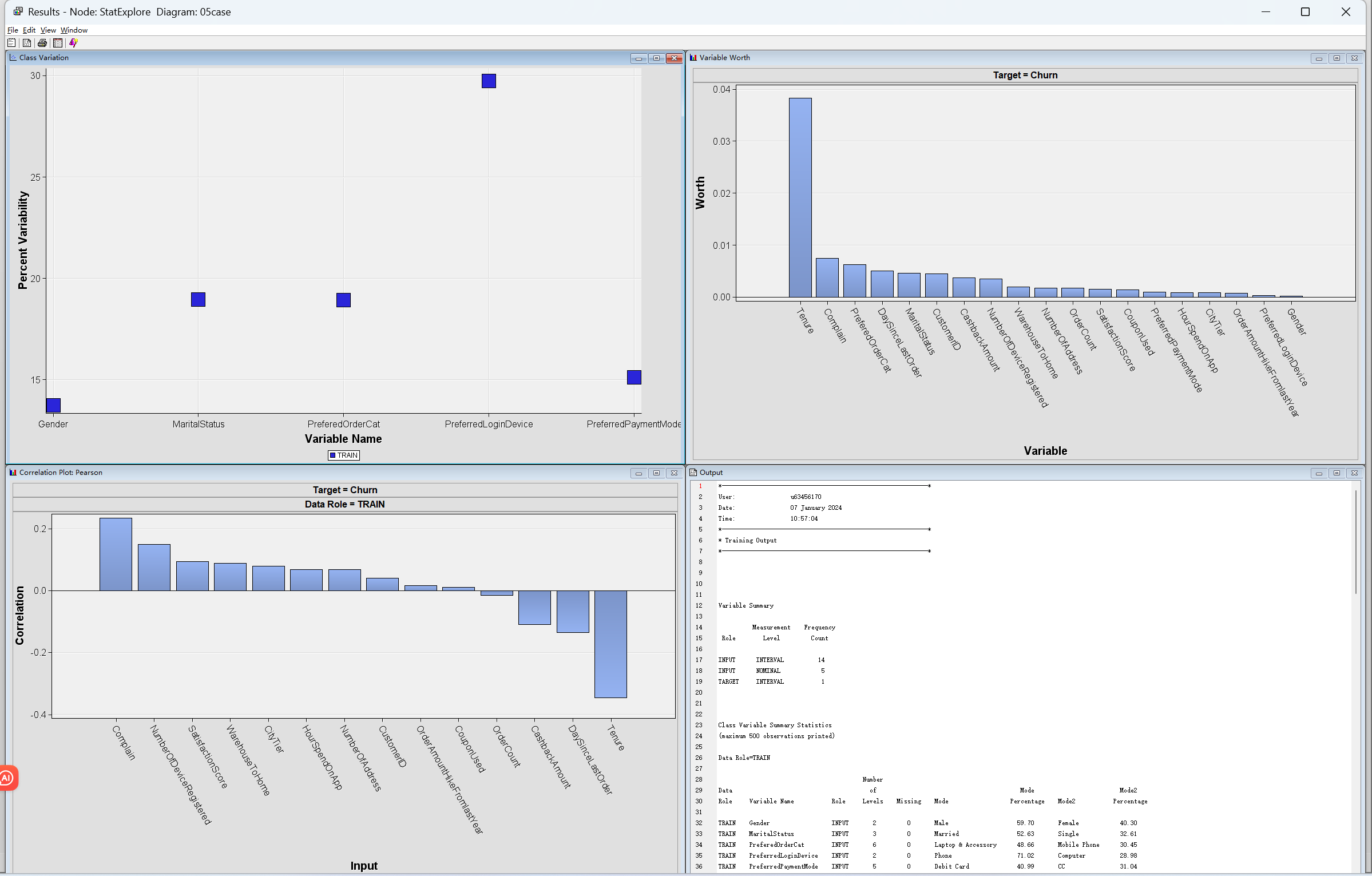


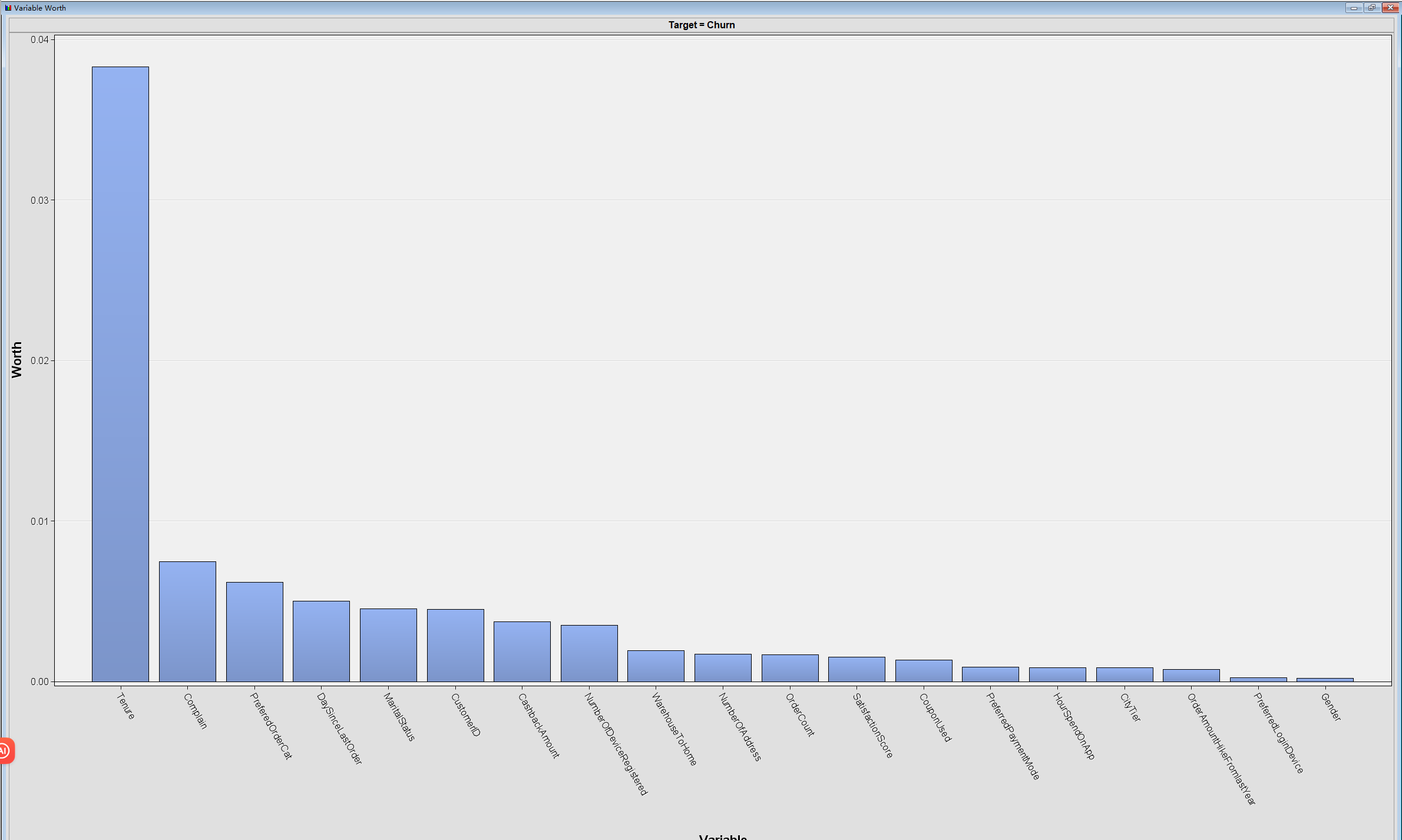
Run statistic explore:

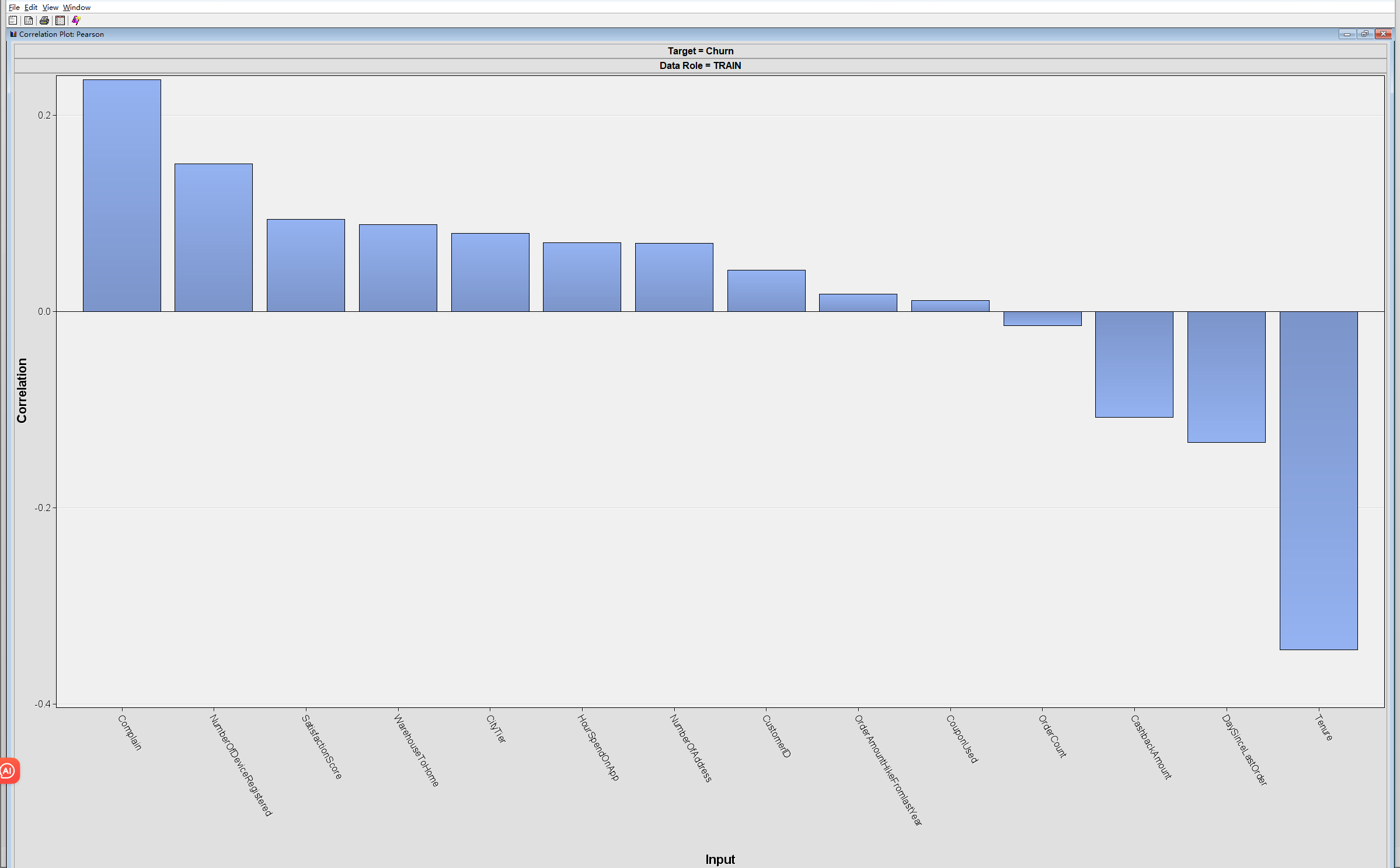
I set the training to 70%, and validation to 30%



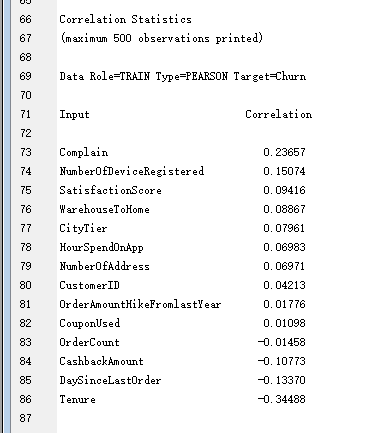


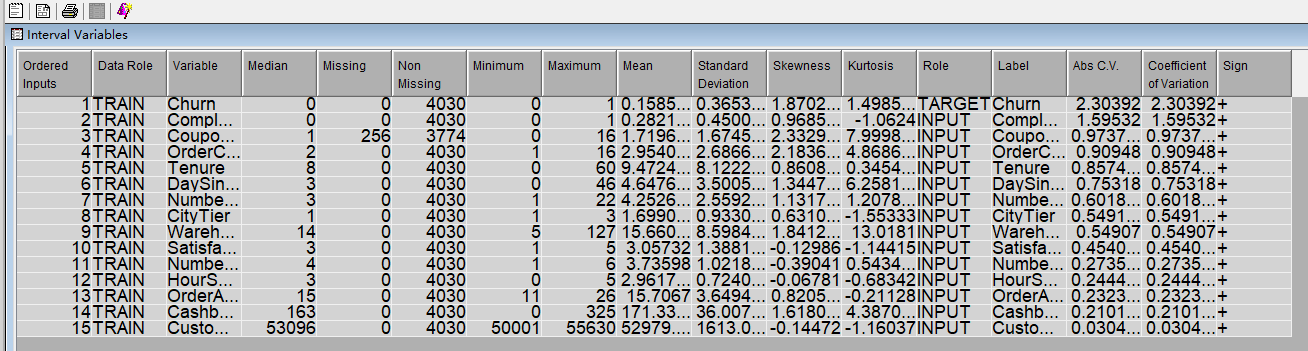


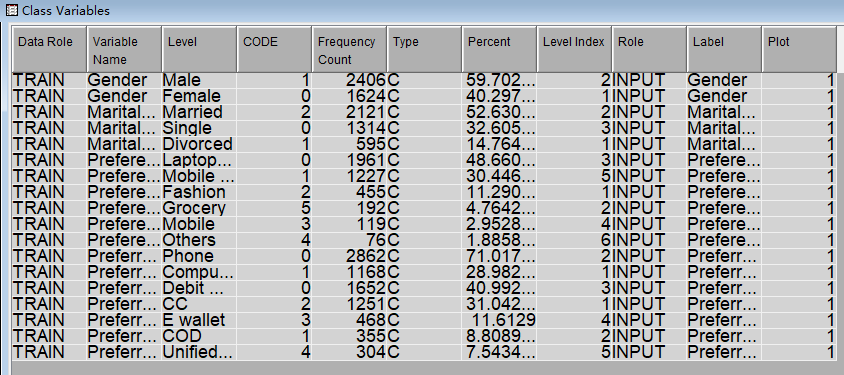




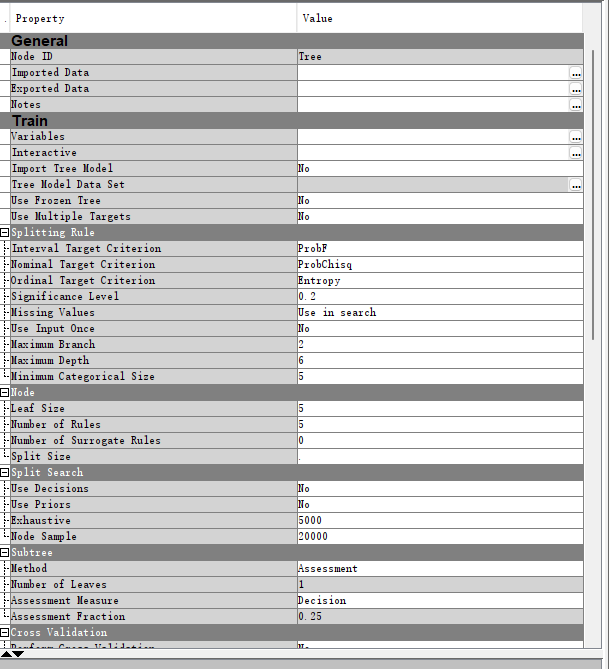


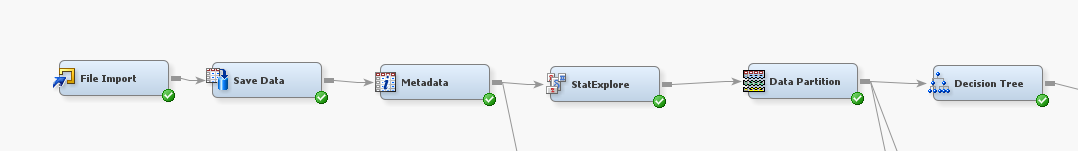


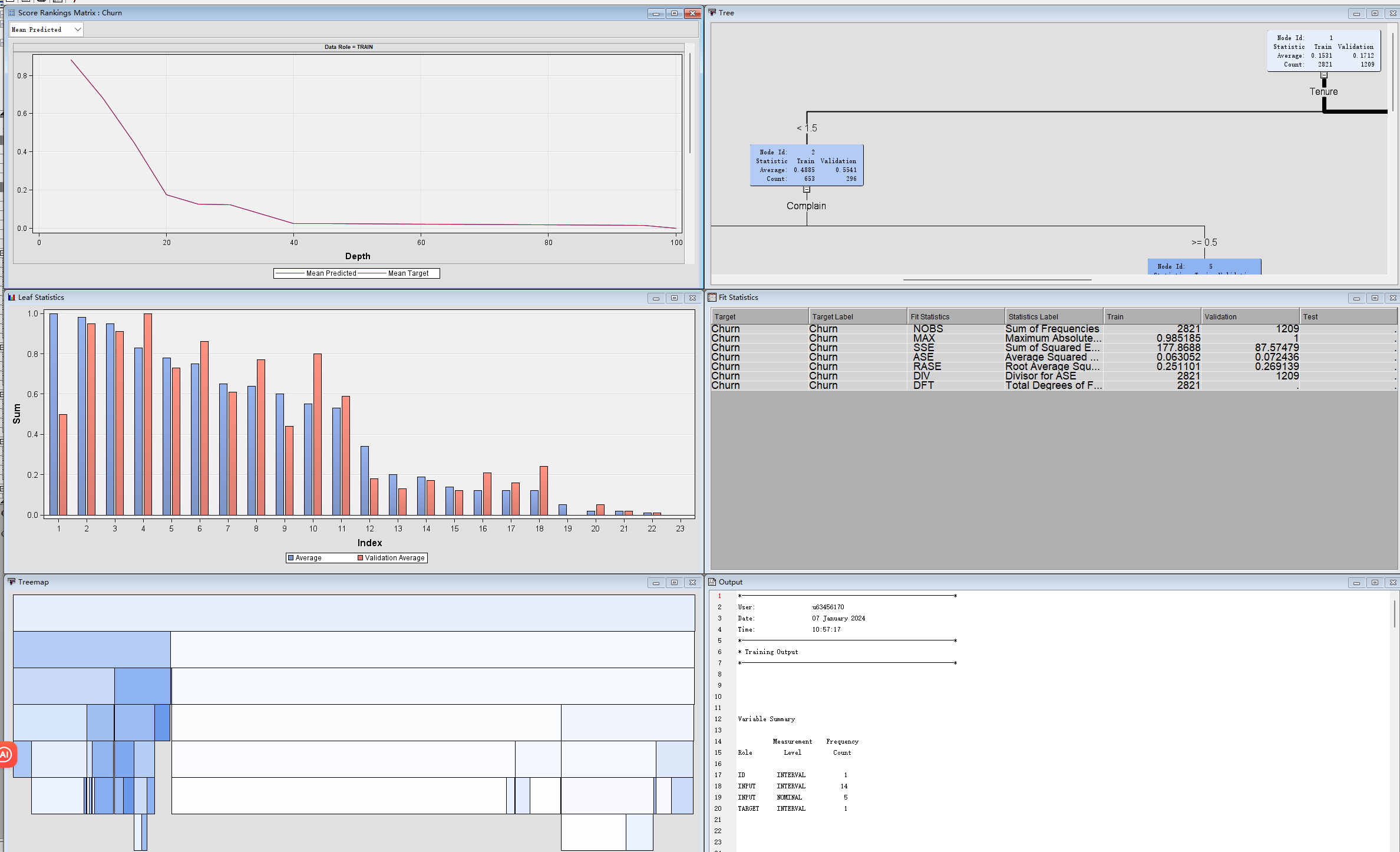


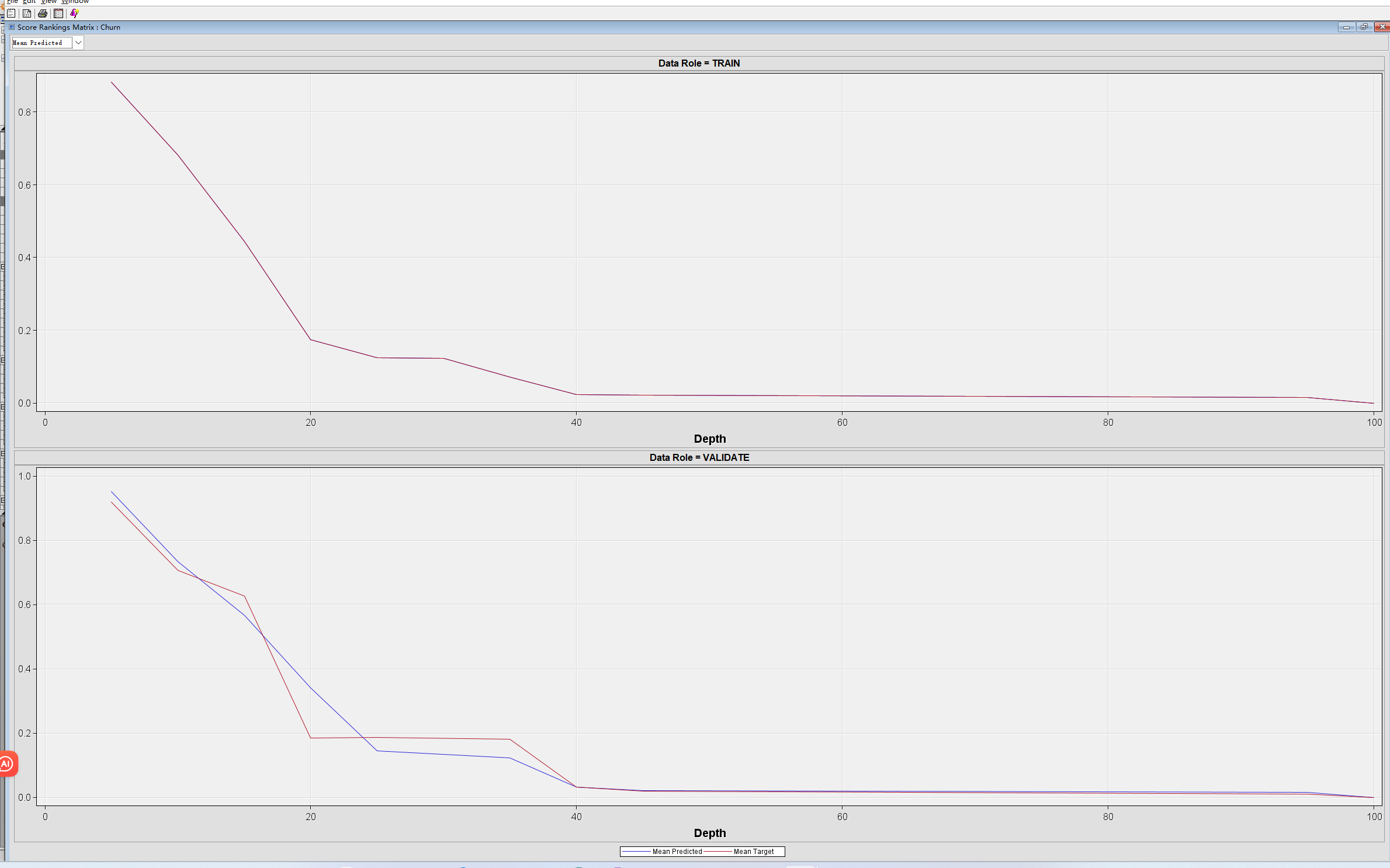


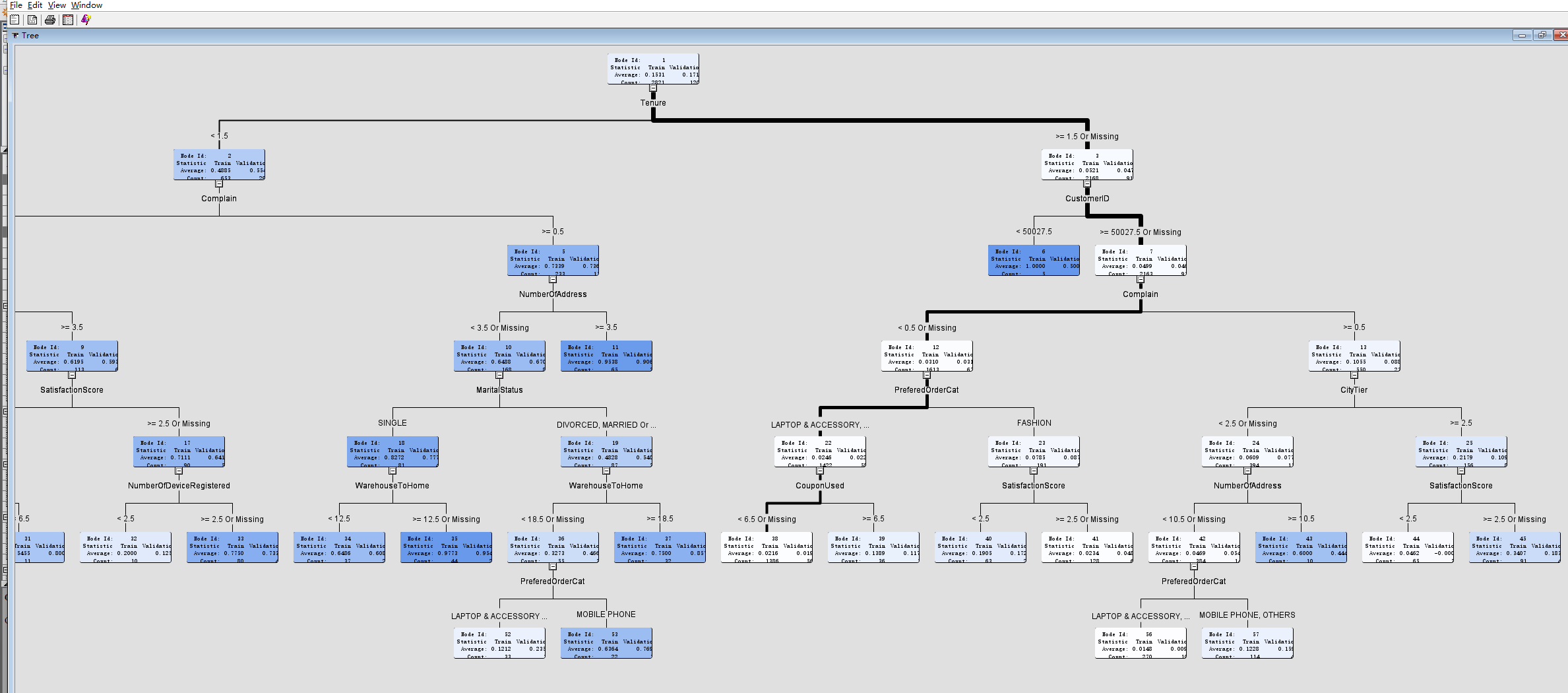
Run decision tree:

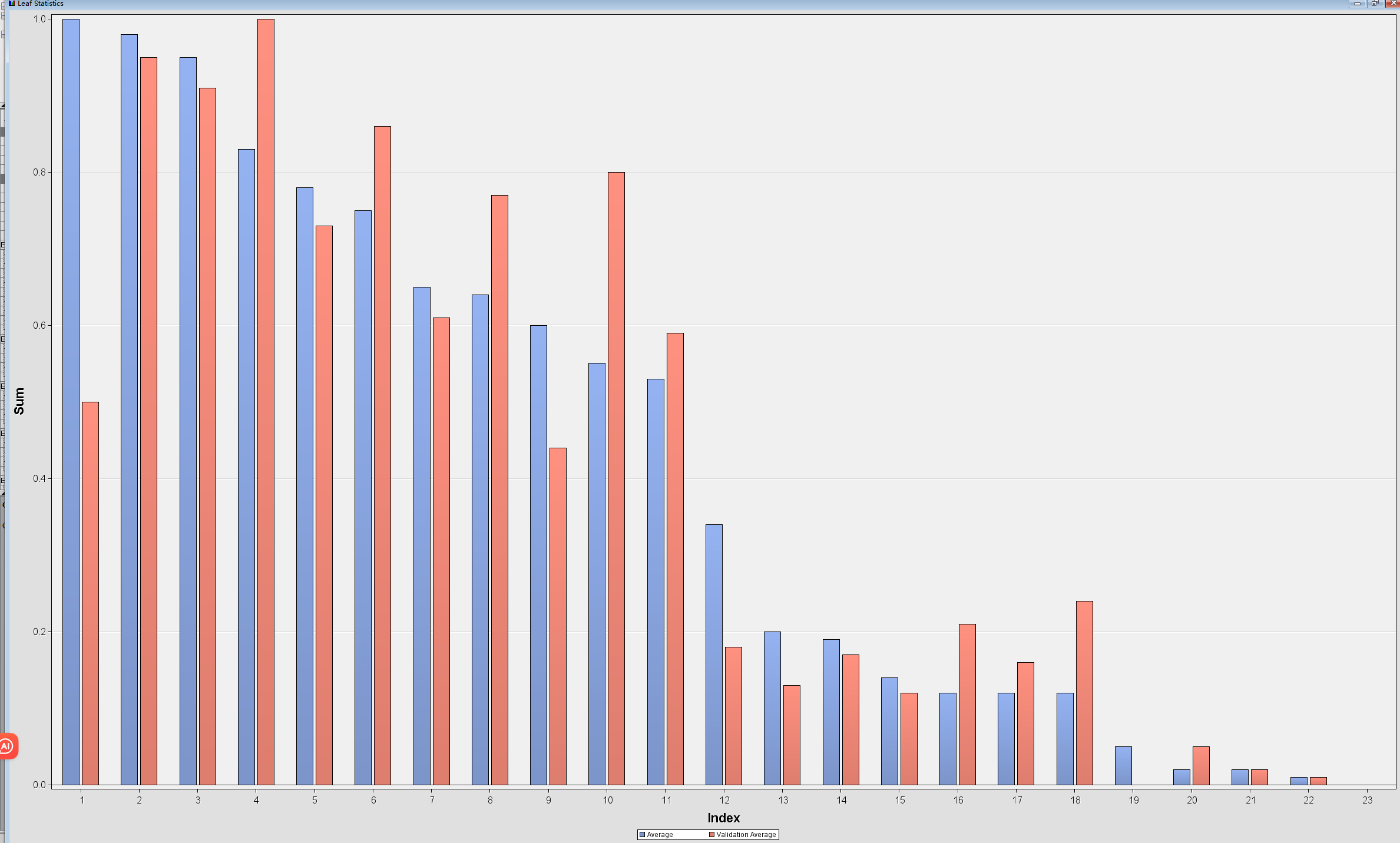


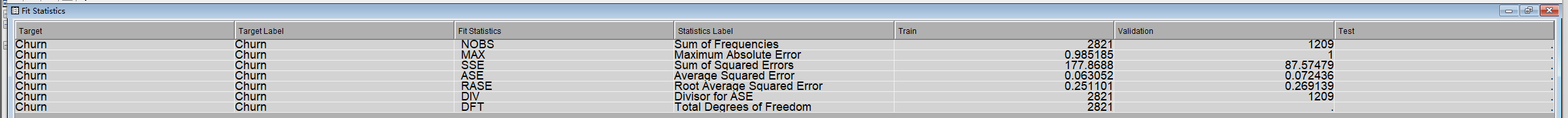


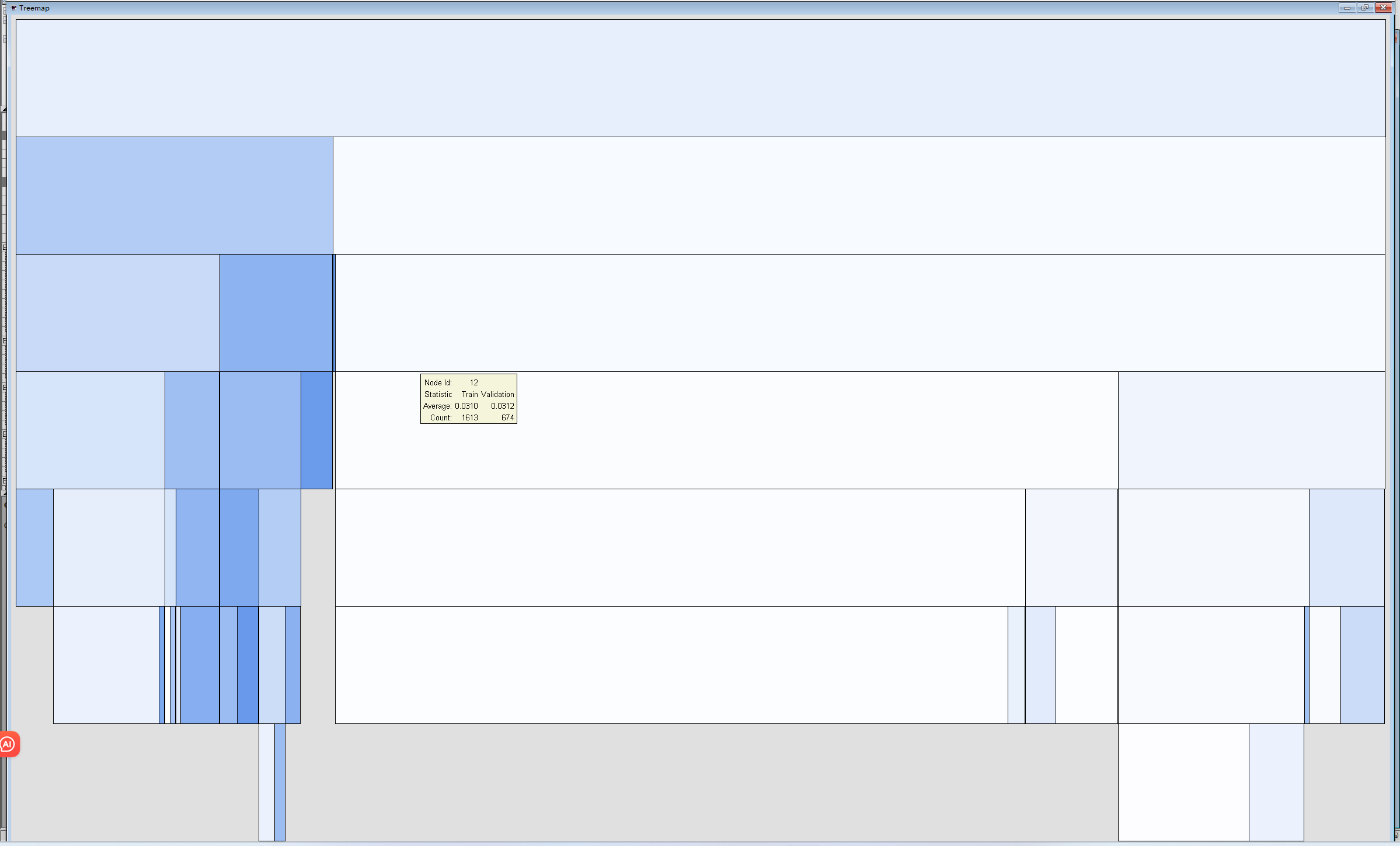


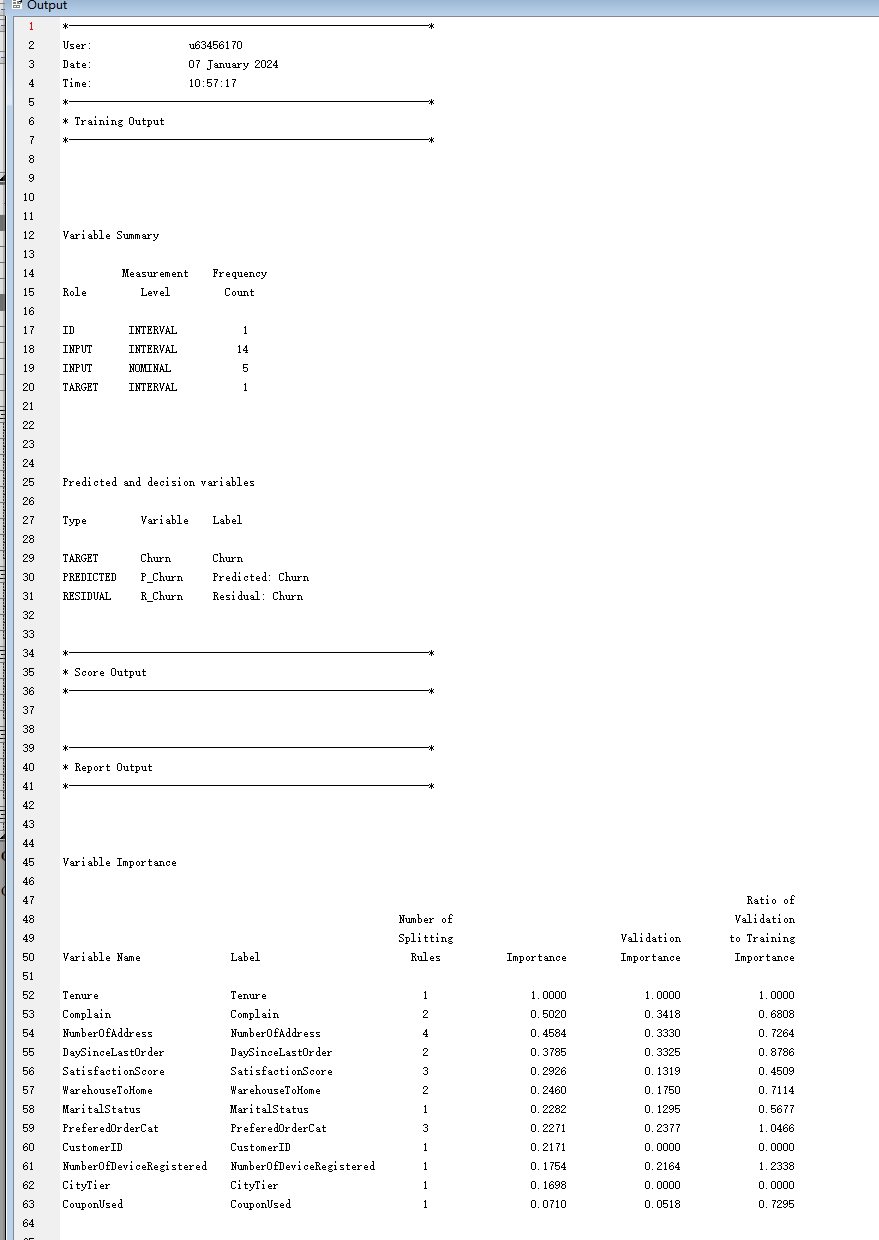


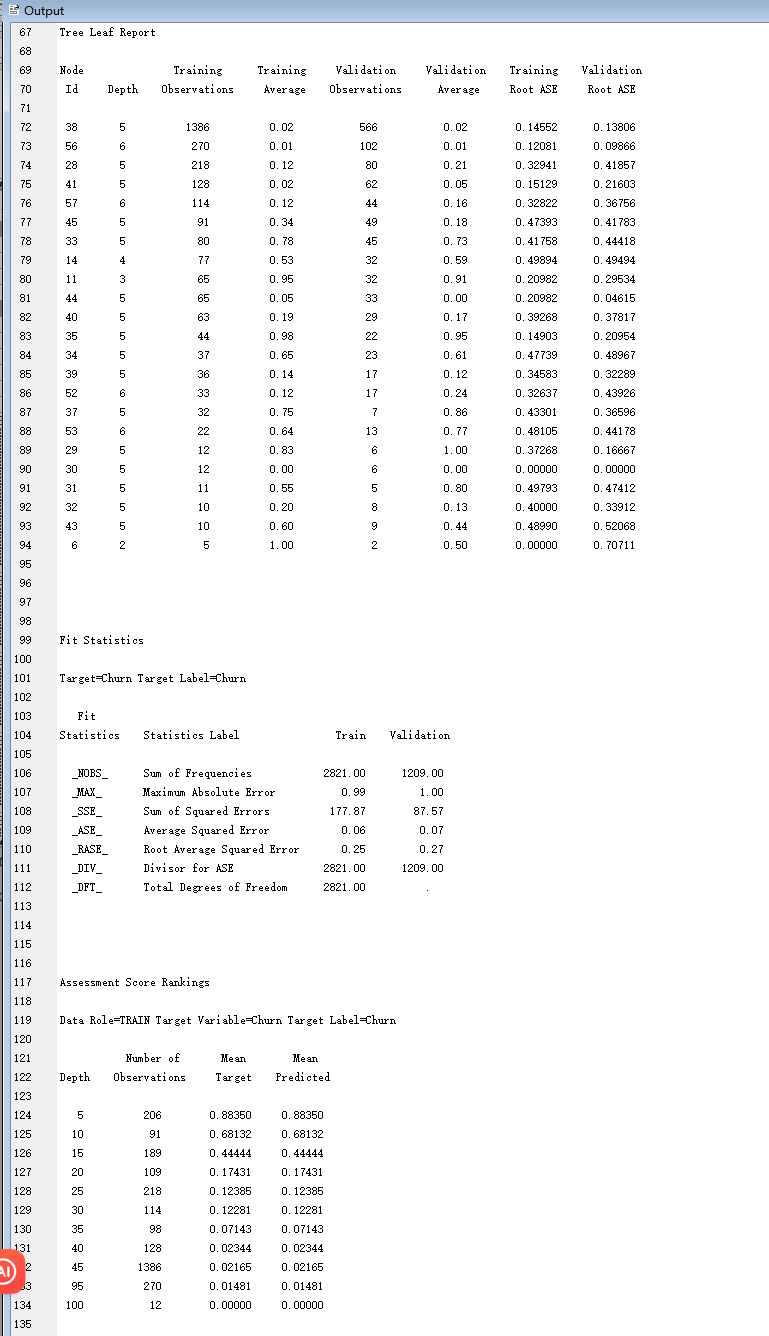


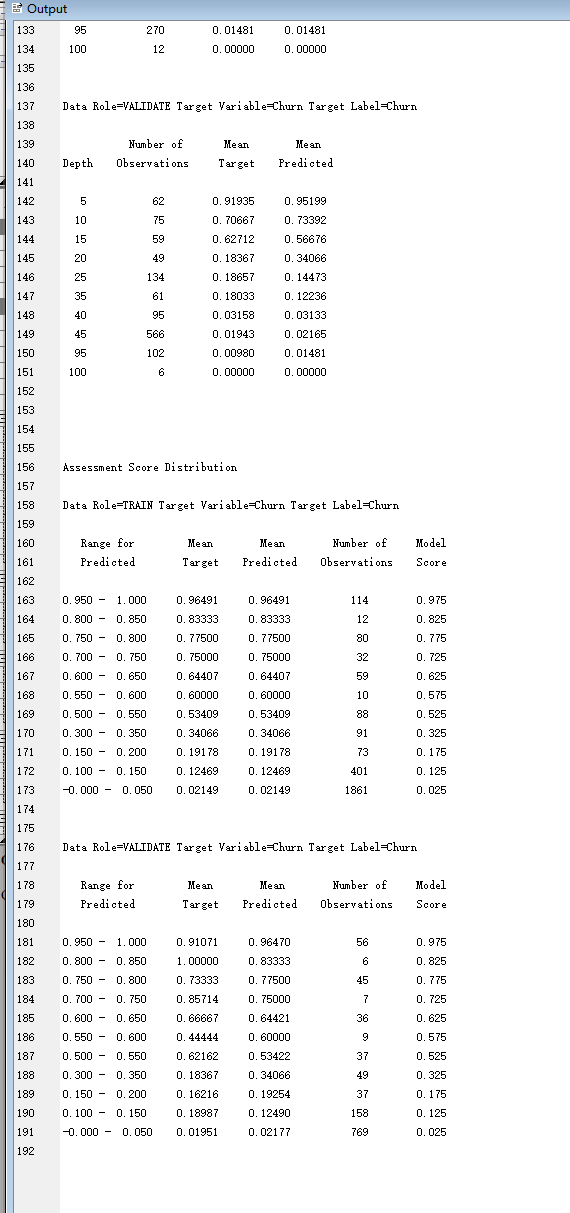




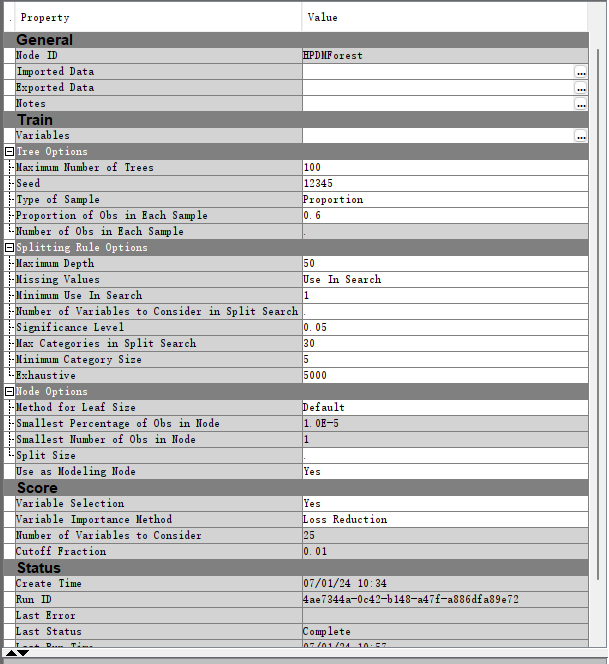


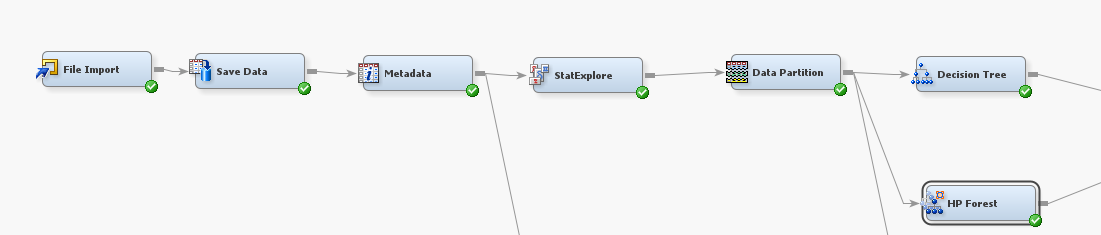


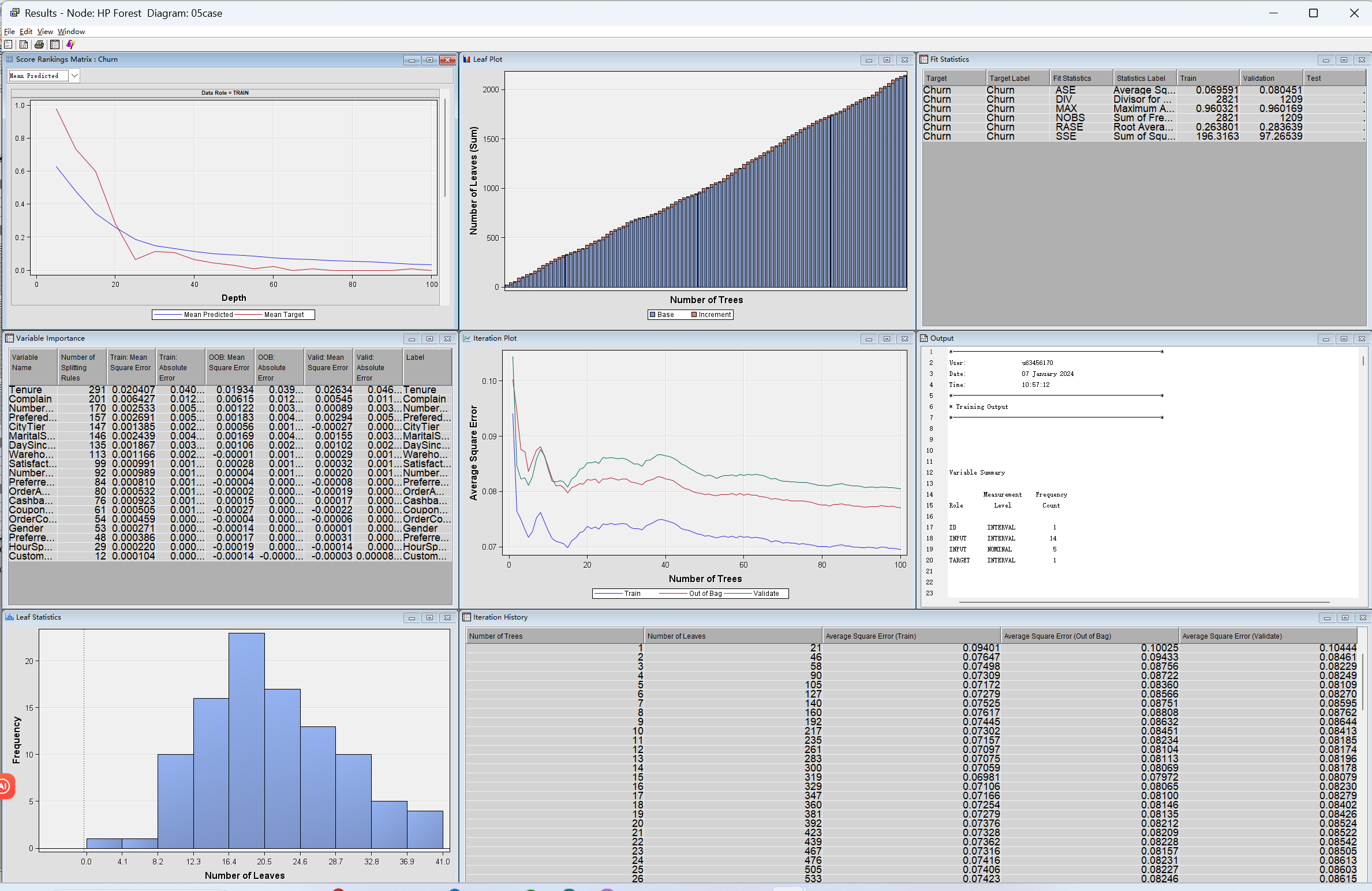


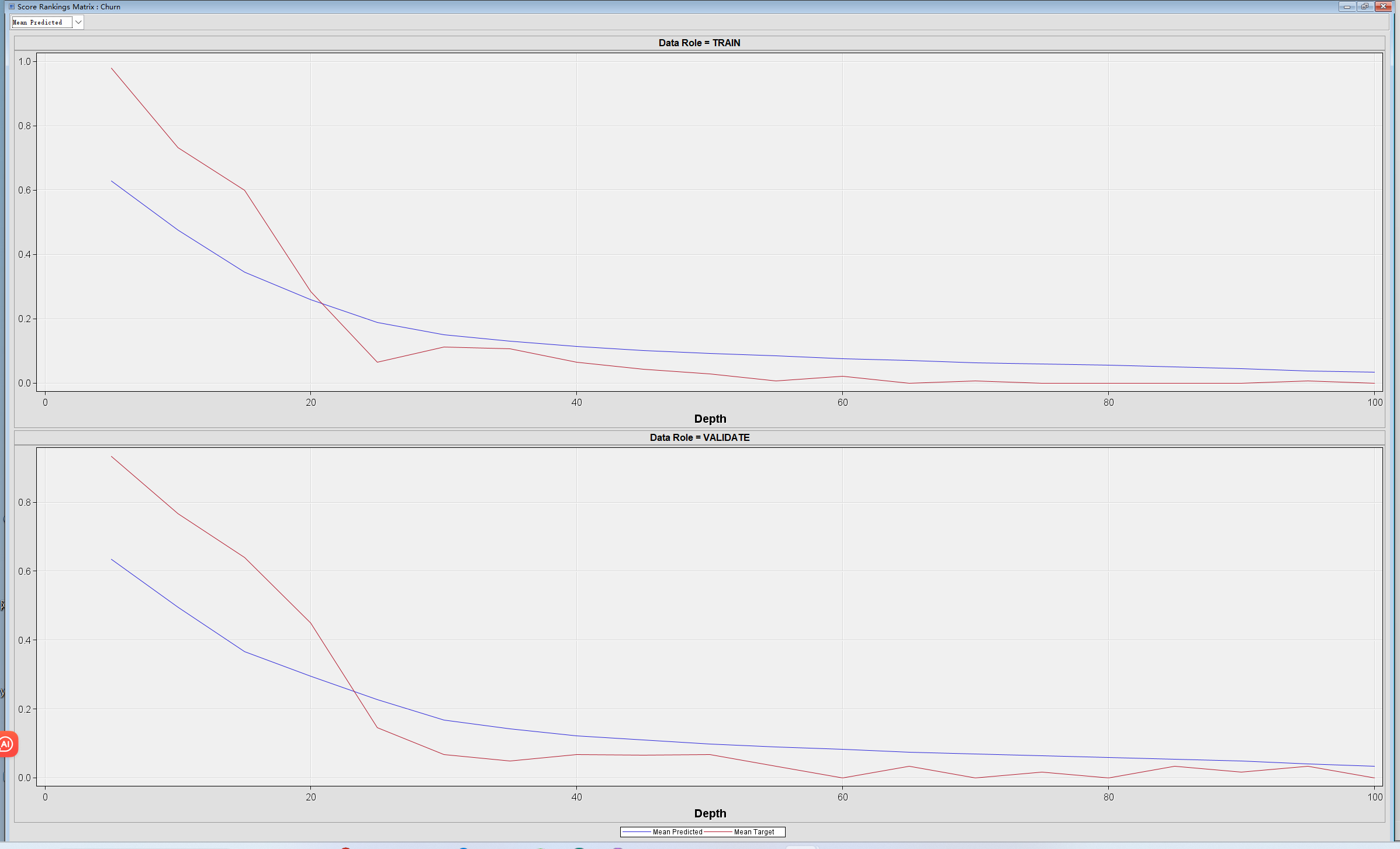


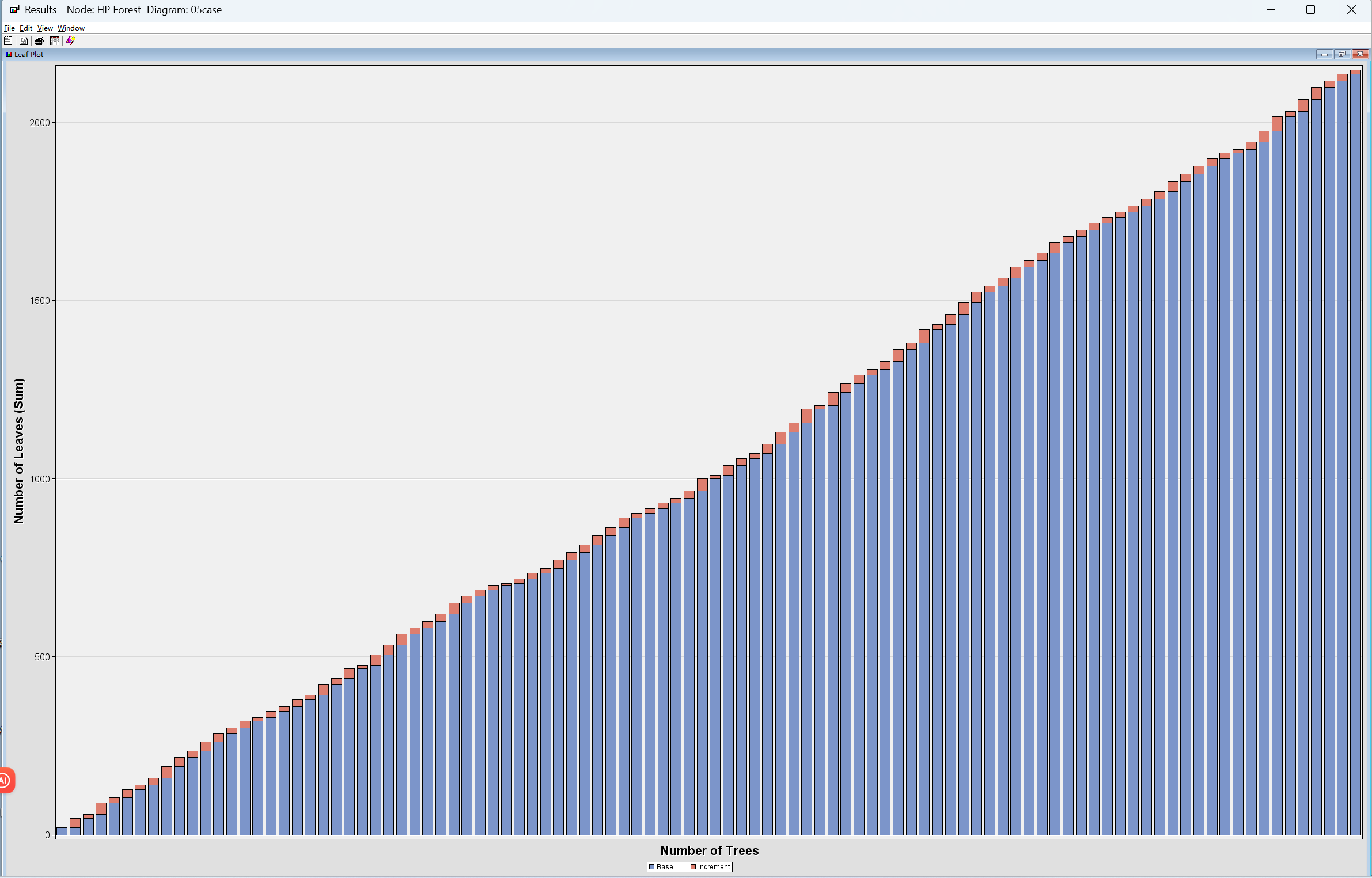
Run HP Forest:

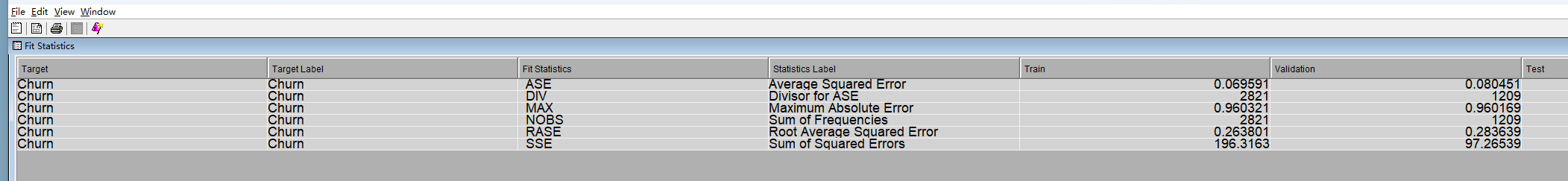


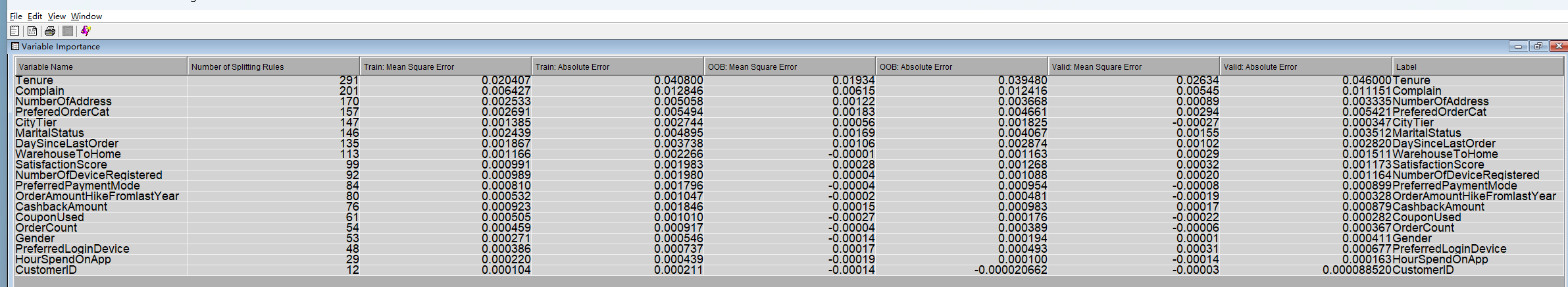


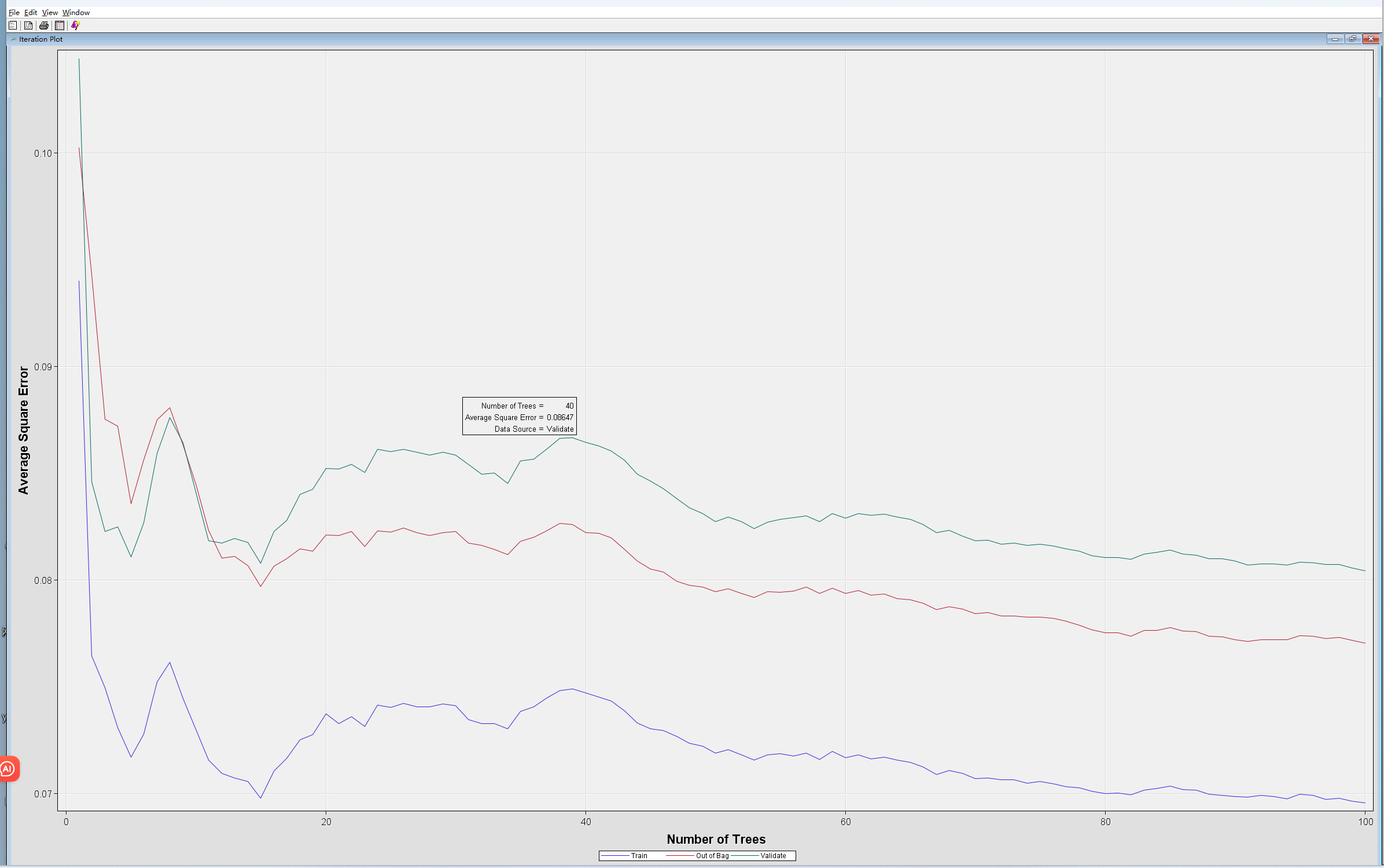


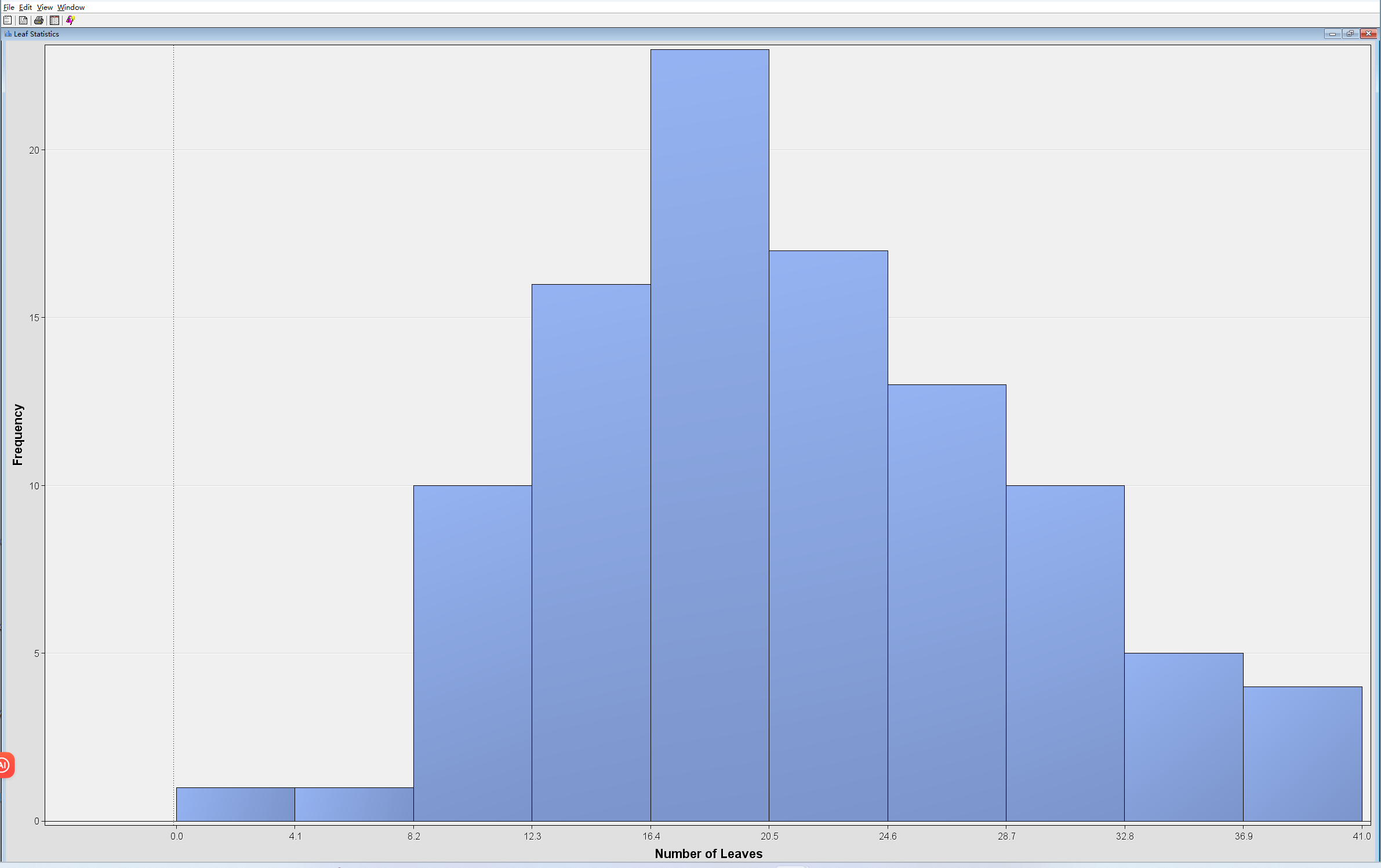




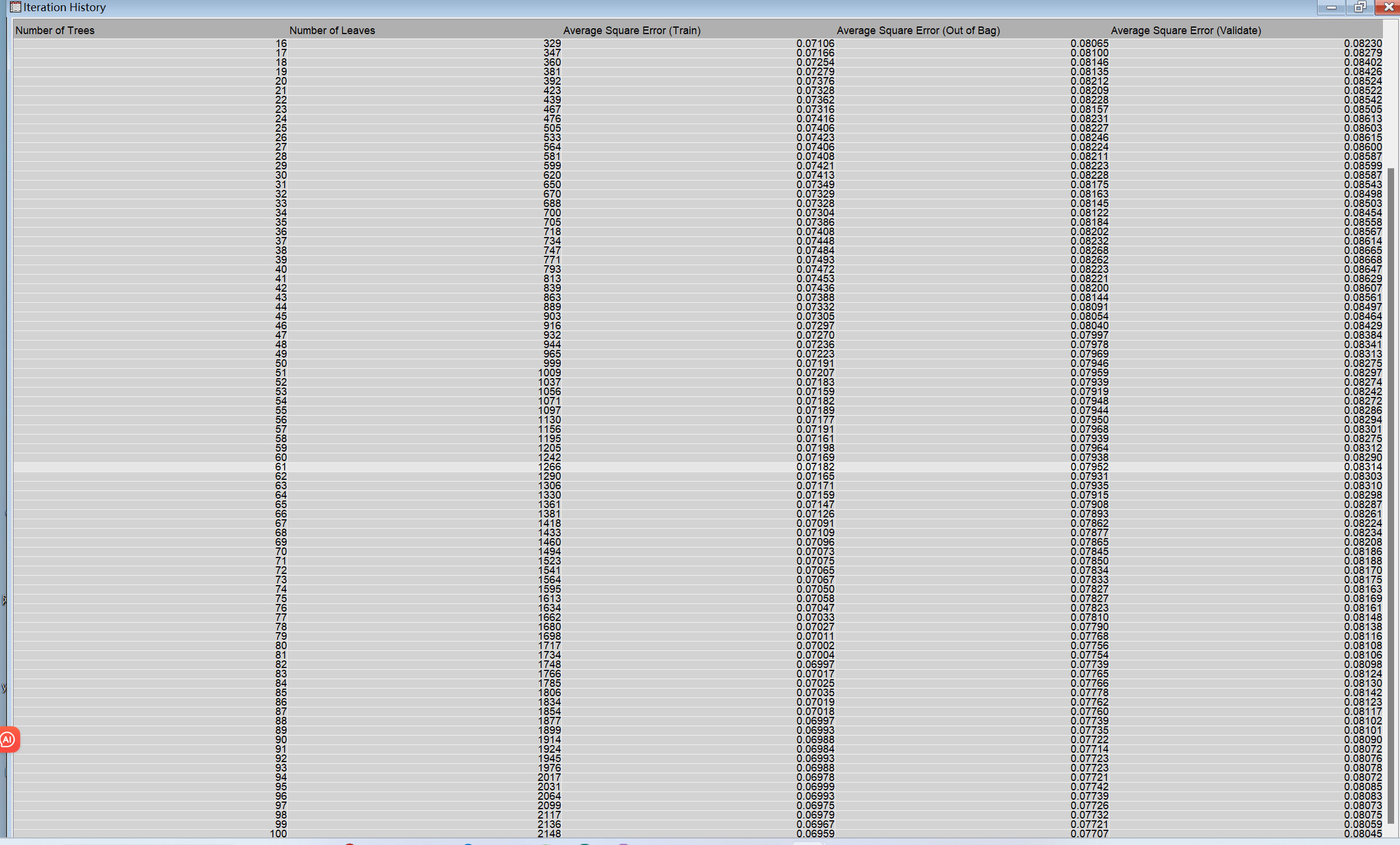


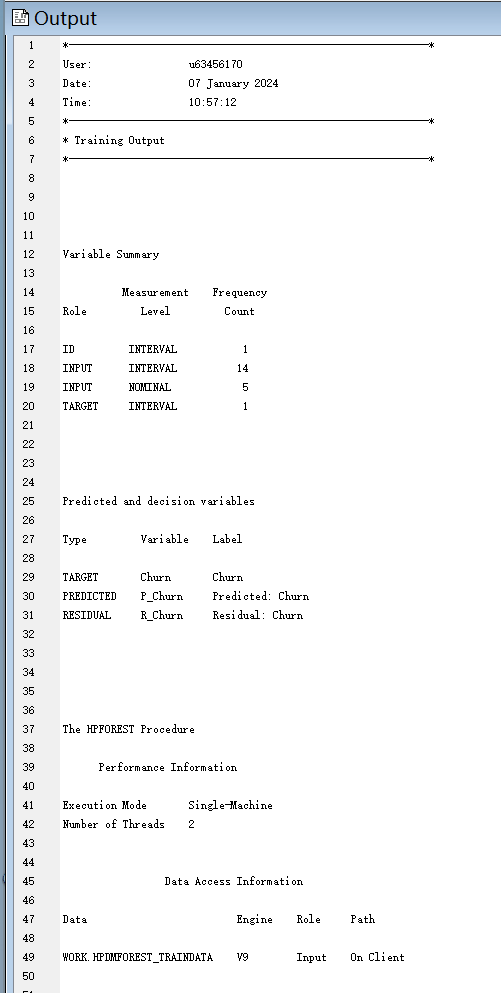


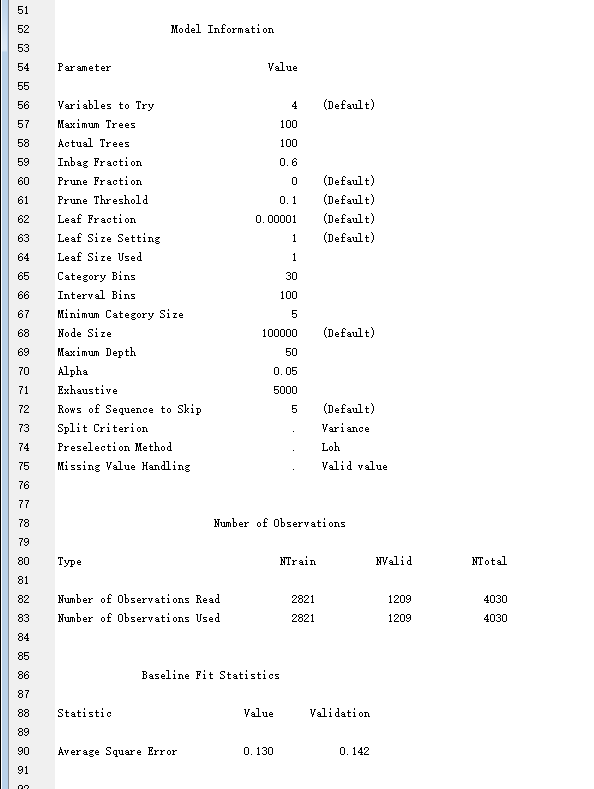


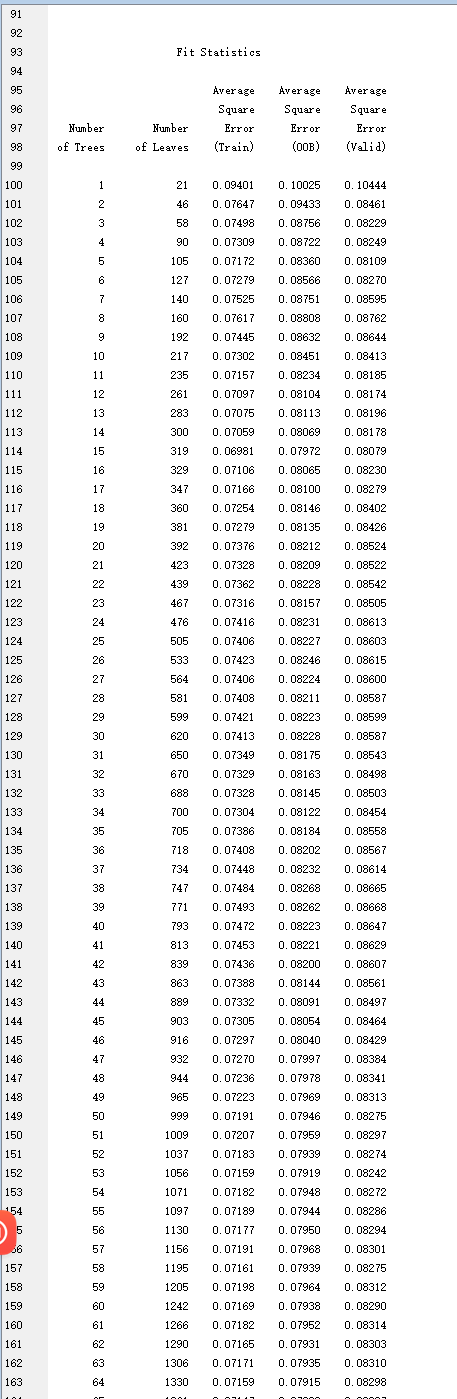


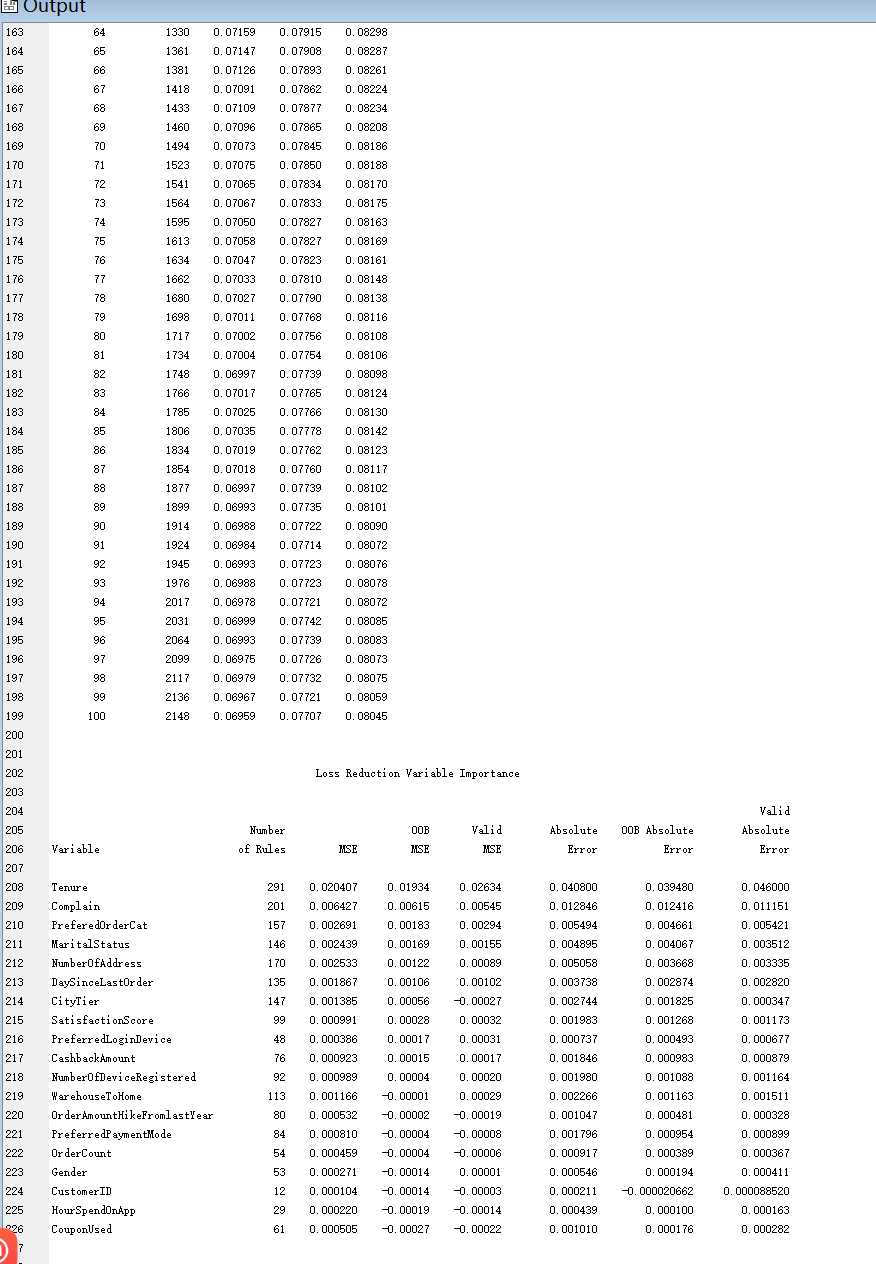


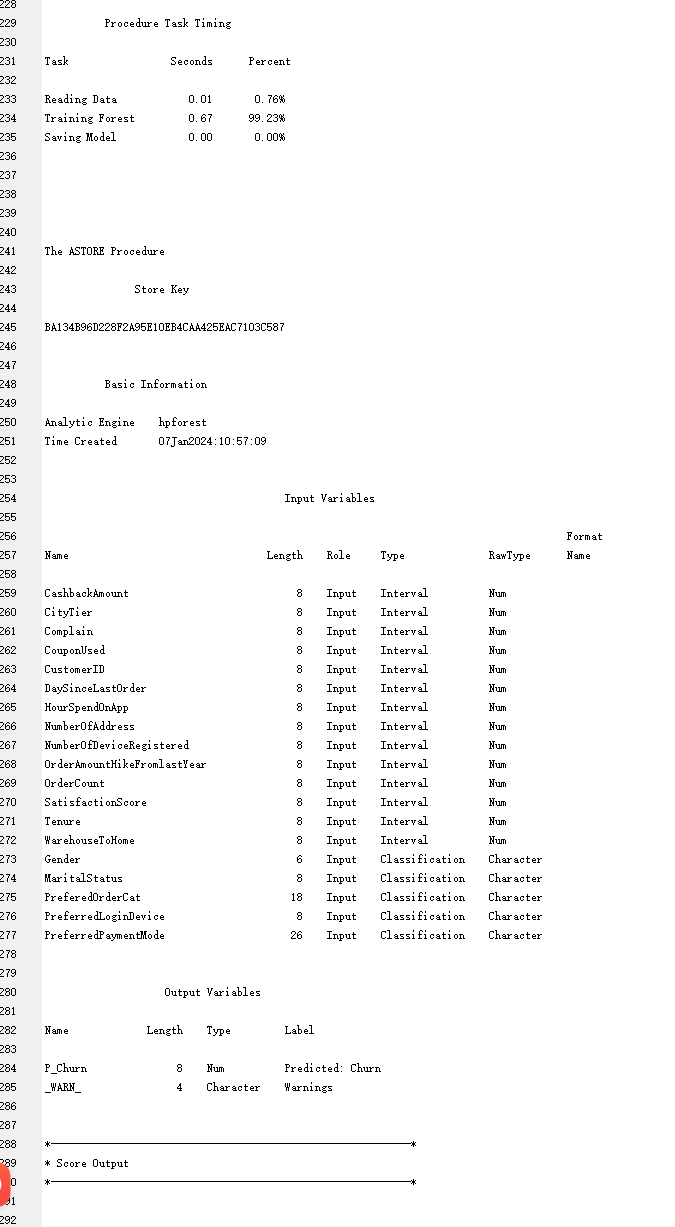


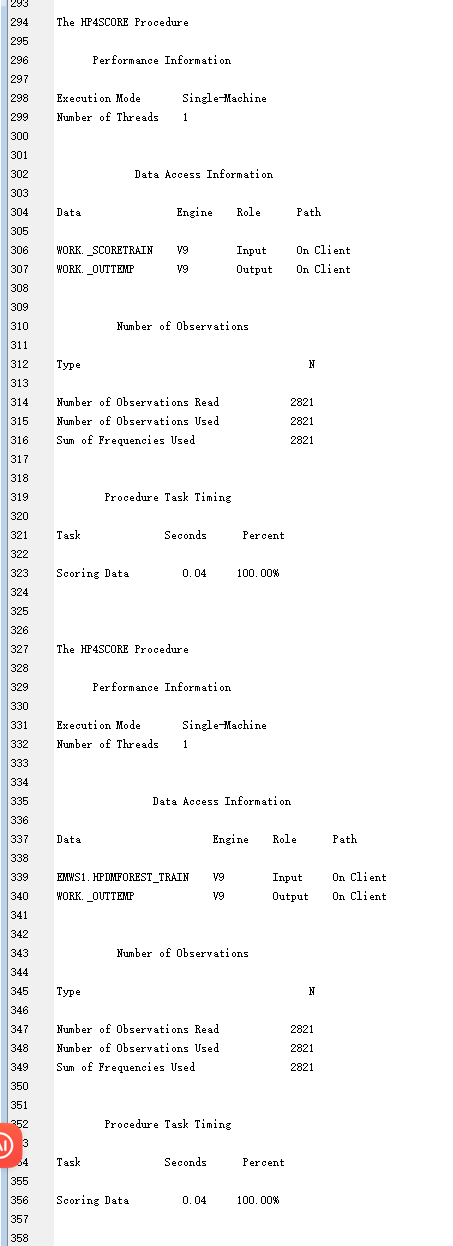


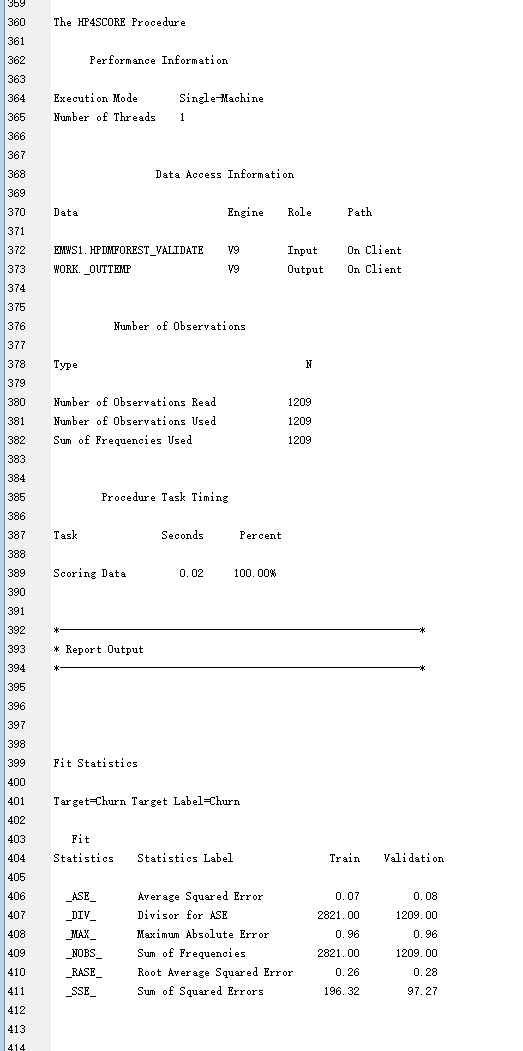


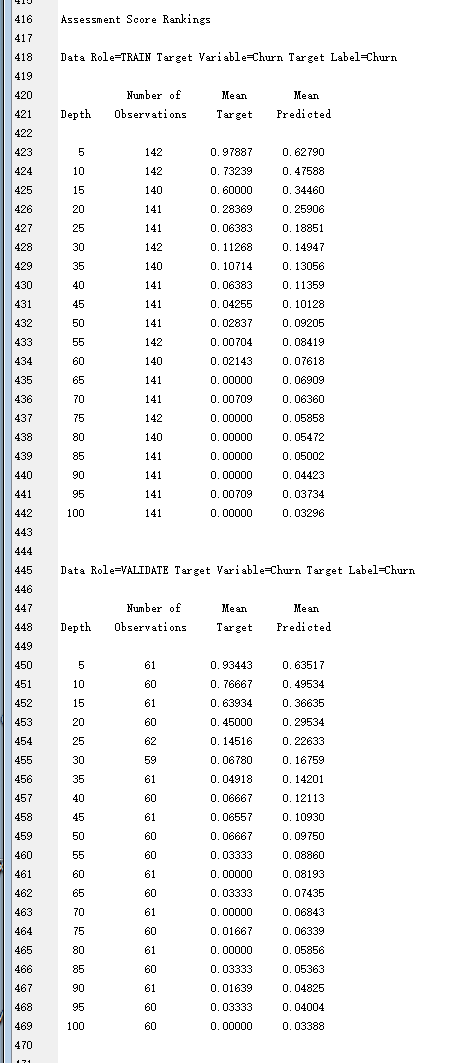


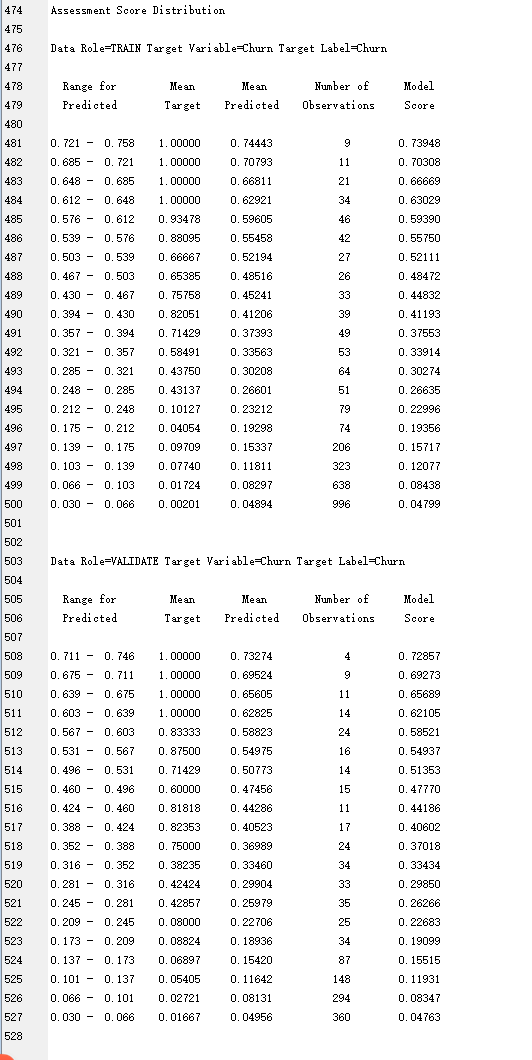




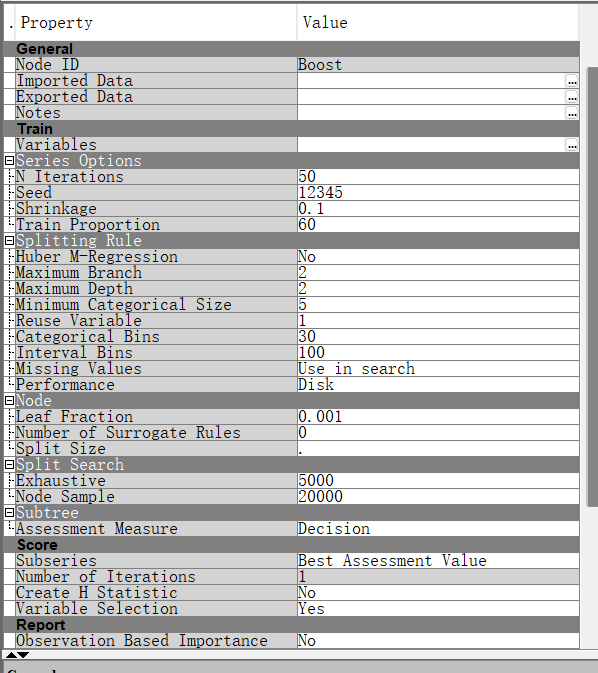


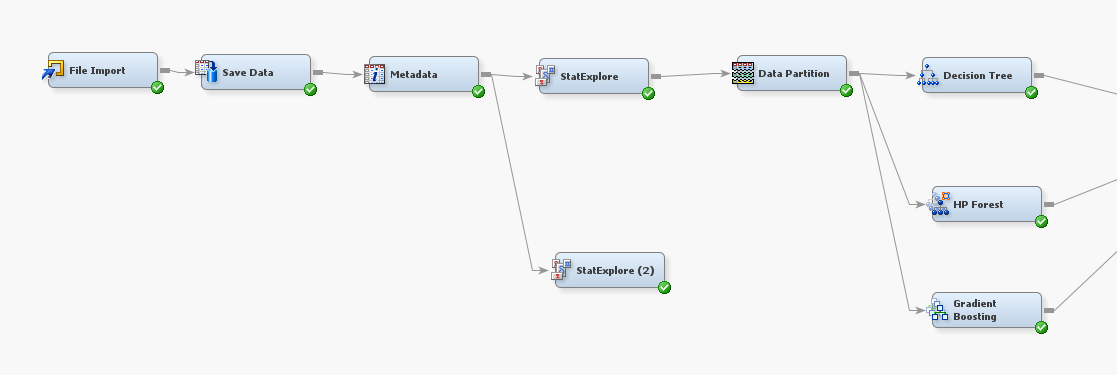


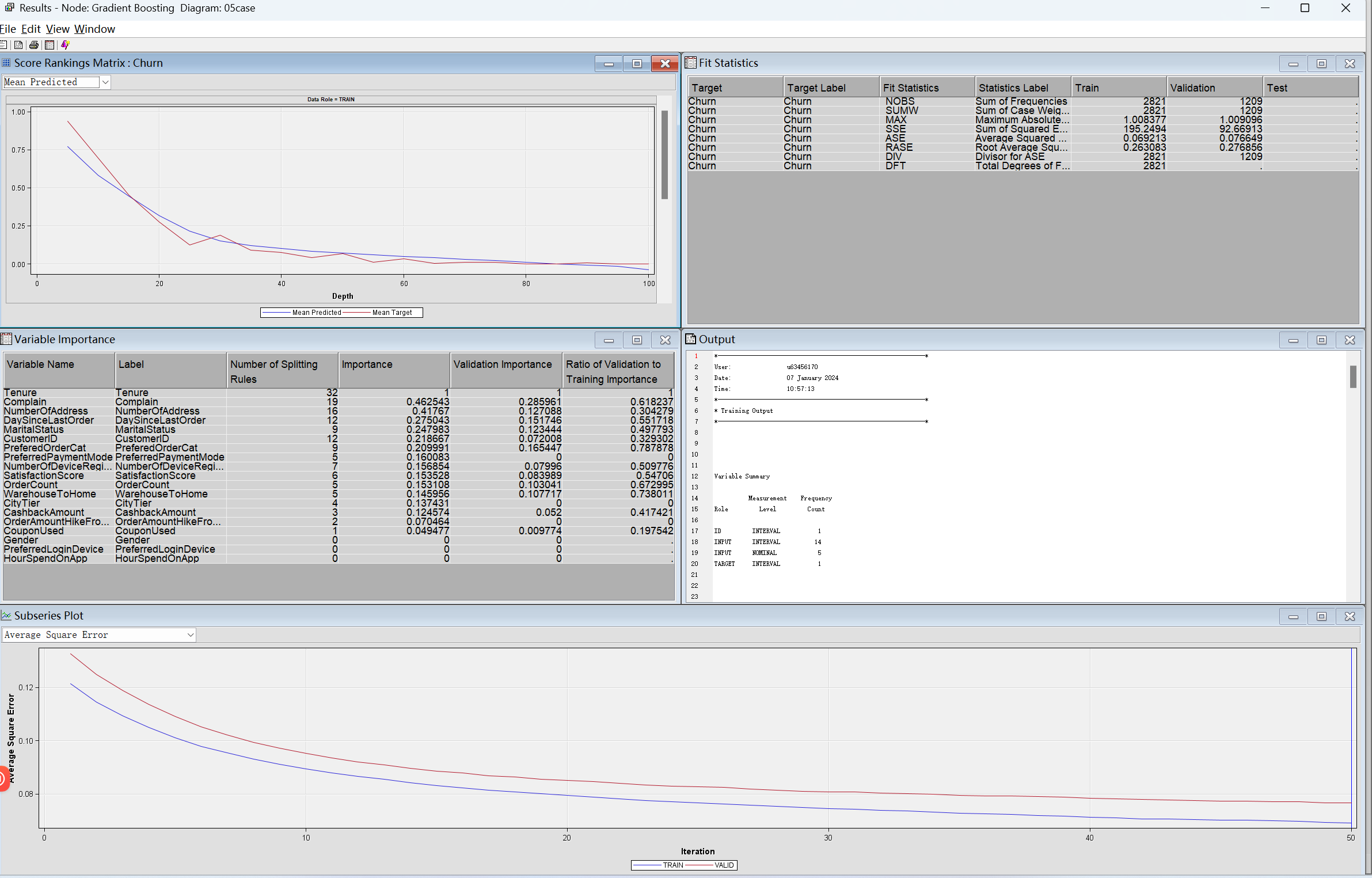


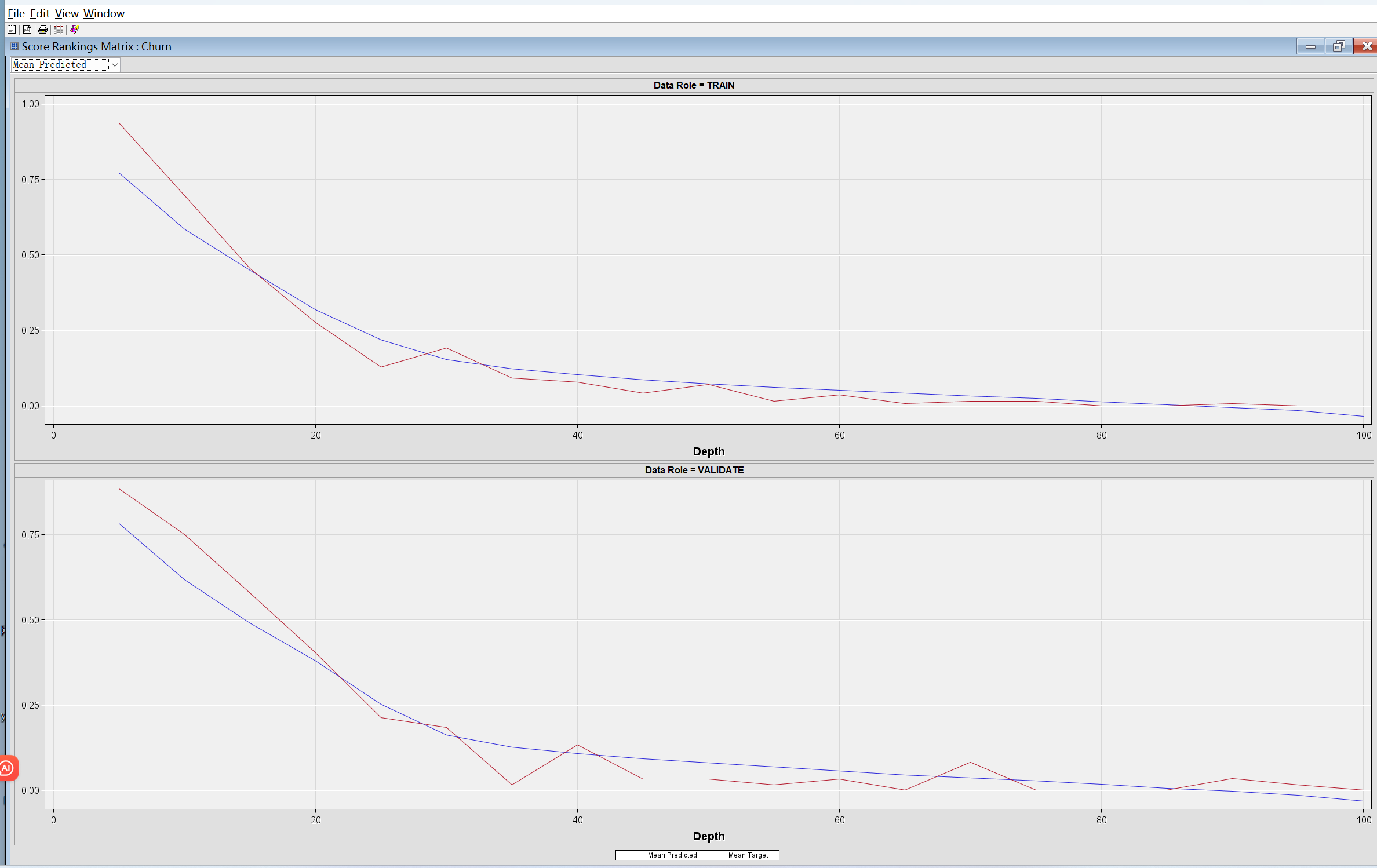


Run gradient boosting

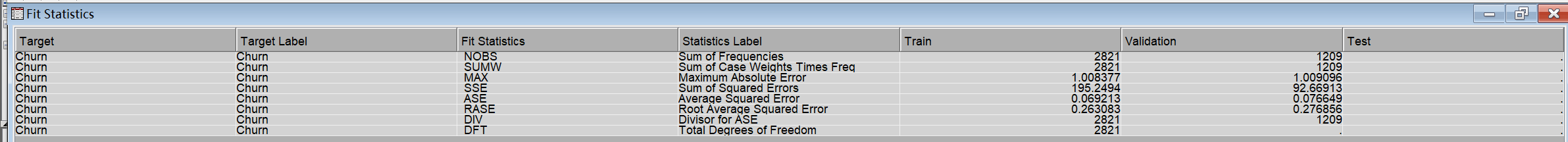


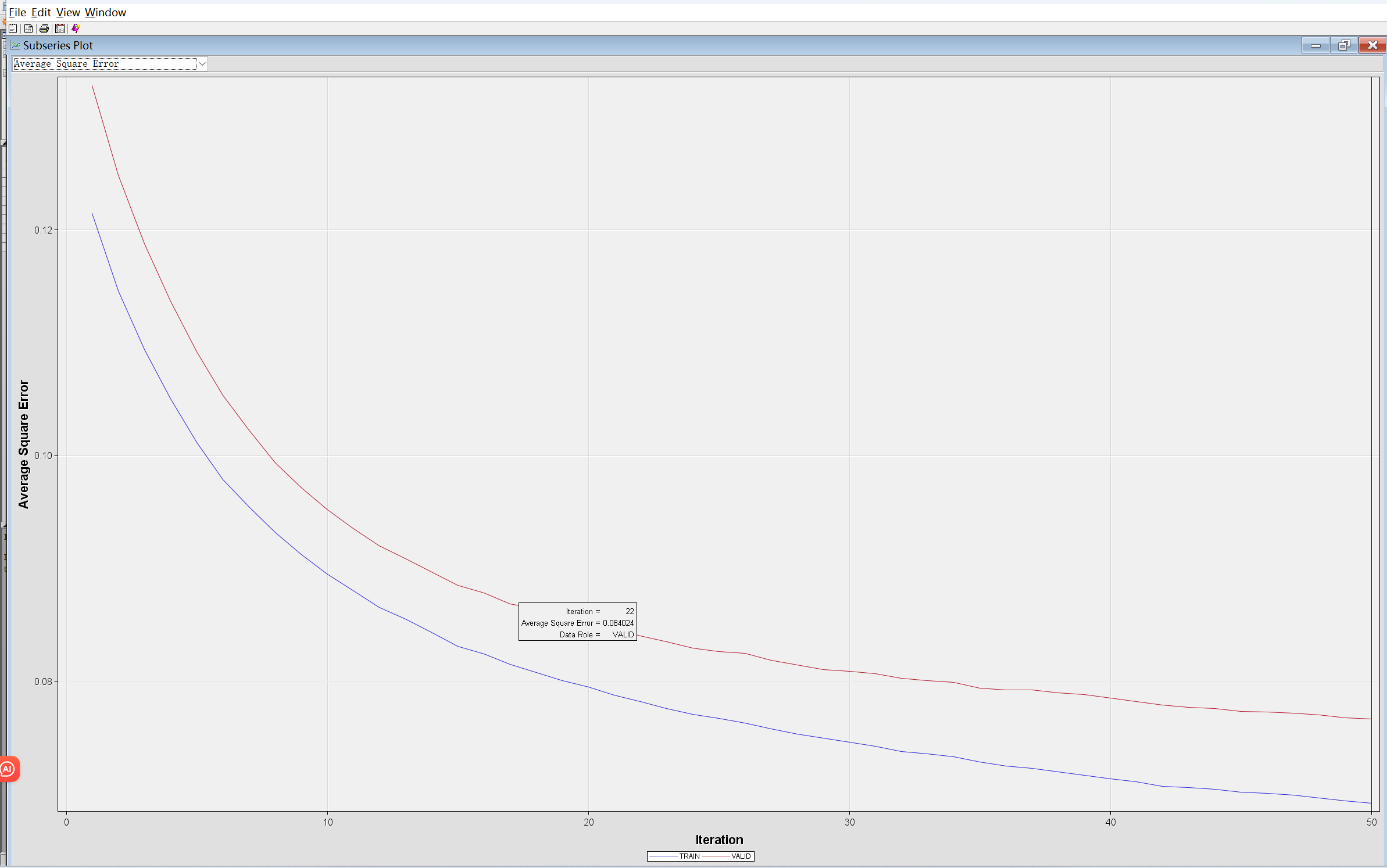




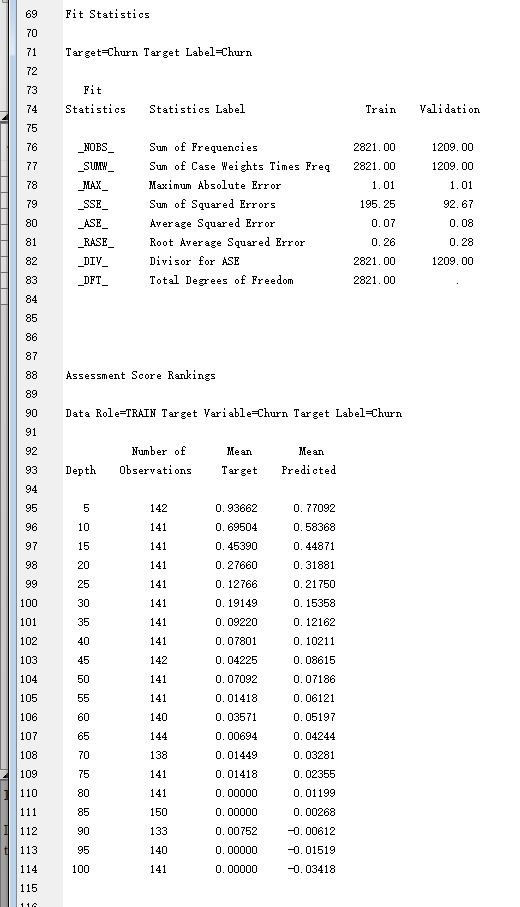


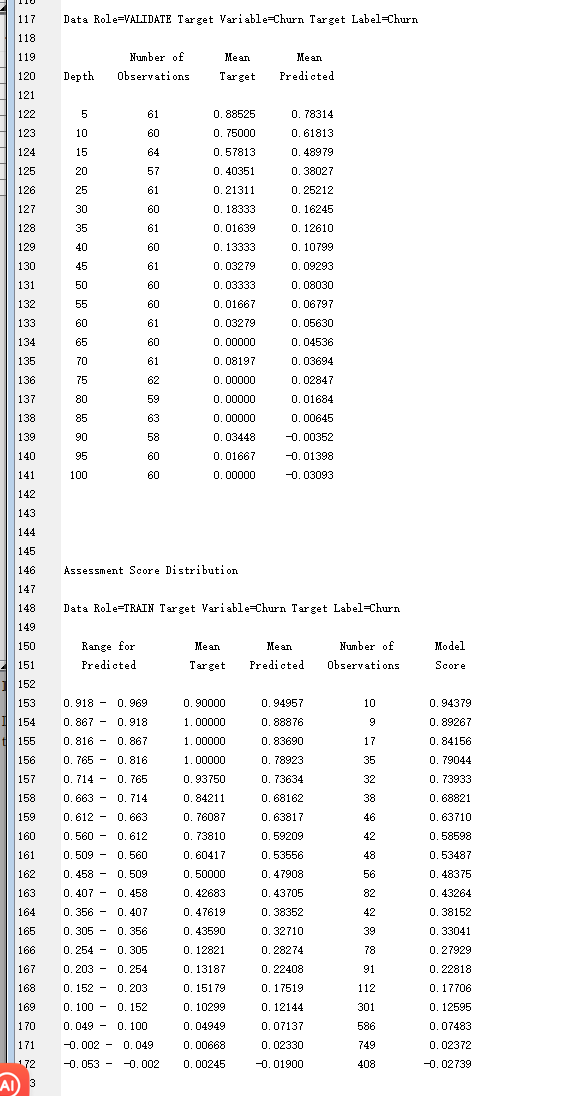


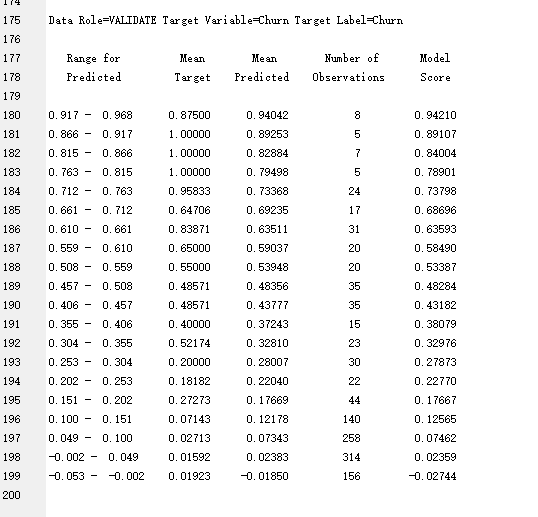




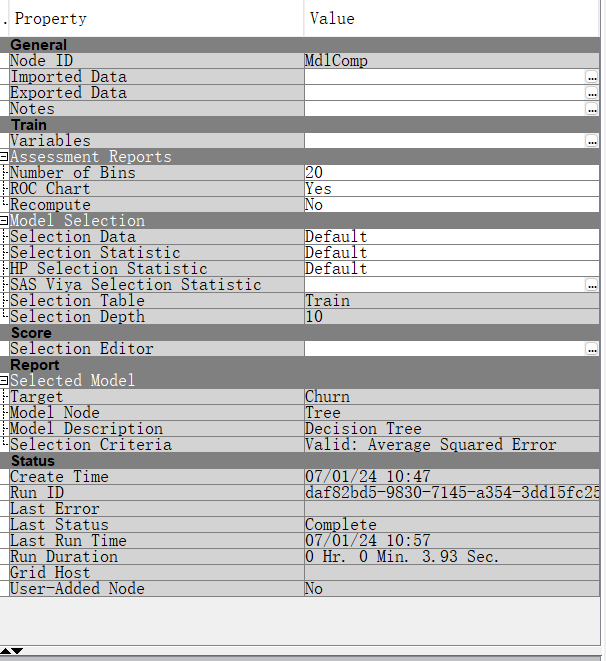


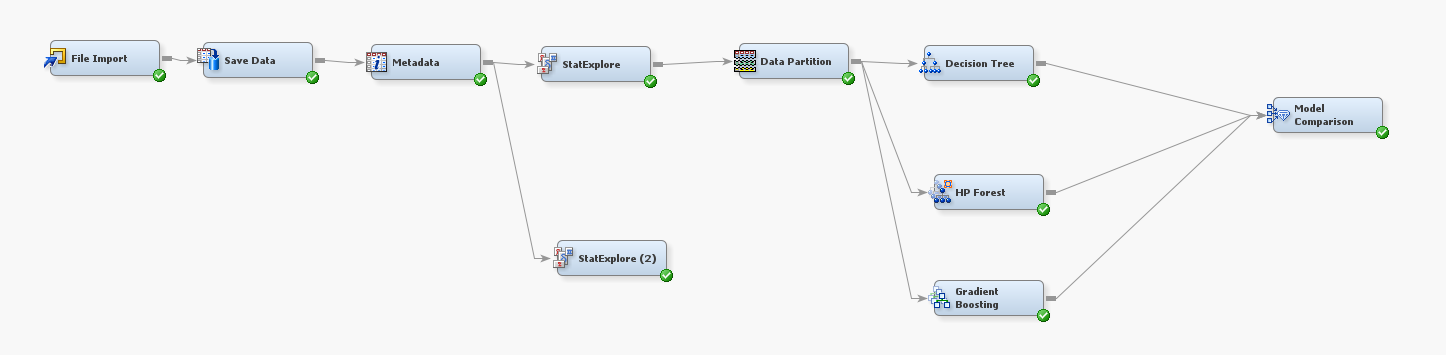


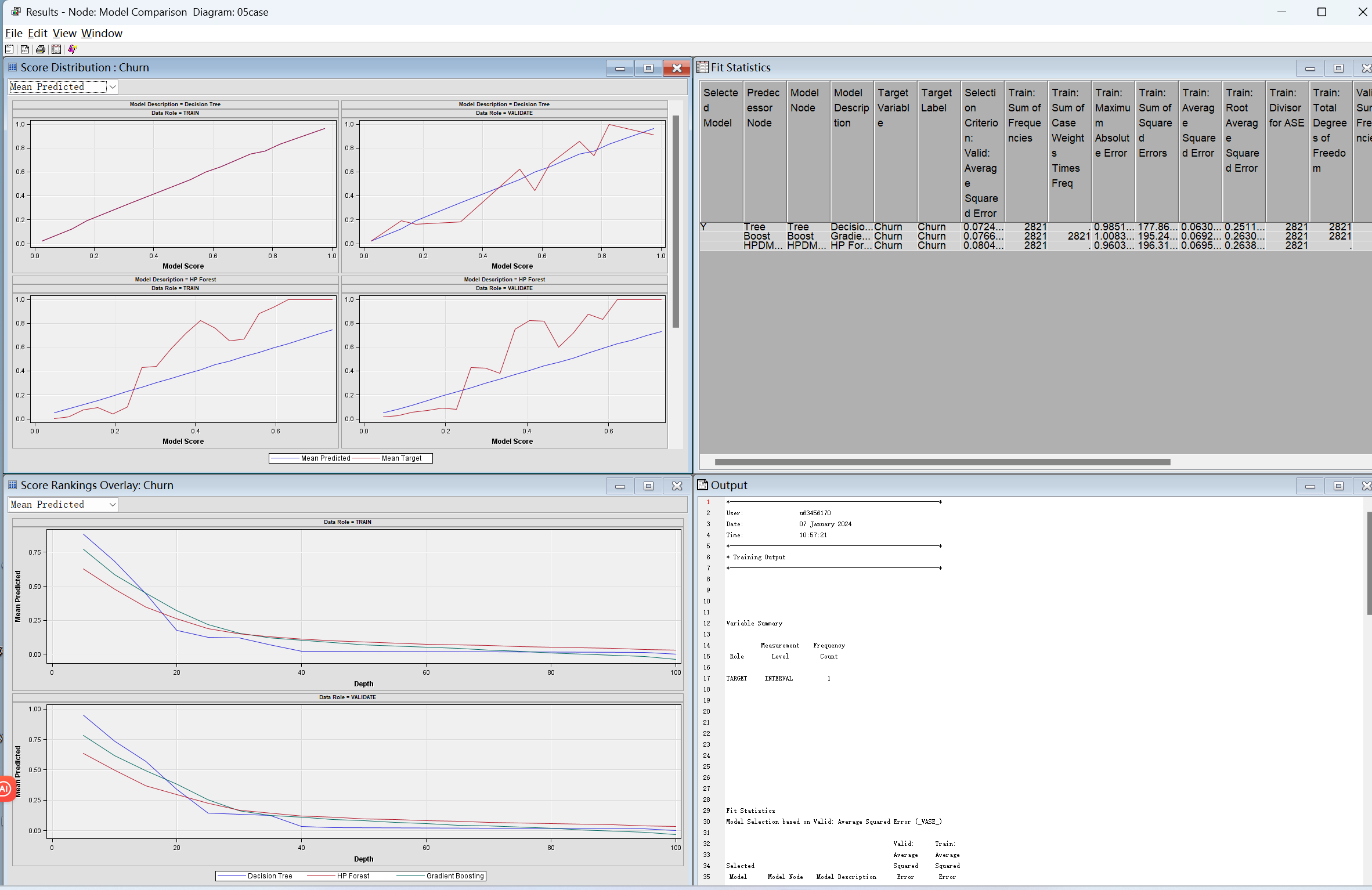


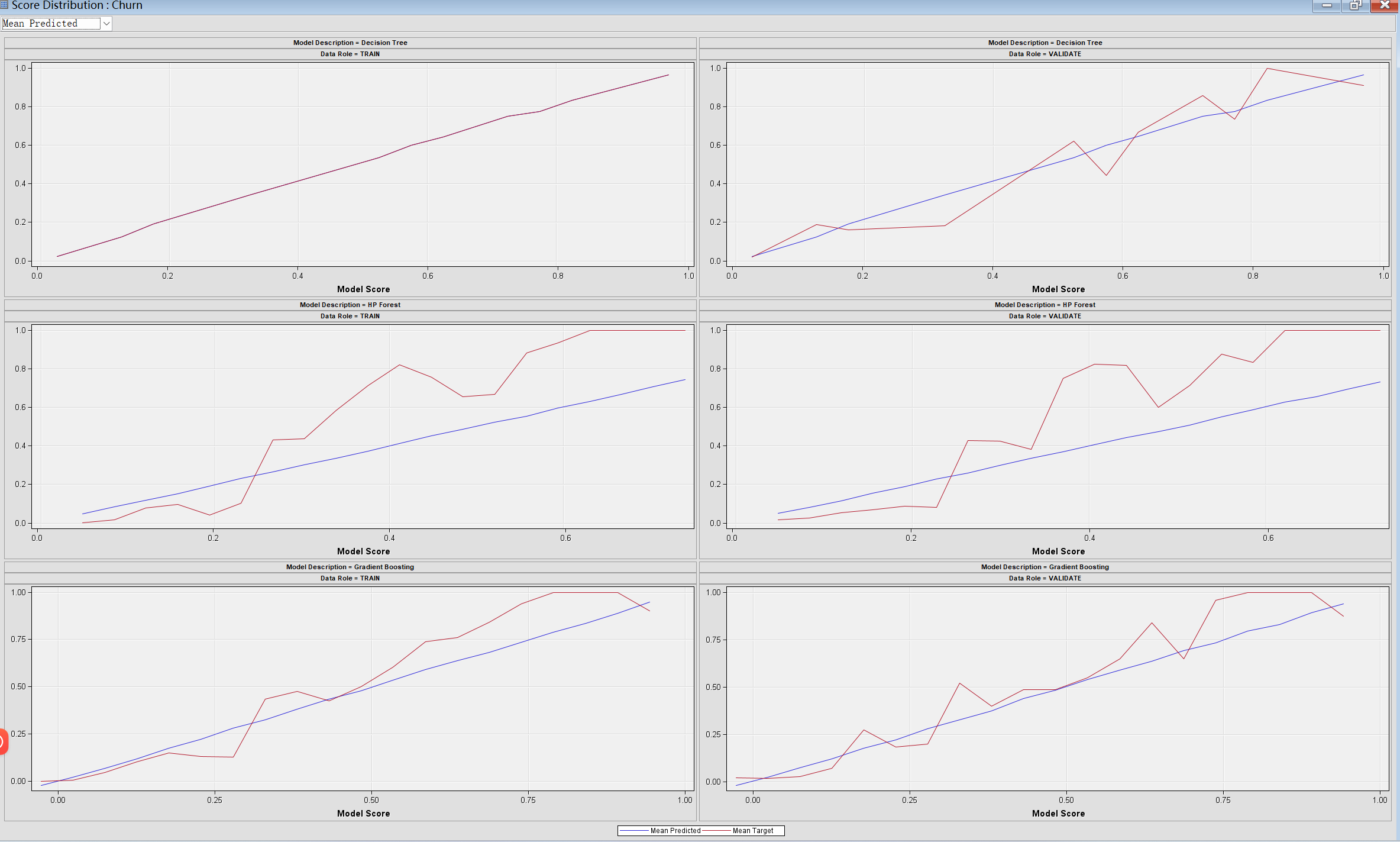


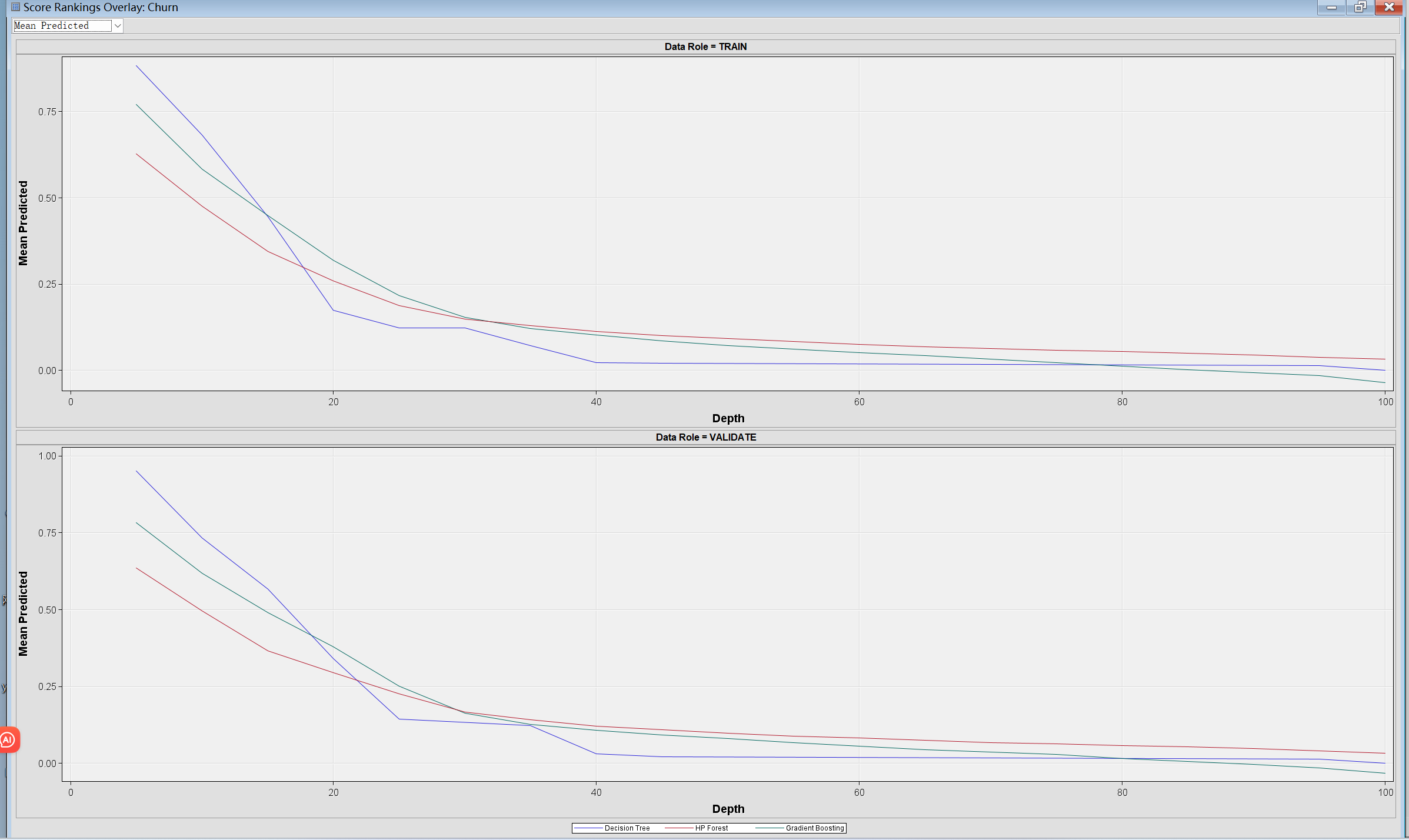
Run model comparison:

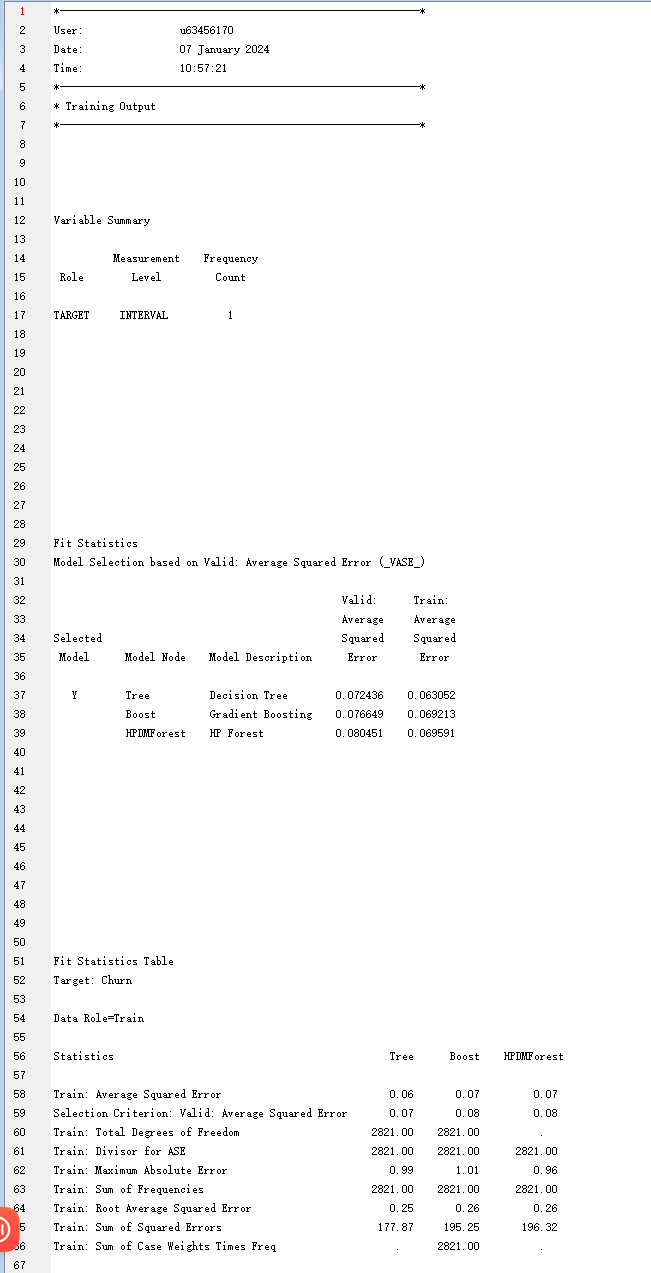


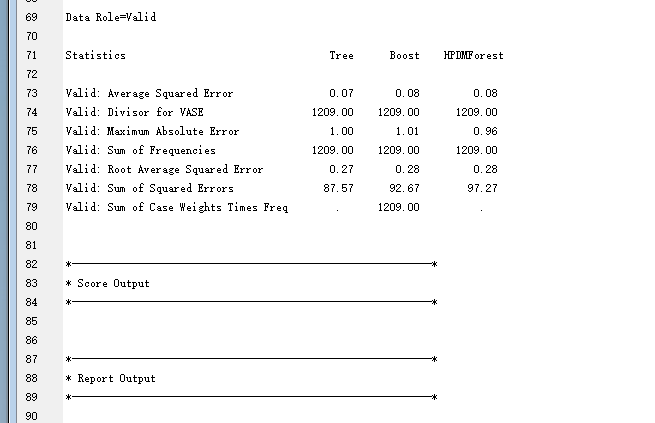


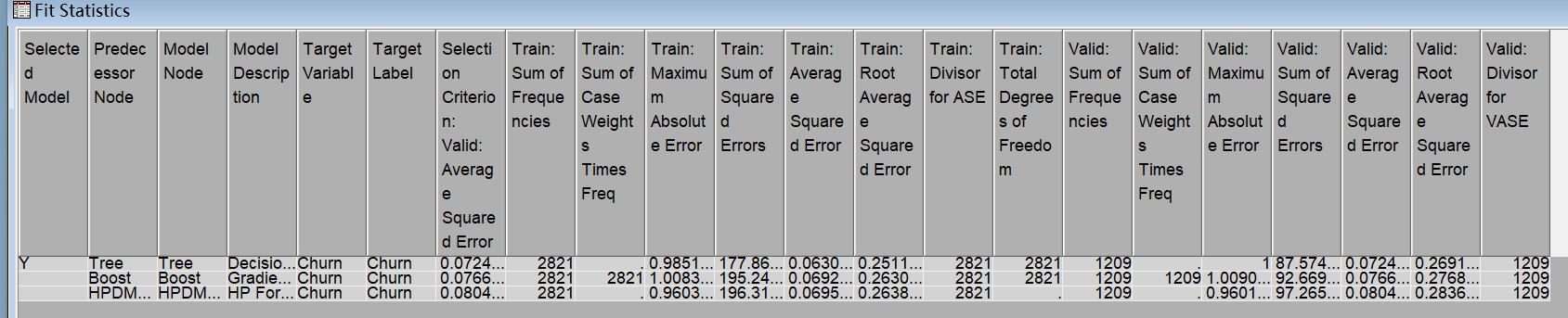












Here are three models (Decision tree, Boost, HPDM) and this table shows the statistics of three models for training, validation, and testing.

The Decision Tree has the lowest RASE and SSE, suggesting better predictive accuracy compared to the other two models.

The Boosting model has a slightly higher RASE and SSE compared to the Decision Tree.

The HPDM, which may involve hyperparameter tuning, has the highest RASE among the three models, indicating a higher average prediction error.

Advantage of the project

By delving into the relationships between these variables, we can better understand the factors contributing to customer churn, enabling the development of targeted marketing strategies, enhancing customer loyalty, and ultimately optimizing the organization's business performance.

Insights into customer behavior and suggestion for business strategy.

The analysis of the decision tree and ensemble models, particularly in the context of customer behavior, yields valuable insights for strategic business decisions. The Decision Tree model, with a low Root Average Squared Error (RASE) and Sum of Squared Errors (SSE), provides a solid foundation for understanding factors influencing customer churn and loyalty. Key features such as "Tenure," "Preferred Login Device," "Satisfaction Score," and more play crucial roles in the decision-making process. The Boosting model, while showing slightly higher RASE and SSE, uncovers nuanced patterns in customer behavior, offering additional depth to the analysis. On the other hand, the HPDM, likely a hyperparameter-tuned Decision Tree, exhibits higher RASE, suggesting increased complexity. Strategic recommendations include prioritizing insights from the Decision Tree, leveraging the nuanced findings from the Boosting model, and carefully evaluating the benefits of hyperparameter tuning. Businesses can optimize customer retention, tailor promotional strategies based on "Coupon Usage" and "Order Frequency," and implement targeted engagement approaches considering "Days Since Last Order."

In summary, a comprehensive analysis of the decision tree and ensemble models provides actionable insights for businesses to enhance customer retention strategies, optimize promotional efforts, and improve overall customer engagement.