Week 5: Decision Trees. Part 2: Fuzzy inference of DT

Seve Martinez

Grand Canyon University

Abstract

Decision tree boundaries can appear unfair or unjust due to their harsh cutoffs. This can be rectified by blurring the boundaries between classes using fuzzy logic. By creating triangular or trapezoidal spaces using IF-THEN statements, the boundaries can be much more forgiving and more human-like in their application.

Keywords: decision trees, fuzzy logic

Week 5: Decision Trees. Part 2: Fuzzy inference of DT

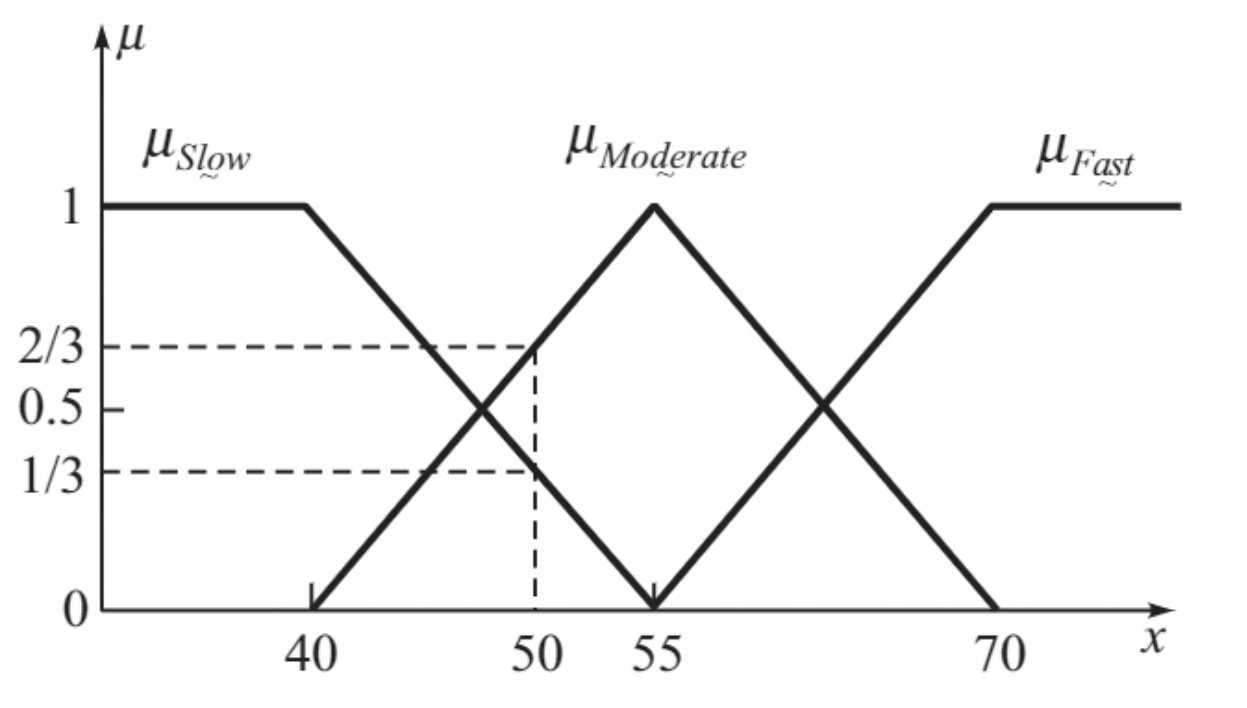
Decision tree boundaries can appear unfair or unjust due to their harsh cutoffs. This can be rectified by blurring the boundaries between classes using fuzzy logic. By creating triangular or trapezoidal spaces using IF-THEN statements, the boundaries can be much more forgiving and more human-like in their application.

Fuzzify Attributes

Each attribute has a crisp linguistic variable associated with it. For example, wind has two: Strong and Weak. While this may be understandable to humans, it ultimately creates a vagueness for machines. However, it is known that internally, humans subconsciously pass through a series of IF-THEN statements to decide between the two. Instead of returning a particular class, instead we seek to return a value between 0 and 1 that leans toward one class or the other.

X1 - Temperature

Temperature has three possible values: hot, mild and cool. These can be separated through temperatures rather than linguistic verbiage. In order to do this, a good way would be to use continuous variables for temperature and divide evenly. Gopal provides an example of this for speed:



Assuming a temperature set of X = [0, 100] F then,

Temp = 50

μcool = 1/3

μmild = 2/3

μhot = 0

Indicating that the third fuzzy set does not have an end point so that we know for sure that anything above is the proper class. This is a triangular style of inference.

X2 – Wind

Wind only has two options: Weak and Strong. In order to fuzzify this, a catesian product of the two crisp sets will be used.

This will be much simpler and we can use a singleton set (Gopal, 2020).

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{otherwise} \end{cases}$$

This will take the form of a an impulse function.

X3 – Traffic jam

This will also be a singleton function as represented above.

Membership Functions

In order to implement the fuzzy logic for each variable, membership functions must be built to apply the logic. These would take the form of IF-THEN statements that will result in the membership of a combination of X_j values. These functions will output a value in the interval [0, 1] which will assist in determination of the class.

$$\tilde{A}_{1k_1} = \{(x_1, \mu_{\tilde{A}_{1k_1}}(x_1)) \mid x_1 \in X_1, \mu_{\tilde{A}_{1k_1}} \in [0,1]\};\ k_1 = 1, 2, 3$$

$$\tilde{A}_{2k_2} = \{(x_2, \mu_{\tilde{A}_{2k_2}}(x_2)) \mid x_2 \in X_2, \mu_{\tilde{A}_{2k_2}} \in [0,1]\};\ k_2 = 1, 2$$

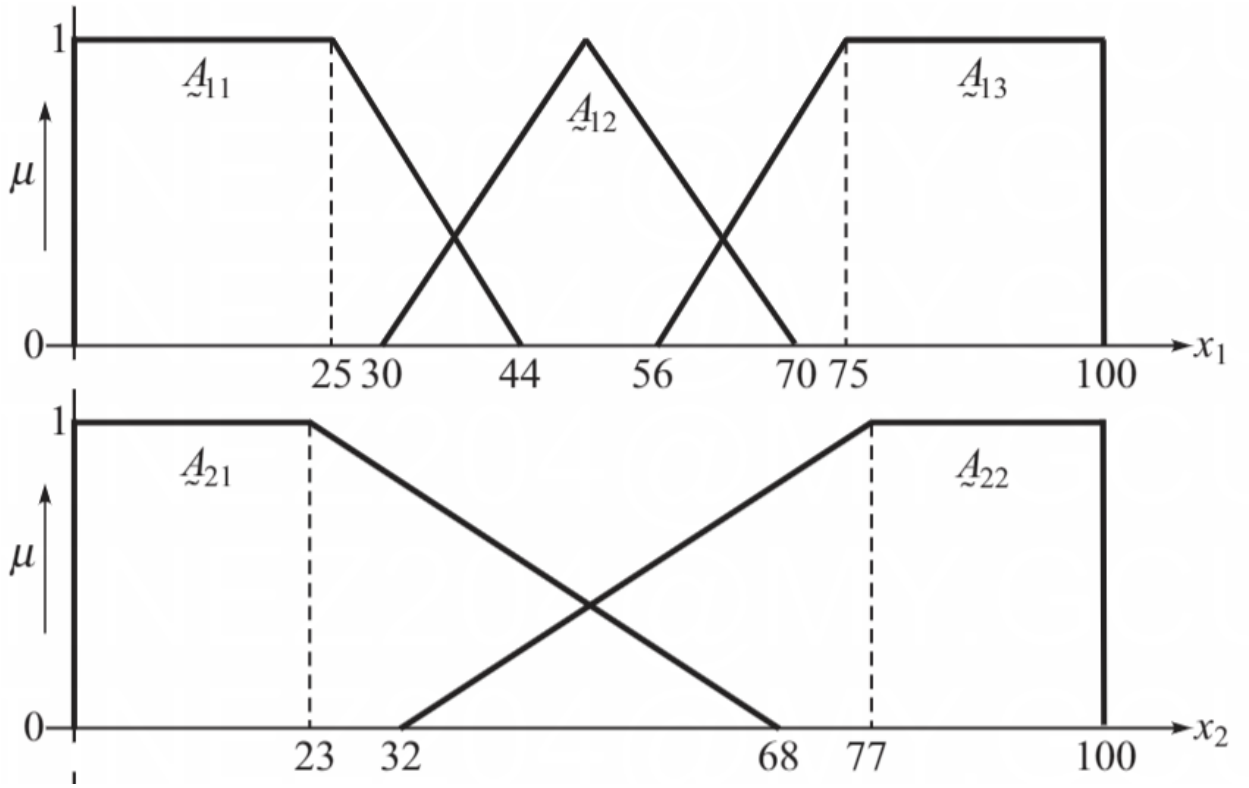
$$\tilde{B}_k = \{(y, \mu_{\tilde{B}_k}(y)) \mid y \in Y, \mu_{\tilde{B}_k} \in [0,1]\};\ k = 1, 2, 3$$

(Gopal, 2020)

Where the first two ‘A’ values are the variables Temperature and Wind, and ‘B’ is drive = yes/no (k = 1, 2 in this case).

IF TEMP = ‘HOT’ AND WIND=’COOL’ AND ‘TRAFFIC-JAM’ = ‘SHORT’ THEN = y = ‘yes’.

Gopal also provides a graph that will mirror our efforts:



The first graph represents temperature and the second represents traffic jam and wind. Values are shown in 0 – 100%.

While we do not need rules for every combination of sets, in this small case, they will exist in completion.

Tree Map

The same split was chosen for brevity. Laid out with the new rules based on the prior logic, the tree will look as the following:

|Wind|

<Strong>

|Temperature|

<Hot>

[Driving](0.86)

<Mild>

|Traffic-jam|

<Long>

[Driving](0.85)

<Short>

[Driving](0.82)

<Cool>

[Driving](0.93)

<Weak>

|Traffic-jam|

<Short>

[Driving](0.90)

<Long>

|Temperature|

<Hot>

[Driving](0.89)

<Mild>

[Driving](0.64)

<Cool>

[Driving](0.89)

Respective Rules for the tree are below:

IF Wind IS Strong AND Temperature IS Hot THEN Driving (0.86)

IF Wind IS Strong AND Temperature IS Mild AND Traffic-jam IS Long THEN Driving (0.85)

IF Wind IS Strong AND Temperature IS Cool AND Traffic-jam IS Short THEN Driving (0.82)

IF Wind IS Weak AND Traffic-jam IS Short THEN Driving (0.93)

IF Wind IS Weak AND Traffic-jam IS Long THEN Driving (0.90)

IF Wind IS Weak AND Traffic-jam IS Short THEN Driving (0.89)

IF Wind IS Weak AND Traffic-jam IS Long AND Temperature IS Hot THEN Driving (0.89)

IF Wind IS Weak AND Traffic-jam IS Long AND Temperature IS Mild THEN Driving (0.64)

IF Wind IS Weak AND Traffic-jam IS Long AND Temperature IS Cool THEN Driving (0.89)

Conclusion

By leveraging a fuzzy logic based on an interval system, it is possible to create soft boundaries for trees that would prevent unjust cutoffs. The three input variables were fuzzified using a variety of techniques and then those were applied to IF-THEN statements that showed it was possible to create a more “human-like” tree.

References

Baldwin, James & Xie, Dong. (2004). Simple Fuzzy Logic Rules Based on Fuzzy Decision Tree for Classification and Prediction Problem.. 175-184. Last Name, F. M. (Year). *Book Title.* City Name: Publisher Name.

Banakar, A., Zareiforoush, H., Baigvand, M., Montazeri, M., Khodaei, J., & Behroozi, K. N. (2017). Combined Application of Decision Tree and Fuzzy Logic Techniques for Intelligent Grading of Dried Figs. *Journal of Food Process Engineering*, *40*(3), n/a-N.PAG. https://doi-org.lopes.idm.oclc.org/10.1111/jfpe.12456

Jooyeol Yun, Jun won Seo, and Taeseon Yoon (2014) THE NEW APPROACH ON FUZZY DECISION TREES International Journal of Fuzzy Logic Systems (IJFLS) Vol.4, No.3, July 2014. https://arxiv.org/abs/1408.3002

Olaru, C., Wehenkel, L. (2003). A complete fuzzy decision tree technique. Elsevier. https://services.montefiore.uliege.be//stochastic/pubs/2003/OW03/OW03.pdf