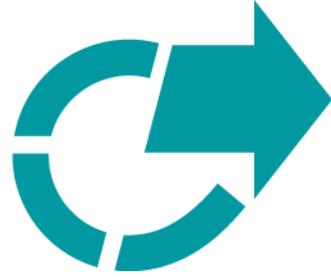


Frischer Wind für ein bekanntes Thema – Klassifikation 4.0

AK Fernerkundung
Heidelberg, 4.-5.10.2018

Ruth Leska, GEOSYSTEMS GmbH

GEOSYSTEMS
THE GEOSPATIAL EXPERTS



GEOSYSTEMS ist Ihr Partner für
Geo-IT Lösungen und **Hexagon
Geospatial Platinum Partner**
mit herausragender Expertise in
Geodatenmanagement, Smart
M.Apps, GIS, Photogrammetrie
und Fernerkundung.





Portfolio



Hexagon Geospatial Portfolio

GEODATEN AUSWERTEN UND VERTEILEN



APPS ERSTELLEN & VERKAUFEN



M.App Portfolio

A simpler way to see your world

M.App Portfolio® ist eine cloudbasierte Plattform, die es den Hexagon Geospatial Partnern ermöglicht, **Hexagon Smart M.Apps zu entwerfen, zu bauen und zu veröffentlichen**. Das M.App Portfolio untergliedert sich in die M.App Foundation, das M.App Studio und den M.App Exchange Marktplatz.

M.App Enterprise® ist eine On-Premise-Plattform, um für Ihre Organisation **geographische Apps zu erstellen**. M.App Enterprise speichert Ihre Bilddaten, Vektoren und Punktwolken, Ihre Workflows, Analysen und Abfragen, und all das ist über eine einfach zu bedienende Nutzeroberfläche zugänglich.



GEOSYSTEMS Produkte

UAV-DATEN

Photogrammetrisch
prozessieren



ATMOSPHERE

korrigieren & Dunst
entfernen



SENTINEL-DATEN

downloaden & prozessieren



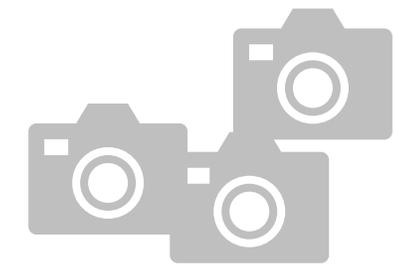
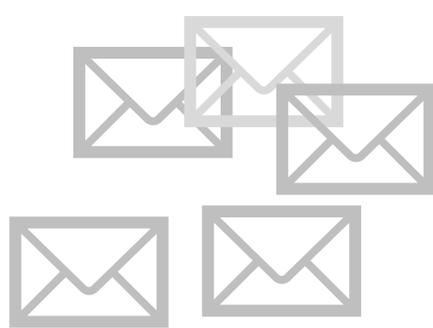
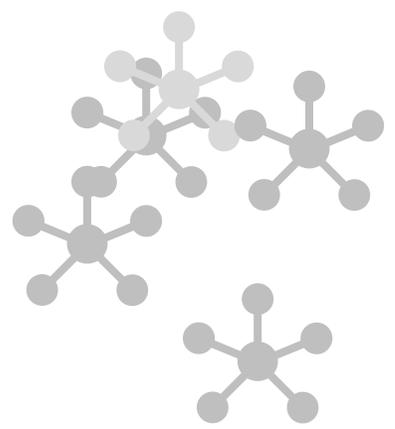


Give it to the machines and let them sort it out

Machine Learning und Deep Learning

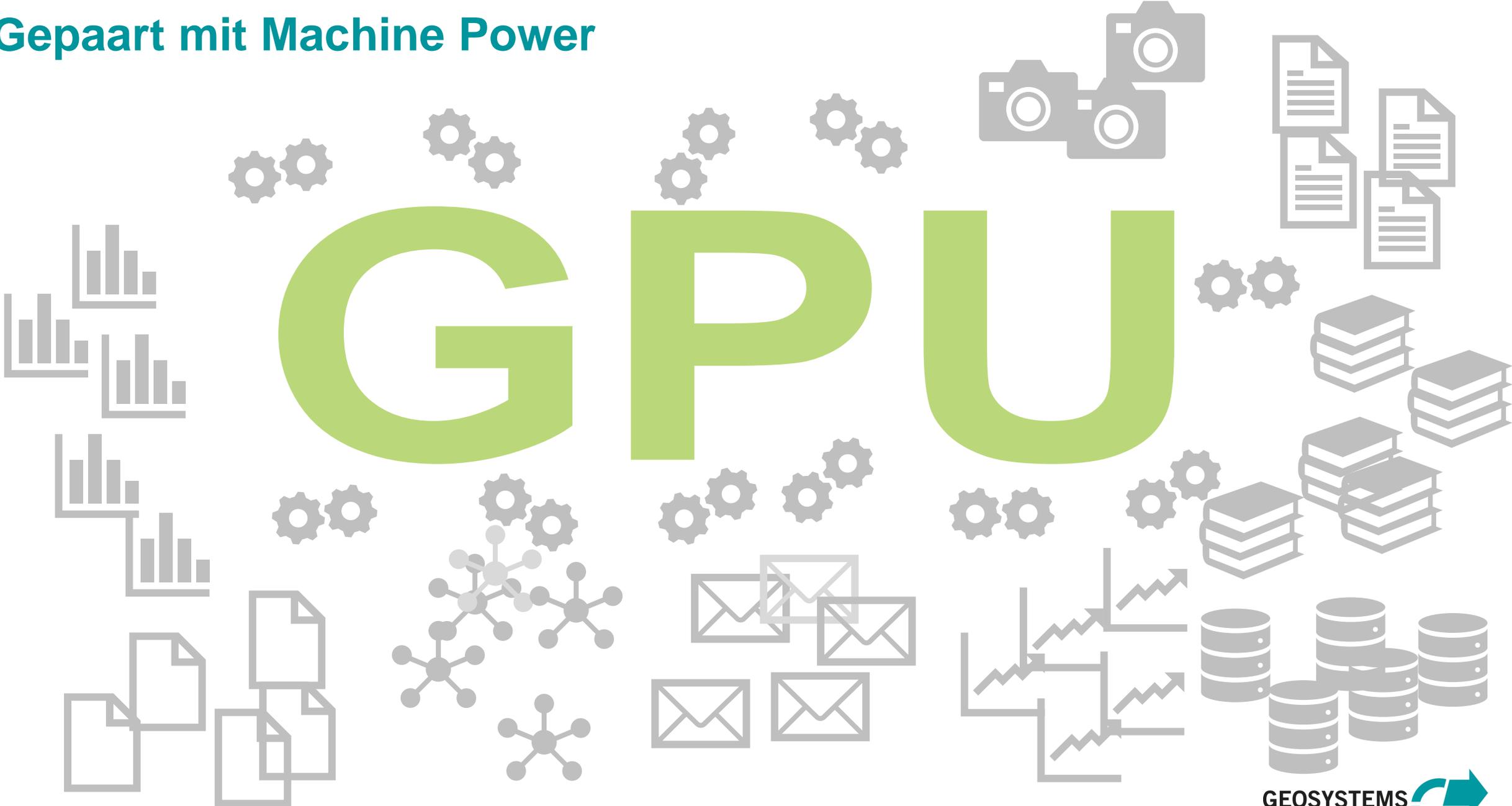
Ein großes Thema der letzten Jahre

BIG DATA



Gepaart mit Machine Power

GPU



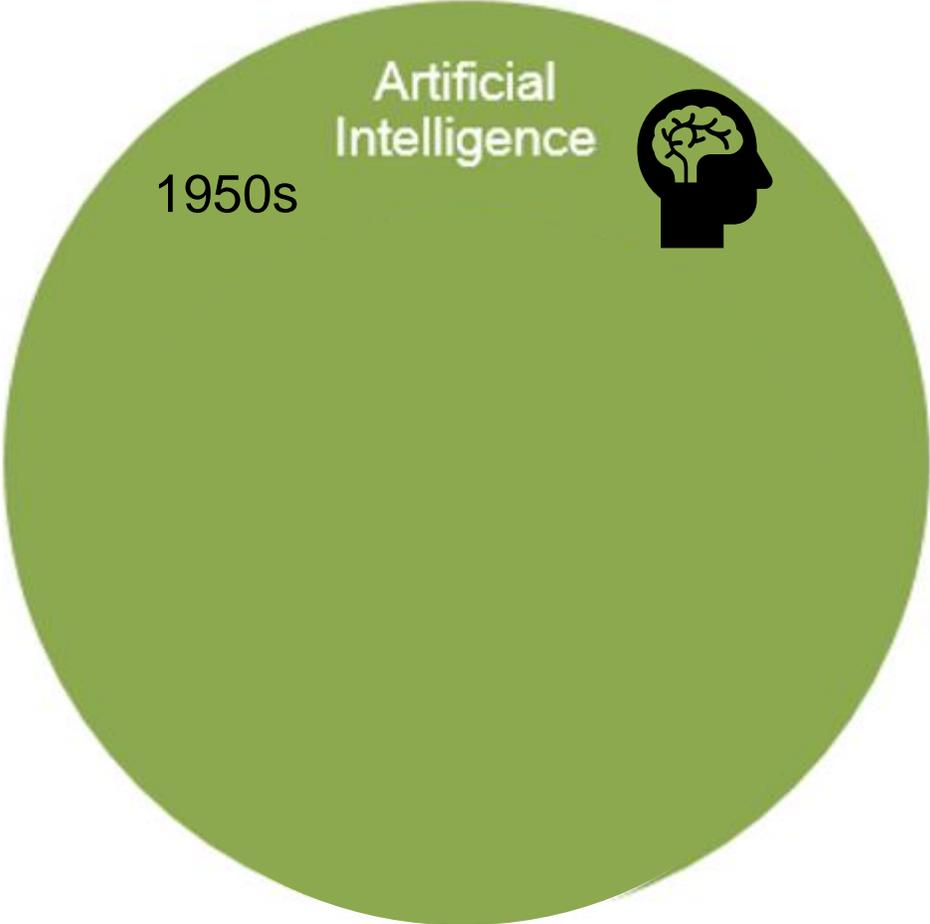
“Give it to the machines and let them sort it out”

➔ MACHINE LEARNING

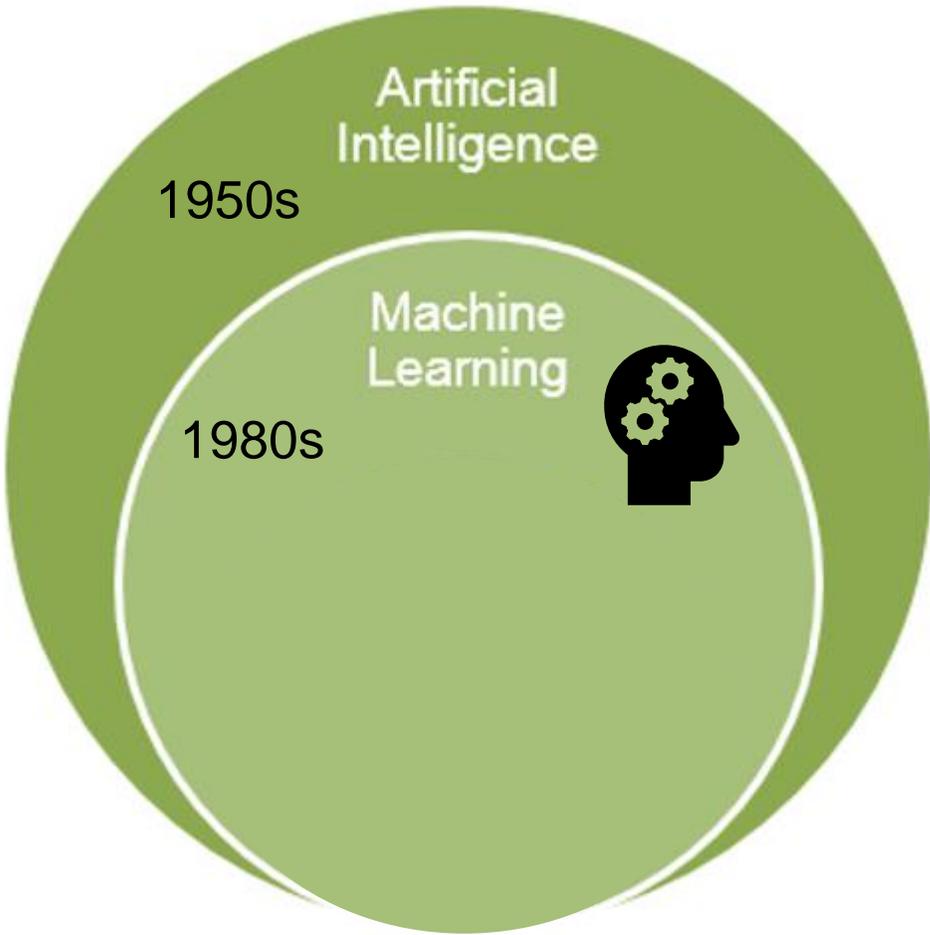


Machine Learning ist ein Ansatz, Daten anhand „unsichtbarer“ Muster zu erkennen – ohne zu wissen wie, wo und nach was zu suchen ist

Künstliche Intelligenz

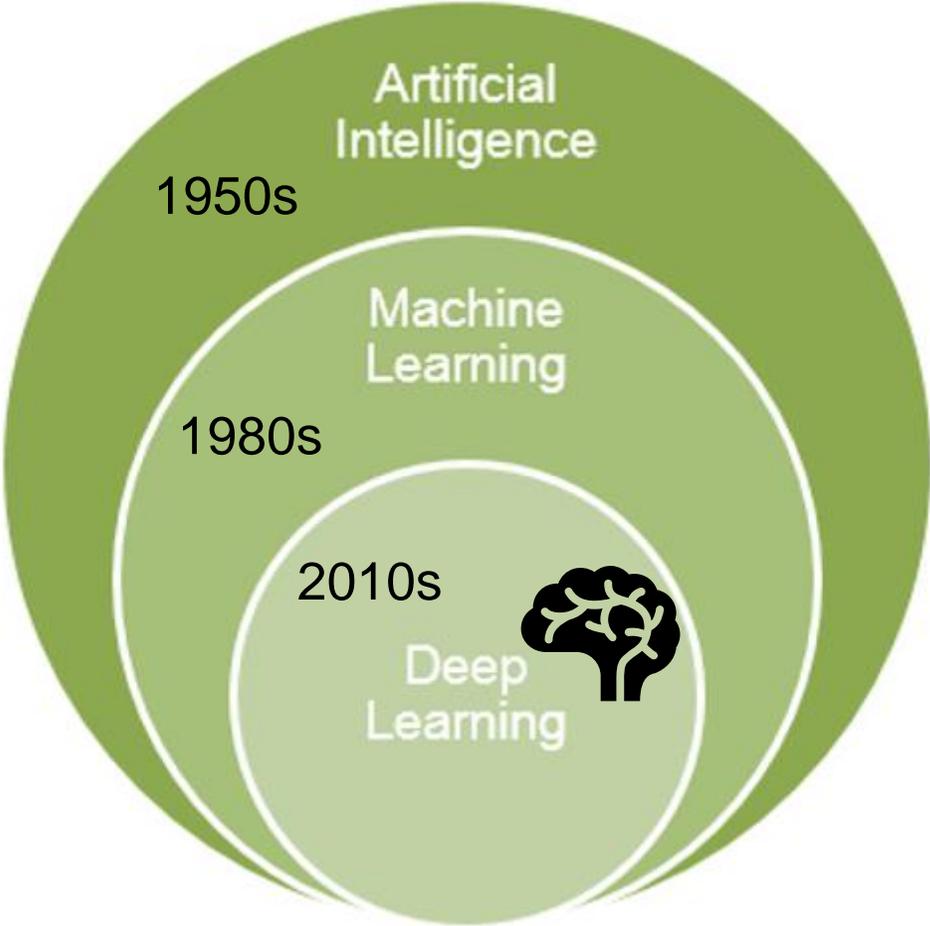


Machine Learning



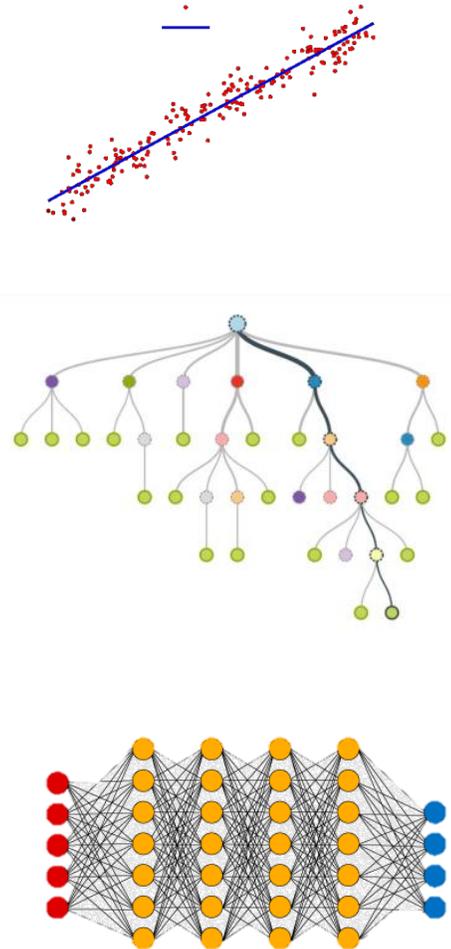
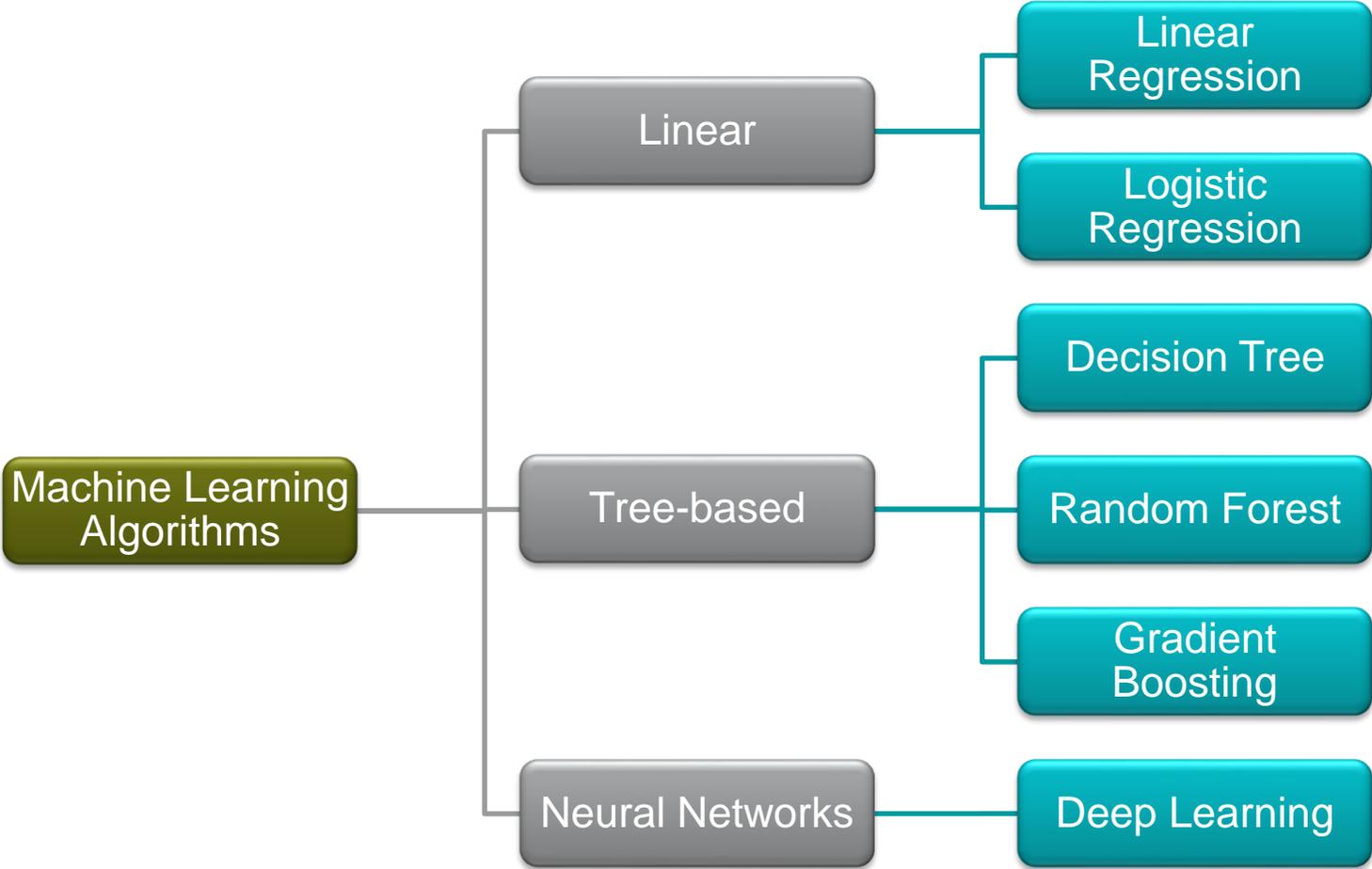
K Means oder ISODATA
Bayesian Networks

Deep Learning



DNNs - Deep Neural Networks

Machine Learning Algorithmen



Wie funktioniert Machine Learning?

Train



Classify



Machine Learning packages



Caffe

The word 'Caffe' is written in a large, red, serif font.

theano

The word 'theano' is written in a large, blue, sans-serif font.

Machine Learning in der Geospatial Welt?



Die meisten Machine Learning Bibliotheken sind nicht für die Geospatial-Welt gemacht.



Komplexe Machine Learning Ansätze sind noch in der Forschung und brauchen einen Experten.



Eigene Abläufe und Lösungen entwickeln ist nicht einfach.

Machine learning-Bedarf in der Geospatial Welt

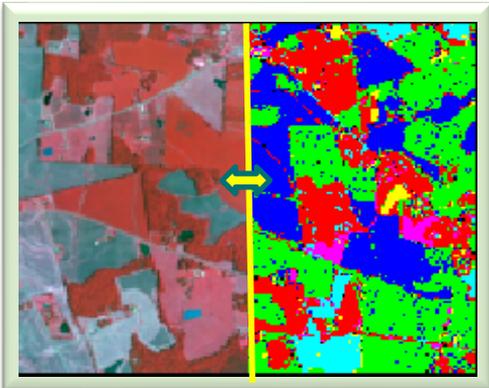


Image Classification



Change Detection



Feature Extraction



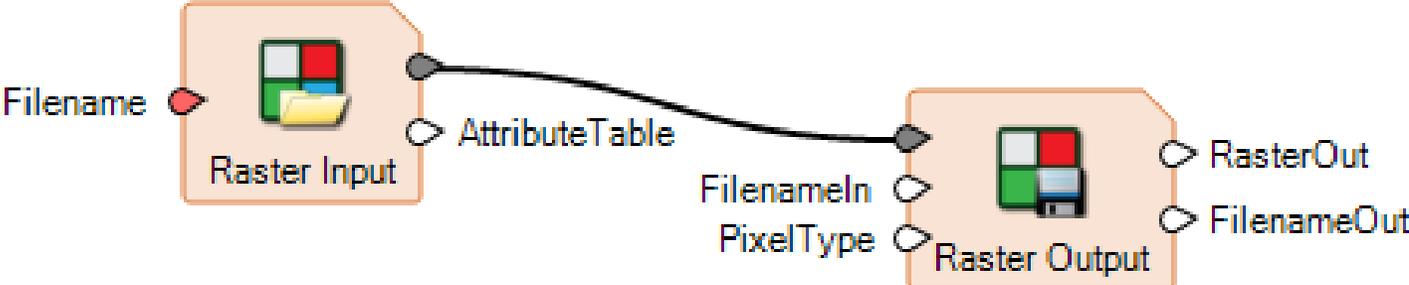
Prediction



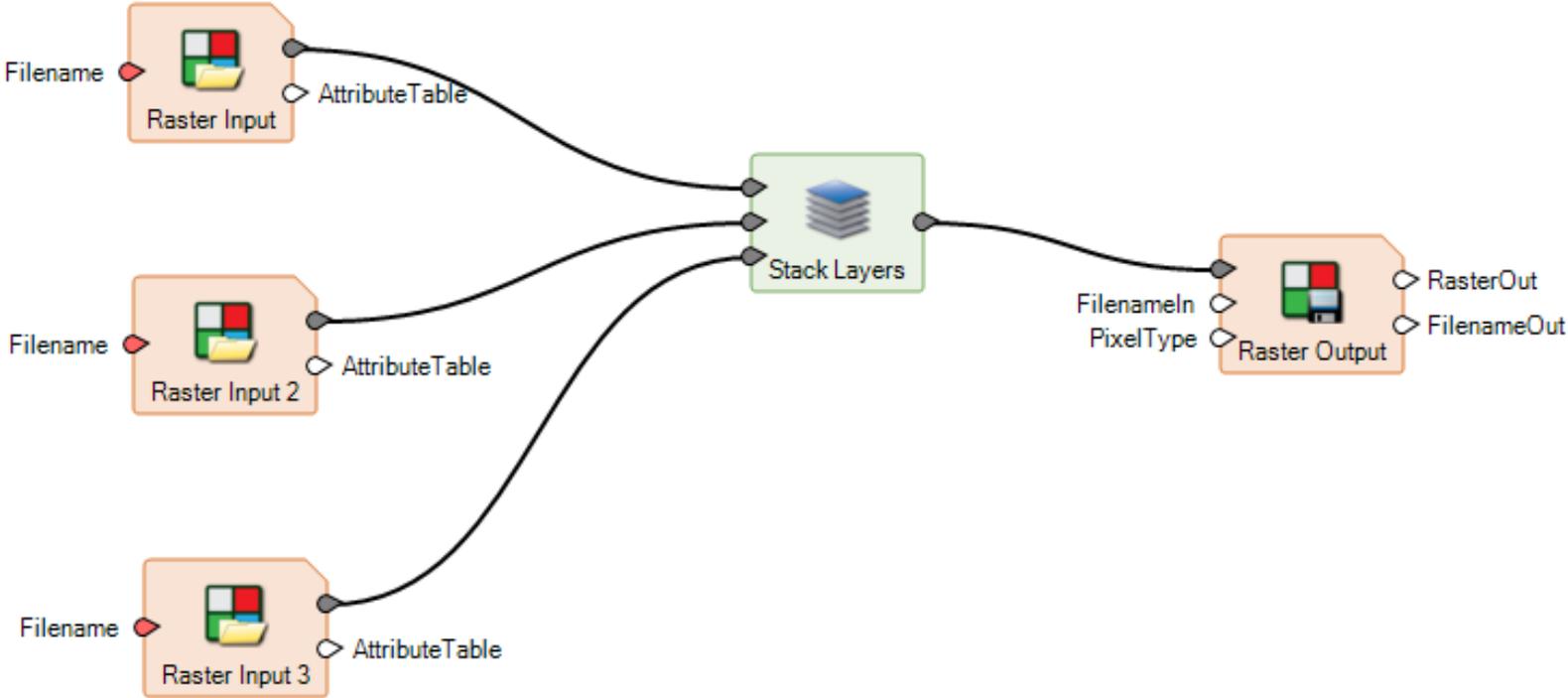
Machine Learning mit dem Spatial Modeler



Einfachstes Spatial Model



Layer Stack im Spatial Model



Wo findet sich der Spatial Modeler?

- IMAGINE Professional
- IMAGINE Essentials (Ausführen von Modellen)
- GeoMedia Professional
- ERDAS APOLLO Professional
- M.App Studio
- M.App Enterprise

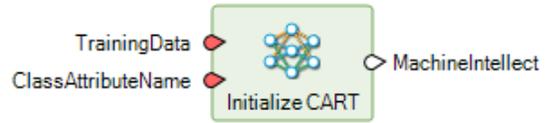


Verwendete OpenSource-Bibliotheken

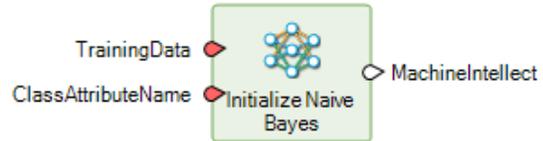


Caffe

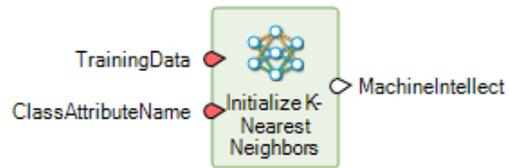
Machine Learning Operatoren im Spatial Modeler



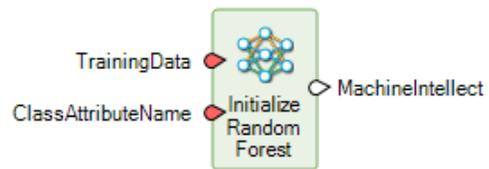
Classification and Regression Trees



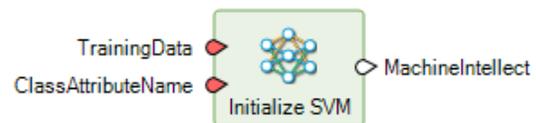
Naive Bayes



K-Nearest Neighbors

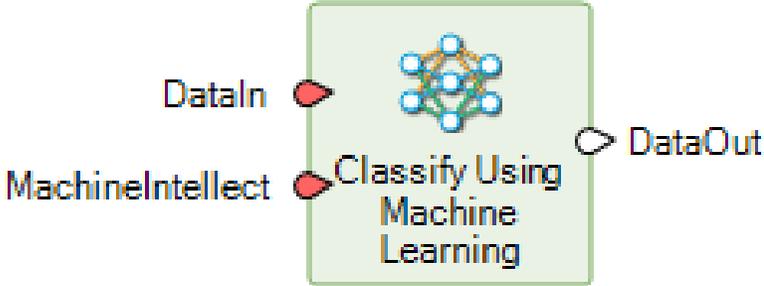


Random Forests



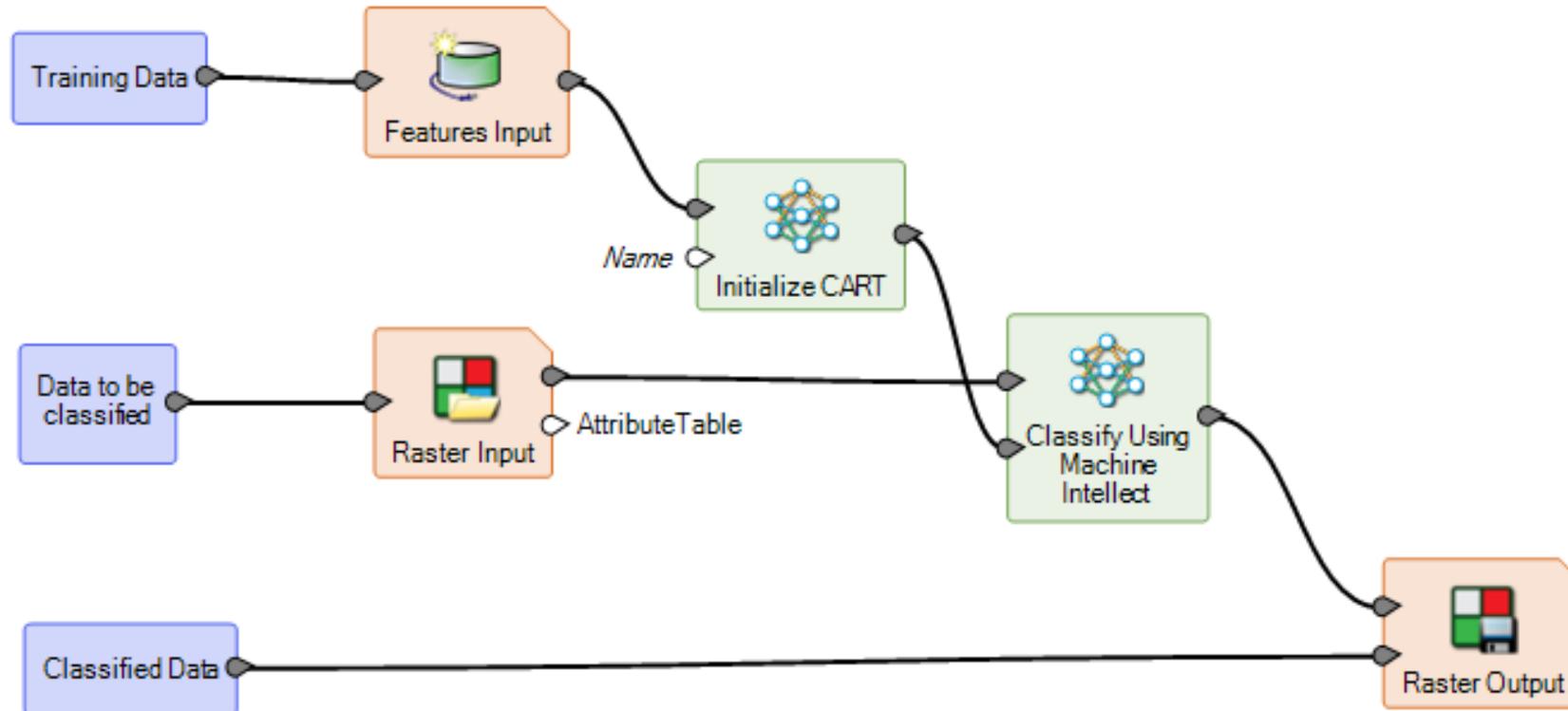
Support Vector Machines

Classify Using....



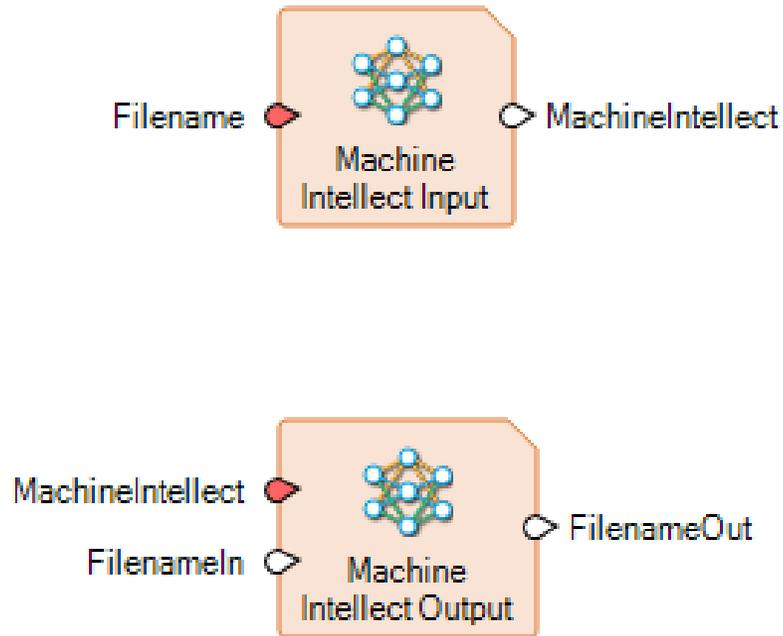
Unsupervised Classifier

Beispiel-Model für den CART-Algorithmus

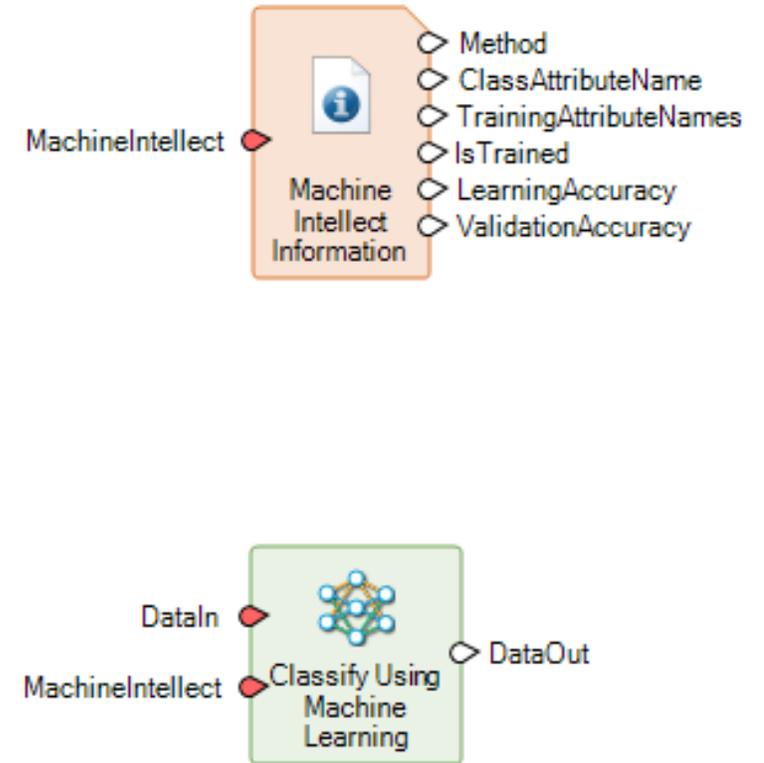


Machine Intellect

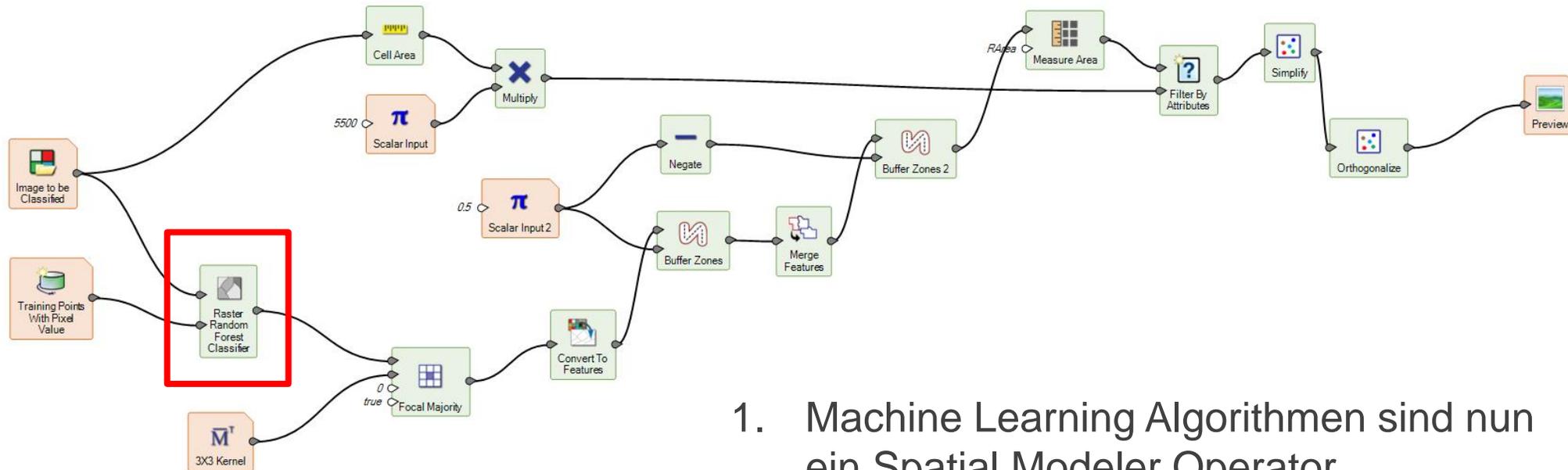
Gut beschrifteten!



Machine Intellect

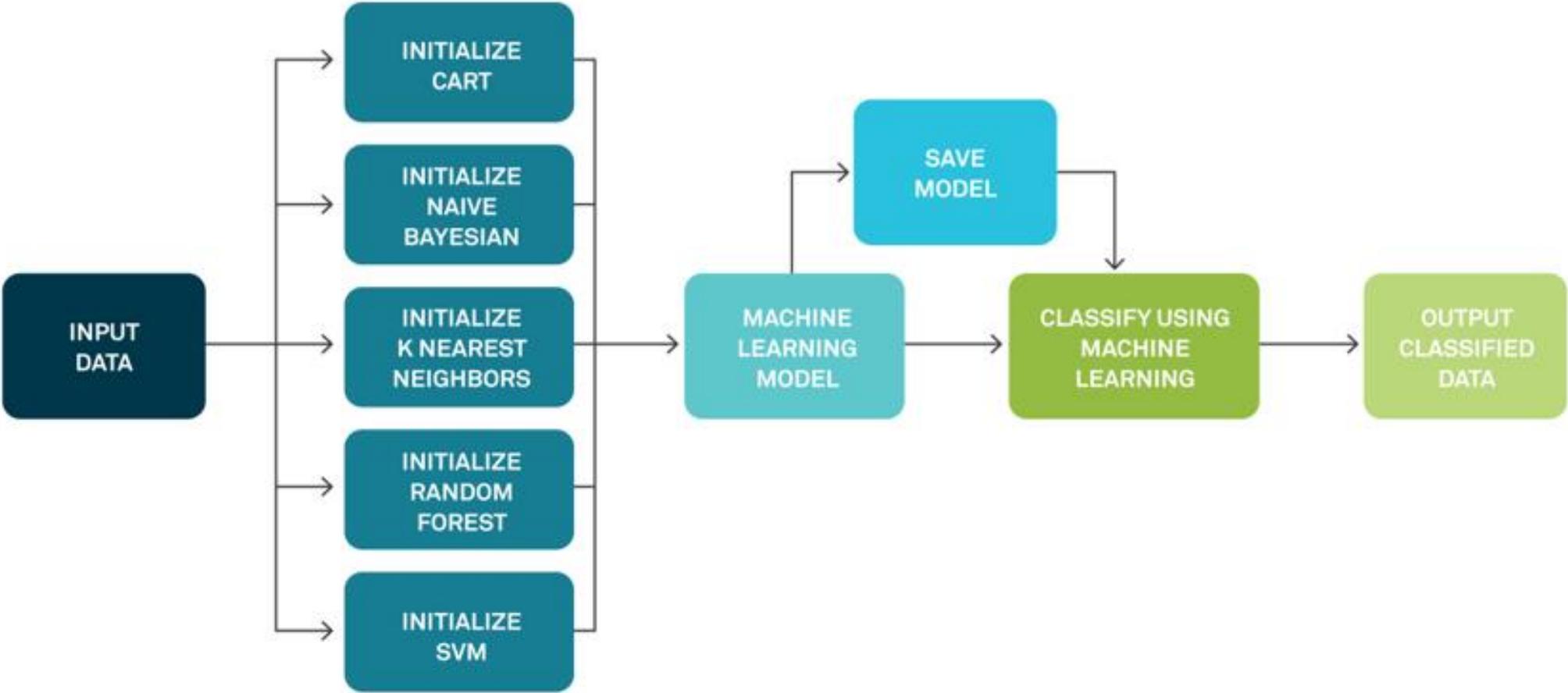


Machine Learning mit dem Hexagon Geospatial Spatial Modeler



1. Machine Learning Algorithmen sind nun ein Spatial Modeler Operator.
2. Kombination von Machine Learning mit den Operatoren des Spatial Modelers.
3. Einbauen/Integrieren von Machine Learning Algorithmen mit vorhandenen Workflows.

Ablaufschema Machine Learning in IMAGINE



Phasen eines Machine Learning Projekts

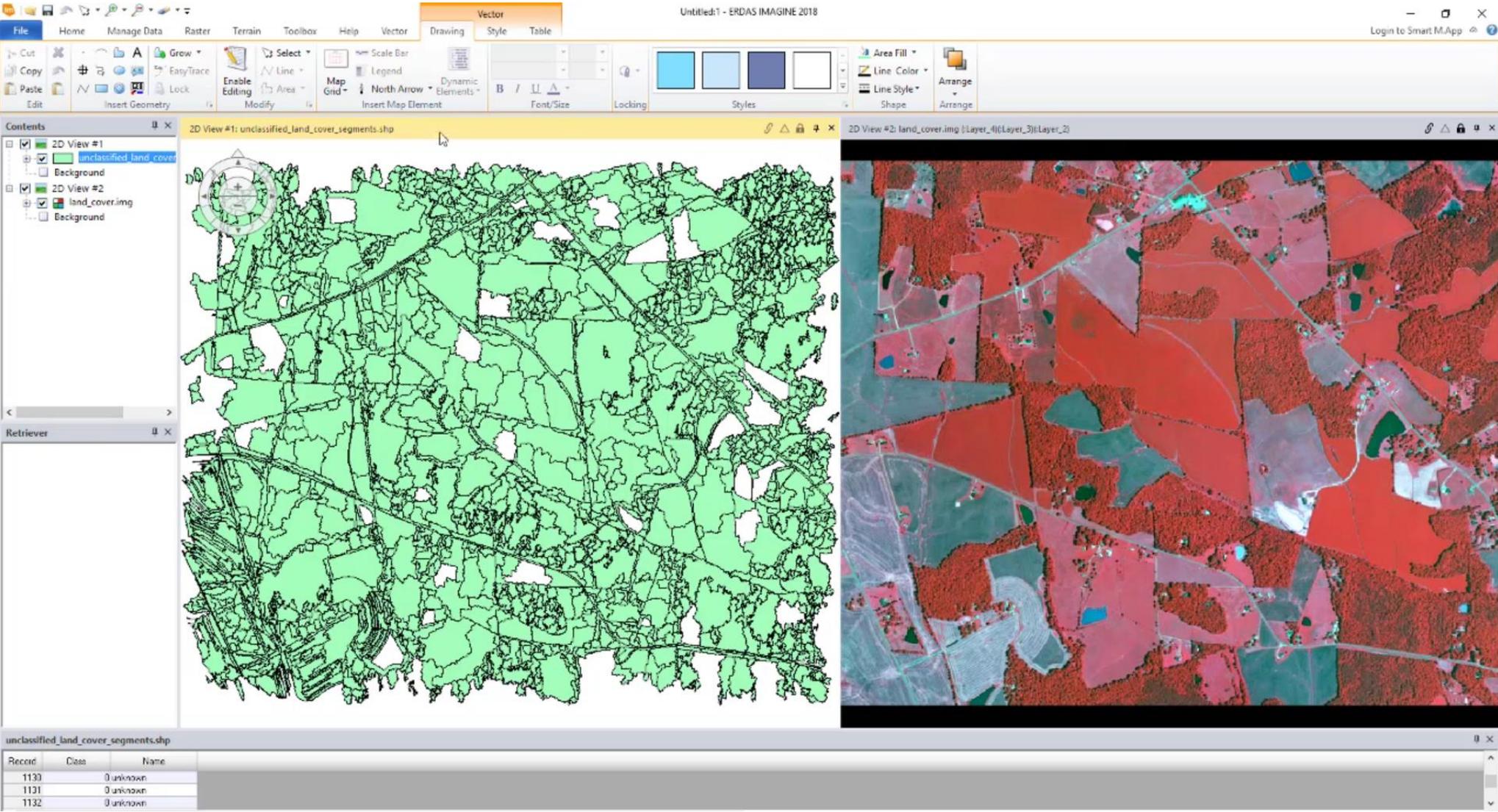
- Vorbereitung der Trainingsdaten
- Auswahl des Machine Learning Modells
- Modell trainieren
- Testen und Validieren
- Anwenden



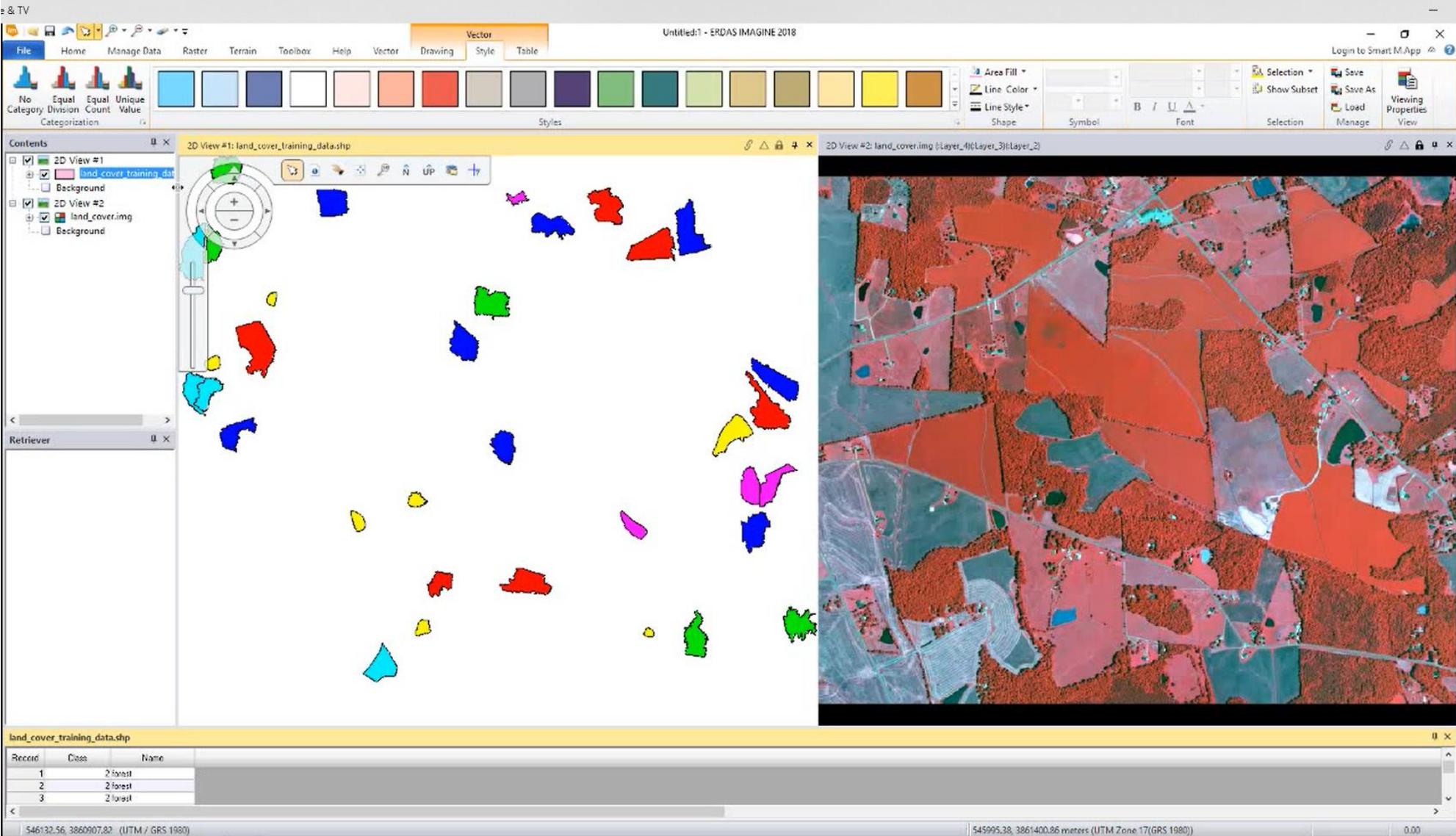
Machine Learning Beispiel - Vektorklassifikation



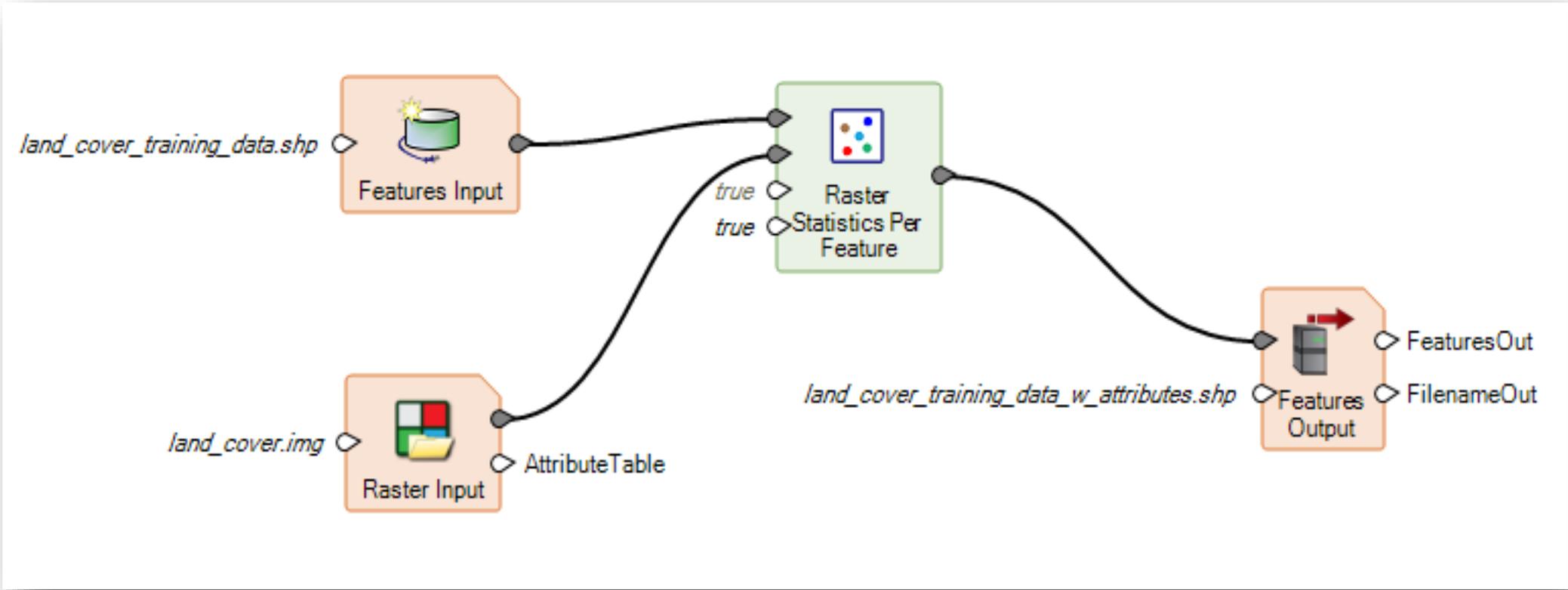
Zu klassifizierende Vektoren und Rasterdatensatz für das Training



Vektor-Trainingsgebiete



Trainingsdaten aufbereiten – Statistik als neues Attribut



Trainingsdaten aufbereiten – Mean + StdDev

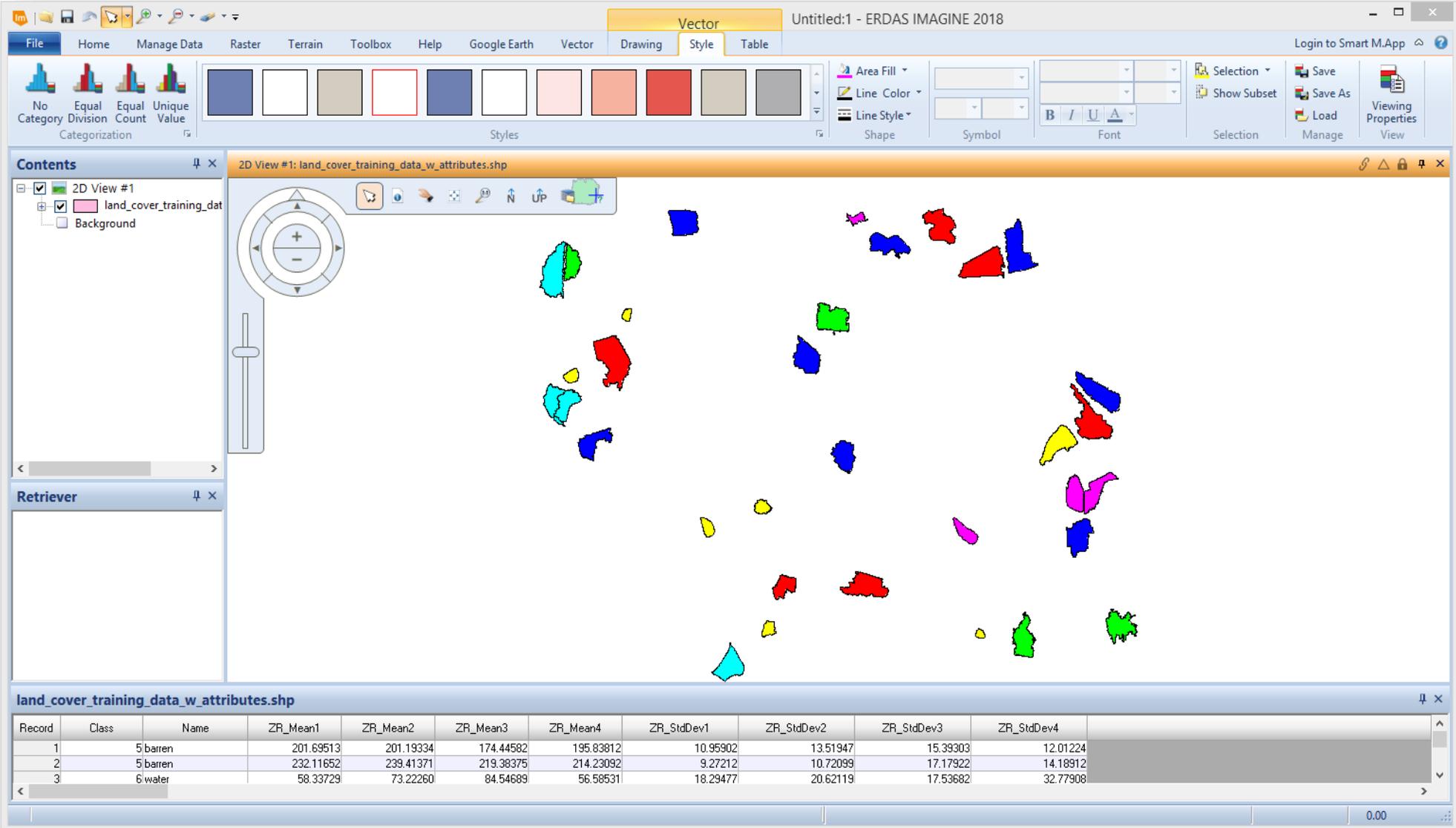
The diagram illustrates a GIS workflow for preparing training data. It starts with two input files: *land_cover_training_data.shp* and *land_cover.img*. The *land_cover_training_data.shp* file is processed by a **Features Input** tool, and the *land_cover.img* file is processed by a **Raster Input** tool. Both tools feed into a **Statistics** tool. The **Statistics** tool has the following settings:

- Raster**: true
- Statistics Features**: true

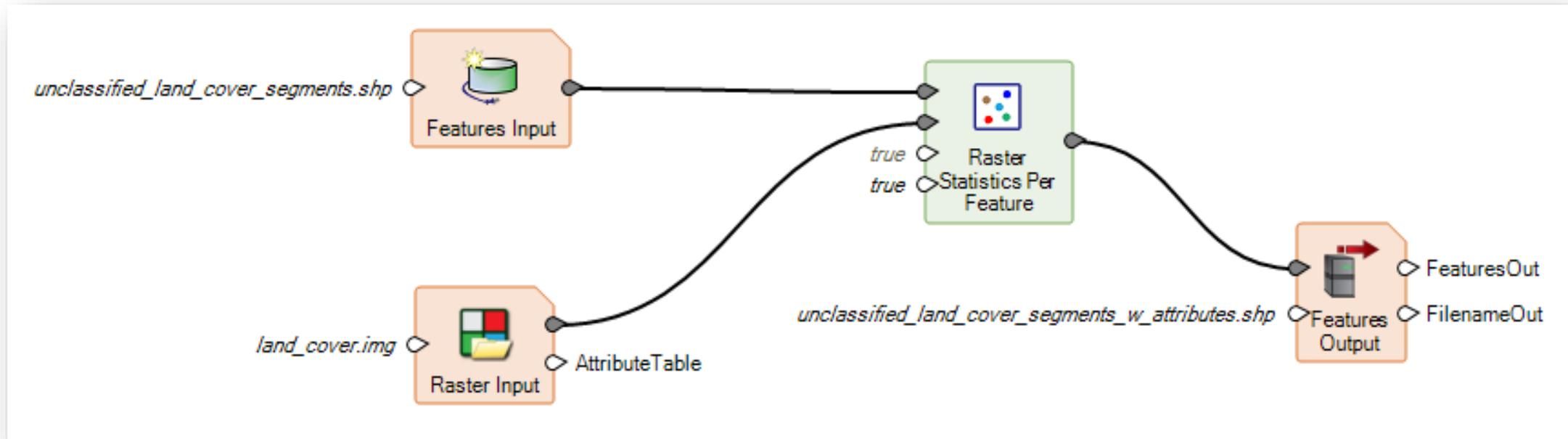
The **Properties** window for the **Statistics** tool is open, showing a list of statistical operations. The **ComputeStdDev** operation is checked.

Show	Name	Value	Objects Supported	Required
<input checked="" type="checkbox"/>	Raster		IMAGINE.Raster	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	ComputeMean	true	IMAGINE.Boot	
	MeanAttribute	"Mean"	IMAGINE.String	
	ComputeMax	false	IMAGINE.Boot	
	MaxAttribute	"Max"	IMAGINE.String	
	ComputeMin	false	IMAGINE.Boot	
	MinAttribute	"Min"	IMAGINE.String	
	ComputeMode	false	IMAGINE.Boot	
	MedianAttribute	"Median"	IMAGINE.String	
	ComputeMajority	false	IMAGINE.Boot	
	ModeAttribute	"Mode"	IMAGINE.String	
<input checked="" type="checkbox"/>	ComputeStdDev	true	IMAGINE.Boot	
	StdDevAttribute	"StdDev"	IMAGINE.String	
	ComputeDiversity	false	IMAGINE.Boot	
	DiversityAttribute	"Diversity"	IMAGINE.String	
	ComputeMajority	false	IMAGINE.Boot	
	MajorityAttribute	"Majority"	IMAGINE.String	
	ComputeMinority	false	IMAGINE.Boot	
	MinorityAttribute	"Minority"	IMAGINE.String	
	ComputeSum	false	IMAGINE.Boot	
	SumAttribute	"Sum"	IMAGINE.String	
	ComputeCount	false	IMAGINE.Boot	
	CountAttribute	"Count"	IMAGINE.String	
	AttributeNan	"ZR_"	IMAGINE.String	
<input checked="" type="checkbox"/>	FeaturesOut		IMAGINE.Features	

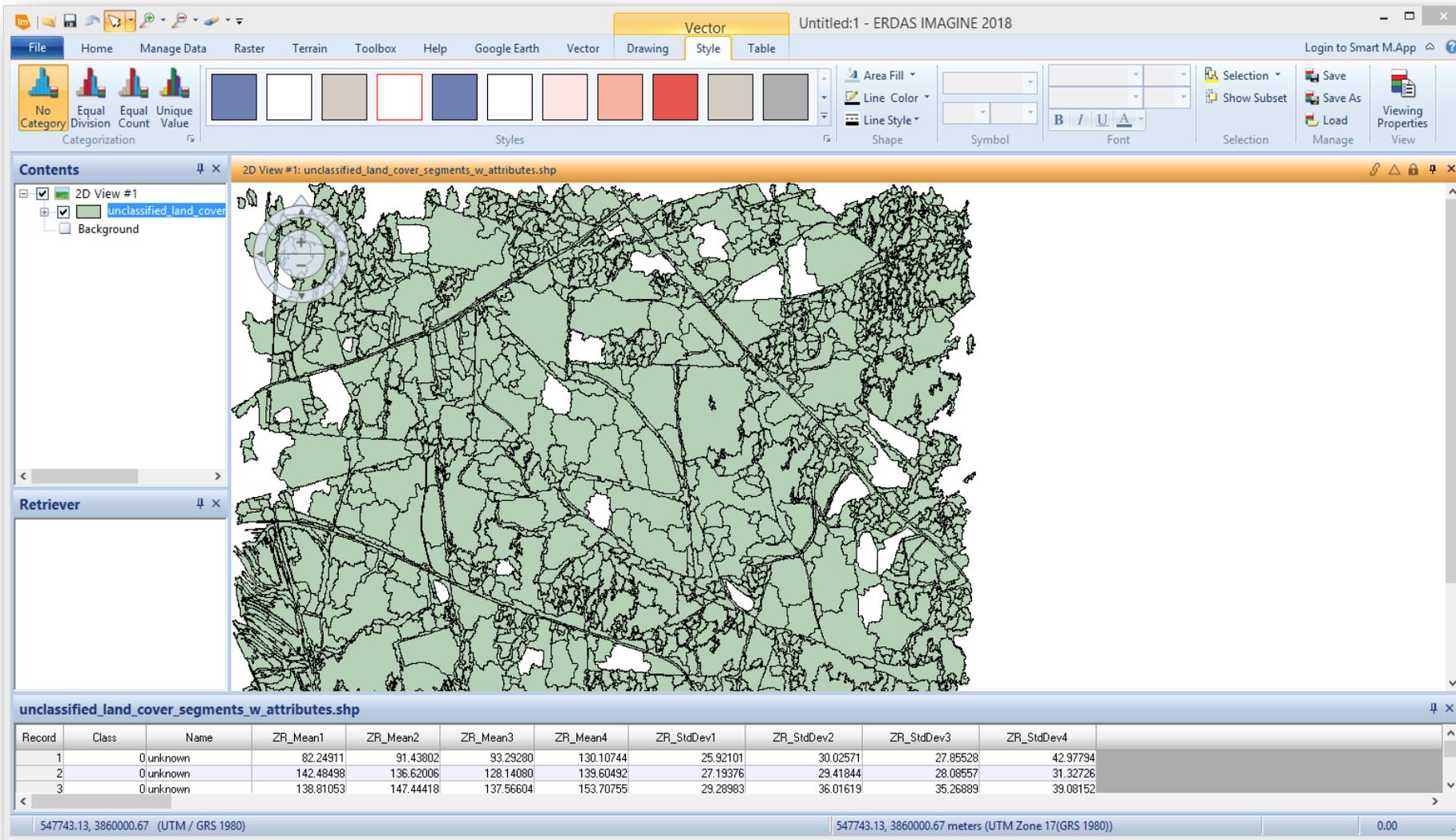
Trainingsdaten aufbereiten – Mean + StdDev



Eingangsdaten aufbereiten – Mean + StdDev



Eingangsdaten aufbereiten – Mean + StdDev wurden berechnet

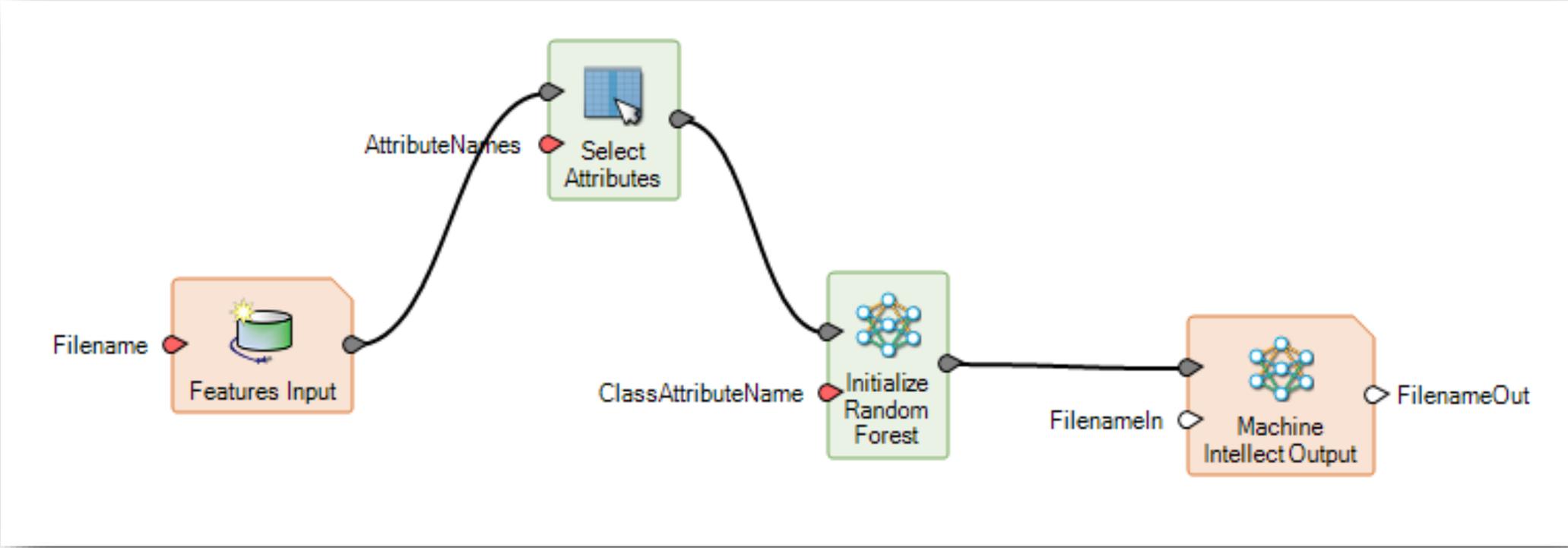


The screenshot displays the ERDAS IMAGINE 2018 software interface. The main window shows a 2D view of a land cover map with a green background and black outlines. The interface includes a menu bar (File, Home, Manage Data, Raster, Terrain, Toolbox, Help, Google Earth, Vector, Drawing, Style, Table), a toolbar with various icons, and a 'Contents' panel on the left. The 'Contents' panel shows a tree view with '2D View #1' expanded, containing 'unclassified_land_cover' and 'Background'. Below the map is a 'Retriever' panel. At the bottom, a data table is visible, titled 'unclassified_land_cover_segments_w_attributes.shp'. The table has columns for Record, Class, Name, and four ZR_Mean and ZR_StdDev columns. The data is as follows:

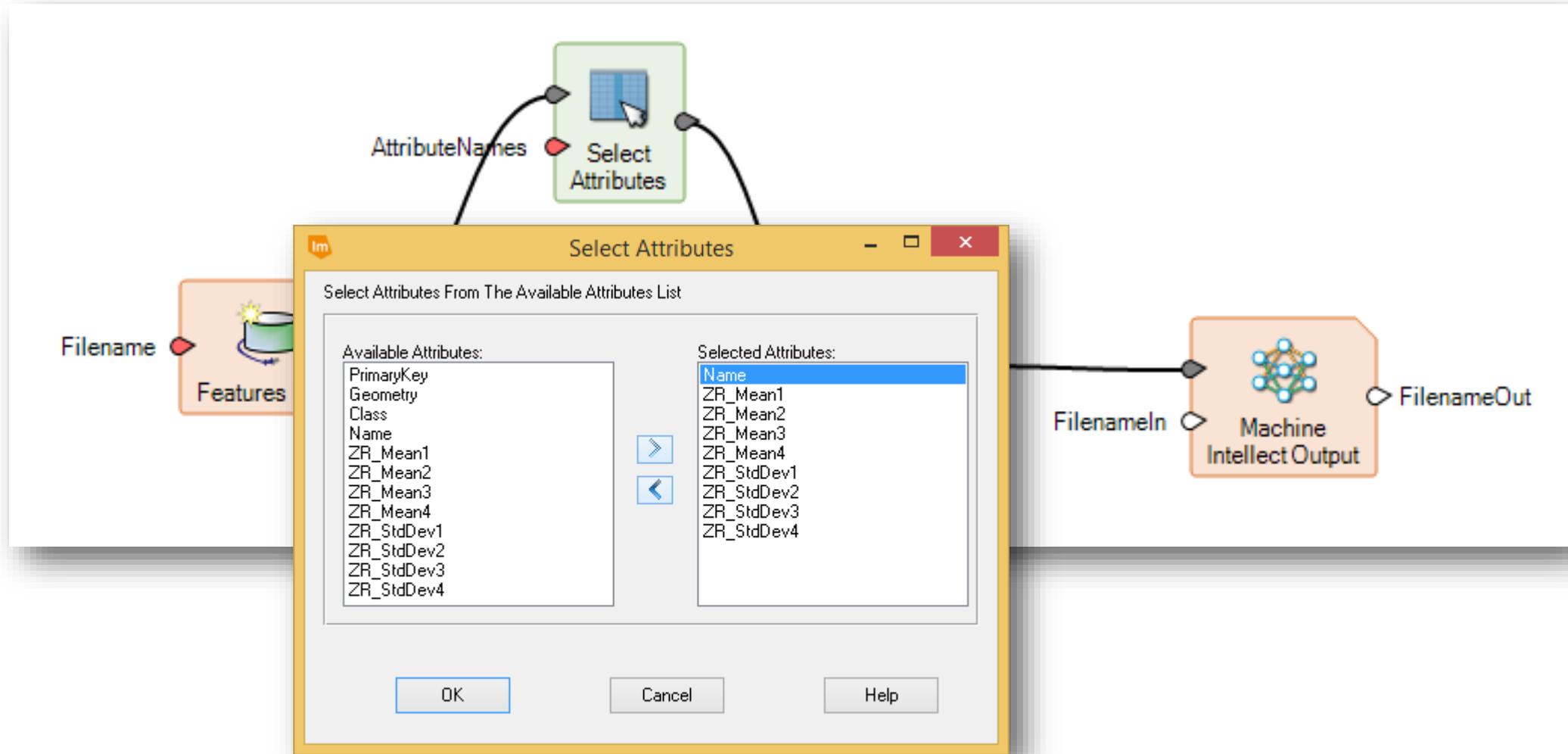
Record	Class	Name	ZR_Mean1	ZR_Mean2	ZR_Mean3	ZR_Mean4	ZR_StdDev1	ZR_StdDev2	ZR_StdDev3	ZR_StdDev4
1	0 unknown		82.24911	91.43802	93.29280	130.10744	25.92101	30.02571	27.85528	42.97794
2	0 unknown		142.48498	136.62006	128.14080	139.60492	27.19376	29.41844	28.08557	31.32726
3	0 unknown		138.81053	147.44418	137.56604	153.70755	29.28983	36.01619	35.26889	39.08152

The status bar at the bottom shows coordinates: 547743.13, 3860000.67 (UTM / GRS 1980) and 547743.13, 3860000.67 meters (UTM Zone 17(GRS 1980)).

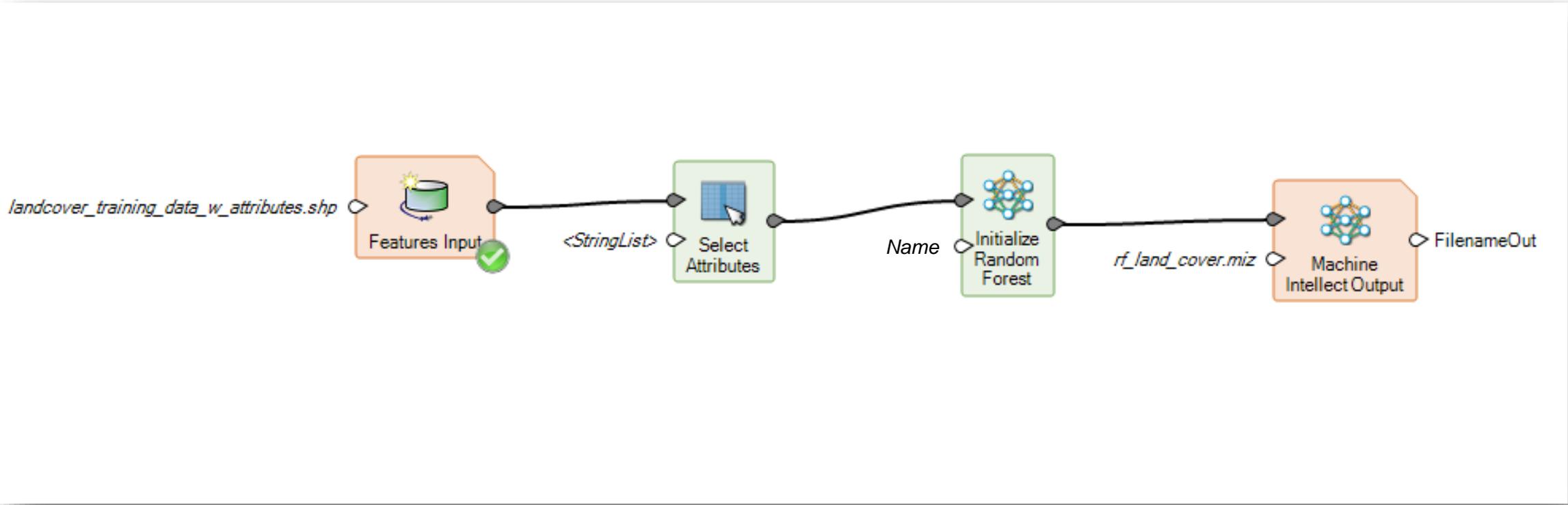
Random Forest Algorithmus trainieren – Initialize



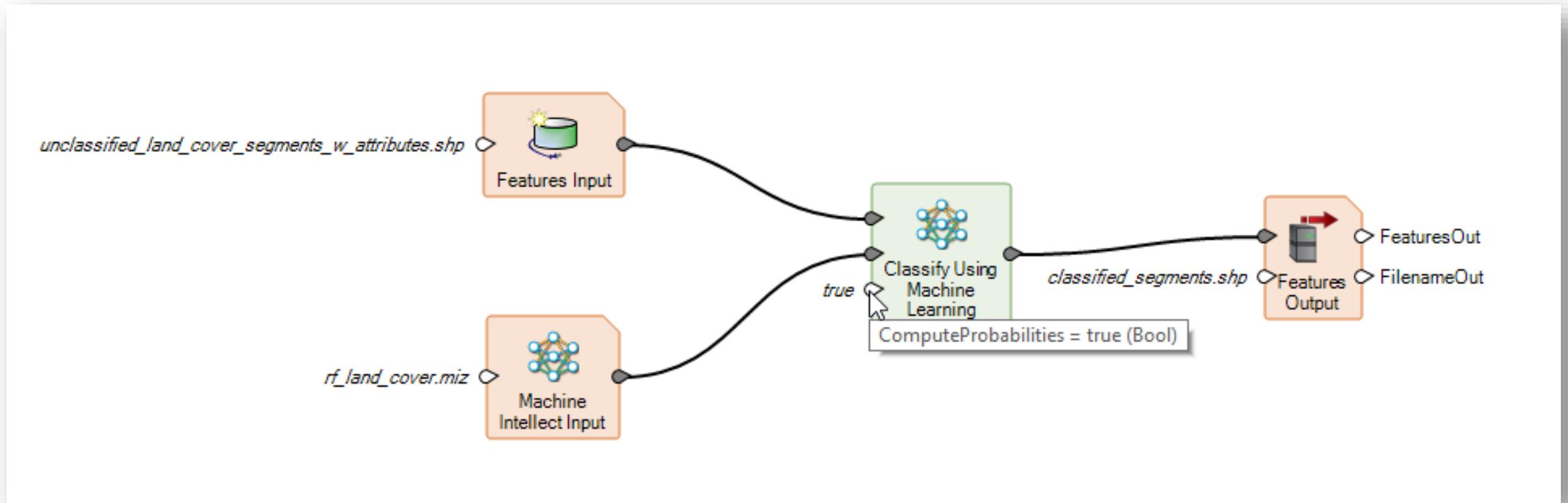
Random Forest Algorithmus trainieren – Select Attributes



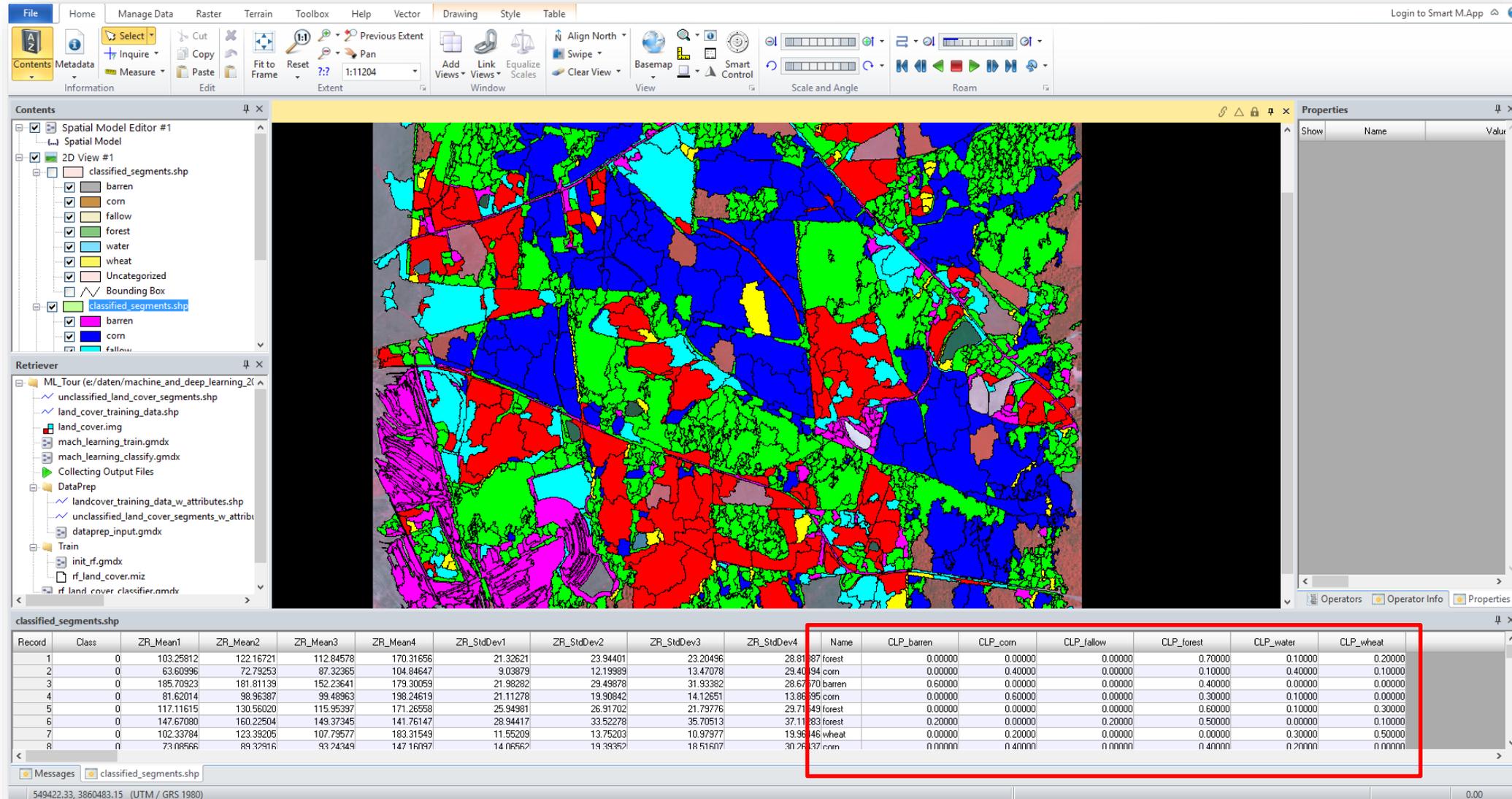
Random Forest Algorithm – Ausgabe des Machine Intellect



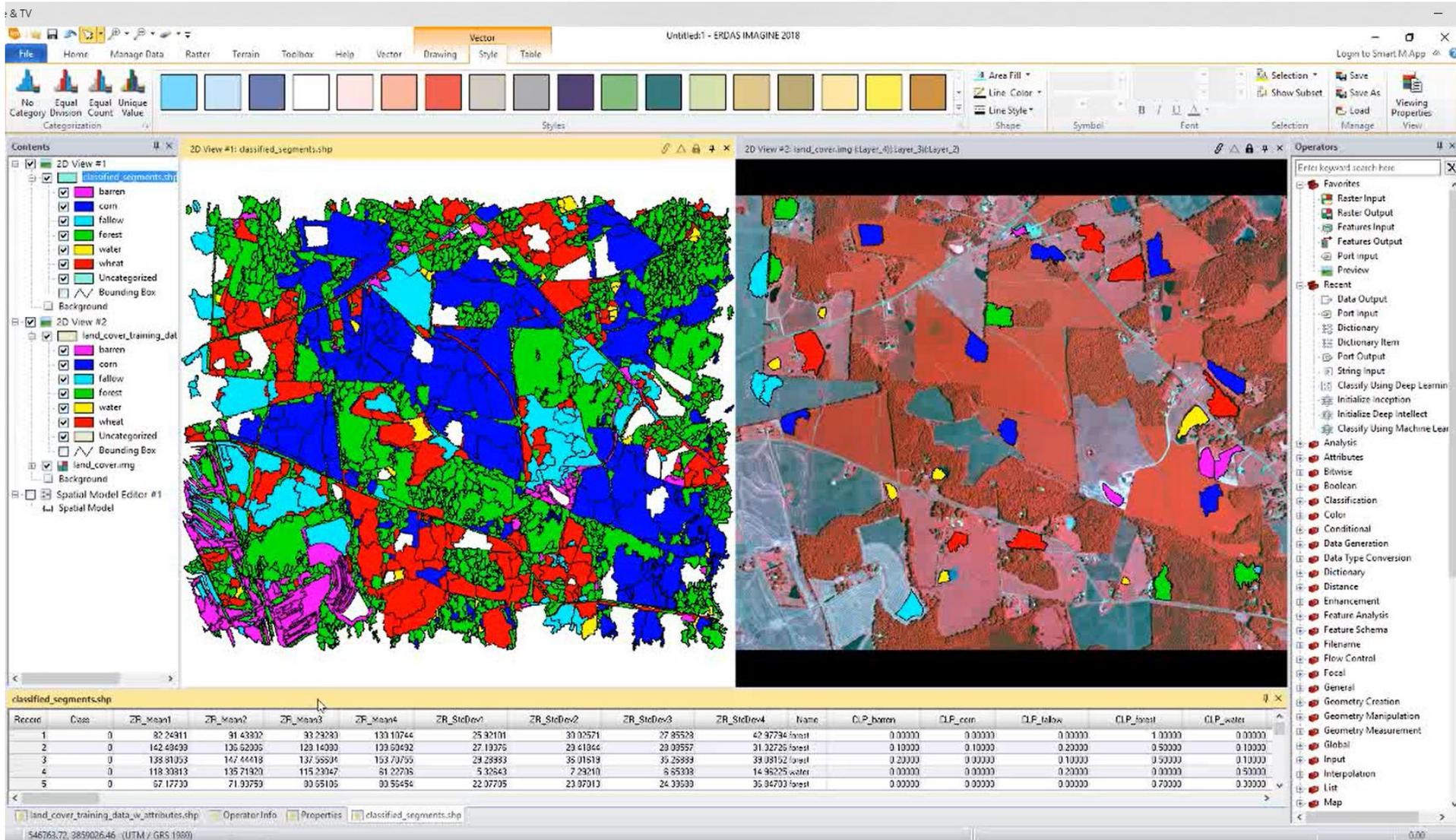
Mit trainiertem Machine Intellect klassifizieren



Ergebnisanalyse



Ergebnis



Herausforderungen

- Overfitting – Modell zu komplex

Modell funktioniert gut auf Trainingsdaten,
läßt sich aber nicht gut generalisieren.

- Modell vereinfachen (weniger Parameter oder einfacheres Modell)
- Reduktion der Eingangsinformation (weniger Attribute)
- Mehr Trainingsdaten

Herausforderungen des Modells

- Overfitting – Modell zu komplex

Modell funktioniert gut auf Trainingsdaten, lässt sich aber nicht gut generalisieren.

- Modell vereinfachen (weniger Parameter oder einfacheres Modell)
- Reduktion der Eingangsinformation (weniger Attribute)
- Mehr Trainingsdaten

- Underfitting – Modell zu einfach

Modell funktioniert auch auf Trainingsdaten nicht gut.

- Einschränkungen untersuchen und verringern
- Komplexeres Modell mit mehr Parametern wählen
- Bessere Trainingsdaten verwenden

Herausforderungen der Trainingsdaten

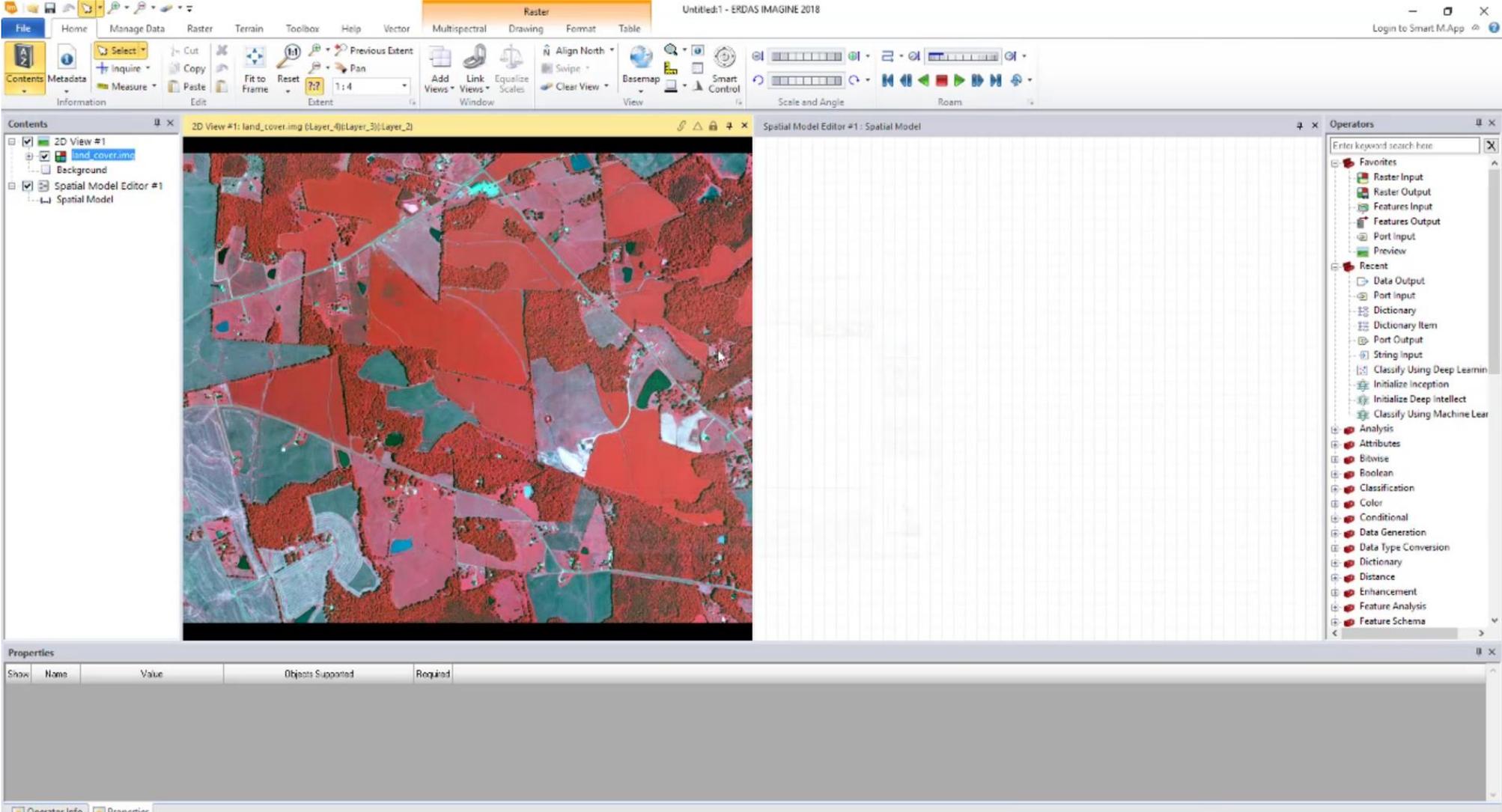
- Ausreichend hohe Quantität
- Repräsentiv
- Gute Qualität
- Ausreichend viel relevante Attribute
- Wenig irrelevante Attribute



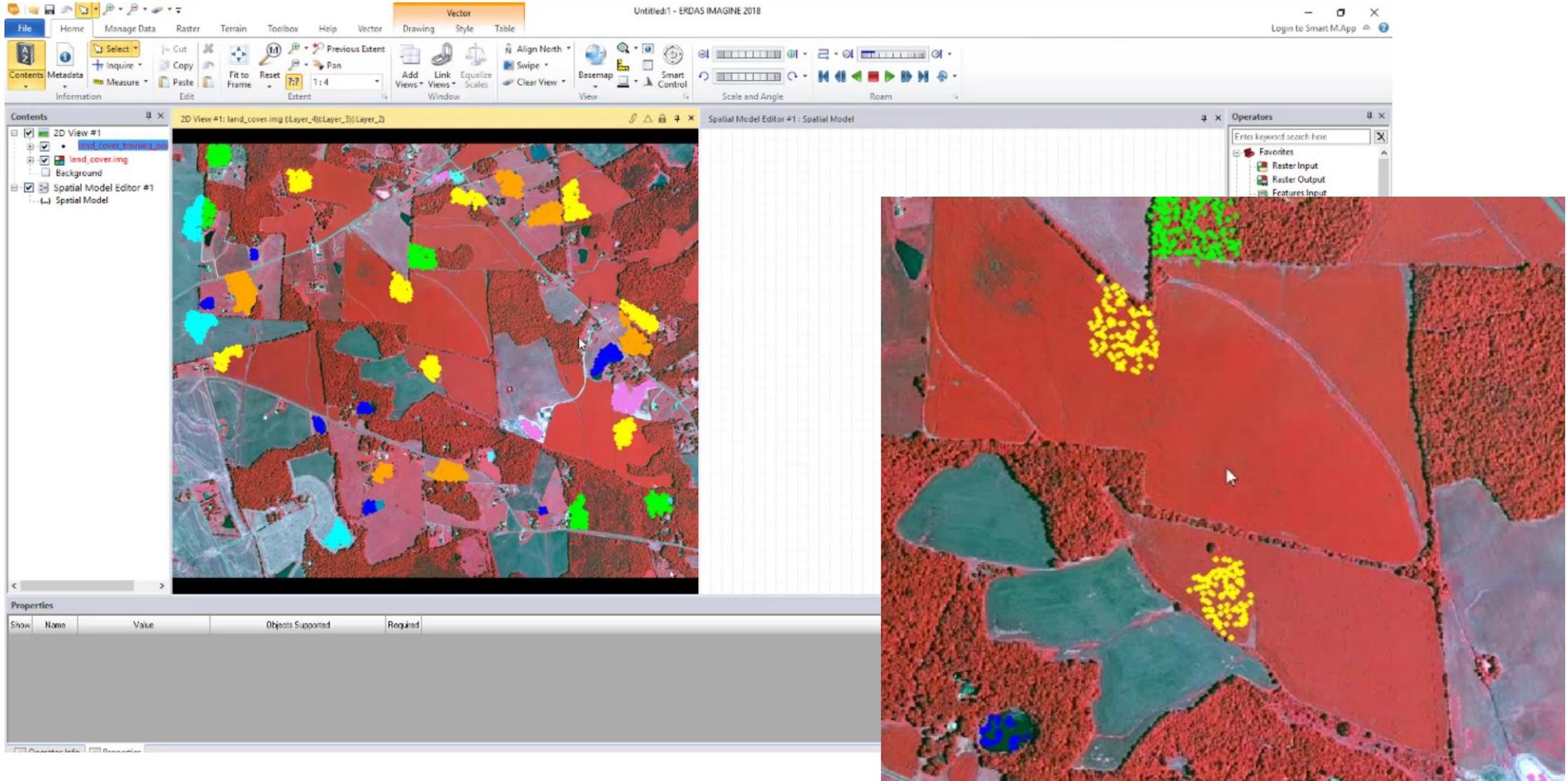
Machine Learning Beispiel - Rasterklassifikation



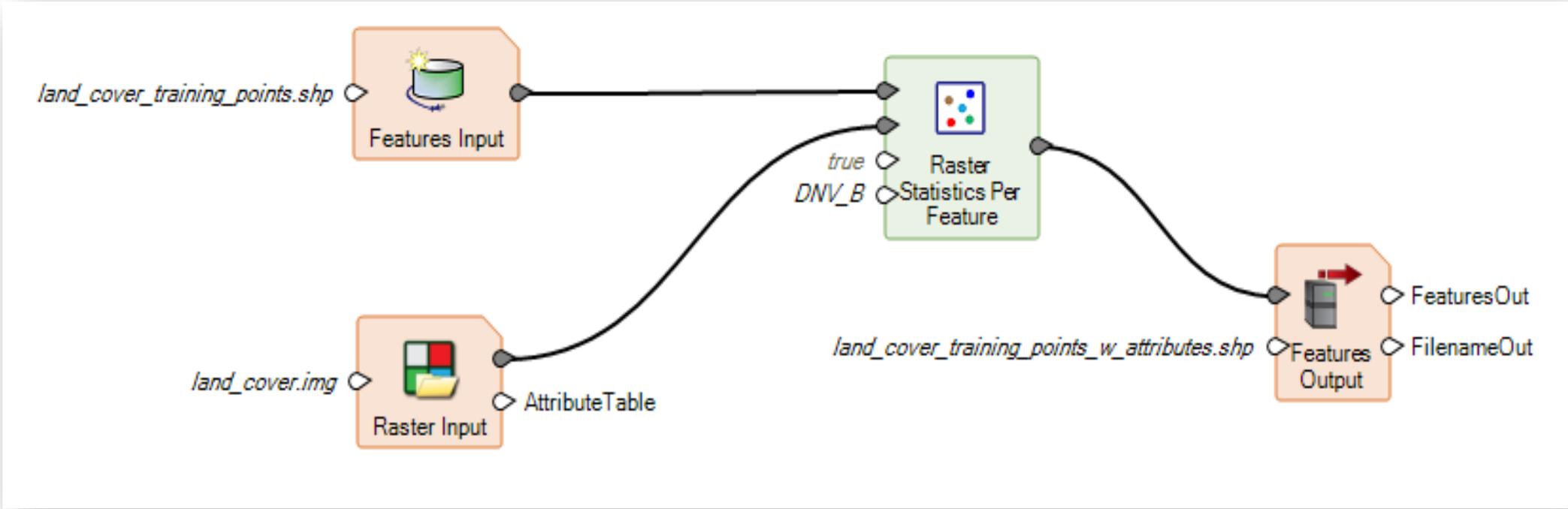
Zu klassifizierendes Rasterbild



Rasterbild mit Trainingsdaten (Punkt-Vektoren)



Trainingsdaten aufbereiten – DN als neues Attribut

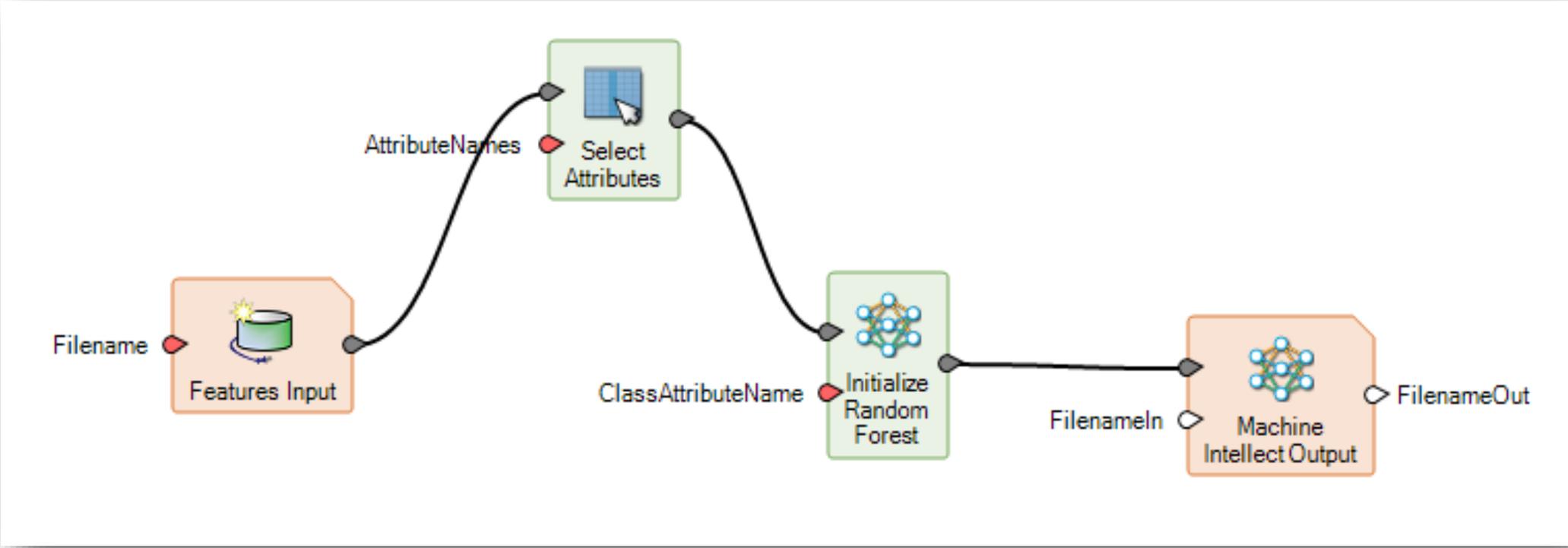


Trainingsdaten aufbereiten – DN Value per Band

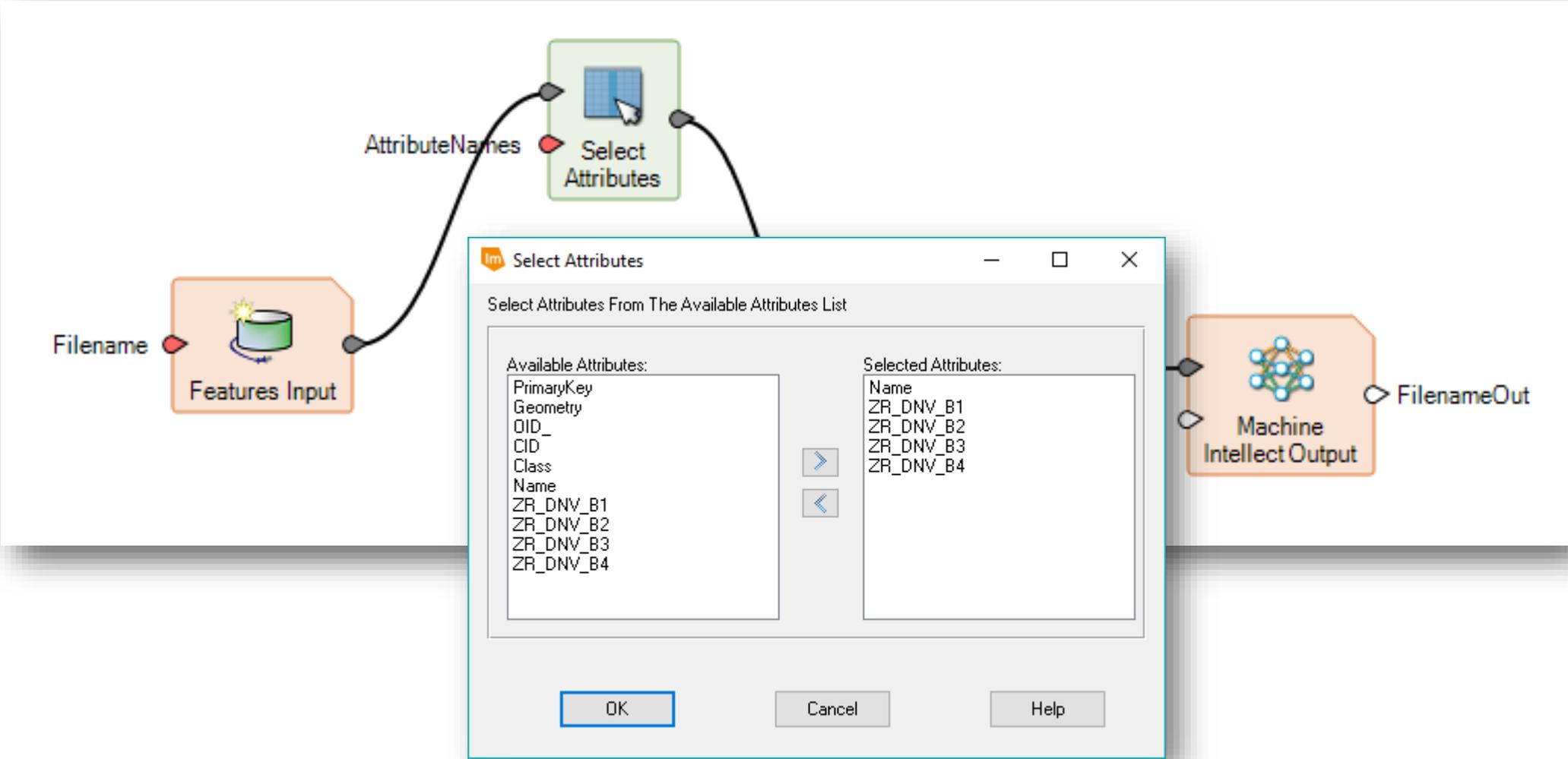
The screenshot shows the ERDAS IMAGINE 2018 interface. The main window displays a map titled "2D View #1: land_cover_training_points_w_attributes.shp" with various colored points. The interface includes a menu bar (File, Home, Manage Data, Raster, Terrain, Toolbox, Help, Google Earth, Vector, Drawing, Style, Table), a toolbar with categorization tools, and a "Contents" pane on the left. At the bottom, an attribute table for "land_cover_training_points_w_attributes.shp" is visible, with a red box highlighting the columns ZR_DNV_B1 through ZR_DNV_B4.

Record	OID_	CID	Class	Name	ZR_DNV_B1	ZR_DNV_B2	ZR_DNV_B3	ZR_DNV_B4
1	0	13	4 wheat		139.00000	147.00000	124.00000	171.00000
2	0	13	4 wheat		142.00000	148.00000	125.00000	171.00000
3	0	13	4 wheat		132.00000	143.00000	123.00000	185.00000

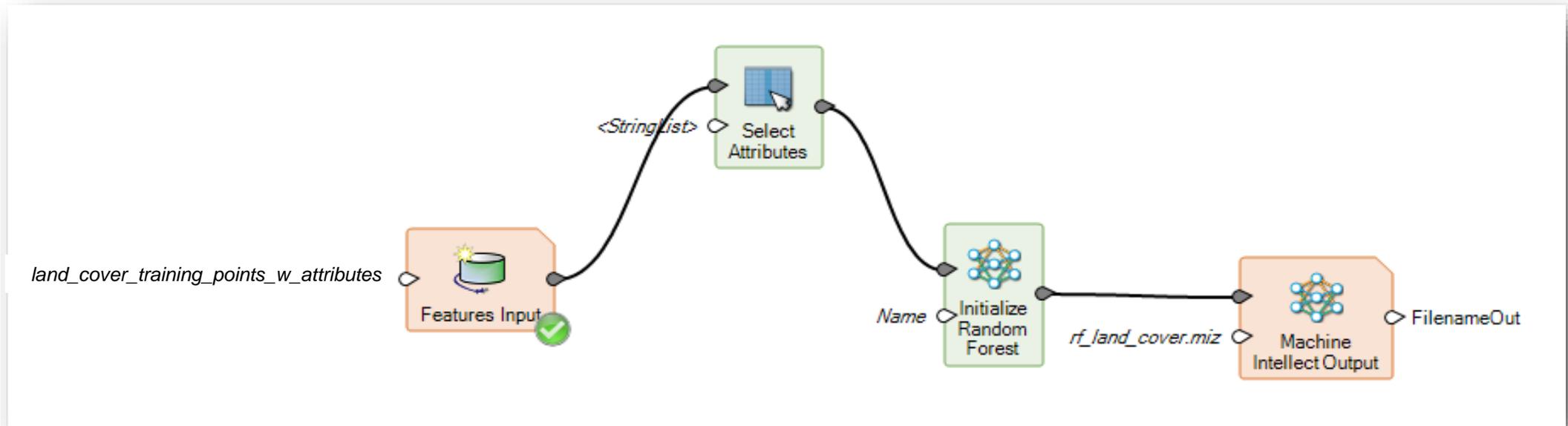
Random Forest Algorithmus trainieren – Initialize



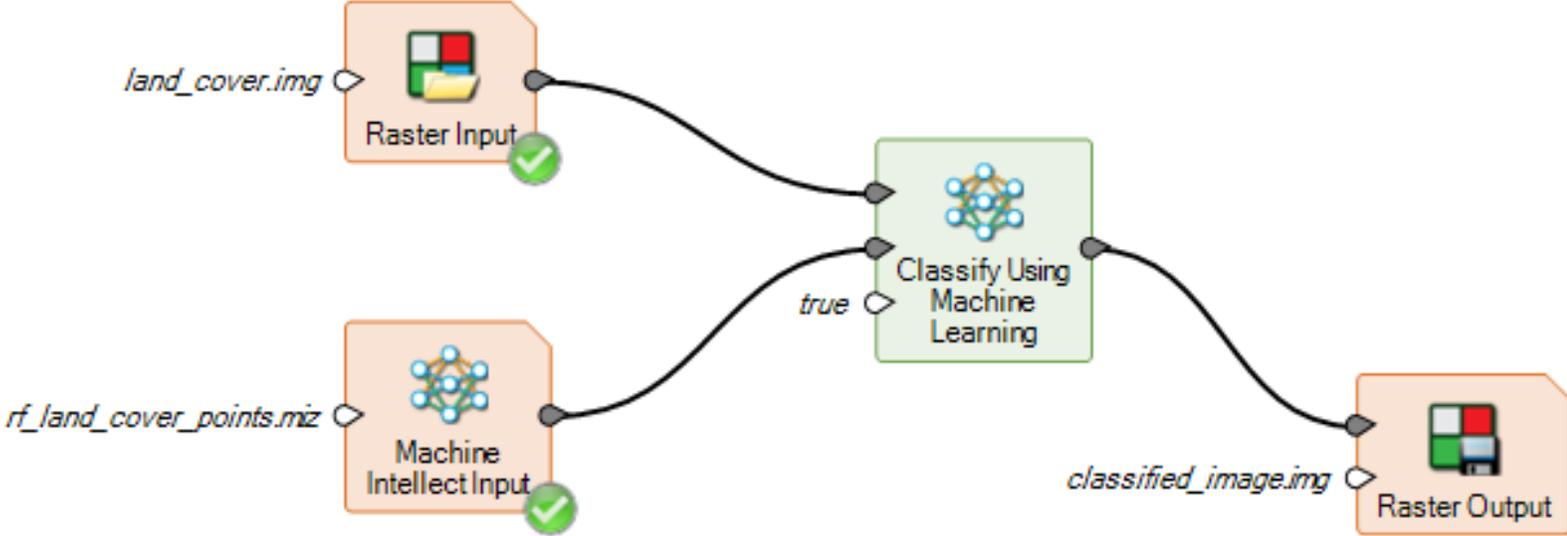
4. Random Forest Algorithmus trainieren – Initialize



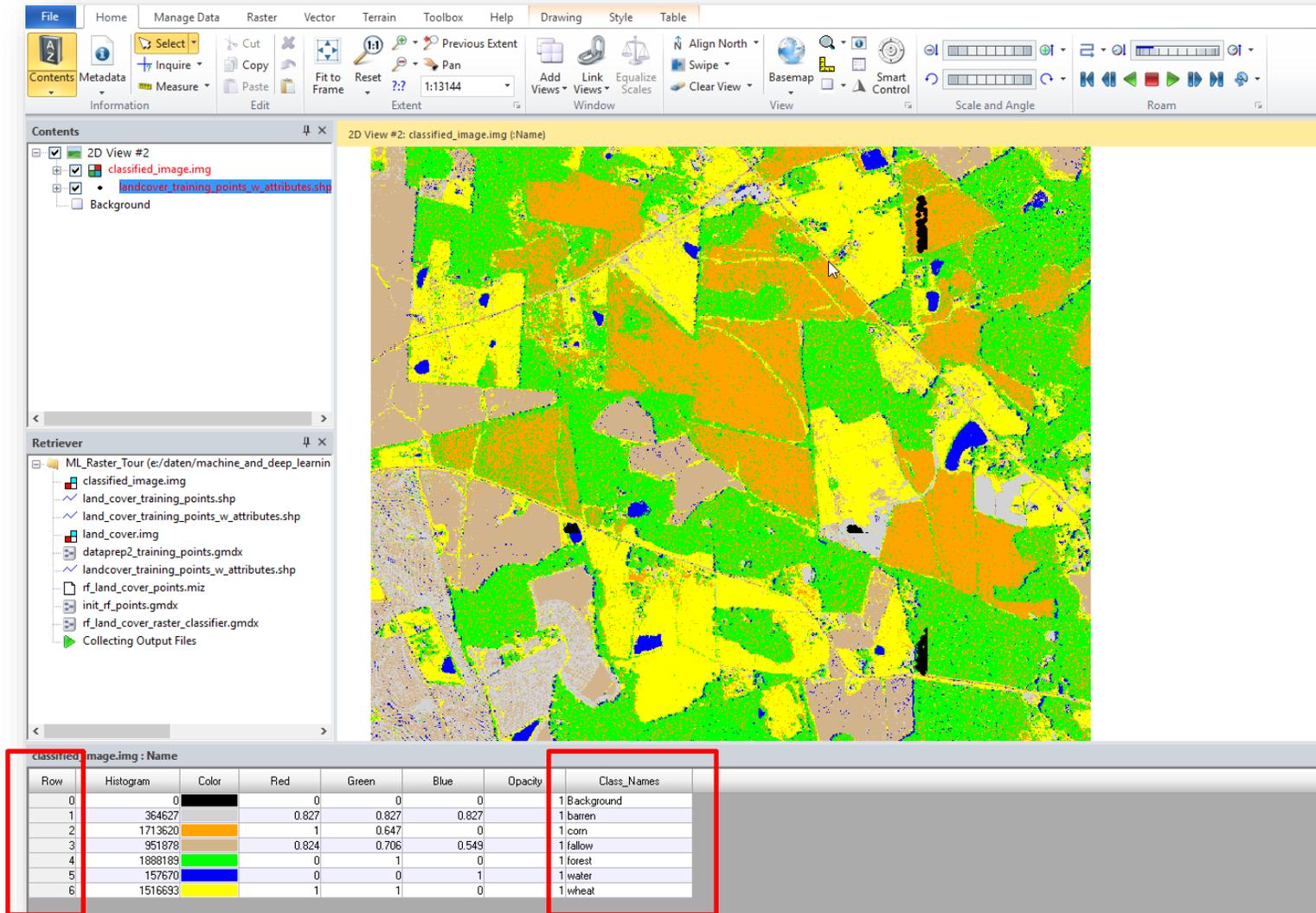
Random Forest Algorithm – Machine Intellect ausgeben



Mit trainiertem Machine Intellect klassifizieren



Ergebnisanalyse



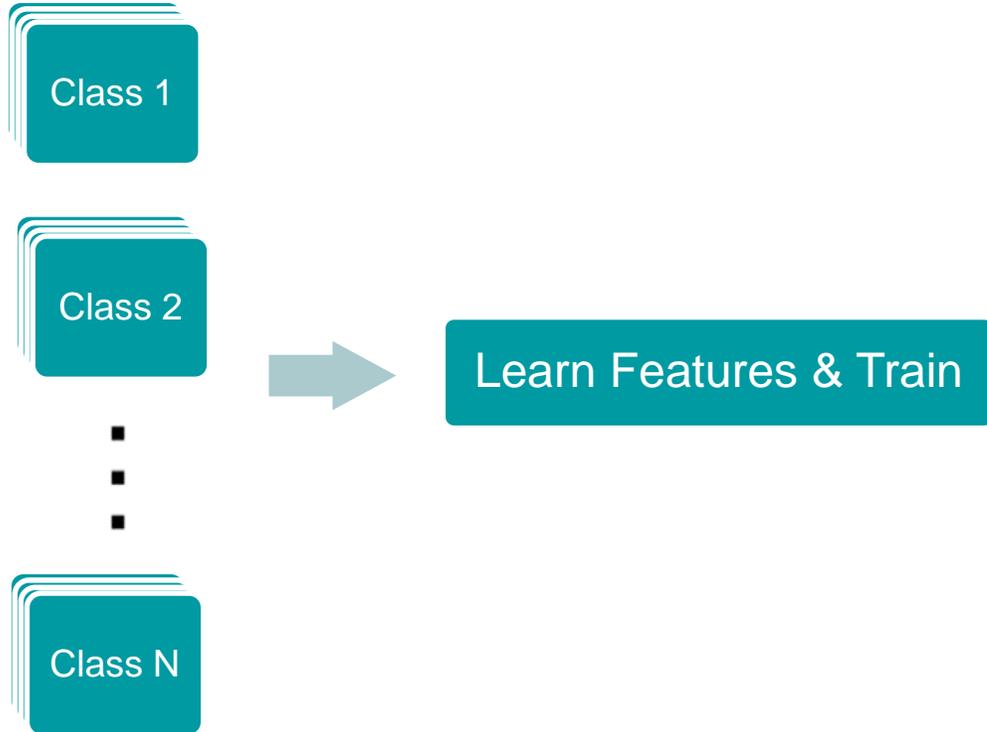


Deep Learning

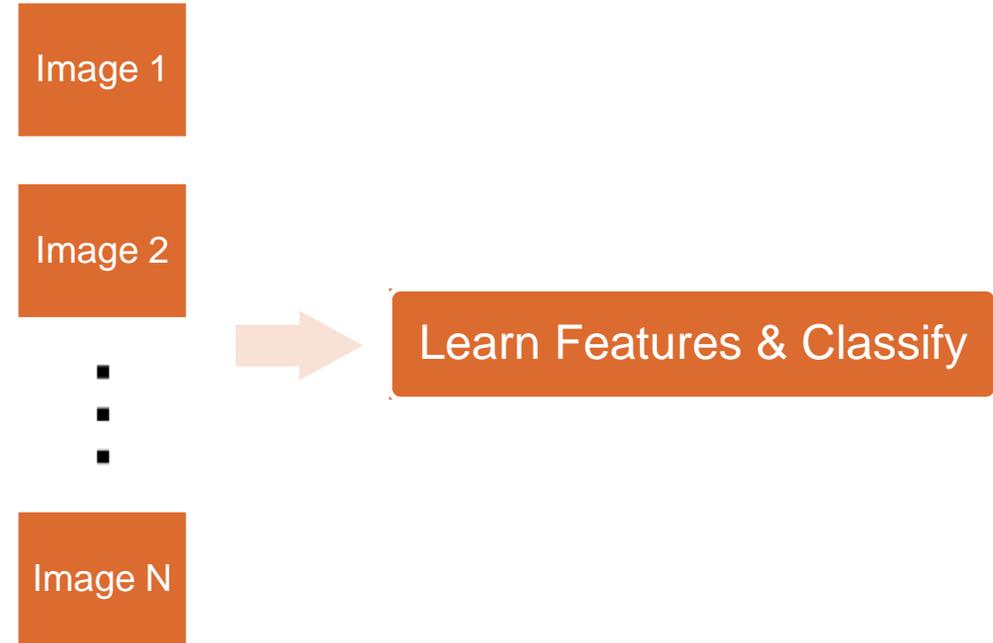


Deep Learning

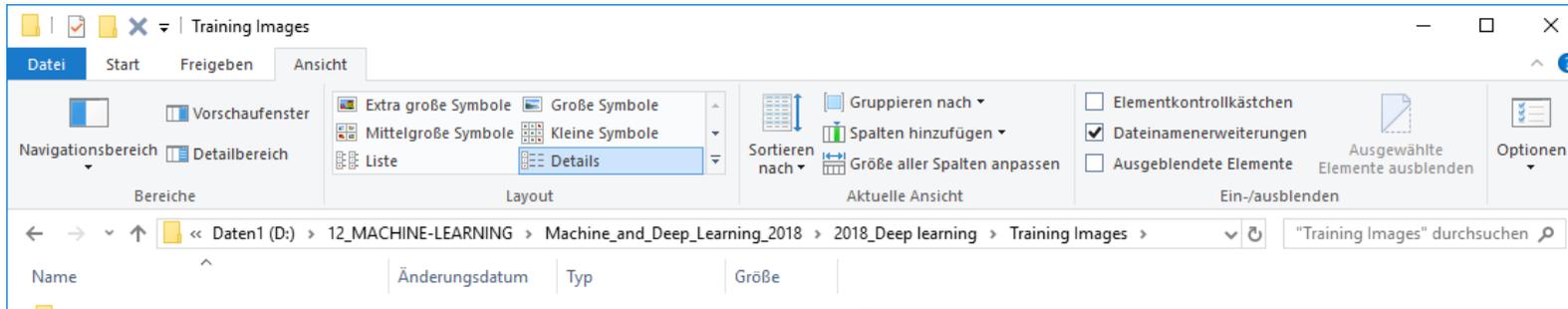
Training Images



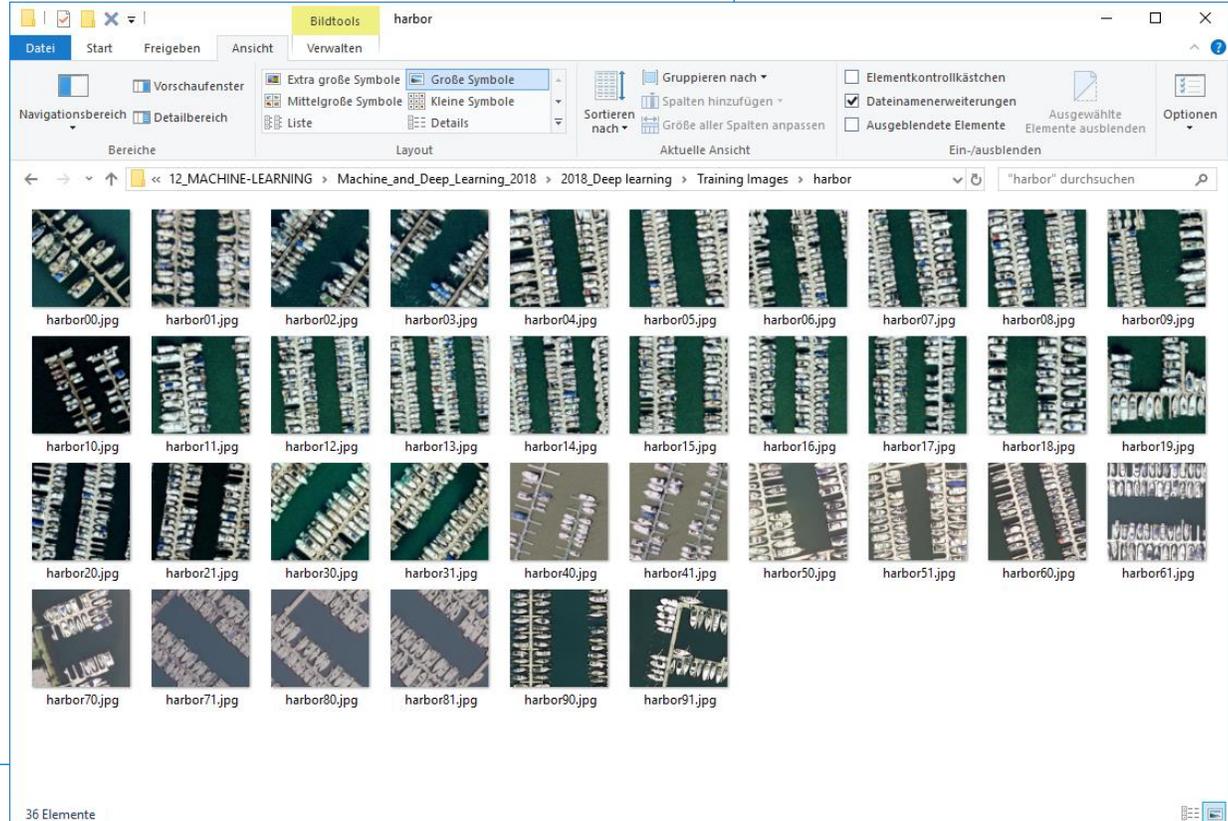
Images to Classify



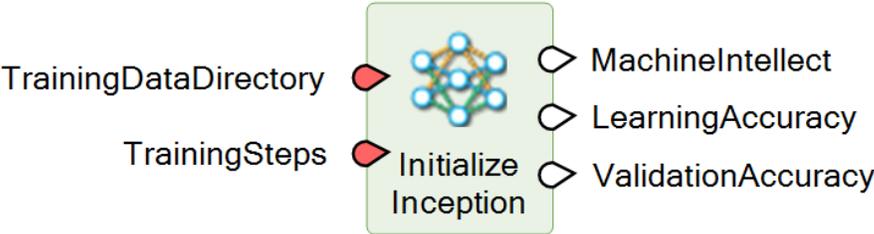
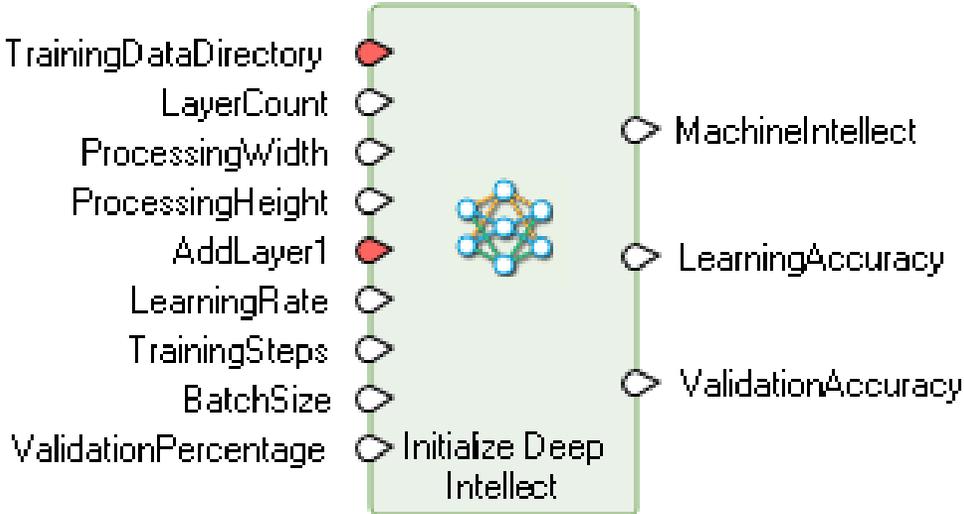
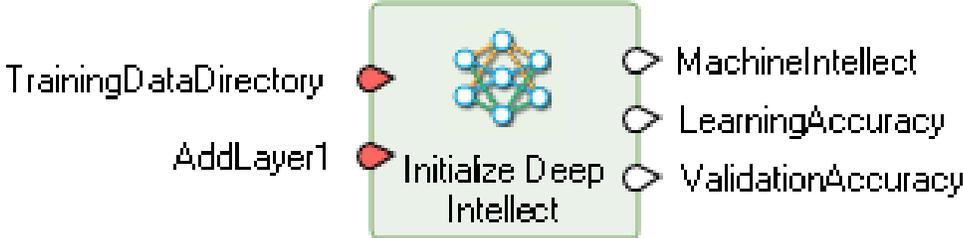
Umfangreiche Trainingsbilder werden benötigt



Name	Änderungsdatum	Typ	Größe
agricultural	02.10.2018 15:28	Dateiordner	
airplane	02.10.2018 15:28	Dateiordner	
baseballdiamond	02.10.2018 15:29	Dateiordner	
beach	02.10.2018 15:29	Dateiordner	
buildings	02.10.2018 15:29	Dateiordner	
chaparral	02.10.2018 15:29	Dateiordner	
denseresidential	02.10.2018 15:29	Dateiordner	
forest	02.10.2018 15:29	Dateiordner	
freeway	02.10.2018 15:29	Dateiordner	
golfcourse	02.10.2018 15:29	Dateiordner	
harbor	02.10.2018 15:29	Dateiordner	
intersection	02.10.2018 15:30	Dateiordner	
mediumresidential	02.10.2018 15:30	Dateiordner	
mobilehomepark	02.10.2018 15:30	Dateiordner	
overpass	02.10.2018 15:30	Dateiordner	
parkinglot	02.10.2018 15:30	Dateiordner	
river	02.10.2018 15:30	Dateiordner	
runway	02.10.2018 15:30	Dateiordner	
sparseresidential	02.10.2018 15:30	Dateiordner	
storagetanks	02.10.2018 15:31	Dateiordner	
tenniscourt	02.10.2018 15:31	Dateiordner	

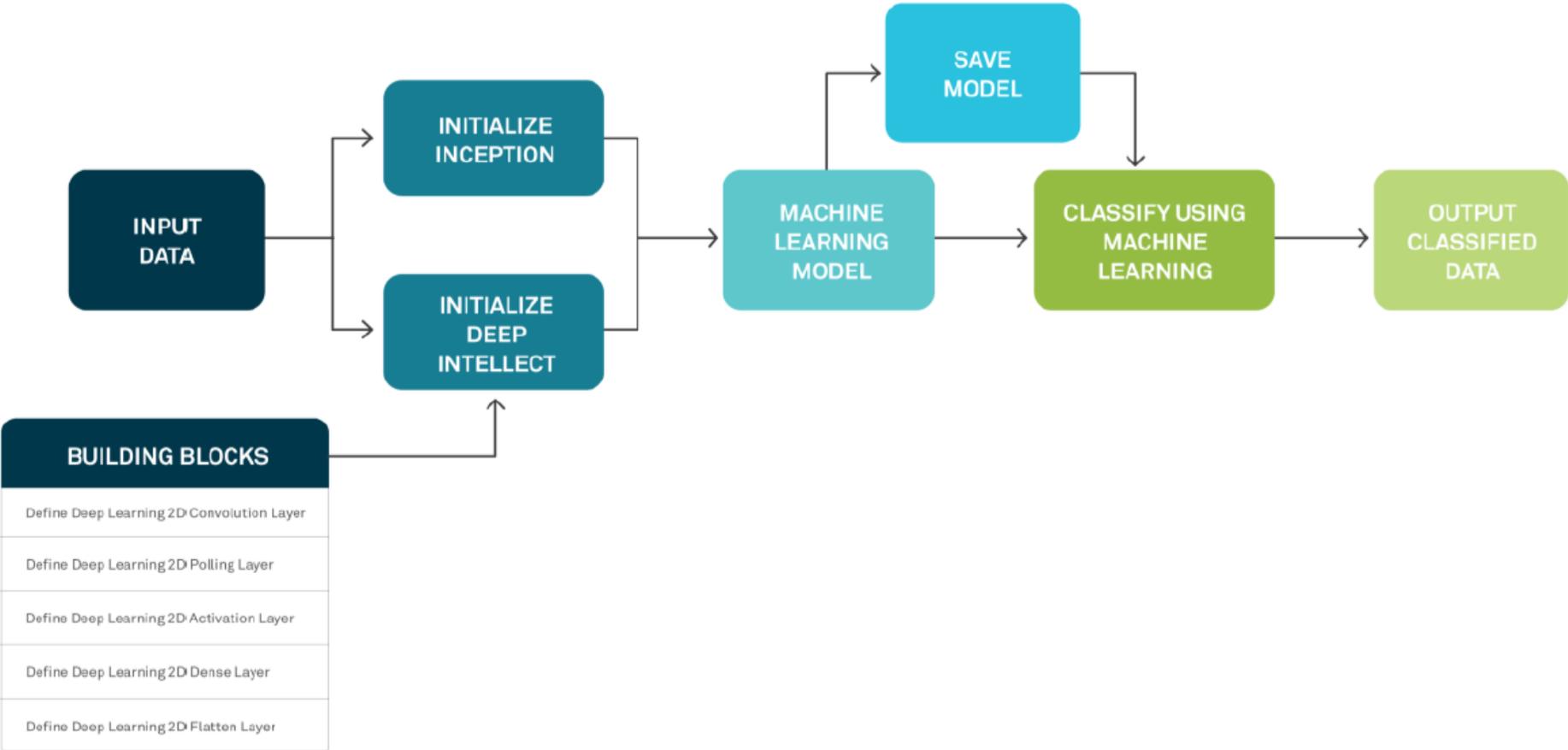


Initialize Deep Intellect



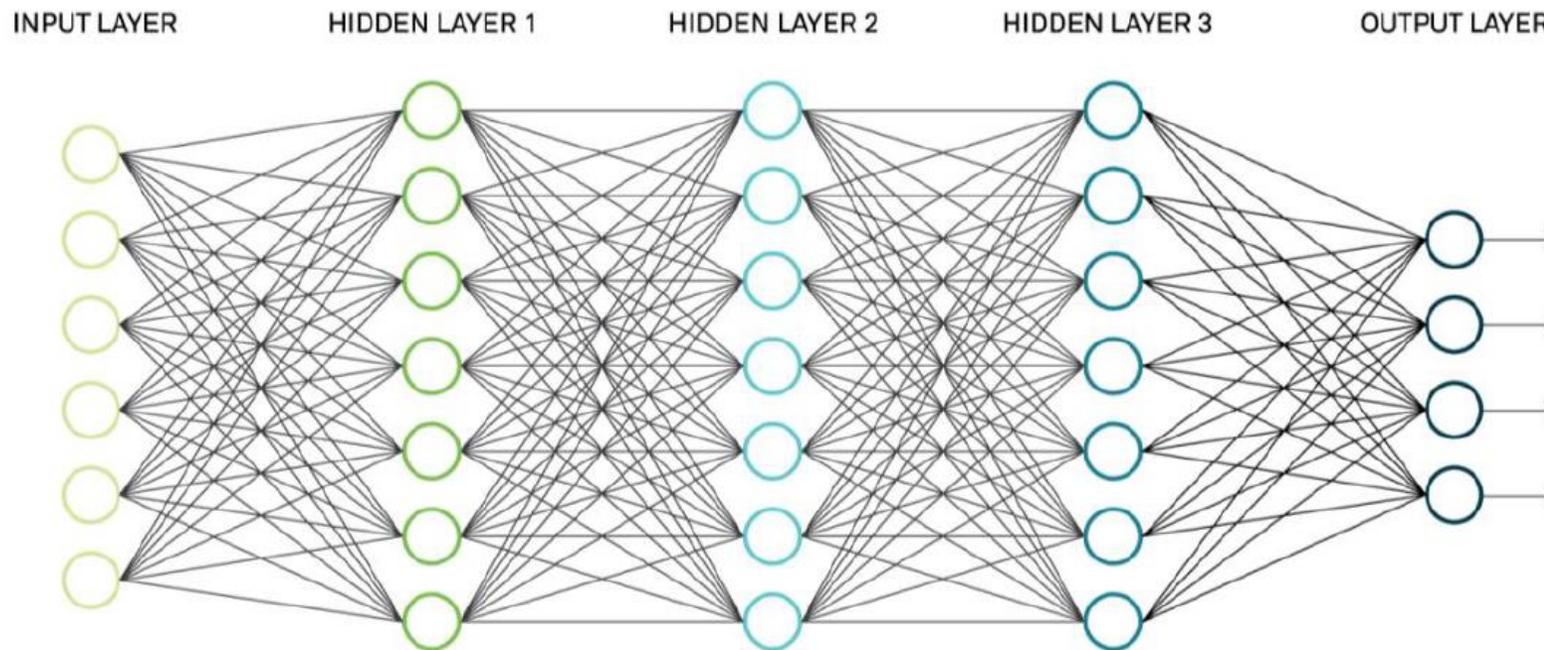
Trained by Google

Ablaufschema Deep Learning in IMAGINE



Deep Neural Networks

- Input layer – This is an input to the network. No computation is done in this layer.
- Hidden layers – All the computations are performed here. The input data is transformed through a series of operations.
- Output layer – Represents probability scores



A neural network with three hidden layers

CNN – Convolutional Neural Networks

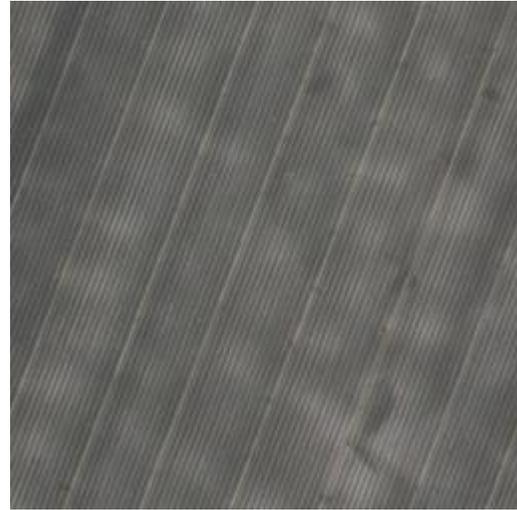
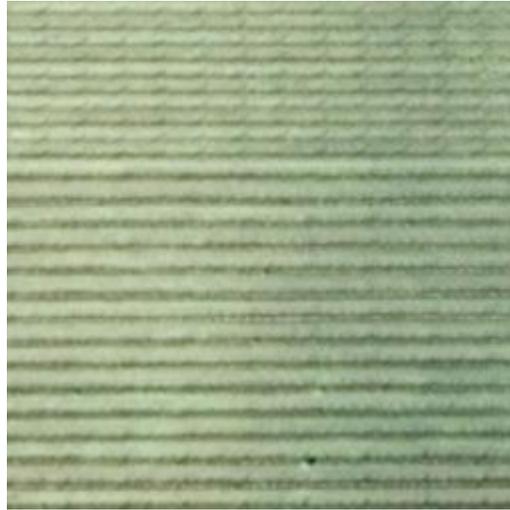
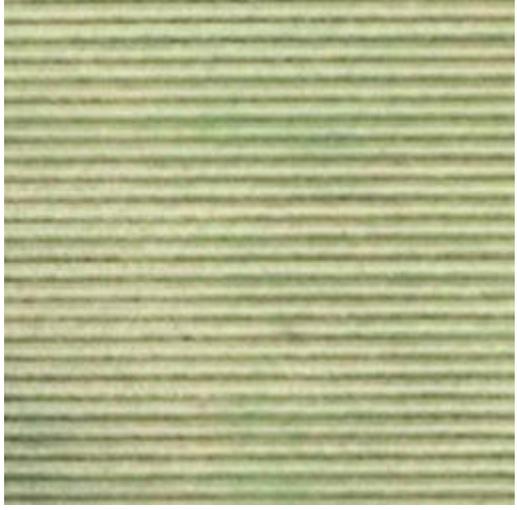
Convolutional Neural Networks (CNN) sind gut geeignet für bildrelevante Aufgaben

Der **Input Layer** wird durch die bereitgestellten Bilddaten abgebildet.

Hidden Layers wenden eine Reihe von Filtern auf das Bild an, um Features zu extrahieren und zu lernen, welche davon für die Klassifikation verwendet werden sollen.

Die Werte im **Output Layer** repräsentieren die Wahrscheinlichkeit eines Bildes zu einer Klasse zu gehören.

Trainingsdaten Acker



Testmosaik zum Klassifizieren



Deep Learning results

- 2120 images classified
 - Size: 8960 rows, 15360 columns
 - Divided into 2120 grids
 - 1941 classified correctly
 - 91.56% accuracy

Land Cover Type	Omissions	Commissions
agricultural	6	1
airplane	6	3
baseballdiamond	5	4
beach	1	0
buildings	15	19
chaparral	1	0
denseresidential	26	32
forest	1	0
freeway	8	7
golfcourse	5	6
harbor	1	2
intersection	5	17
mediumresidential	29	12
mobilehomepark	11	23
overpass	2	11
parkinglot	2	3
river	5	6
runway	11	4
sparseresidential	12	5
storagetanks	14	7
tenniscourt	14	18

Machine Learning vs. Deep Learning

	Machine Learning	Deep Learning
Training Dataset	Small	Large
Choose your own training attributes	Yes	No
Training time	Short	Long
CPU/ GPU/ Mem	Standard	Big

Fazit

- Machine Learning können alle => Open Source Bibliotheken stehen zur Verfügung
- ABER: Hexagon Geospatial macht das Machine Learning einfach
- Machine Learning-Operatoren können in Prozessketten eingebunden und gebatcht werden
- IMAGINE bietet die Werkzeuge, Trainingsdaten und zu klassifizierende Daten geeignet aufzubereiten.
- IMAGINE bietet die Werkzeuge, Daten nach der Klassifikation zu bearbeiten (Funktionale Attribute)
- Der Spatial Modeler wird kontinuierlich weiterentwickelt.
- Entwicklungen auf dem Sektor des Machine Learning werden von Hexagon Geospatial verfolgt und ggfs. implementiert
- Weitere Bibliotheken können über Python adressiert werden (Python-Anbindung des Spatial Modelers)

Mehr dazu auf dem UGM 2018 !



06. - 07. November 2018, Germering bei München

Vielen Dank für Ihre Aufmerksamkeit!



GEOSYSTEMS
THE GEOSPATIAL EXPERTS 

Dipl.-Ing. (FH)
Ruth Leska
Projektingenieurin
Project Engineer

Telefon | +49 89 894343-26
Fax | +49 89 894343-99
E-Mail | r.leska@geosystems.de

GEOSYSTEMS GmbH
Riesstraße 10
82110 Germering
Germany
www.geosystems.de