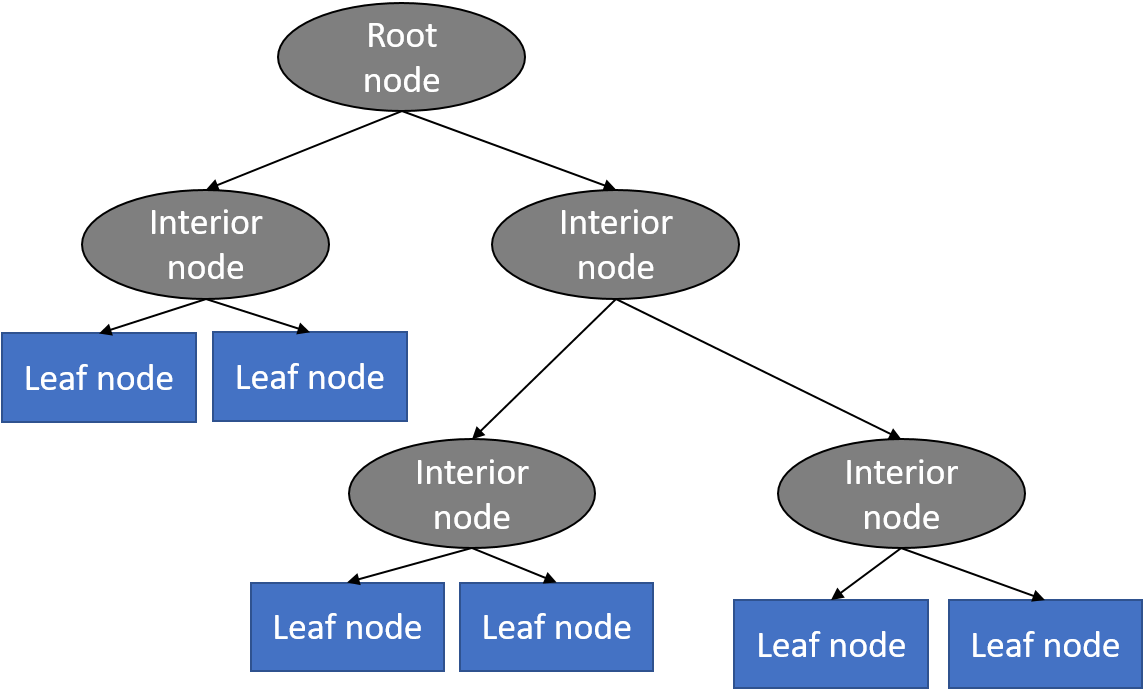
**What is Regression Tree**

Regression trees are a type of decision tree algorithm used for regression analysis, aiding with visualizations. The benefit of a regression tree over a linear regression model is that regression trees can ignore the assumptions that a linear regression model has. This is because they are non-linear, non-parametric models that are more robust than a traditional linear regression model. The result of a regression tree is a tree-like structure that shows the predictor variables (numerical or categorical) used to optimize a continuous outcome variable. Regression trees offer a flexible and interpretable approach for regression analysis, providing insights into the relationships between predictors and the response variable. They can be applied to various domains and are particularly useful when dealing with non-linear relationships or when interpretability is desired.

**How Regression Trees are Executed**

Regression trees are executed with splits of the nodes. There are three different nodes in a regression tree: the root node, the internal node, and the leaf nodes. The process starts with the root node and is the topmost node of the tree. It represents the entire dataset or a subset of it, and it's split into internal nodes based on the best feature to separate the data into more homogeneous groups. At each step, the algorithm selects the best variable and splits point by the smallest sum of squares residual to create two child nodes. This process continues recursively until a stopping criterion is met, such as a maximum tree depth or a minimum number of data points per terminal node. This results in the leaf nodes. Leaf nodes are the endpoints of the decision tree, and they represent the final predicted outcome or value. A leaf node contains the predicted class label in a classification tree or the predicted value in a regression tree. Each leaf node corresponds to a specific decision or prediction made based on the path taken through the tree. This is a significant advantage of regression trees, making the model easy to interpret. When leveraging the model to make a prediction, you follow the nodes down to where the data point lands.



**Final Thoughts**

Regression trees are easy to understand and interpret, making them popular for data exploration and visualization. However, they can suffer from overfitting if the tree is too complex or if the data is noisy.

Reading List

1. <https://medium.com/analytics-vidhya/regression-trees-decision-tree-for-regression-machine-learning-e4d7525d8047>
2. <https://www2.stat.duke.edu/~rcs46/lectures_2017/08-trees/08-tree-regression.pdf>
3. <https://www.youtube.com/watch?v=g9c66TUylZ4>
4. <https://pages.stat.wisc.edu/~loh/treeprogs/guide/wires11.pdf#:~:text=Classi%EF%AC%81cation%20and%20regression%20trees%20are%20machine-learning%20methods%20for,can%20be%20represented%20graphically%20as%20a%20decision%20tree>.
5. <https://towardsdatascience.com/the-only-guide-you-need-to-understand-regression-trees-4964992a07a8>
6. <https://www.datacamp.com/tutorial/decision-trees-R>
7. <https://www.projectpro.io/recipes/build-classification-trees-r>
8. <https://www.statology.org/classification-and-regression-trees-in-r/#:~:text=Example%201%3A%20Building%20a%20Regression%20Tree%20in%20R,4%3A%20Use%20the%20tree%20to%20make%20predictions.%20>
9. <https://www.nature.com/articles/nmeth.4370>
10. <https://www.digitalvidya.com/blog/classification-and-regression-trees/>

Group Members:

Logan Barger

Joseph Beightol

Aaron Brommers

Kurt Meyer

Matt Ostendorf