

# ENVI Tutorial: Decision Tree Classification

---

Decision Tree Classification .....	2
Files Used in this Tutorial .....	2
Background .....	2
Decision Tree Input .....	2
Displaying Images for Decision Tree Classification .....	4
Entering Decision Tree Rules .....	5
Pairing the Expression Variable with a File .....	5
Entering Additional Rules .....	5
Executing the Decision Tree .....	7
Viewing the Decision Tree Results .....	8
Modifying the Decision Tree .....	9
Adding New Decisions .....	9
Changing the Class Colors and Names .....	9
Using Band Indices in Decision Expressions .....	9
Pruning the Decision Tree .....	10
Saving Tree Survivors to a Mask .....	11

# Decision Tree Classification

This tutorial is designed to introduce you to the capabilities of ENVI's decision tree classifier. You will implement a decision tree classifier, explore the various display options for decision trees, prune your decision tree, modify the class characteristics resulting from the tree, and more.

## Files Used in this Tutorial

ENVI Resource DVD: Data\decision

File	Description
boulldr_tm.dat	Landsat 5 TM image of Boulder, Colorado
boulldr_tm.hdr	ENVI header for above
boulder_dem.dat	Spatial subset of USGS DEM of Boulder, Colorado
boulder_dem.hdr	ENVI Header for above

## Background

A decision tree is a type of multistage classifier that can be applied to a single image or a stack of images. It is made up of a series of binary decisions that are used to determine the correct category for each pixel. The decisions can be based on any available characteristic of the dataset. For example, you may have an elevation image and two different multispectral images collected at different times, and any of those images can contribute to decisions within the same tree. No single decision in the tree performs the complete segmentation of the image into classes. Instead, each decision divides the data into one of two possible classes or groups of classes.

ENVI provides a decision tree tool designed to implement decision rules, such as the rules derived by any number of excellent statistical software packages that provide powerful and flexible decision tree generators. Two examples that are used commonly in the remote sensing community include CART by Salford Systems and S-PLUS by Insightful. The logic contained in the decision rules derived by these software packages can be used to build a decision tree classifier with ENVI's interactive decision tree tool.

Even if you have not used one of these packages to derive any decision rules, you may find ENVI's new decision tree tool to be a useful way to explore your data, or to find areas in your data that fulfill certain criteria. You do not need to start by having a tree that has been generated outside of ENVI in order to learn useful information about your dataset using the decision tree tool.

## Decision Tree Input

Within ENVI, the input to the decision tree classification can be one image, or a set of images of the same area. If the images are georeferenced, they do not need to be in the same projection or pixel size. ENVI will stack them on-the-fly. In ENVI, decision trees can be applied to multiple datasets.

The decision rules for the decision tree that we will be implementing in this tutorial can be described as follows: First, the image pixels are divided into two groups: those with NDVI values above 0.3 and those with NDVI values less than or equal to 0.3. The pixels with higher NDVI values are then divided into two groups: those with slope higher than or equal to 20 degrees and those with slope lower than 20

degrees. The pixels with lower slope are divided into these two groups: north facing slopes, and non-north-facing slopes. The pixels that had high NDVI and slopes greater than or equal to 20 are not subdivided. The pixels that have NDVI less than or equal to 0.3 are divided into these two groups: those with band 4 values below 20, and those with band 4 values greater than or equal to 20. However, the band 4 values equal to 0 are different from the values below 20 and band 1 values below the mean of band 1 are different from NDVI less than or equal to 0.3, so subgroups are also derived for these pixels.

Or the same tree can be described using the following criteria:

- Class 1: NDVI above 0.3, slope greater than or equal to 20 degrees
- Class 2: NDVI above 0.3, slope below 20 degrees, non-north-facing aspect
- Class 3: NDVI above 0.3, slope below 20 degrees, north-facing aspect
- Class 4: NDVI less than or equal to 0.3, band 4 greater than or equal to 20
- Class 5: NDVI less than or equal to 0.3, band 4 below 20
- Class 6: band 4 equal to 0
- Class 7: band 1 less than the mean of band 1

## Displaying Images for Decision Tree Classification

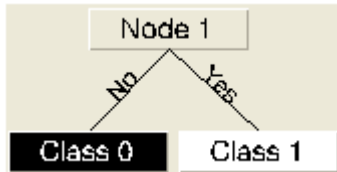
The images used in this tutorial include a subset of a Landsat 5 TM image of Boulder, Colorado. This image has been warped to match a SPOT image that has 10 m pixels, and so the Landsat image now has 10 m pixels also. You will also use a spatial subset of a USGS DEM that includes the area of the Landsat subset. The elevation image has 30 m pixels. The two images are in different projections, with the Landsat being in State Plane, and the DEM being in UTM.

Before attempting to start the program, ensure that ENVI is properly installed as described in the Installation Guide that shipped with your software.

1. From the ENVI main menu bar, select **File > Open Image File**.
2. Navigate to the `Data\decision` directory, select the file `bouldr_tm.dat` from the list, and click **Open**. The Available Bands List appears on your screen, and the color infrared (CIR) image is loaded into a display group.
3. From the ENVI main menu bar, select **File > Open Image File**.
4. Navigate to the `Data\decision` directory, select the file `boulder_dem.dat` from the list, and click **Open**. The Available Bands List appears on your screen.
5. Click **Display #1** and select **New Display**.
6. From the Available Bands List, select **Band 1** of the `boulder_dem.dat` file from the list, and click **Load Band** to load the image into a new display group.
7. Examine the images in the display groups. Visualizing the image data before the tree is built can be helpful. When you are finished examining the images that will be used in the decision tree, close both display groups by selecting **File > Cancel** from each Display group menu bar.

## Entering Decision Tree Rules

1. From the ENVI main menu bar, select **Classification > Decision Tree > Build New Decision Tree**. The ENVI Decision Tree dialog appears. By default, the decision tree tool starts with one empty decision node that will divide the pixels in the dataset into two groups, using whatever binary decision expression is entered into that empty node.



2. The first decision will be based on the Landsat image. To define the decision, click on the decision node labeled **Node 1**. The Edit Decision Properties dialog appears.
3. In the **Name** field, type **NDVI > 0.3**. This is the text that will appear on the decision node in the graphical view of the decision tree.
4. In the **Expression** field, type **{ndvi} gt 0.3**, then click **OK**. The Variable/File Pairings dialog appears.

The expression entered into the field tells ENVI to divide the pixels into those that have green vegetation and those that do not. NDVI is a commonly used index of green vegetation abundance calculated from the red and near-infrared bands of a multispectral image. The decision tree will calculate the NDVI vegetation index for each pixel, and find all of the pixels that have values higher than 0.3. The pixels with NDVI above 0.3 should have some green vegetation.

## Pairing the Expression Variable with a File

1. In the Variables/Files Pairing dialog, click once on the **{ndvi}** variable. The Select File to Associate with Variable dialog appears.
2. Select the `bouldr_tm.dat` image and click **OK**. This step tells the decision tree that when evaluating this decision expression, the NDVI values should be calculated from the `bouldr_tm.dat` image.

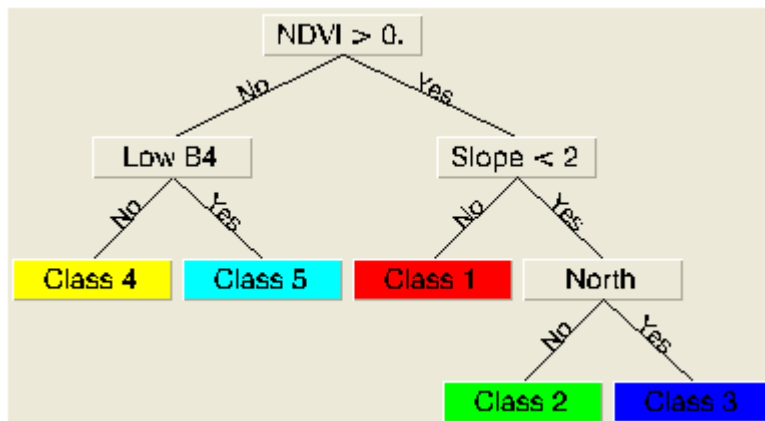
Because the wavelengths are known, ENVI will figure out which bands are needed in the NDVI calculation. If an image without wavelengths in the header had been chosen, then you would have been prompted to identify the red and near infrared bands for the NDVI calculation.

## Entering Additional Rules

This is now a very simple decision tree classifier. Pixels with NDVI greater than 0.3 will be included in Class 1 (the white class), and pixels with NDVI less than or equal to 0.3 will be included in Class 0 (the black class). You can enter additional decision rules to develop a more complicated classification.

1. Right-click on the **Class 1** node and select **Add Children** to subdivide the pixels with high NDVI into subcategories. Two new classes are automatically created beneath Class 1.
2. The node previously labeled Class 1 is now blank. Click once on this node. The Edit Decision Properties dialog appears.

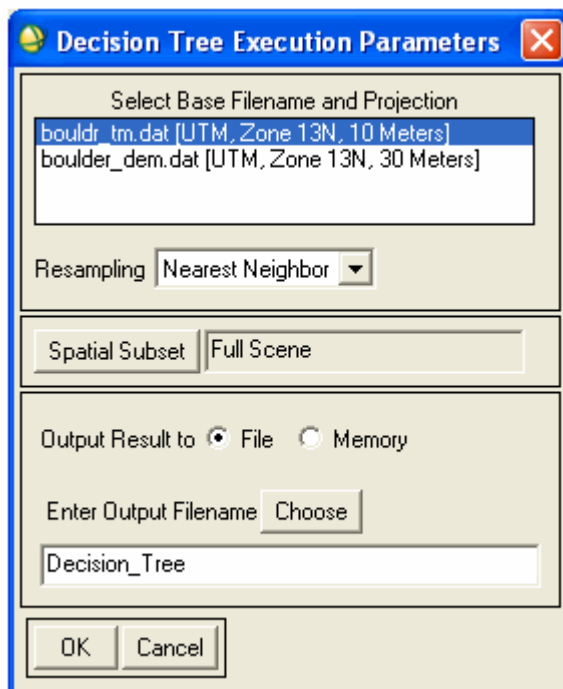
3. In the **Name** field, type **Slope < 20**.
4. In the **Expression** field, type: **{slope} lt 20**, then click **OK**. This decision will divide the high NDVI pixels based on the steepness of the slope.
5. Right-click on **Class 2** (the green terminal node) and select **Add Children** to subdivide the pixels with high NDVI and low slopes into those with north facing aspects and those without strongly north facing aspects.
6. The node previously labeled Class 2 is now blank. Click once on this node. The Edit Decision Properties dialog appears.
7. In the **Name** field, type **North**.
8. In the **Expression** field, type: **{aspect} lt 20** and **{aspect} gt 340**, then click **OK**.
9. Right-click on **Class 0** (the black terminal node) and select **Add Children** to subdivide the pixels with low NDVI values into those with values less than 20 in band 4 (these will mainly correspond to water).
10. The node previously labeled Class 0 is now blank. Click once on this node. The Edit Decision Properties dialog appears.
11. In the **Name** field, type **Low B4**.
12. In the **Expression** field, type: **b4 lt 20**, then click **OK**.



## Executing the Decision Tree

The tree is now finished, but before it can be executed, all of the variables used in the tree expressions must be paired with image files.

1. In the Variables/Files Pairing dialog, click once on the **{b4}** variable. The Select File to Associate with Variable dialog appears.
2. Select **Band 4** from the `boulldr_tm.dat` image and click **OK**.
3. Pair the following variables and files or bands as described in steps 1 and 2 above:  
`{slope} = boulder_dem.dat`  
`{aspect} = boulder_dem.dat`
4. From the ENVI Decision Tree dialog menu bar, select **Options > Execute**, or right-click in the blank background area of the ENVI Decision Tree dialog and choose **Execute**. The Decision Tree Execution Parameters dialog appears.
5. Select the `boulldr_tm.dat` image to use as the base. The map projection, pixel size and extent of the other image will be adjusted to match.
6. Enter an output filename for the classified image, and click **OK**.



While the calculation is taking place, the node that is currently being used to group pixels is temporarily colored light green. Therefore, it is possible to watch the classification process proceed from node to node. When the classification is finished, the result is automatically loaded into a new display.

## Viewing the Decision Tree Results

The color of a given pixel in the output from the decision tree is determined by the color of the terminal node into which that pixel is classified. Thus, for example, the yellow pixels in the decision-tree result are the pixels that answered no at every decision point. So, they have low NDVI values and high values in band 4. This means that they are not water and they are not vegetation.

1. Examine the ENVI decision tree once again. Notice that the default view does not include all of the information that could be shown.
2. In the blank background area of the ENVI Decision Tree dialog, right-click and select **Zoom In**. Now each node label displays the number of pixels and the percentage of the entire image that progressed to each node of the tree. You can resize the Decision Tree window to view the complete window.
3. Place the mouse cursor over any node in the decision tree, and notice that information about that node appears in the status bar at the bottom of the ENVI Decision Tree dialog. This is another useful way to quickly get information about nodes in the tree, especially when the tree view is not zoomed in to show node details.



## Modifying the Decision Tree

After you have executed a decision tree and examined the results, you may find that additional decision rules may be useful. For example, in this decision tree some of the pixels with Band 4 values below 20 are the border pixels, which all have values of 0. They're showing up as cyan because their values are below 20, but really they have low band 4 values for a completely different reason than the other cyan pixels.

### Adding New Decisions

1. Right-click on the **Class 5** terminal node (pixels with band 4 values below 20), and select **Add Children**.
2. Rename the blank node previously labeled Class 5 to **B4 = 0**, type the expression **b4 eq 0** in the field provided, and click **OK**.
3. Execute the tree again by right-clicking in the Decision Tree dialog background and selecting **Execute**. In the output image, the border pixels are now in their own class, but the magenta border looks strange.

### Changing the Class Colors and Names

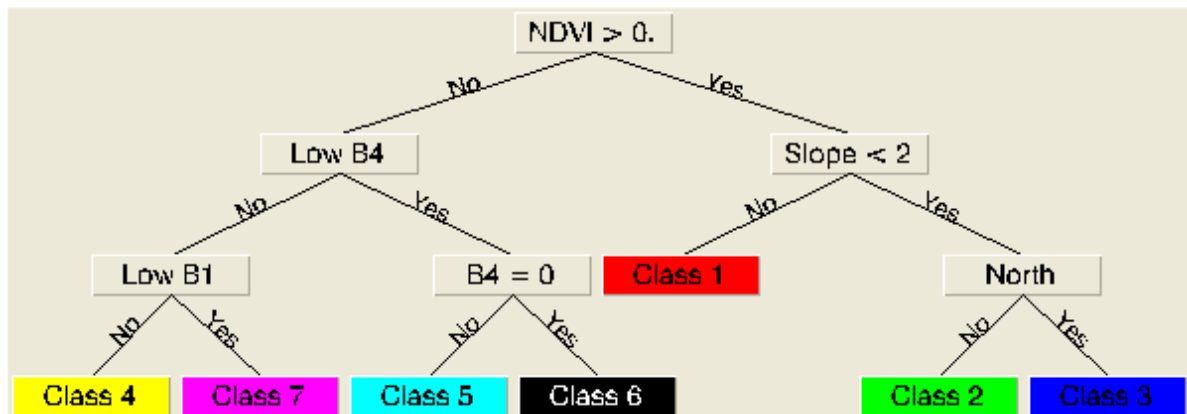
1. Click on **Class 6** (the magenta terminal node). The Edit Class Properties dialog appears.
2. In the **Name** field, type **Border**.
3. Click the **Color** button to display a list of colors and select **Colors 1-20 > Black**.
4. Click **OK**.
5. To apply the color change to the tree result, execute the tree again. The border should now appear black.

### Using Band Indices in Decision Expressions

Several built-in decision tree variables require band indices to be used in the decision expression.

1. Right-click on **Class 4** (the yellow terminal node) and select **Add Children**. This node includes pixels with low NDVI values but high Band 4 values.
2. Rename the blank node previously labeled Class 4 to **Low B1**, type the expression **b1 lt {mean[1]}** in the field provided, and click **OK**. This expression asks whether the pixels have a value in band 1 less than the mean value for band 1. The "1" inside the square brackets indicates that band 1 is to be used for the mean calculation. Some of the variables that ENVI is able to calculate on-the-fly for decision trees, including the mean calculation, have to be associated with a whole file rather than individual bands. In those cases you need to specify which band in the file you want to use for the calculation, in square brackets.
3. In the Variable/File Pairings dialog, click on the **b1** variable, pair it with **Band 1** of `bouldr_tm.dat`, then click **OK**.
4. Pair the **{mean}** variable with the `bouldr_tm.dat` file, then click **OK**.

5. To apply the change to the decision tree, execute the tree again. Some of the yellow pixels that have low Band 1 values are now magenta.



## Pruning the Decision Tree

Often when using decision tree, it is desirable to test whether a specific set of children are useful additions to a tree. ENVI's Decision Tree tool provides two methods for removing children once they have been added:

- **Delete** permanently removes the children from the tree.
- **Prune** temporarily removes them and allows them to be reinstated without redefining the decision rule or node properties.

In this exercise, you will prune children from the decision tree, then restore them. Pruning and restoring the children allows you to compare the tree results with and without particular sets of children.

1. Right-click on the **Low B1** node and select **Prune Children**. Notice that, while the children are still visible, they are no longer colored nor connected to the tree. This indicates that they have been pruned, and will not be used when the tree is executed.
2. Right-click on the **Class 8** node and select **Restore Pruned Children**.

## Saving Tree Survivors to a Mask

The option to save survivors to a mask is available for every node in the tree, so you can make a mask of the pixels that have survived to any position within the tree.

1. Right-click on **Class 1** (the red node) in the decision tree and select **Save Survivors to Mask**. This option will make a binary mask image that assigns a value of 1 to pixels that are included in that class and a value of 0 to pixels that are not included in that class. The Output Survivors to Mask Filename dialog appears.
2. Type an output filename for the survivor mask, then click **OK**. The resulting mask image is listed in the Available Bands List.
3. Load the new mask image into a new gray scale display. Notice that the white pixels in the mask correspond to the red pixels in the decision tree output image.

You can save decision trees, including all of the pairings between variables and files. Saving a decision tree allows you to restore it in a later ENVI session.

1. From the Decision Tree dialog menu bar, select **File > Save Tree**. The Save Decision Tree dialog appears.
2. Type an output filename for the tree, and click **OK**.
3. When you are finished, you can quit your ENVI session by selecting **File > Exit** from the ENVI main menu bar.