Regression Trees

Not purely functional, but functional enough!

What are they?

- Explain & predict outcomes
- Building the tree
  - Detect Variance and purify by recursively partitioning data until more homogenous than parent nodes
- Find relationships between input values and target values
  - When the value of the input improves the ability to predict the target value, a strong relationship is formed
- Continuous Variables - mean, std dev, variance in each node
- Terminal Nodes - numerical values

What can you do with them?

- Data Mining
  - Which customer group brings in the most revenue?
- Machine Learning
  - Random Forests - A collection of trees that take inputs, randomly sample the data, and reweigh the trees to optimize predictions.
  - How should we modify our portfolio of products and services to maximize revenue?

The Problem

- **Santa's been slacking and needs to figure out how many and which kinds of toys he needs to make before the big day!**

```
children :: [Child]
  children =
    [ Child {
        isGood = .75
      , age = 7
      , location = (52.52, 13.41)
      ,  height = 1.35
      , weight = 37
      }
    ]
```

The Solution

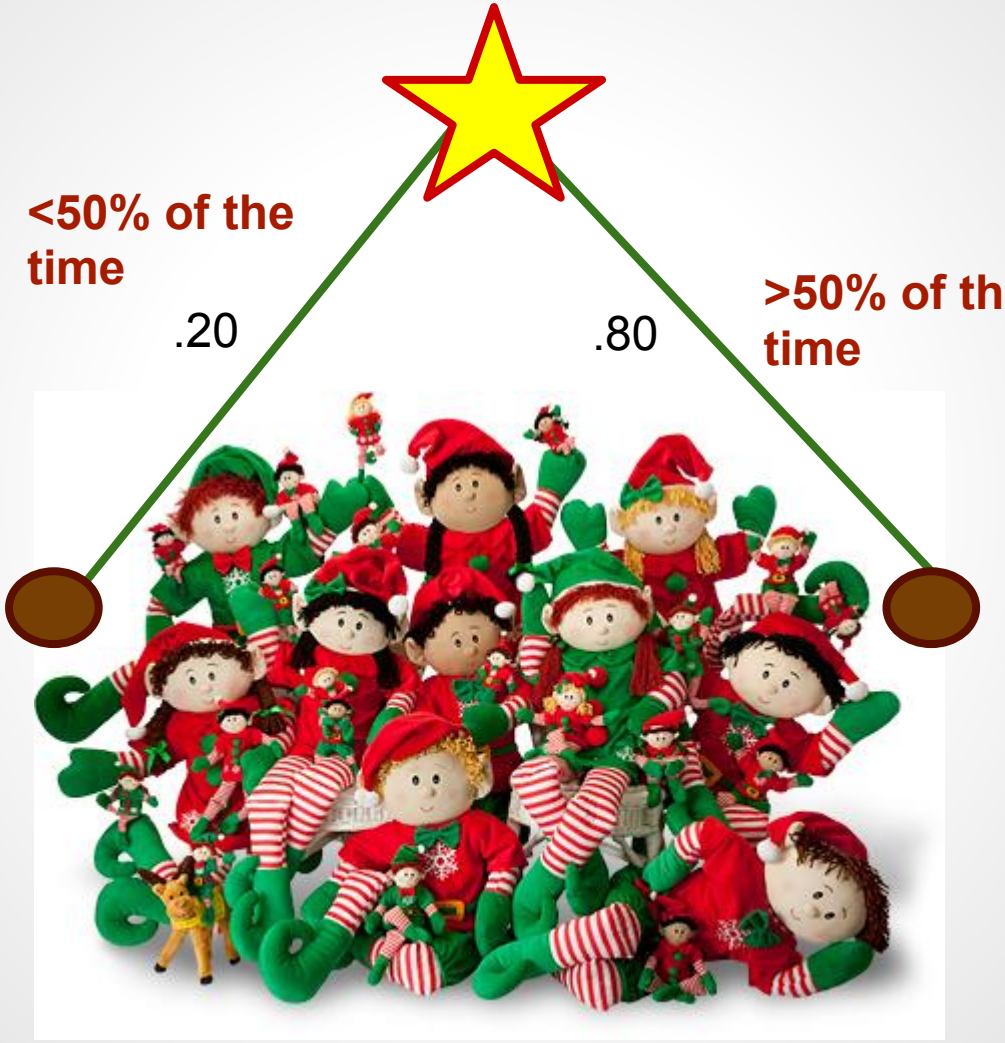
Take each record, map its information to its respective categories and value. Take the result, and map it over the entire list of data. Then, starting with an empty tree, apply a foldl over the list in order to build up the tree.

```
data RegressionTree = RegressionTree Category Value
[RegressionTree] deriving (Show)

makeRegressionTree :: [a] -> [a -> Category] -> (a -> Value)
-> RegressionTree

makeRegressionTree xs catFuncs valueFunc =
  let catsAndValues = map (\x -> (map ($ x) catFuncs,
valueFunc x)) xs in
    foldl (\rtree (cats, value) -> add (rootCategory:cats)
value rtree) empty catsAndValues
```

Santa's Workshop



Result of printing the tree: All: value=3B: likelihood=1: Good: value=2.4B: likelihood=0.8: Bad: value=600M: likelihood=0.2

Growing the Tree

```
add :: [Category] -> Value -> RegressionTree -> RegressionTree
add (c1:c2:cs) v (RegressionTree c' v' xs)
    | c1 == c' = RegressionTree c1 (v + v') $ map (add (c2:cs)
v) (maybeAddChild c2 xs)
add (c1:[]) v (RegressionTree c' v' xs)
    | c1 == c' = RegressionTree c1 (v + v') xs
add [] _ _ = error "No categories"
add _ _ rtree = rtree
```

Starting at the top of the tree, the function goes down the tree adding values to categories along the way.

1. If category doesn't exist, add it and place the value in that category
2. If terminal node, then add the value to that category
3. Error
4. Just return the tree

To Add or Not to Add...

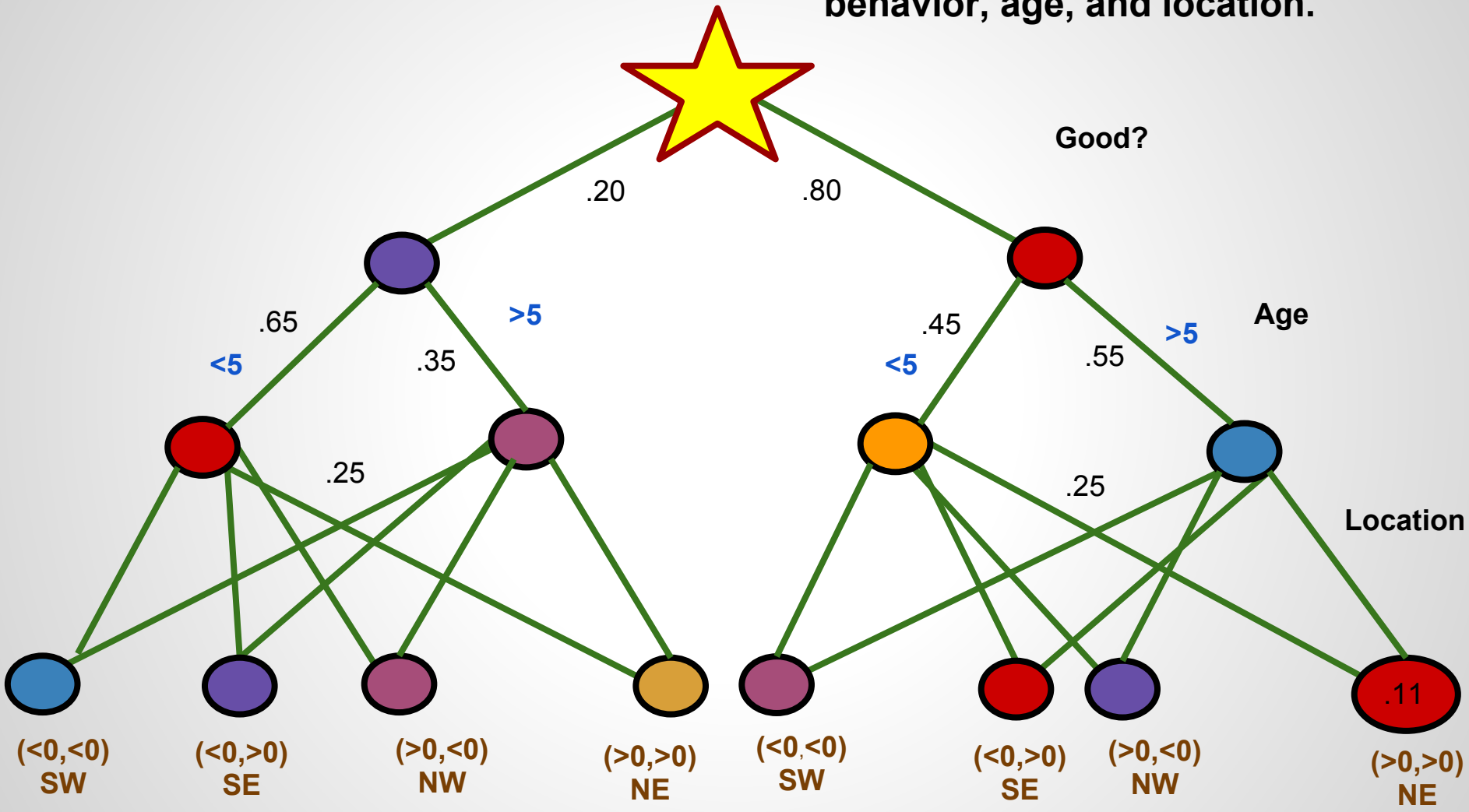
Check if supplied category matches existing category. If true, then add category to tree as terminal node with corresponding value of zero.

```
maybeAddChild :: Category -> [RegressionTree] ->
[RegressionTree]

maybeAddChild c xs
  | not $ any (\(RegressionTree c' _ _) ->   c ==
c') xs = RegressionTree c 0 [] : xs
  maybeAddChild _ c xs = xs
```

Santa's Workshop

Suppose gifts are gender neutral: iPhone, Laptop, coal, rocks, mittens, socks. 16 items depending on behavior, age, and location.



Ex. 3 billion kids => Santa should make 330 million iPhones to take to kids who are good >50% of the time, older than 5, living in the Northeastern hemisphere.

Things to be careful about

- Splitting rule: When subsequent split results in little overall improvement in predictions, stop
- In the case of ML
  - Overfitting
    - Noisy Data
    - Incorrect Predictions


MERRY CHRISTMAS
swedish
GOD JUL
norwegian
slovene
VESEL BOŽIČ
croatian
sicilian
BON NATALI
JOYEUX NOËL
french
albanian
slovak
LINKSMŲ KALĖDŲ
lithuanian
swiss german
FROHE WEIHNACHT
portuguese
FELIZ NATAL
urdu
KRISMAS MUBARAK
chinese
finnish
圣诞快乐
HYVÄÄ JOULUA
polish
danish
spanish
WESOŁYCH ŚWIĄT
FELIZ NAVIDAD
english (uk)
NADOLIG LLAWEN
HAPPY CHRISTMAS
welsh
german
italian
ЩАСЛИВОГО РІЗДВА
BUON NATALE
MALIGAYANG PASKO
japanese
filipino
メリークリスマス
SHNORHAVOR SOORP TSANOONO
dutch
armenian
GELUKKIG KERSTFEEST
greek
irish
ΚΑΛΑ ΧΡΙΣΤΟΥΓΕΝΝΑ
NOLLAIG SHONA DHUIT
turkish
NOELINIZ KUTLU OLSUN
СЧАСТЛИВОГО РОЖДЕСТВА
russian
romanian
hungarian
bulgarian
ВЕСЕЛА КОЛЕДА
german
FROHE WEIHNACHTEN