FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE
SMART SURVEYORS FOR LAND
AND WATER MANAGEMENT
CHALLENGES IN A NEW REALITY

Pattern Mining in Sentinel 2A Satellite Images Using Knime Analytics Platform

ID 10937

Rudiney Pereira, Elisiane Alba, Juliana Marchesan,
Mateus Schuh and Roberta Fantinel

Federal University of Santa Maria
Brazil

UFSM

Introduction

> The importance of research is to allow image processing without human interference to efficiently determine land use and land cover changes.

> Land use patterns and land cover mining tool designed to manage knowledge from Sentinel 2A series satellite image data.

> It was elaborated on Knime Analytics platform through the selection of configured and connected nodes constituting a workflow composing all the methodological phases.

Federal University of Santa Maria
Brazil
UFSM

Introduction

> The graphic platform allows different configurations for the nodes with the main phases: loading and visualization of images; feature extraction (non-redundant numeric vector) that characterizes land use and cover; creation of attributes for each target (class of land use and land cover; test phase and evaluation of the predictive model.

Introduction

> The peculiarity in this case is that we must extract a vector of numerical resources from the image before starting the machine learning of the classifier algorithm in the implementation phase.

> Two reasons justify:

1. volume of images available
2. is a variety of tools that are easy to use for data mining

Material and Methods

1. *Images dataset:* 1,370 files with 30 files per spectral band with 10 X 10 pixels of 10 spectral bands

2. *Building the Workflow in Knime:* was elaborated by placing nodes with their different functions and interconnected with each other to flow processing in main steps

. preparation of data such as reading, extracting characteristics and filtering data;
. data partitioning, machine learning in the decision tree and random forest algorithms and classification prediction;
. performance analysis of the classification algorithms.

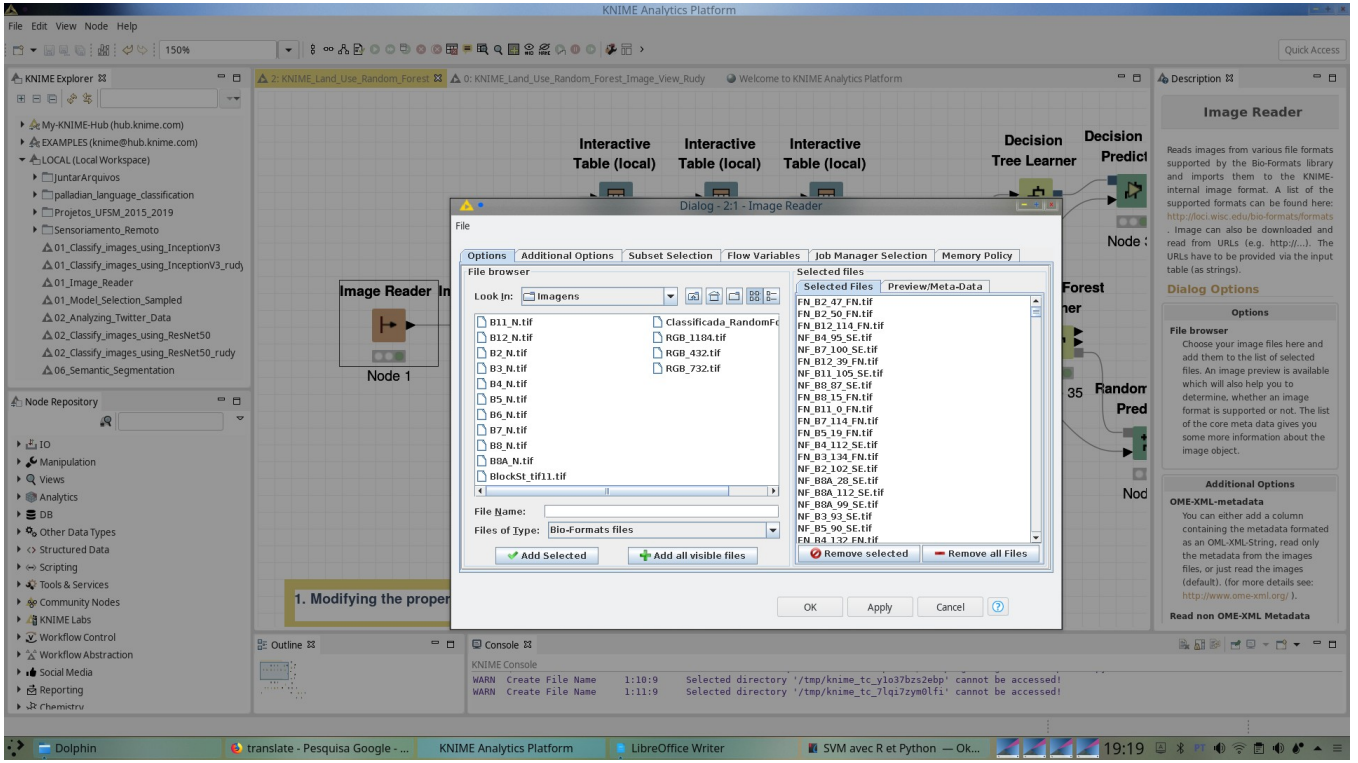
Federal University of Santa Maria
Brazil
UFSM

FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE

SMART SURVEYORS FOR LAND AND WATER MANAGEMENT

CHALLENGES IN A NEW REALITY

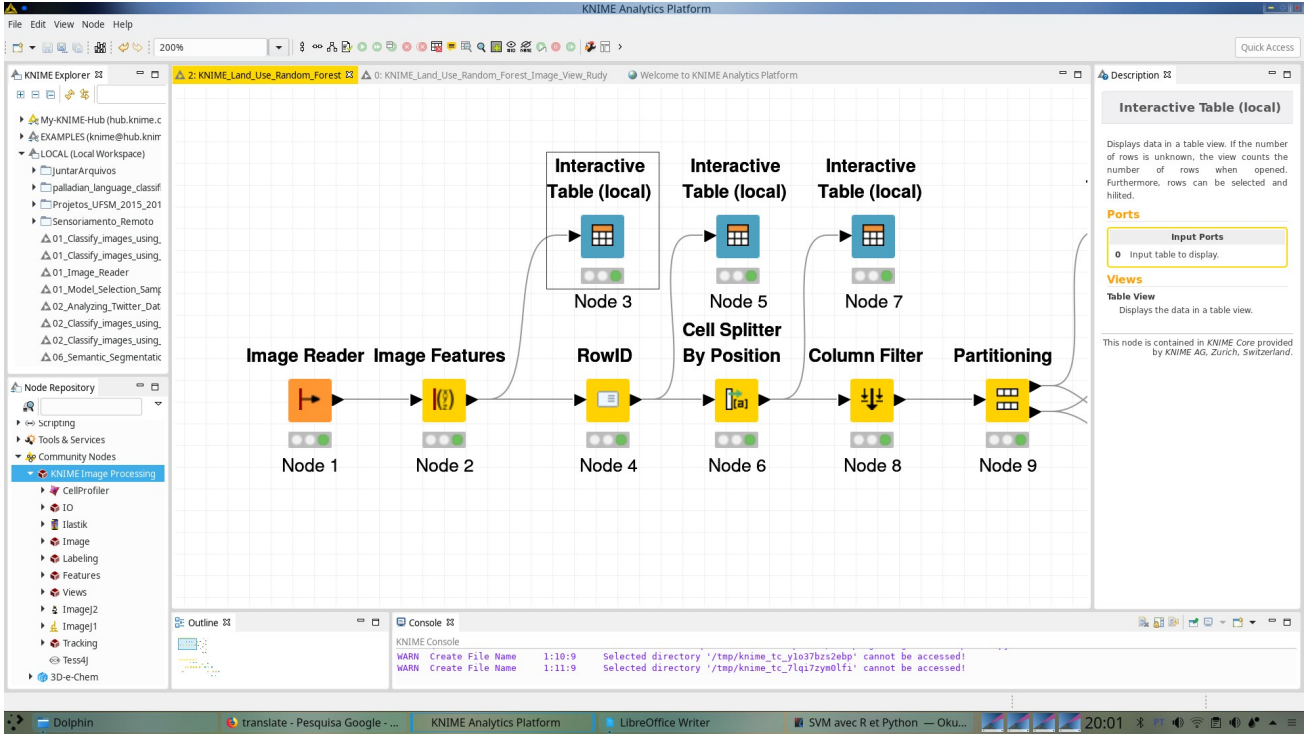


Data sets:

1.370 images Tiff format file

Federal University of Santa Maria
Brazil
UFSM

FIG NL THE NETHERLANDS
e-WORKING WEEK 2021 20-25 JUNE
SMART SURVEYORS FOR LAND AND WATER MANAGEMENT
CHALLENGES IN A NEW REALITY



Building the workflow in Knime Analytics Platform, partial view

Federal University of Santa Maria
Brazil
UFSM

Material and Methods

3. *Feature extraction:* was performed by the node called "Image Features" and consisted of configuring extraction options

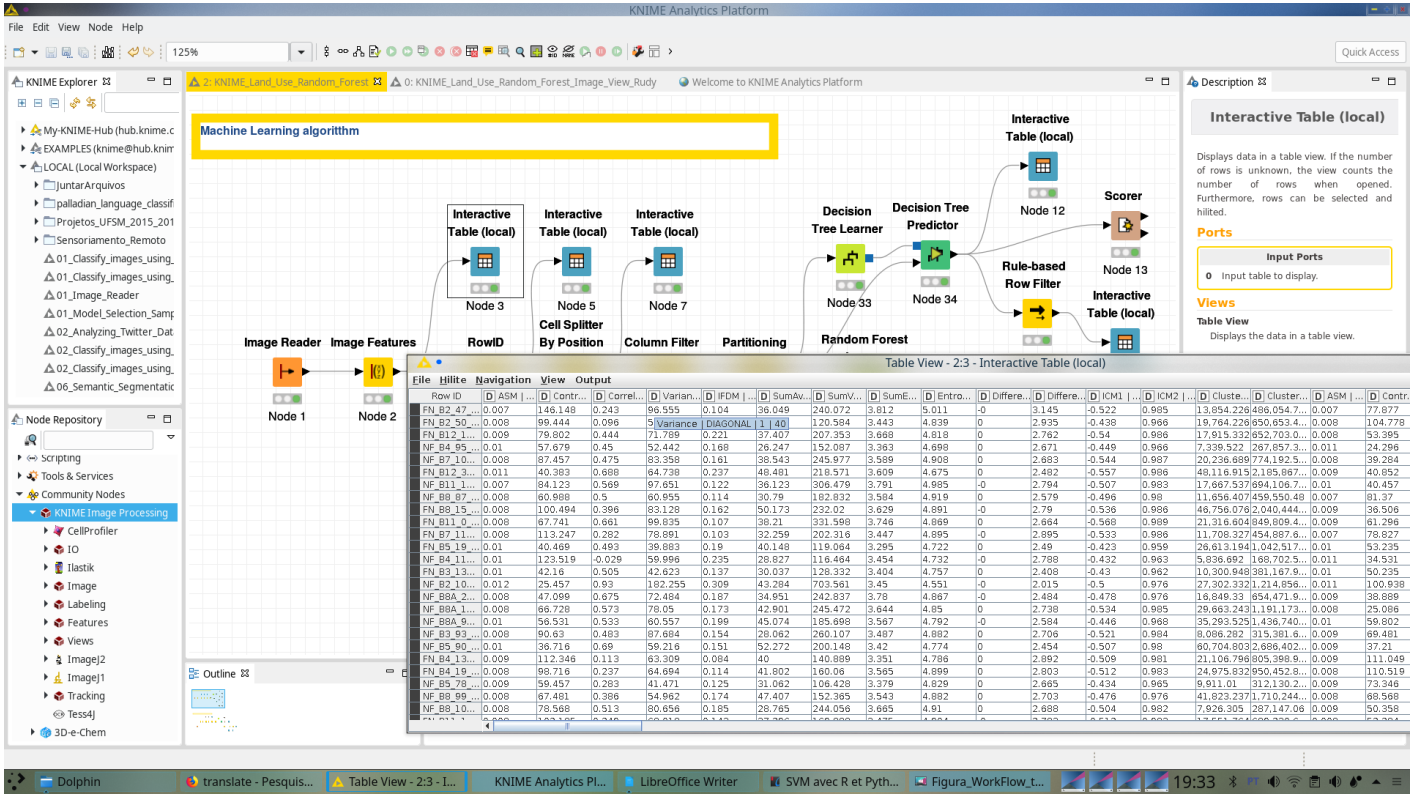
4. *Creating the target attibute and learning and evaluating th predictive model*

FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE

SMART SURVEYORS FOR LAND AND WATER MANAGEMENT

CHALLENGES IN A NEW REALITY



Feature extraction of images:

. values minimum, maximum, average, geometric mean, statndard deviation, variance, contrast, correlation, entropy and other parameters

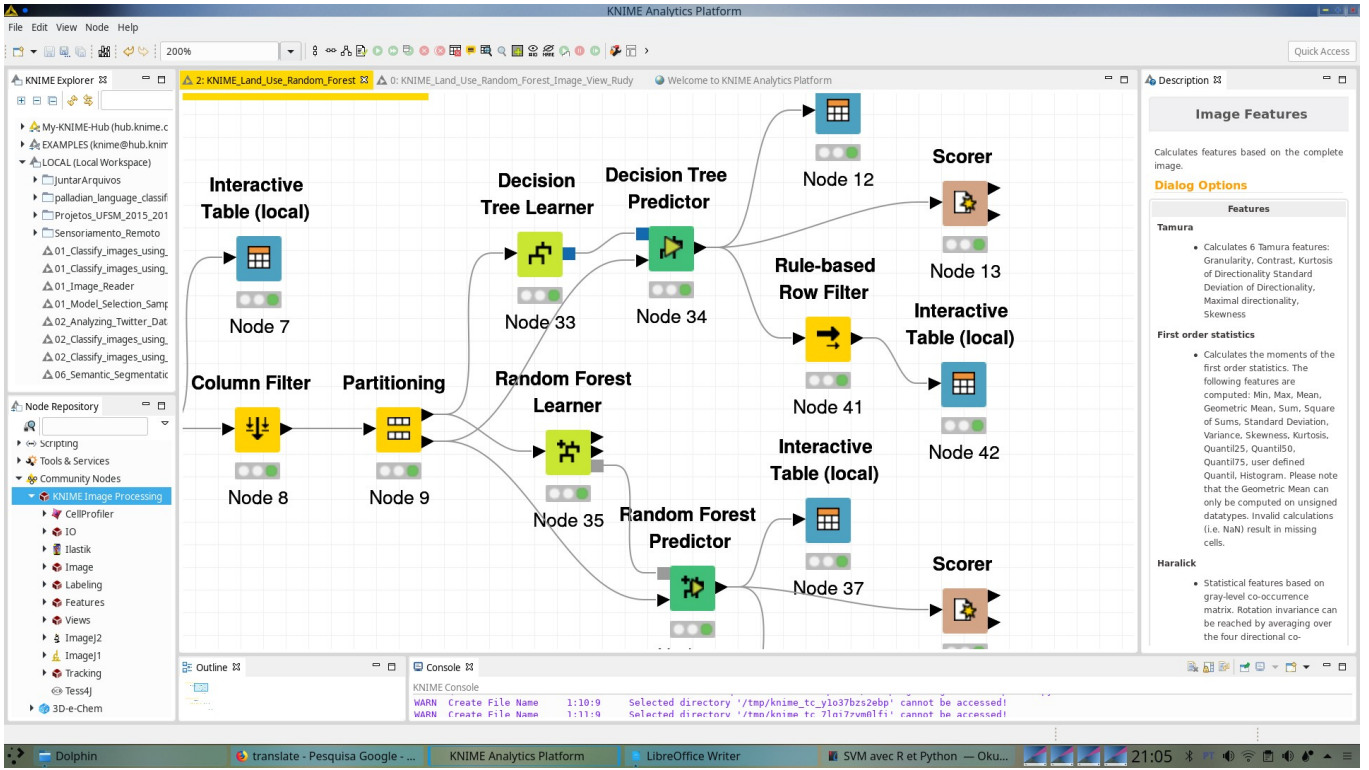
Federal University of Santa Maria
Brazil
UFSM

FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE

SMART SURVEYORS FOR LAND AND WATER MANAGEMENT

CHALLENGES IN A NEW REALITY



Data set was partitioned into two other sets, training and testing

The data partitioning had two flows:

1. Decision Tree Predictor algorithm;

2. Randon Forest Predictor algorithm

Federal University of Santa Maria
Brazil
UFSM

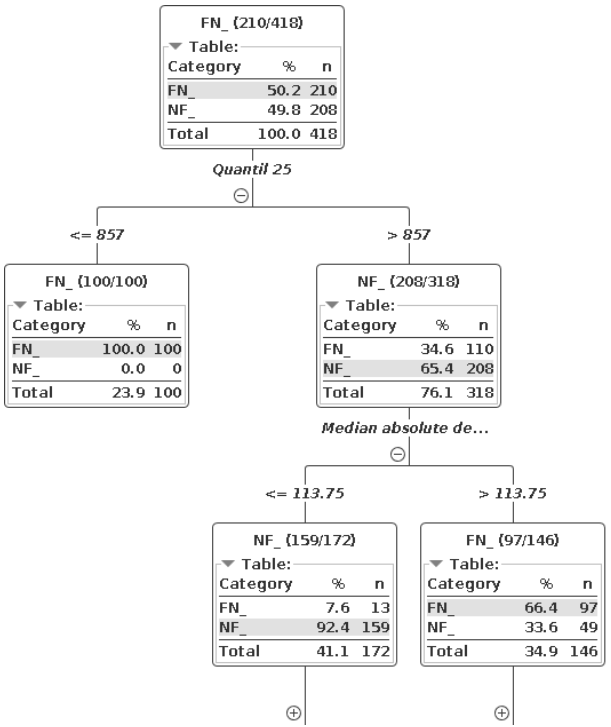
Results and Discussion

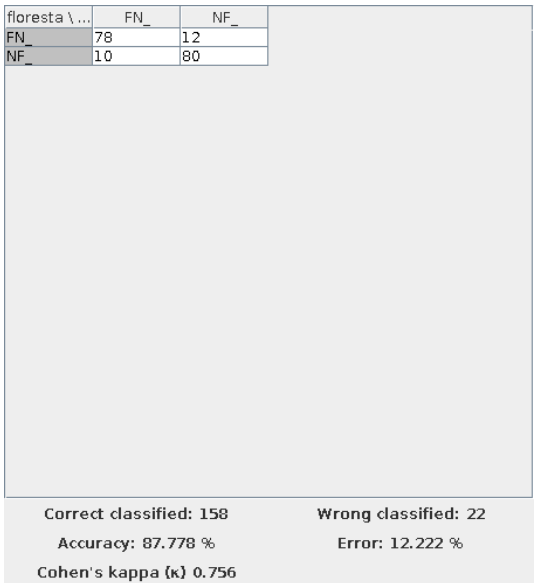
1. *Decision Tree algorithm:* the node produces two information to analyze the classifier's performance.

The first consists of a graphical information in the form of a tree in which it presents the weights and percentages of each of the parameters, used in the characterization process of data image.

The second possible information is represented which shows the performance values of the classifier such as: samples classified correctly, classification errors and accuracy.

The Decision Tree classification algorithm had an accuracy of 87.778%, an error (incorrectly classified samples) of 12.22%, with the Cohen's Kappa k = 0.756 index, considered satisfactory.





The **left figure** represent the classifier performance with weights and percentages of each of the parameters.

The **right figure** present Confusion Matrix.

Federal University of Santa Maria
Brazil
UFSM

Results and Discussion

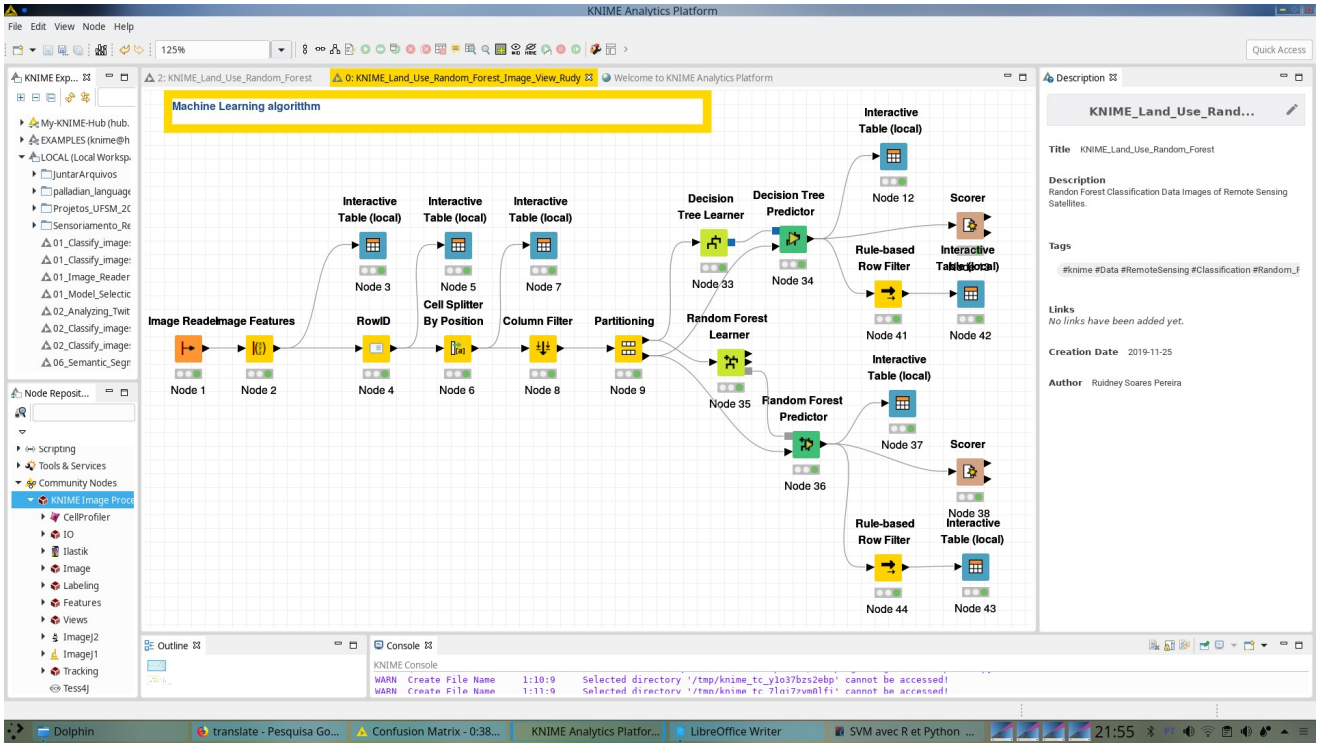
2. Random Forest algorithm: shows the performance values of the classifier such as: samples classified correctly, classification errors and accuracy. The Random Forest classification algorithm had an accuracy of 93.333% accuracy, an error (incorrectly classified samples) of 6.667%, with the Cohen's Kappa k = 0.867 index, considered satisfactory.

FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE

SMART SURVEYORS FOR LAND AND WATER MANAGEMENT

CHALLENGES IN A NEW REALITY



Overview of Workflow Knime for image classification

floresta \ ...	FN_	NF_
FN_	80	10
NF_	2	88

Correct classified: 168 Wrong classified: 12

Accuracy: 93.333 % Error: 6.667 %

Cohen's kappa (κ) 0.867

Randon Forest Confusion Matrix Performance

Federal University of Santa Maria
Brazil
UFSM

Conclusion

1. The Knime Analytics platform appears as a high-performance tool for complex analyzes without requiring a single line of code with a programming language;

2. The flow model used allows it to be improved, as it is possible to export, edit and adapt it to the interests of each user.

3. The analysis of the performance of the Decision Tree and Random Forest algorithms allowed us to conclude that it is possible to classify the images with the necessary precision.

4. Random Forest was the one that presented the best performance in the classification of images with the target of interest Forests.

5. Then, it is concluded that it is possible to make an intelligent knowledge management.

Federal University of Santa Maria
Brazil
UFSM

FIG NL THE NETHERLANDS

e WORKING WEEK 2021 20-25 JUNE

SMART SURVEYORS FOR LAND AND WATER MANAGEMENT

CHALLENGES IN A NEW REALITY

Thanks for your attention!

Contact: rudiney.s.pereira@ufsm.br

Federal University of Santa Maria
Brazil

UFSM