

# ENVI MACHINE LEARNING

## 概要とチュートリアル

Harris Geospatial株式会社

---

2022.12



# 概要

動作要件やDeep Learning Moduleとの比較について

# 一般論: Machine Learning と Deep Learningの比較



## 機械学習(MACHINE LEARNING)

- 使いやすい
- 数分で学習し、結果を得ることができる
- データ管理はあまり要求されない
- CPUでトレーニングが可能
- 解像度に影響されない
- 主にスペクトルに注目する
- トレーニングデータの量より質が重要となる

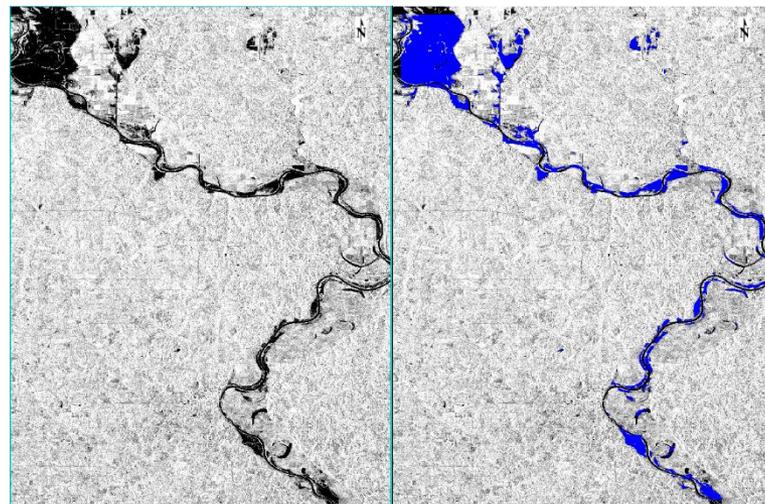
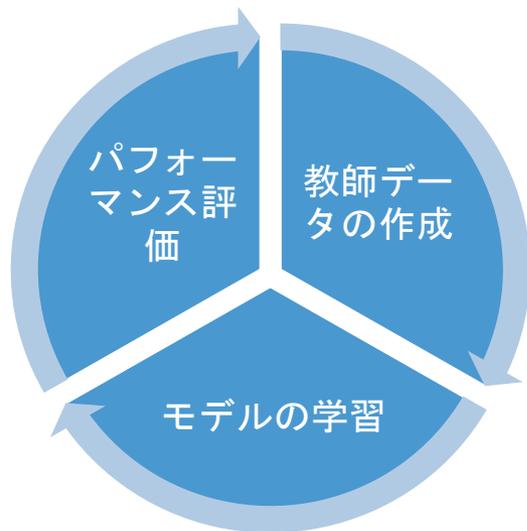
## 深層学習(DEEP LEARNING)

- 使いにくい（理解が難しい）
- 学習と結果の取得に数日かかる場合がある
- 大量のデータを管理する必要がある
- 主にGPUが使用される
- 解像度の影響を受ける
- 形状に注目するが多い
- 教師データの質と量の両方が重要となる

# 参考: ENVI Deep Learning Moduleとは



- TensorFlowおよびKerasに基づくディープラーニングを使用した画像解析のためのENVIのオプションモジュールです。
- プログラミングの知識なしに、GUIベースでディープラーニングの処理を実行することが可能です。
- このモジュールの利用には、NVIDIA GPUを搭載しているといったハードウェア要件があります。



ENVI Deep Learningの処理フロー

事例: SARを用いた洪水エリアの自動抽出

# 比較: ENVI Machine Learning と ENVI Deep Learning



## ENVI Machine Learning

- ENVI Deep Learningの中に含まれているが、ENVIのライセンスのみで実行できる
- ENVI 5.6.3から利用可能
- CPUを使用する
- 搭載メモリ16 GB以上を推奨
- Intelを想定してライブラリが最適かされているが、AMDでの動作もサポートする
  - 特定のアルゴリズムについては、Intelのほうがパフォーマンスで優位となる（後述）



## ENVI Deep Learning

- 実行にはENVI Deep Learningのライセンスが必要となる
- 動作にはNVIDIA社製の要件を満たしたGPUが搭載されている必要がある
- 搭載メモリ8 GB以上を推奨



# ツールのインテル最適化について



- 機械学習ではCPUリソースが使用され、GPUは利用されません。
- ENVI Machine LearningはIntel社製のCPUに最適化されています。
  - 一部のアルゴリズムでは、他社製品のCPU下で速度が著しく遅くなる場合があります。

インテルに最適化されたライブラリによるアルゴリズムの性能向上例  
※「RBF SVM Classify」は処理が終了していないため、おおよその時間

タスク名	処理時間 [秒]	Intel CPUでの 処理時間[秒]	差分[%]
RBF SVM Train	94.5	6.7	-93%
RBF SVM Classify	15000.0	69.4	-100%
Nearest Neighbors Classify	257.3	6.6	-97%
Random Forest Train	2.3	2.1	-9%
Random Forest Classify	18.5	13.1	-29%

# ENVI Machine Learningの具体的な動作要件

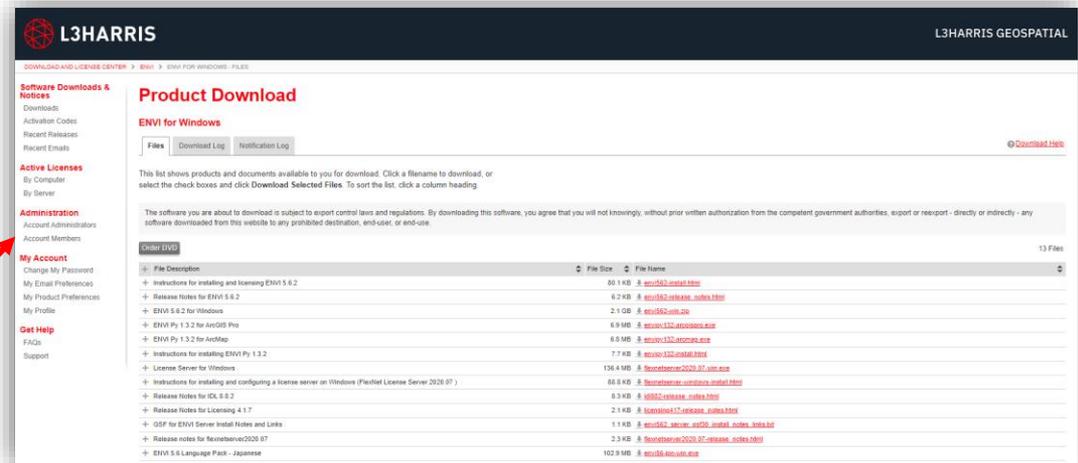
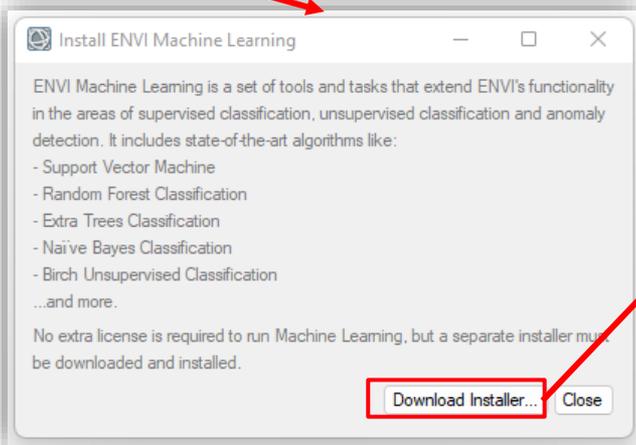
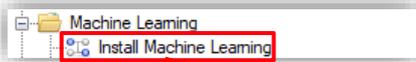


- ベースソフトウェア
  - ENVI 5.6.3
  - ENVI Deep Learning Module 2.0
- OS:
  - Windows 10 and 11 (Intel/AMD 64-bit)
  - Linux (Intel/AMD 64-bit, kernel 4.18.0 or higher and glibc 2.28 or higher)
- ハードウェア要件:
  - Advanced Vector Extensions (AVX) が搭載されていること。

# ENVI Machine Learningの利用方法



- ENVI 5.6.3をインストールすると、ENVIツールボックスにMachine Learningというフォルダがデフォルトで存在します。
- ENVI Deep Learning Module 2.0がインストールされていない場合、[Install Machine Learning]というメニューのみが存在します。
- これをダブルクリックで起動するとメッセージが表示され、Harris Download and License Centerに遷移します
  - ENVIのライセンスがある場合、Harris Download and License CenterからENVI Deep Learning Module 2.0をダウンロードすることが可能です

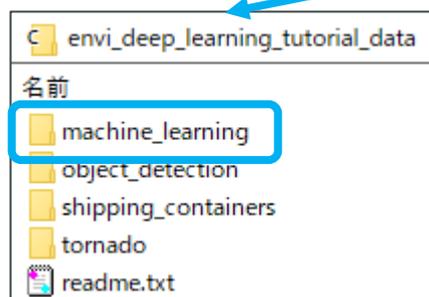


# ENVI Machine Learningのサンプルデータ



- Harris Download and License CenterからENVI Machine Learningのサンプルデータを手に入れることが可能です。
  - ENVI Deep Learning Module 2.0インストーラと同じ場所に「envi\_deep\_learning\_tutorial\_data.zip」の名前で登録されております。
  - 後述するチュートリアルでは、ここで入手したデータを使用します(以下配置場所)
    - 特徴量抽出: machine\_learning¥anomaly
    - 教師付き分類: machine\_learning¥supervised
    - 教師なし分類: machine\_learning¥unsupervised

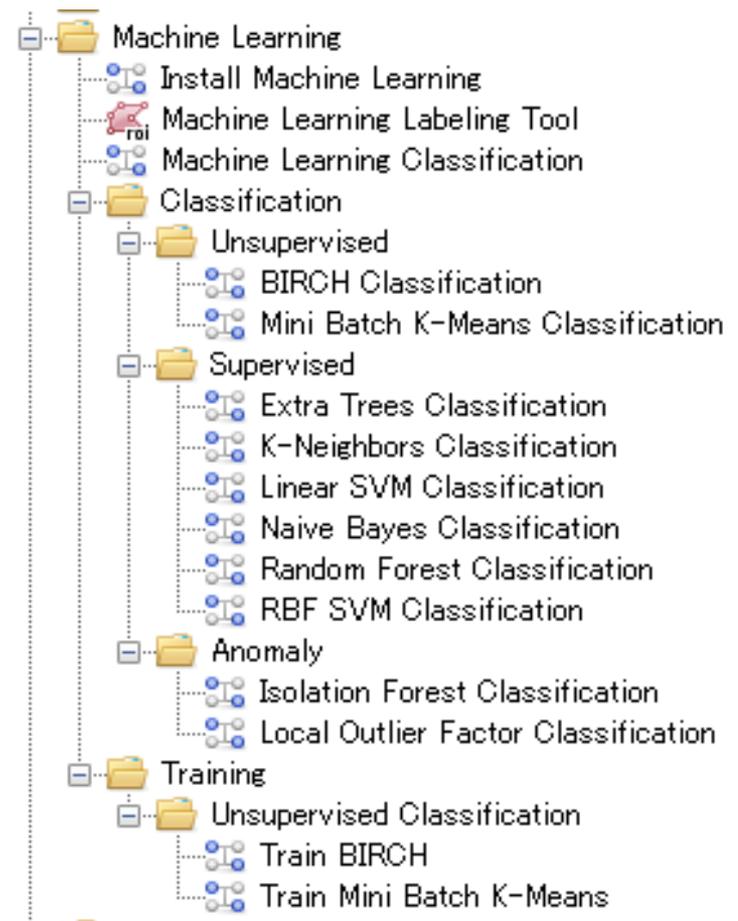
File Description	File Size	File Name
+ Instructions for installing ENVI Deep Learning 2.0	73 KB	<a href="#">envideeplearning-install.html</a>
+ ENVI Deep Learning 2.0 Release Notes	3.8 KB	<a href="#">envideeplearning20-release_notes.html</a>
+ ENVI Deep Learning 2.0 for Windows (for ENVI 5.6.3 and ENVI Machine Learning)	1.3 GB	<a href="#">envideeplearning20-win.exe</a>
+ ENVI Deep Learning 2.0 for Linux (for ENVI 5.6.3 and ENVI Machine Learning)	2.5 GB	<a href="#">envi_deep_learning20-linux.tar.gz</a>
- ENVI Deep Learning 2.0 Tutorial Data	2.2 GB	<a href="#">envi_deep_learning_tutorial_data.zip</a>



# ENVI Machine Learningの実行



- ENVI Deep Learning Module 2.0をインストールすると、右図に示すように複数のツールが実行できるようになります。
- ツールはENVIツールボックスのMachine Learningフォルダにまとまっています。



# ENVI Machine Learningのアルゴリズム



ENVI Machine Learningには、以下の3つの機械学習アルゴリズムが含まれます。

## 教師付き

- **概要**
  - ENVI の従来の分類アルゴリズムと同様、カラーテーブルとクラスルックアップを含むラスタを作成します。
  - 1つ以上のラスタと、ラスタごとに1つ以上のROIを使用します。
- **アルゴリズム**
  - Random Forest
  - Extra Trees
  - Naïve Bayes
  - Nearest Neighbors (K-Neighbors)
  - SVM (Linear, RBF)
    - SVMは実行速度やパフォーマンスの観点から他のアルゴリズムに劣ります
    - SVM(RBF)については、パフォーマンスの観点からintel CPU上での利用を推奨します

## 特異点抽出

- **概要**
  - 特異ではない(正常な)教師情報に基づいてモデルを学習し、インプットされたデータの各ピクセル値が異常か否かを[1]/[0]で返す新たなアルゴリズムです。
  - 教師となる正常なスペクトルを持った一つ以上のラスタを使用します。
- **アルゴリズム**
  - Isolation Forest
  - Local Outlier Factor
    - これらの両アルゴリズムのパフォーマンスは同程度です。

## 教師なし

- **概要**
  - ISO Data Classificationと同様に、指定した個数のクラスを検出する分類器を生成します。
  - 入力として学習データを持たない1つまたは複数のラスタを使用します。ラベルを指定する必要がないため、ROIは不要です。
- **アルゴリズム**
  - KMeans
  - BIRCH

# ENVIの従来のアプローチと、 Machine Learningによるアプローチとの比較



## 従来のENVIツールを使用したアプローチ

- 手法例:
  - Spectral Angle Mapper
  - Mahalanobis Distance
- アルゴリズム固有のパラメータが性能を決定します
- 多くの場合、ルールイメージや補助的なファイルによってメトリックスが提供されます
- 見つけたいものだけを指定します
- すべてのピクセルを分類するわけではありません
- 長時間の、あるいは、CPU/GPUのリソースを集中的に使用した学習は不要です
- ROIをサンプルとして統計情報を取得することができます

## MACHINE LEARNINGによるアプローチ

- 手法例:
  - Random Forest
  - SVM
- 閾値がありません
- 分類画像(結果)のみが提供されます
- 検出しない(=背景となる)特徴量の教師を示す必要があります
- すべてのピクセル毎に分類が行われます
- 長時間の、あるいは、CPU/GPUのリソースを集中的に使用した学習は不要です
- ROIをサンプルとして統計情報を取得することができます



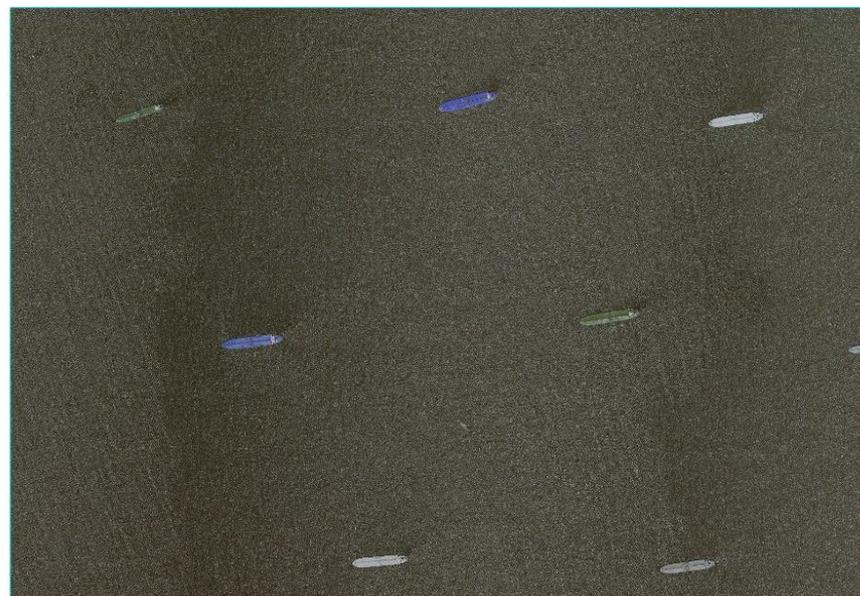
# チュートリアル① 特異点抽出(Anomaly Detection)

<https://www.l3harrisgeospatial.com/docs/envimltutorial/anomalydetection.html>

# 使用するデータ



- このチュートリアルで使用される画像は、National Agriculture Imagery Program (NAIP)によって提供されました。
- この画像は、複数の船を含むロサンゼルス港のサブセットです。これは、1メートルの地上サンプル距離 (GSD) の空間解像度を持つ4バンドデータ (赤/緑/青/NIR) です。



NAIP\_LAHarbor\_Subset.dat



- 特異点抽出(Anomaly detection)はデータセット内の外れ値を検出するためのツールです。
  - 外れ値とは、データセット内の既知の特徴量と比較して正常でないと考えられる特徴量を指します。
    - 例えば「水」が既知の特徴とすると、水以外のものが外れ値(=異常値)とみなされます。
- 特異点抽出ツールは、学習時にひとつの「背景(=background)」という特徴量を指定します。
  - この指定した特徴量を、ツールのなかで正常値であるとみなします。
  - 分類の結果、正常と判定されなかったピクセルがすなわち特異点と判定されます
    - 上記の例で言えば、水が背景(=正常値)であり、最終的に水と判定されなかったものが特異点です。
- データに対して背景のラベルを与える過程は、特異点抽出のうえで非常に重要なプロセスです。
  - 意図せず特異点が背景に含まれて指定された場合、分類の精度が低下します。

# 特異点抽出の処理フロー



インプットとなる1つ以上のラスタに共通した背景特徴を特定し、その周囲にROIを描画する (ラベリング処理)



ラベル付けされたデータでモデル学習を実施する



学習したモデルで分類を行い、分類ラスタを生成する



偽陽性のクリーンアップを実施する(オプション)

## Isolation Forest

マッチングアルゴリズムであり、特異点と判定されるピクセルは少ない、あるいは稀であると考えます。結果的に、特異点は正常値から離れた値のピクセルに対して割り当てられます。

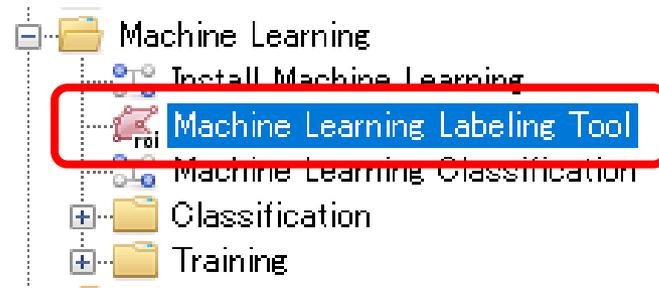
## Local Outlier Factor

与えられたピクセルの近隣に対する局所的な偏差を測定するアルゴリズムです。局所外れ値(Local Outlier)は、それを取り囲むピクセルの近傍に基づいて、ピクセル値の差異を評価することによって決定されます。

# ROI でラスタにラベルを付ける



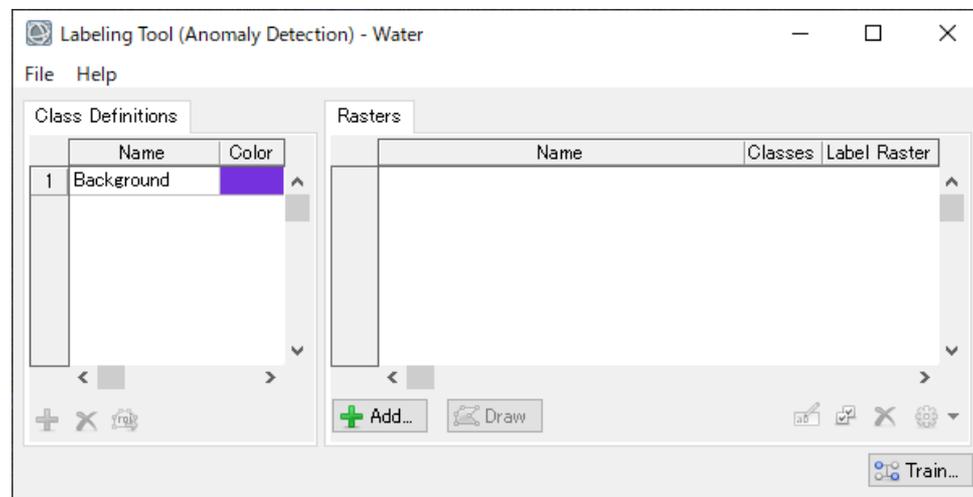
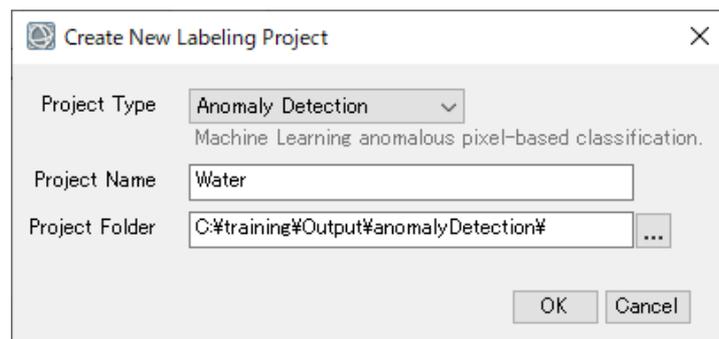
- ラベル付けプロセスを開始するには、特徴のサンプルを収集するために少なくとも1つのラスタデータ(画像)が必要となります。
  - 関心のあるオブジェクトまたはピクセルを選択する従来のラベリングとは異なり、特異点抽出には反対のアプローチが必要です。
  - つまり、特異点抽出では「背景」にあたる情報にラベルをつけます。
- ラベルをつける対象は、一貫したスペクトルプロパティと空間プロパティを持っている必要があります。
- このラベル取得を行う最も簡単な方法は、Machine Learning Labeling Tool を使用することです。



# プロジェクトの新規作成



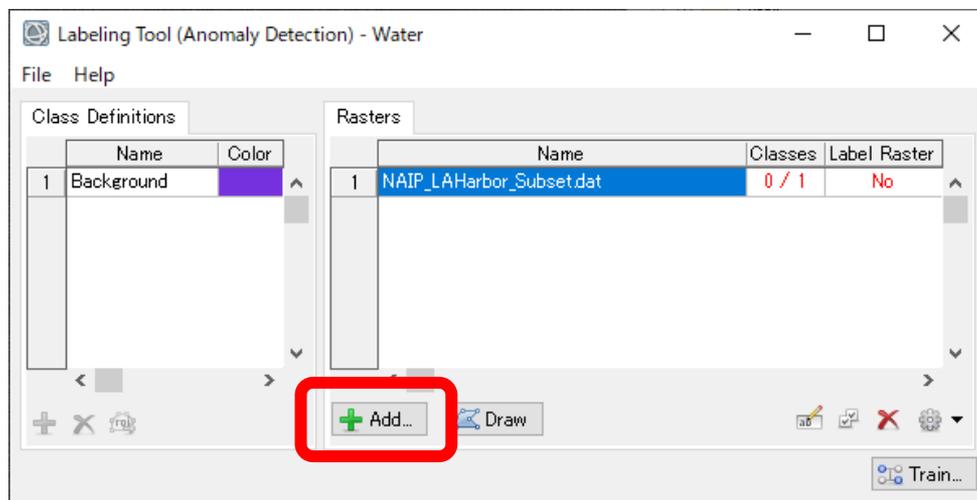
1. ENVIツールボックスのMachine Learning/Machine Learning Labeling Toolをダブルクリックで起動します
2. メニューバー > File > New Projectを選択します
3. Project Typeを[Anomaly Detection]、Project Nameを[Water]、Project Folderを適当な出力先に設定し、OKボタンを押下します
4. 新しい特異点抽出プロジェクトを作成すると、クラス定義Backgroundが自動的に作成されます。
  - 追加のラベルを追加することはできず、ラベルの背景(ここでは水)を識別するために使用されます。



# 教師を取得するラスタデータを追加する



1. Labeling ToolのAddボタンを押下します
2. ラベルを設定する対象のラスタを選択するよう求められるので、ここではサンプルデータanomaly¥NAIP\_LAHarbor\_Subset.datを選択します
3. 選択したデータが[Rasters]セクションに追加されます。
  1. [Classes]列には、赤色の分数 (0/1) が表示されます。
    1. 左の0は、どのクラス ラベルもまだ描画されていないことを意味します。
    2. 右の1は、定義されたクラスの総数(ここではBackgroundというクラス)を表します
  2. [Label Raster]列には、学習の際の教師となるデータ (Label Raster)がまだ作成されていないことを示す赤色の [No] が表示されます。



# 水のピクセルにラベルを付ける



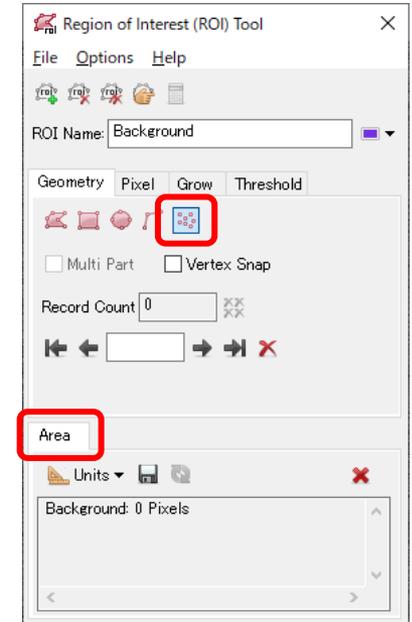
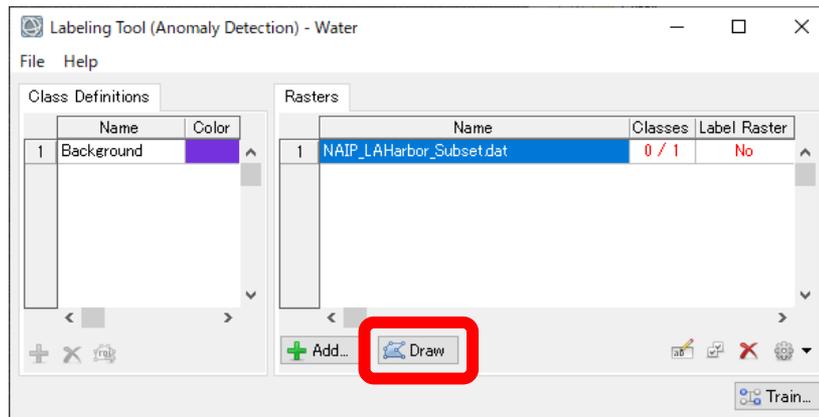
特異点抽出器の学習データは、量より質が重要です。

- ラベル付けされたピクセルが少なすぎる場合:
  - すべての特異なターゲットを見つけるのに十分な情報が得られない可能性があります。
- ラベル付けされたピクセルが多すぎる場合:
  - 分類の実行時間が長くなる可能性があります。
- 特異なターゲットに属するピクセルを教師に含んでしまった場合:
  - 学習が思うように進まず、期待した結果が得られない可能性があります。

# 水のピクセルにラベルを付ける



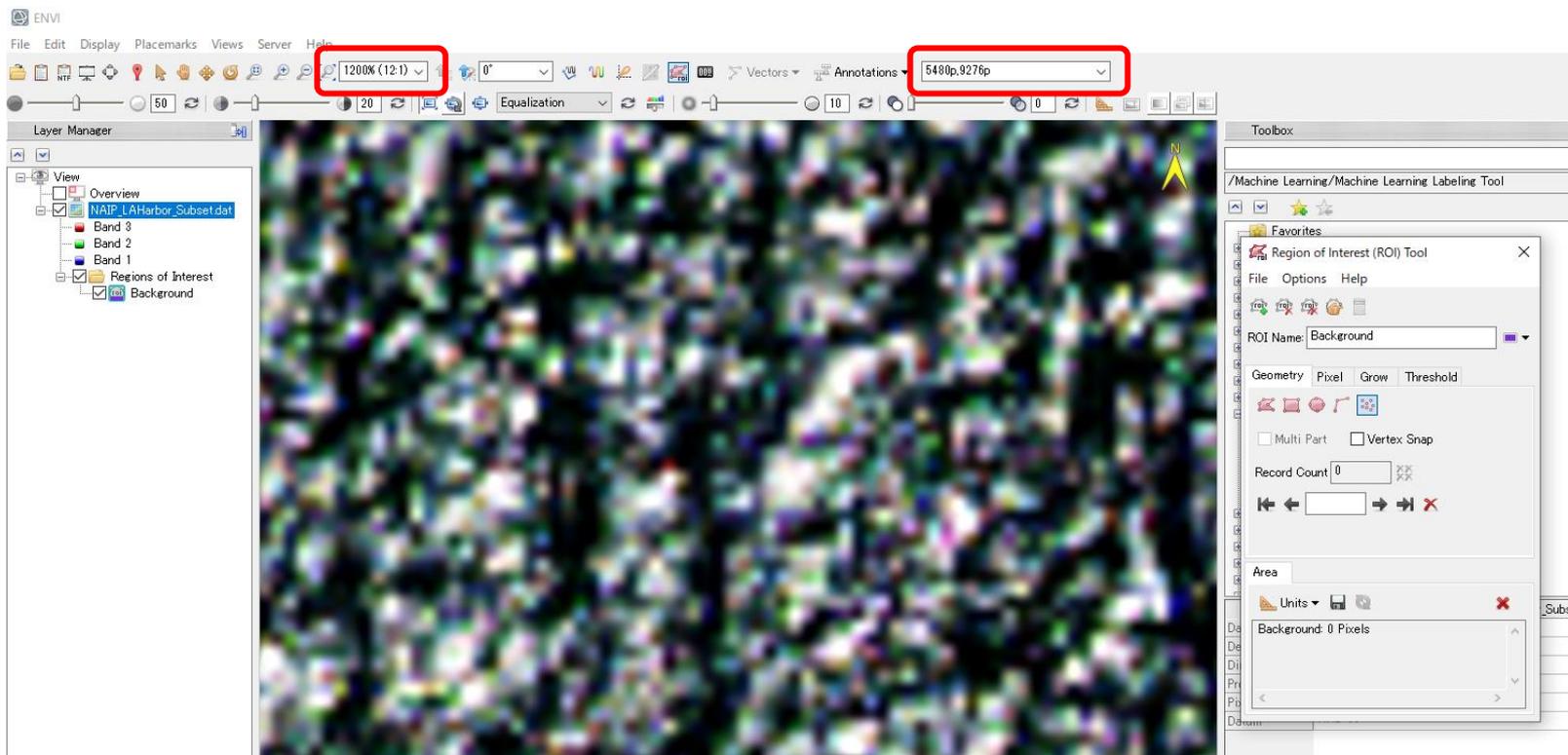
1. Drawボタンを押下すると、ROIツールが起動します。
2. レイヤマネージャで、NAIP\_LAHarbor\_Subset.datを選択してアクティブレイヤにします。
3. ENVI ツールバーのStretch Typeドロップダウンリストをクリックし、[Equalization]を選択します。
4. ROIツールの[Geometry] タブで、[Point]ボタンを選択します。
5. Areaタブを選択します
  - 現在のピクセル数を確認できるようにします。



# 水のピクセルにラベルを付ける



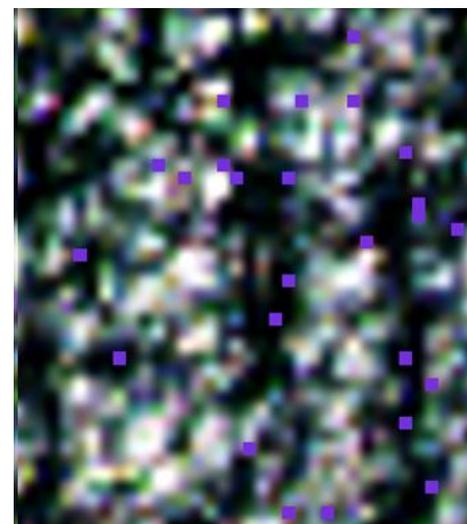
- ENVI ツールバーのGo Toフィールドに、[5480p,9276p] のピクセル座標を入力しEnterで画像の中心のこのピクセル位置にします。
- ENVI ツールバーのZoomドロップダウン リストをクリックし、1200%と入力し12:1に拡大するよう設定して、Enterを押します。



# 水のピクセルにラベルを付ける



8. 関心のあるピクセル上でマウスの左ボタンをクリックして、ピクセルに[背景]の教師ラベルを付けます。
  - 緑と白のピクセルにラベルを付けることは避けてください。これらのピクセルは、画像内の一部の船のピクセル カラー値も表すためです。
  - 白と緑を除いて、できるだけ多くの異なる色のピクセルを取得しようとしながら、異なる色の 25 ピクセルにラベルを付けます。
9. 25 ピクセルにラベルを付けたら、キーボードの Enter キーを押して ROI ポイントを受け入れます。
  - 承認されると、紫色のプラス + アイコンが紫色のボックスになります。

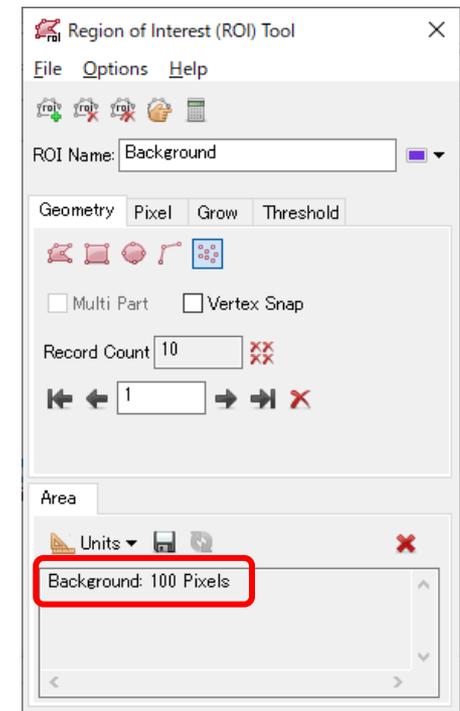


# 水のピクセルにラベルを付ける



10. 以下のピクセル座標を参考に、同様の手順(レイヤ マネージャから NAIP\_LAHarbor\_Subset.dat を選択し、GoTo および Zoom で移動/拡大する)で25 ピクセルずつラベルを付け、ピクセルの総数は 100 にします。
- 目標は、ラベル付けされたピクセルの数を最小限に抑え、船が水に対して特異であると識別するのに十分な情報を提供することです。
  - トレーニング中のピクセル数を最小限に抑えると、分類時間が短縮されます。

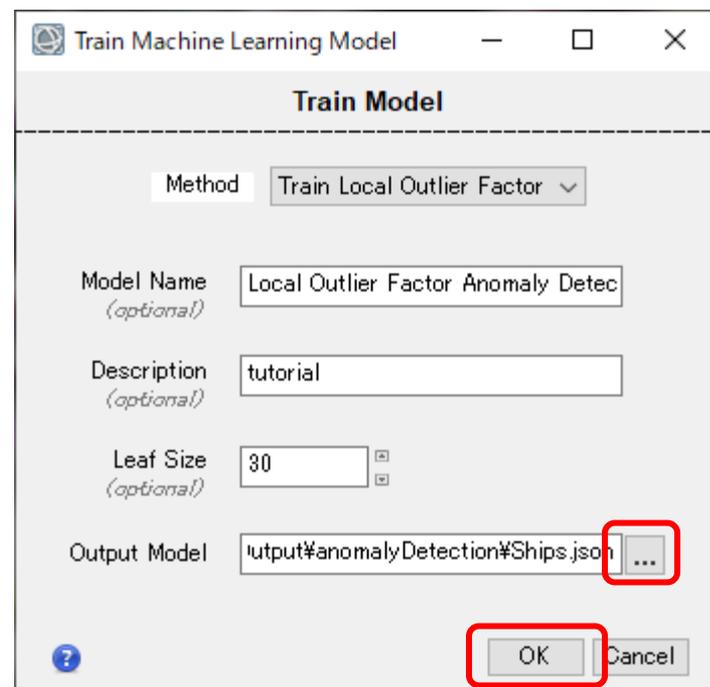
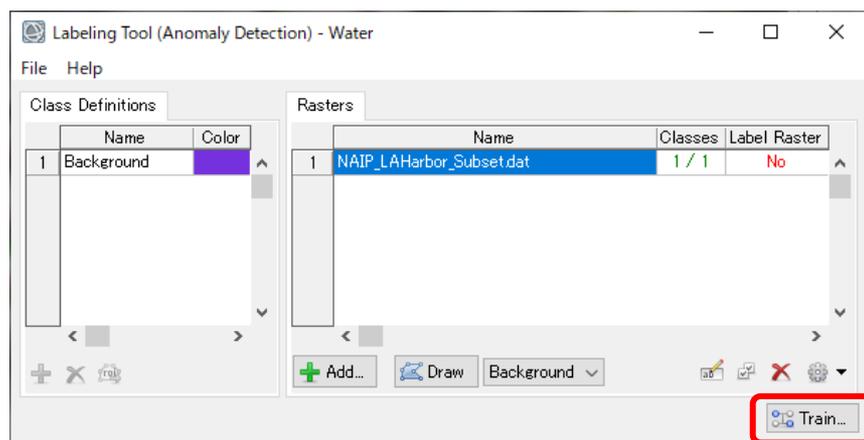
[ 7334p,10161p ]  
[ 9934p,9389p ]  
[ 7920p,11849p ]



# 特異点抽出モデルのトレーニング



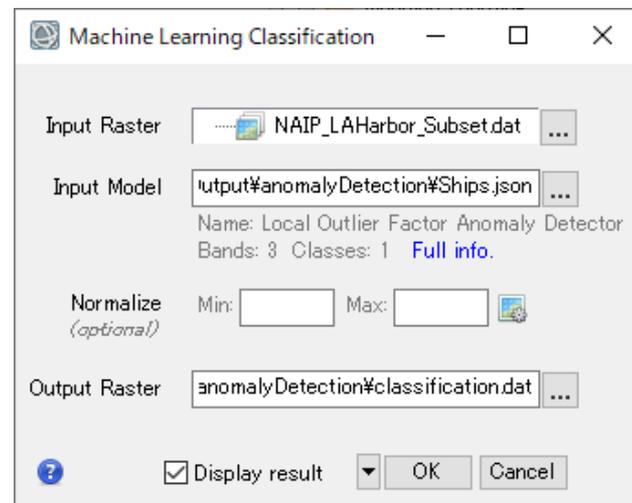
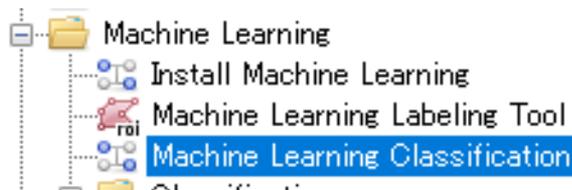
1. [Train]ボタンを押下すると、[Train Machine Learning Model]ダイアログが出現します。
2. MethodやLeaf Size等のパラメータ設定はそのままとし、Output Modelの箇所には出力する学習モデル(json形式)のフルパスを指定します。Browseボタンを押下すると、作成したプロジェクト配下に遷移するので、ここでは[Ships.json]の名前で保存します。
3. OKボタンを押下し学習を開始します。



# 分類の実行



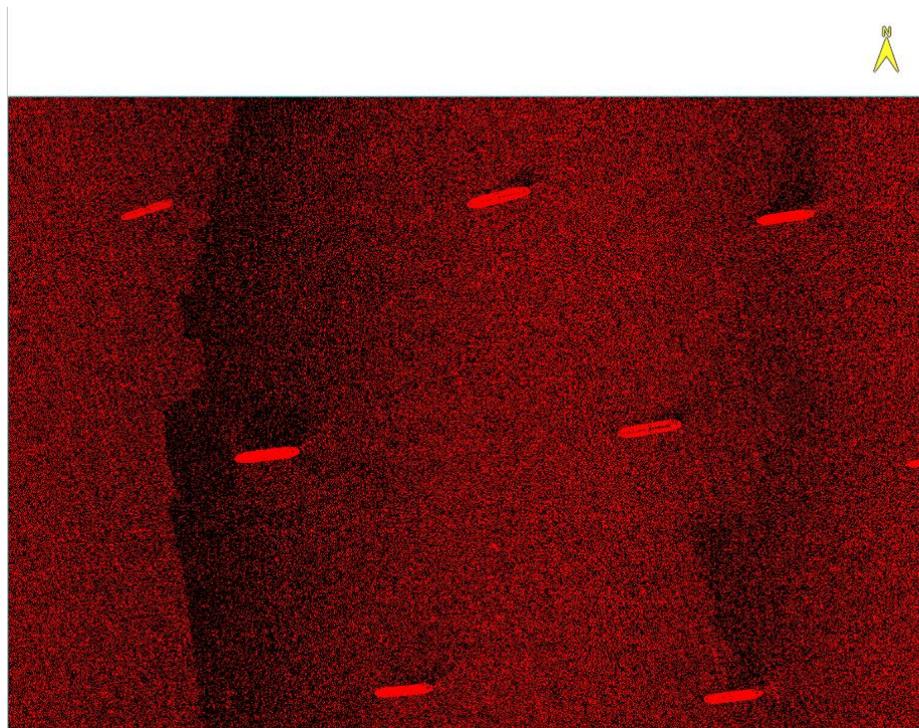
1. 水ピクセルと区別するようにトレーニングされたモデルができたので、同じラスタでモデルを使用して、水ではないピクセルを識別します。ラベリングツールはもう使用しないため、閉じてください。
2. ENVIツールボックスのMachine Learning/Machine Learning Classificationをダブルクリックで起動します。
3. Input Rasterには学習に使用したデータを、Input Modelには先ほど生成された分類器(jsonファイル)を設定します。Normalizeのオプションについてはここでは空欄で問題ありません。Output Rasterに分類結果となるラスタの出力先を設定します。
  - classification.datのような名前でも適切な場所に出力
4. OKボタンで分類処理を実行します。



# 分類の実行



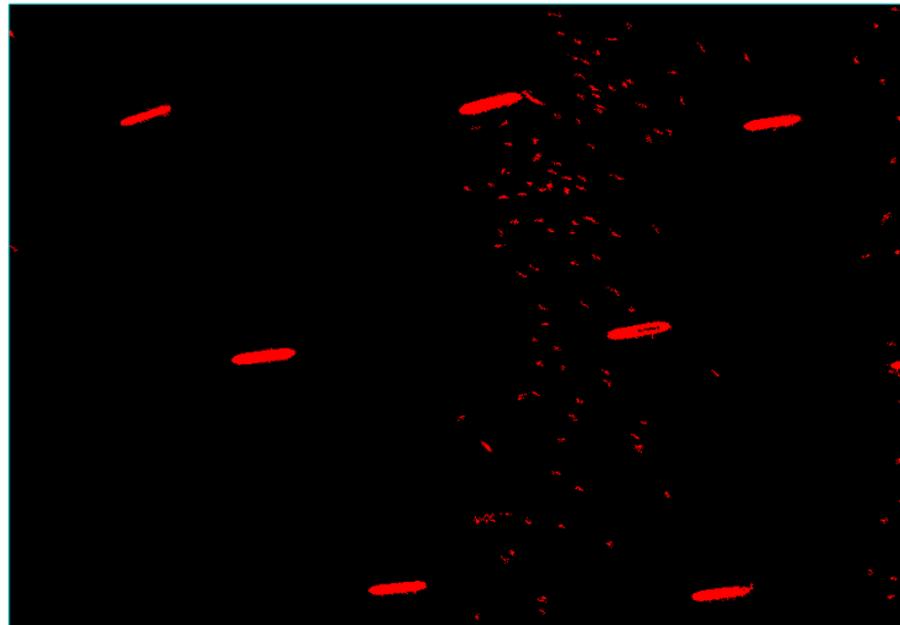
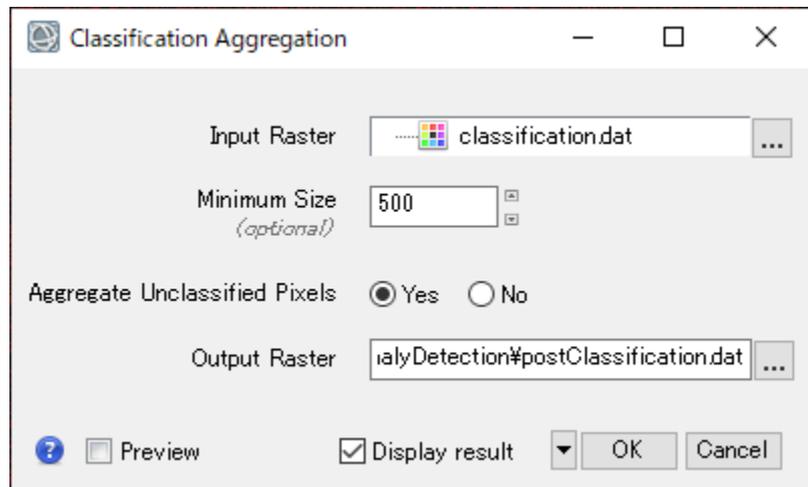
- 結果を一見すると、画像に非常にノイズが多いことがわかります。
  - これは、最小限のデータ (100 ピクセル) を使用した結果です。
  - より多くのラベル付きピクセルをトレーニングへの入力として使用すると、分類時間が大幅に長くなりますが、より鮮明な結果が得られます。
  - より多くのトレーニング済みピクセルを使用しても、結果にはまだノイズが含まれており、それをクリーンアップするために後処理が必要になる可能性があります。



# 後処理の実施



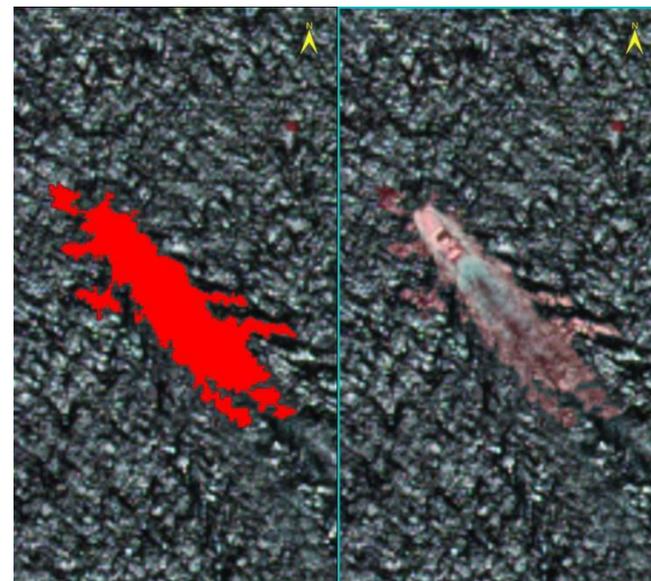
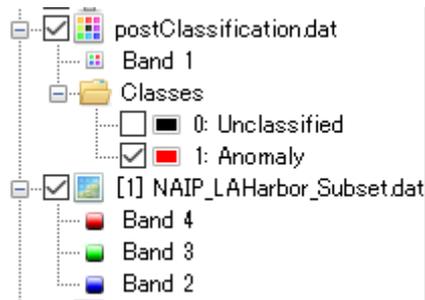
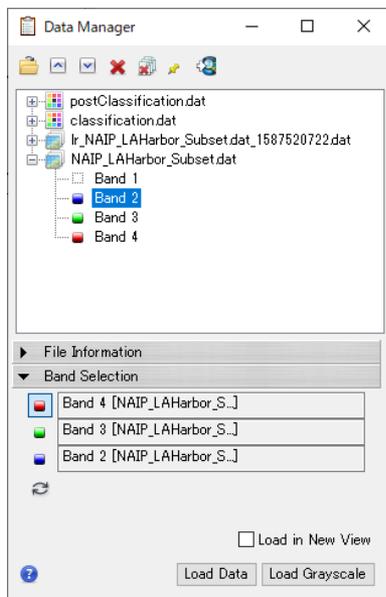
1. ENVIツールボックスの /Classification/Post Classification/Classification Aggregation をダブルクリックで起動します。
2. 分類結果をInput Rasterに設定し、Minimum sizeパラメータを500に設定して、適当な名前で保存し実行します。
3. ノイズが結果から落ちていることを確認します。



# 結果の検証



1. NAIP\_LAHarbor\_Subset.datをR:Band4/G:Band3/B:Band2の構成で開きます。
2. ノイズを落とした結果が開いたラスタの上にくるようにレイヤマネージャの配置順を変えて、Unclassifiedについてはチェックを外し非表示にします。
3. GoToツールから[7389p,10988p]に移動します
4. Zoomレベルを400%にします。
5. ENVIツールバーのTransparencyツールなどを使用し、結果を確認してください。





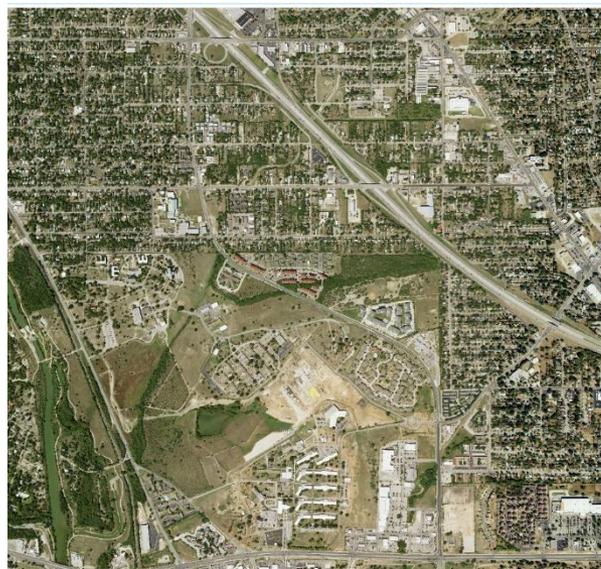
# チュートリアル② 教師付き分類(Supervised Classification)

<https://www.l3harrisgeospatial.com/docs/envimltutorial/supervisedclassification.html>

# 使用するデータ



- このチュートリアルで使用される画像は、National Agriculture Imagery Program (NAIP)によって提供されました。
- この画像は、サンアントニオおよびダラスのそれぞれ異なるエリアに関するサブセットです。これは、1メートルの地上サンプル距離 (GSD) の空間解像度を持つ4バンドデータ (赤/緑/青/NIR) です。



training¥

NAIP\_SanAntonioSE\_2020\_Subset.dat (左: サンアントニオ)

NAIP\_DallasTX\_Oct11\_2020\_Subset.tif (右: ダラス)

# 教師付き分類と本チュートリアルについて

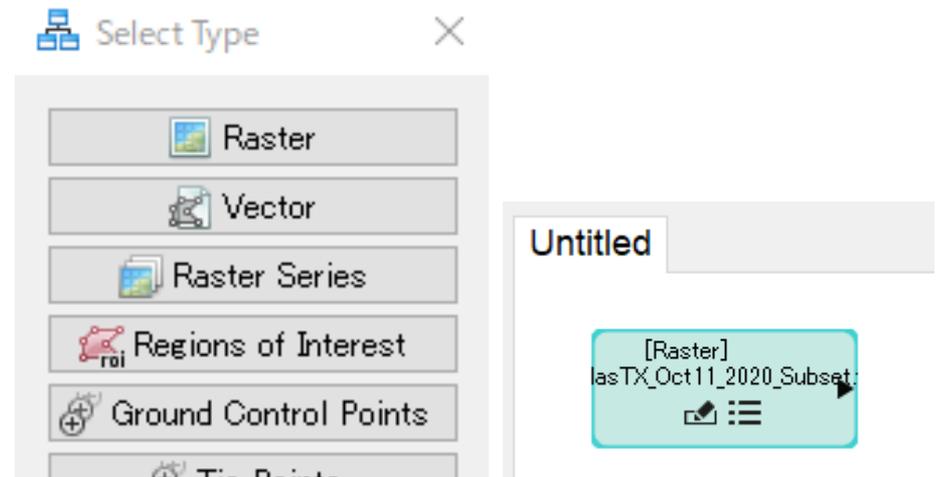


- 教師付き分類は、機械学習アルゴリズムのひとつです。
  - 与えられた教師ラベルを参考に、画像内の学習クラスを識別する関数/モデルを生成します。
- バイナリ分類とマルチクラス分類の2種類があります。
  - このチュートリアルでは、[Manmade(人工物), Trees (樹木), Ground (地面), Water(水域)]の4つのラベルを使用した、都市に関するマルチクラス分類に焦点を当てます。
- このチュートリアルでは、ENVI Modeler を使用してカスタム ワークフローを構築する方法について説明します。
  - ENVI Machine Learning を使用して、簡単に修正された ENVI ワークフローを作成できます。
- このチュートリアルでは、すべてのトレーニング データに事前にラベルが付けられています。
  - 2つのチュートリアルトレーニングデータ(tifおよびdat形式)にそれぞれ対応した教師データがxml形式のROIとして用意されています。
  - これにより、ツールと機械学習ワークフローの作成方法に集中できるようになります。

# ENVI Modeler でモデルを作成する



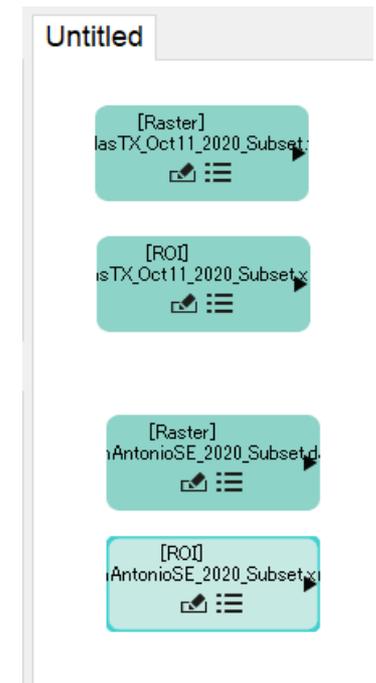
1. メニューバー > Display > ENVI Modelerを選択・起動します。
2. [Basic Nodes]リストから[File]をダブルクリック (またはドラッグ)して、ENVI Modeler の [Untitled] 領域に移動します。
3. [Select Type] ダイアログが出現するので、[Raster]をクリックし、チュートリアルデータを保存した場所に移動し、**NAIP\_DallasTX\_Oct11\_2020\_Subset.tif** をクリックしてから、[Open]をクリックして登録します。



# ENVI Modeler でモデルを作成する



- ふたたび[Basic Nodes]リストから[File]をダブルクリック (またはドラッグ) し、  
今後は[Select Type] ダイアログの[Regions of Interest]をクリックし、  
**NAIP\_DallasTX\_Oct11\_2020\_Subset.xml**を登録します。
- 同様の操作をNAIP\_SanAntonioSE\_2020\_Subset.datというラスタおよび  
NAIP\_SanAntonioSE\_2020\_Subset.xmlというROIについて行い、4つのFileを  
登録します。



# 教師データの準備

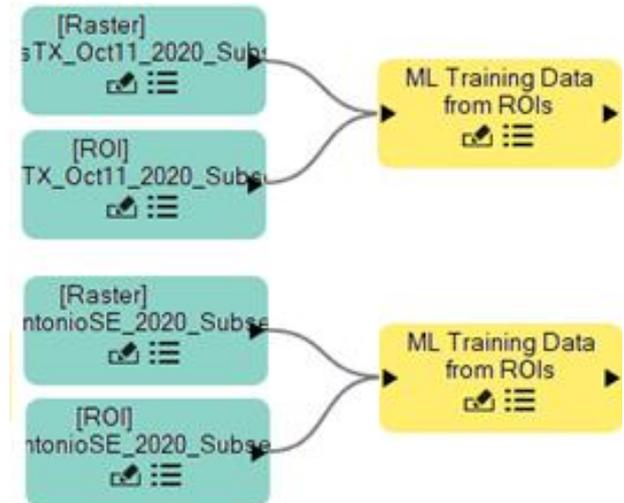
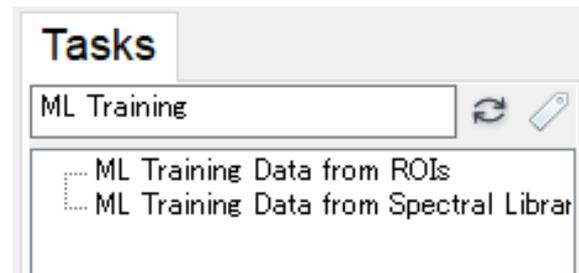


- 教師付き分類を行うためには教師データが必要です。
  - 教師データの作成はENVI Machine Learning ML Training Datafrom ROIsタスクで行います。
  - このタスクのインプットデータは、ラスタと、対象物にラベル付けを行ったROIデータです。
- このチュートリアルでは[Manmade(人工物), Trees (樹木), Ground (地面), Water(水域)]の4つのクラスラベルを使用します。
  - これらには順にクラス値(0,1,2,3)が割り当てられます。

# 教師データの準備



1. ENVI Modelerの[Tasks Search]欄に「ML Training」と入力すると2つのタスクに絞られるので、**ML Training Data from ROIs**を選択します。
2. これはサンアントニオとダラスのそれぞれで使用しますので、Modelerには2つ追加してください。
3. Rasterの出力側を左クリックし、ML Training Data from ROIsの入力側までドラッグします。これによって、ノード同士を繋いでください。この操作を繰り返し、図のようになるまで操作を繰り返してください。

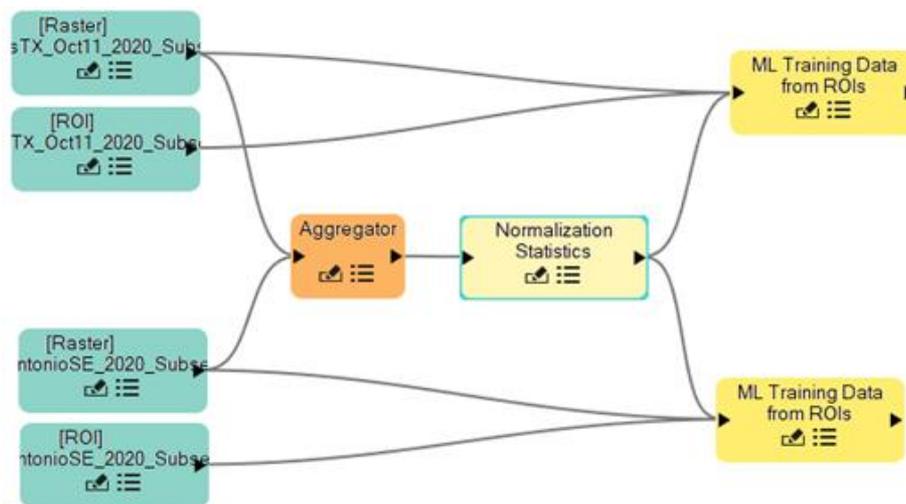


# 教師データの準備



4. ENVI Modelerの[Tasks Search]欄に「Normalization Statistics」と入力し、**Normalization Statistics**を選択・追加します。
5. Basic Nodesに含まれる**Aggregator**を選択・追加します。
6. 2つのRasterの出力側をAggregatorに繋ぎ、AggregatorをNormalization Statistics の入力側に繋がります。Normalization Statistics の出力側を2つのML Training Data from ROIsの入力側に繋ぐと、図のようになります。

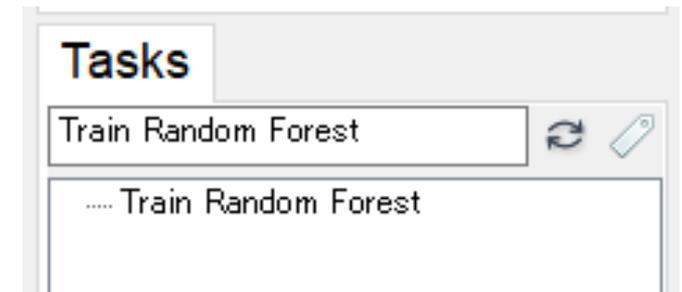
**Normalization Statistics**は、入力ラスタからデータの最小値と最大値を収集します。これはトレーニング時に重要で、データは最小値と最大値を用いて0から1の間でスケールされます。



# トレーニングノードへの接続



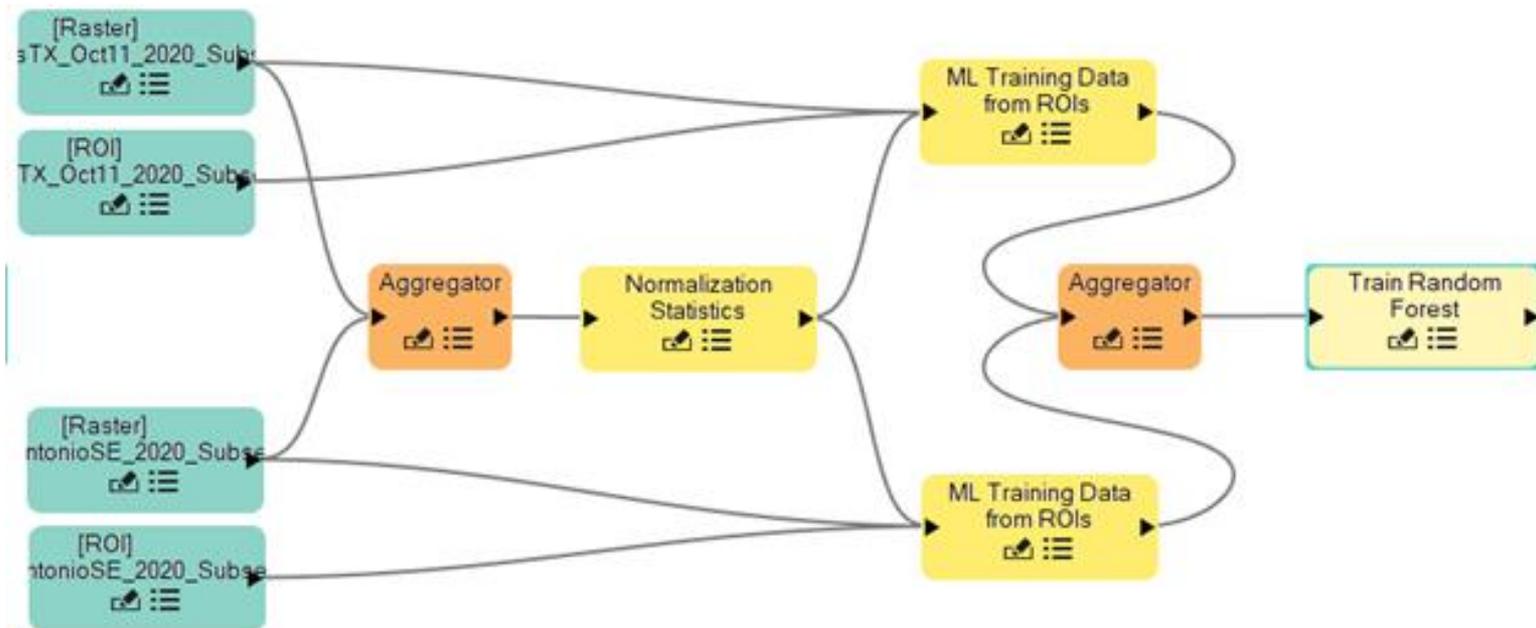
- ML Training Data from ROIで生成された出力データは、教師データです。
    - これを用いて実際に学習を行うには、学習データを集約し、学習タスクノードに接続する必要があります。
    - このチュートリアルでは、Train Random Forestタスクを使用します。
1. ENVI Modelerの[Tasks Search]欄に「Train Random Forest」と入力し、**Train Random Forest**を選択・追加します。
  2. Basic Nodesに含まれる**Aggregator**を選択・追加します。



# トレーニングノードへの接続



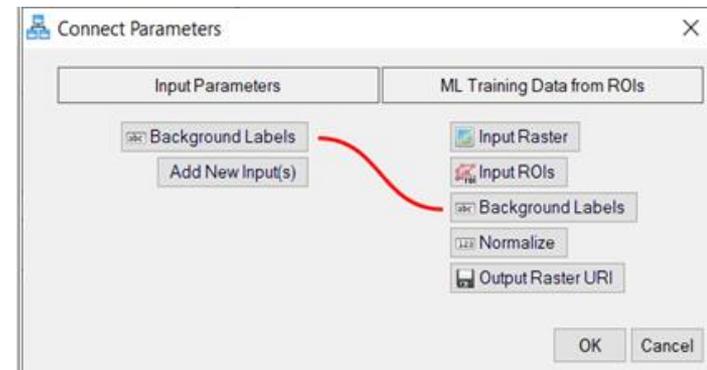
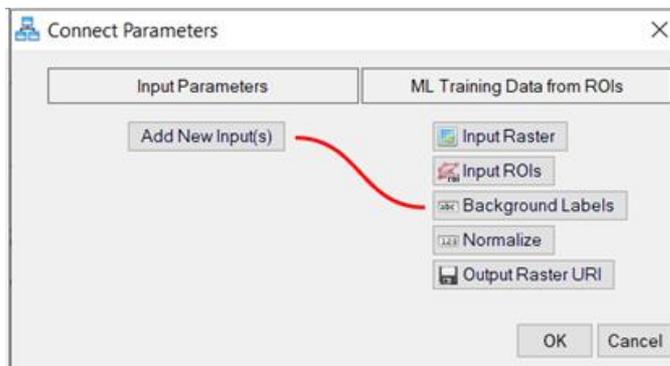
- それぞれのML Training Data from ROIsの出力側を新たに追加したAggregatorに接続し、Aggregatorの出力先をTrain Random Forestの入力側につないで、下図のような形にします。



# Input Parametersの追加



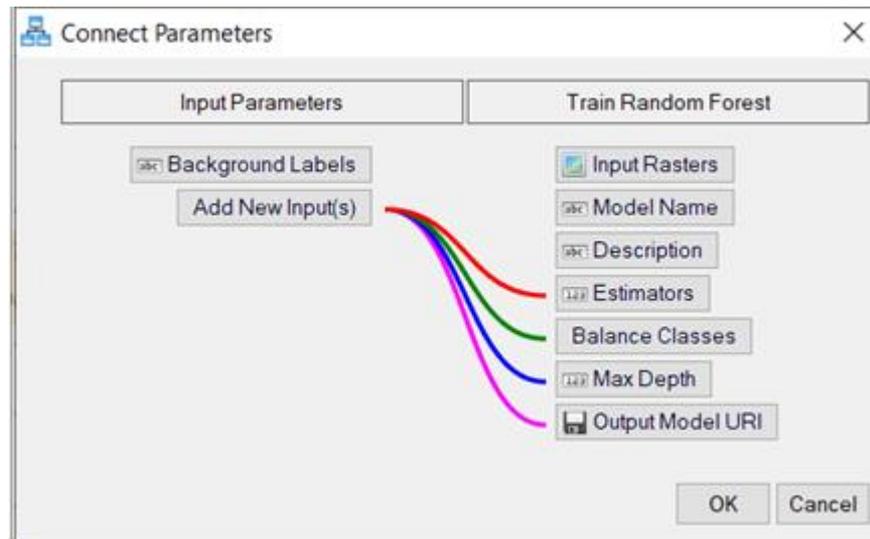
1. Basic Nodesに含まれる**Input Parameters**を選択・追加します。
2. 追加したInput Parametersの出力側を片方のML Training Data from ROIsノードに接続すると、Connection Parameterダイアログが出現します。
3. 左側のInput Parametersは「Add New Input(s)」を、右側のML Training Data from ROIsはBackground Labelsを選択します。この操作により、両要素が赤い線で結ばれたことを確認してOKボタンを押します。
4. 追加したInput Parametersの出力側をもう一方のML Training Data from ROIsノードに接続すると、Connection Parameterダイアログが再び出現します。
5. Input Parameters側のBackground Labelsを選択すると対応する要素が自動で接続されるので、OKボタンで保存して閉じます。



# Input Parametersの追加



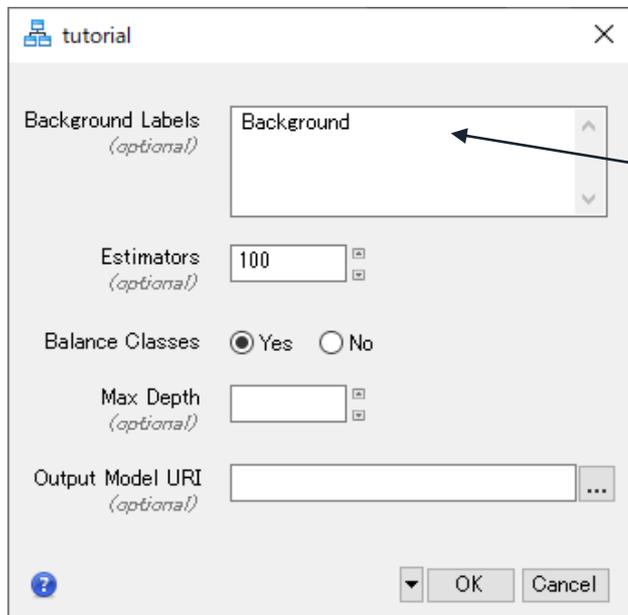
6. Input ParametersとTrain Random Forest nodeを接続します。
7. Train Random Forestの「Estimators」「Balance Classes」「Max Depth」「Output Model URI」をそれぞれ選択すると、これらがAdd New Input(s)に接続されます。図のようになっていることを確認し、OKボタンを押下します。
8. 実行の準備が整ったので、ENVI Modelerメニューバー File > Save Asを選択し、適当な場所に「tutorial.model」の名前で作成したモデルを保存します。



# 学習の実行



1. ENVI ModelerメニューバーのRunボタンを選択すると、パラメータ入力のダイアログが出現します。
2. Background Labelsフィールドで、「Background」というテキストを削除し、このフィールドを空白にします。



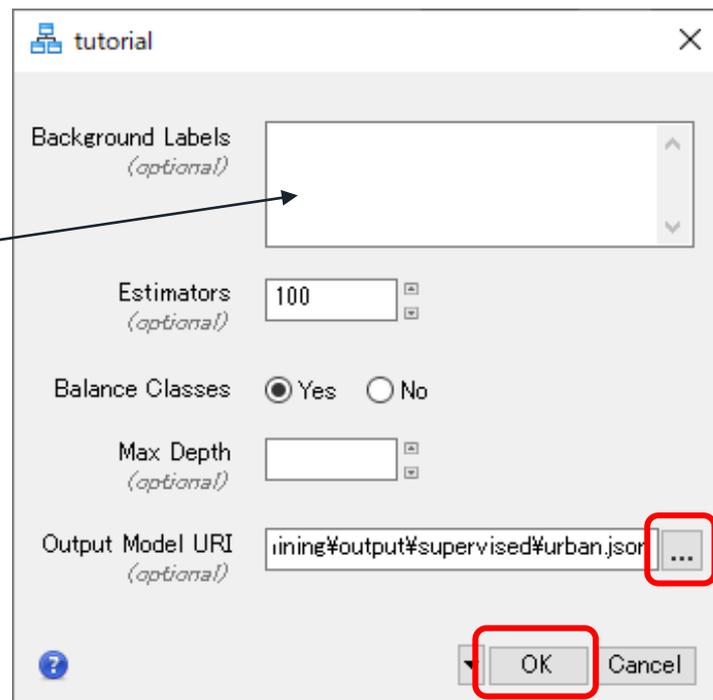
オプションとして、ラベリングプロセスで作成した3つまたは4つのラベル[Manmade、Trees、Ground、Water]を入力することができます。例えば、樹木だけを識別したい場合は、[Manmade、Ground、Water]のラベルを追加します。（つまりこれらを背景とみなします）

# 学習の実行



3. Estimators, Balance Classes, Max Depthパラメータはそのままとし、Output Model URIはブラウズボタンを押下して、適当な名前(ここではurban.json)という名前のjson形式のファイルで保存します。
4. OKボタンを押下するとモデルが学習が開始されます。
5. 学習が終了したら、ENVI Modelerを閉じます。

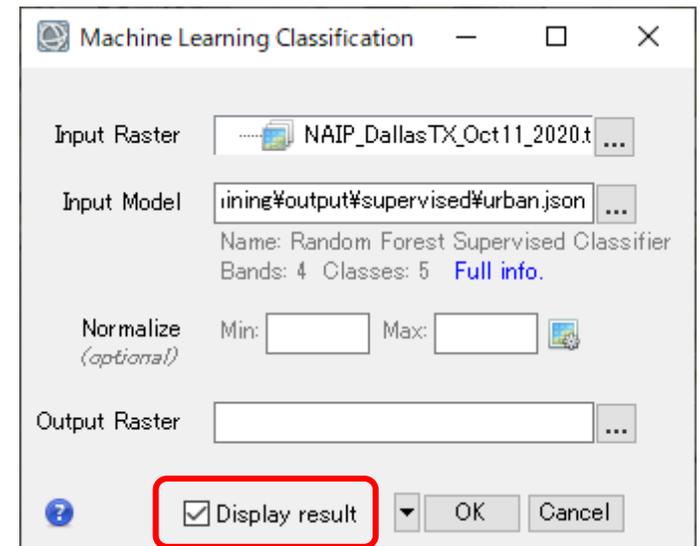
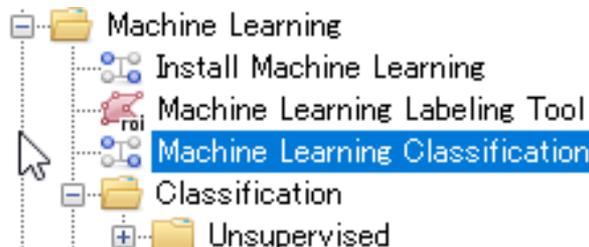
Background Labelsを空欄としているので、学習されたモデルは背景にあたるラベルを持たず、すべてのピクセルが[Manmade、Trees、Ground、Water]のいずれかに分類するモデルとなります。



# 分類の実施



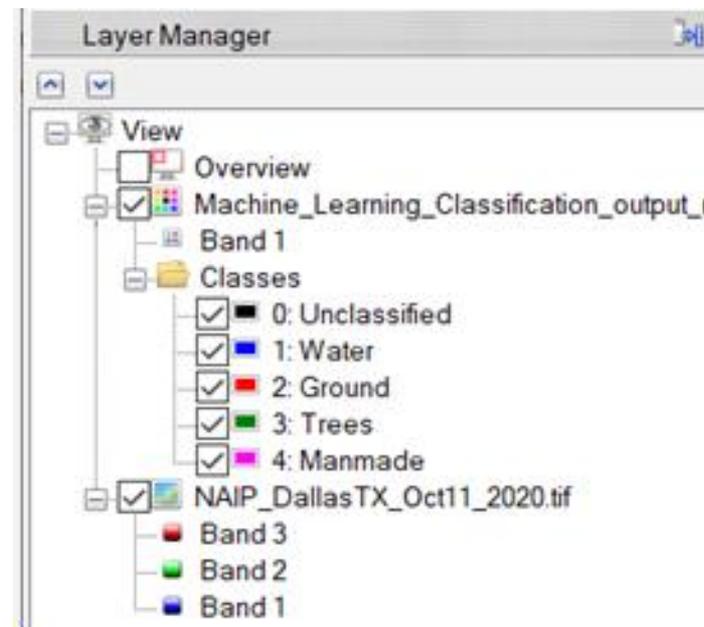
1. ENVIツールボックスの/Machine Learning/Machine Learning Classificationをダブルクリックし起動します。
2. Input Rasterにチュートリアルデータ supervised¥classificationの **NAIP\_DallasTX\_Oct11\_2020.tif**を設定します。
3. Input Modelに先ほど学習した **urban.json**を設定します。
4. NormalizeおよびOutput Rasterを空欄のままとしてOKボタンを押下して分類を実施します。（Display Resultにはチェック入れてください。）



# 分類の実施



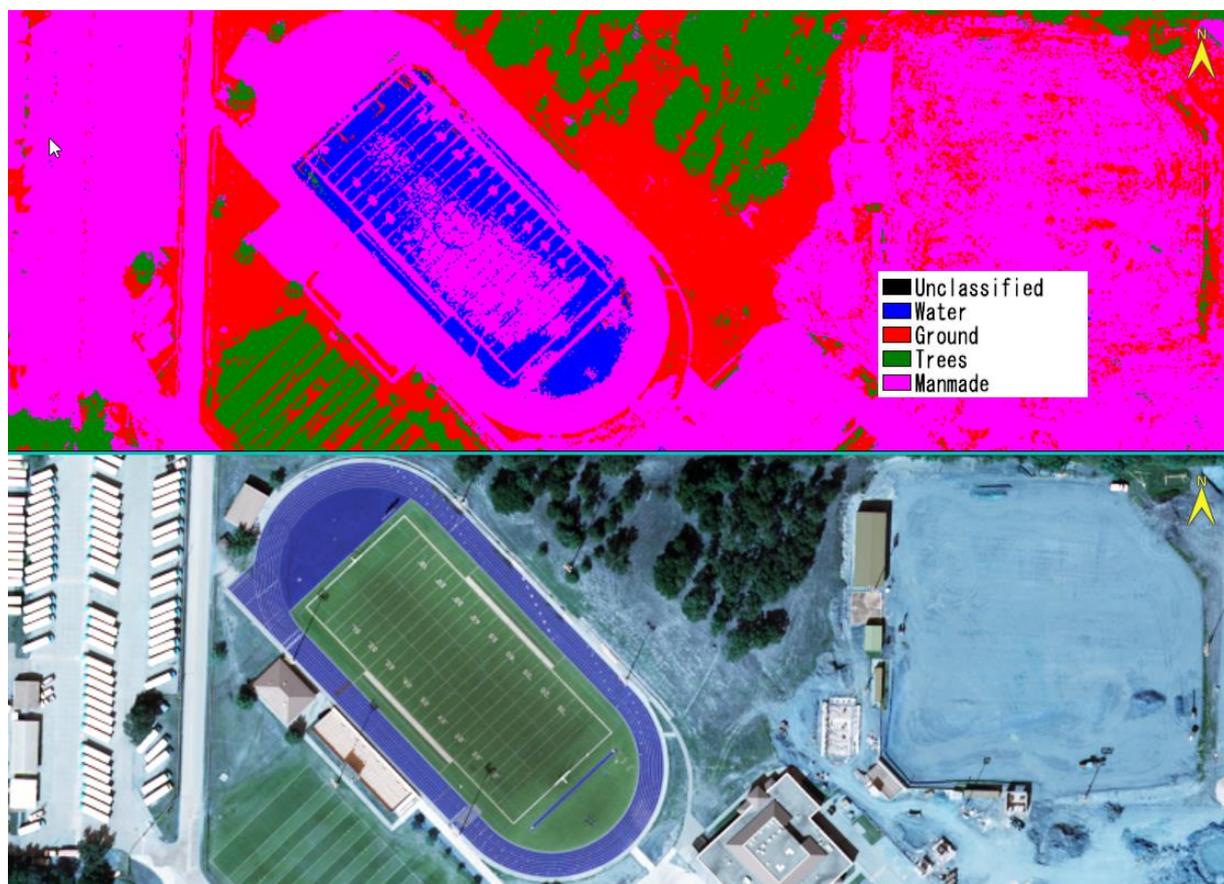
5. Layer Managerに分類結果が表示されます。
6. Data Managerを開き、分類元となったNAIP\_DallasTX\_Oct11\_2020.tifを右クリック > Load Defaultで開きます。
7. Layer Managerで順番を入れ替え、分類結果が上に来るようにしてください。



# 分類の実施



- Transparencyの変更やビューの分割などで結果の比較を行います。
  - 例えば、メニューバーGo Toフィールドに[7784p,7148p]と入力するとスタジアムがビューの中心にくるので結果を確認してみてください。





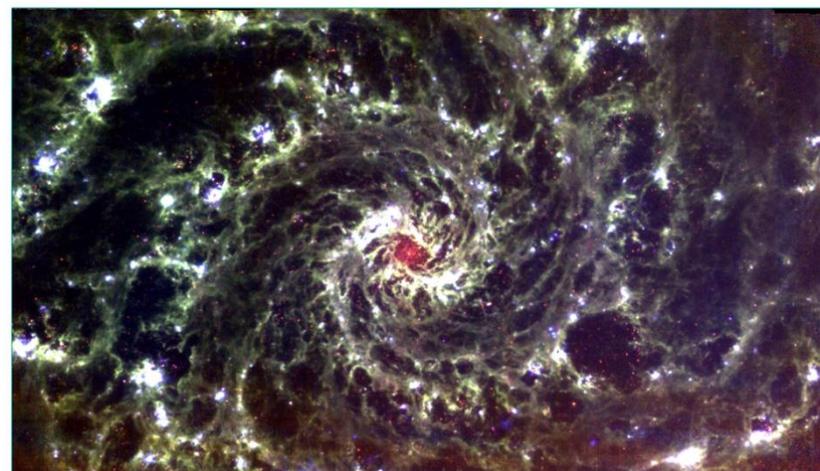
# チュートリアル③ 教師なし分類(Unsupervised Classification)

[https://www.l3harrisgeospatial.com/docs/envimltutorial\\_unsupervisedclassification.html](https://www.l3harrisgeospatial.com/docs/envimltutorial_unsupervisedclassification.html)

# 使用するデータ



- このチュートリアルで使用される画像は、[James Webb Space Telescope \(JWST\)](#) NASA JPL によって収集されました。
  - 画像はパブリック ドメインにあります。ライセンス条項は、[CC BY 4.0](#)によって管理されます。
- この画像は、地球から 3200 万光年離れたメシエ 74 としても知られる渦巻銀河 NGC 628 です。
  - これは、3つのシングルバンド ラスタから構築され、ENVI ファイル形式で保存された 3 バンド イメージです。



JWST\_ngc628.dat



- 教師なし学習は、対象となるシーンやその特徴量になじみのないユーザーにとって有用です。
  - 同じような特徴でグルーピングが行われ、データ内の未知の類似点/相違点の分類や抽出に役に立ちます。
- メリット:
  - 事前に教師ラベルを取得する必要はなく、このため、時間の節約や手動ラベリングによるエラーを回避することができます。
  - 人間の目では認識しにくい特徴やパターンを識別できる可能性があります。
- デメリット:
  - 分類器によって特定された特徴やパターンに意味を付加するのは、ユーザーの役割となります。
  - 要求する分類数が「多い」と結果が飽和し、「少ない」とより多くの特徴を要求した場合に見えていたかもしれない情報を隠してしまう可能性があります。

# 教師なし分類器の学習と分類

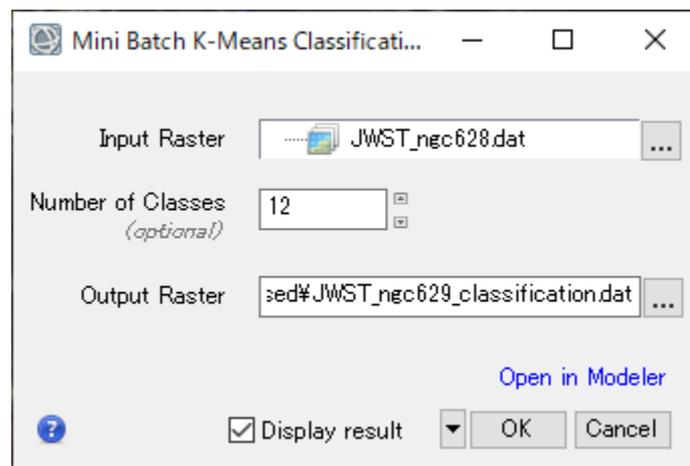


- このチュートリアルでは、単一のラスタでトレーニングと分類を行う **Mini Batch K-Means Classification** タスクを使用します。
  - 単一のラスタを使用することで、トレーニングに使用された画像に特化したモデルが生成されます。
  - より一般化されたモデルを作成するには、ENVITask API または ENVI Modeler を使用し複数のラスタを使用した学習が必要となります。
- Mini Batch K-Means は、KMeans アルゴリズムの変形で、ミニバッチ（入力データのサブセット）を反復使用し、計算時間を短縮しながら、最適な結果を探ろうとするアルゴリズムです。

# 教師なし分類器の学習と分類



1. ENVIツールボックスの /Machine Learning/Classification/Unsupervised/Mini Batch K-Means Classification をダブルクリックしツールを起動します。
2. Input Raster(入力ラスタ)にJWST\_ngc628.dat、Number of Classes(分類数)に3から12の任意の値、Output Raster(出力ラスタ名)では適当なフォルダ配下に JWST\_ngc629\_classification.datの名前で結果を保存します。
3. Display resultにチェックが入っていることを確認し、OKボタンで実行します。



# 教師なし分類器の学習と分類



- 結果がLayer Managerに追加されるので、TransparencyやViewの分割を使用して結果を分類元となったデータと比較してください。

