

# Machine Learning

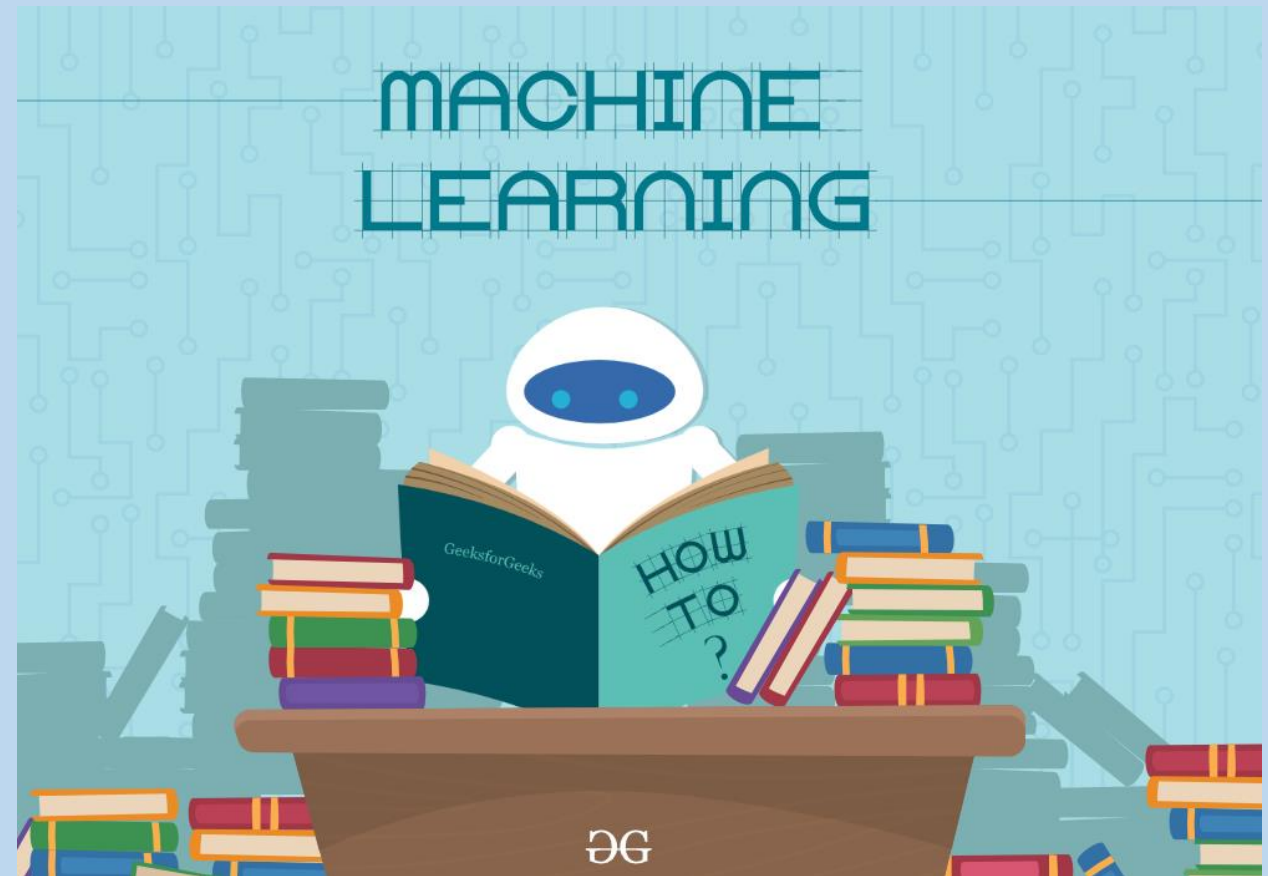




"Field of study that gives computers an ability to learn without being explicitly programmed" - Arthur Samuel  
1959

# Machine Learning

It explores the study and construction of algorithms that can learn from their errors and make predictions about data through inductive rationality, rather than simply following programmed instructions.



# edureka!



**VS**



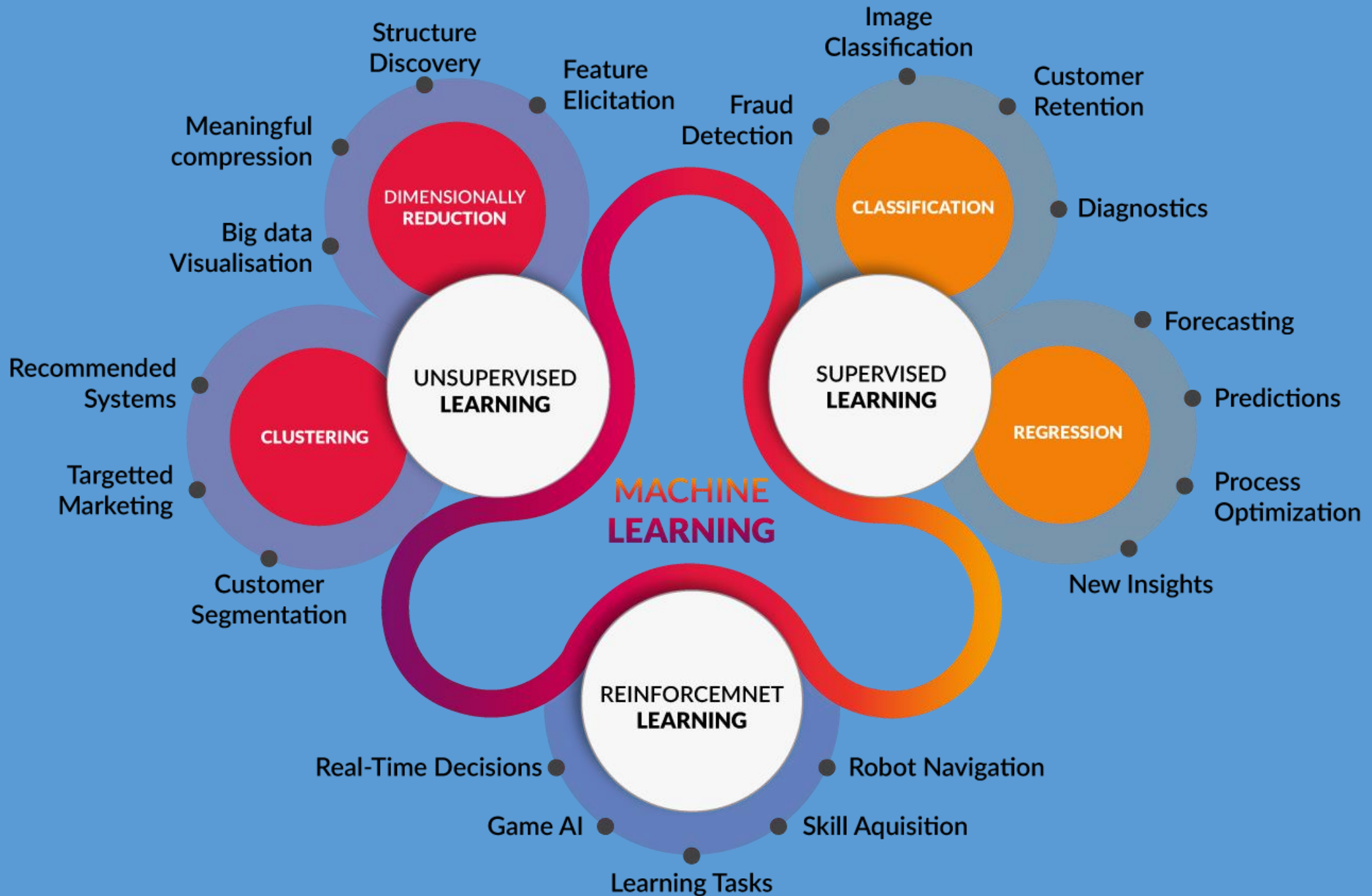
**VS**



**Artificial  
Intelligence**

**Machine  
Learning**

**Deep  
Learning**



# Applications



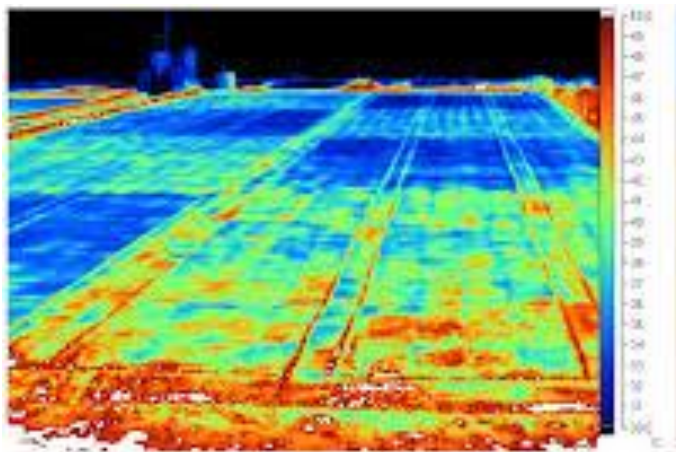
```
role_id => $role_details['role_id'];
resource_id => $resource_details['resource_id'];
is_rule_exists($resource_details['id'], $role_id, $access) {
    // Remove the rule as there is currently no need for it
    $details['access'] = !$access;
    $this->sql->delete('acl_rules', $details);
} else {
    // Update the rule with the new access value
    $this->sql->update('acl_rules', array('access' => $access));
}
foreach($this->rules as $key=>$rule) {
    if ($details['role_id'] == $rule['role_id'] && $details['access'] == $rule['access']) {
        unset($this->rules[$key]);
    } else {
        $this->rules[$key]['access'] = $access;
    }
}
```



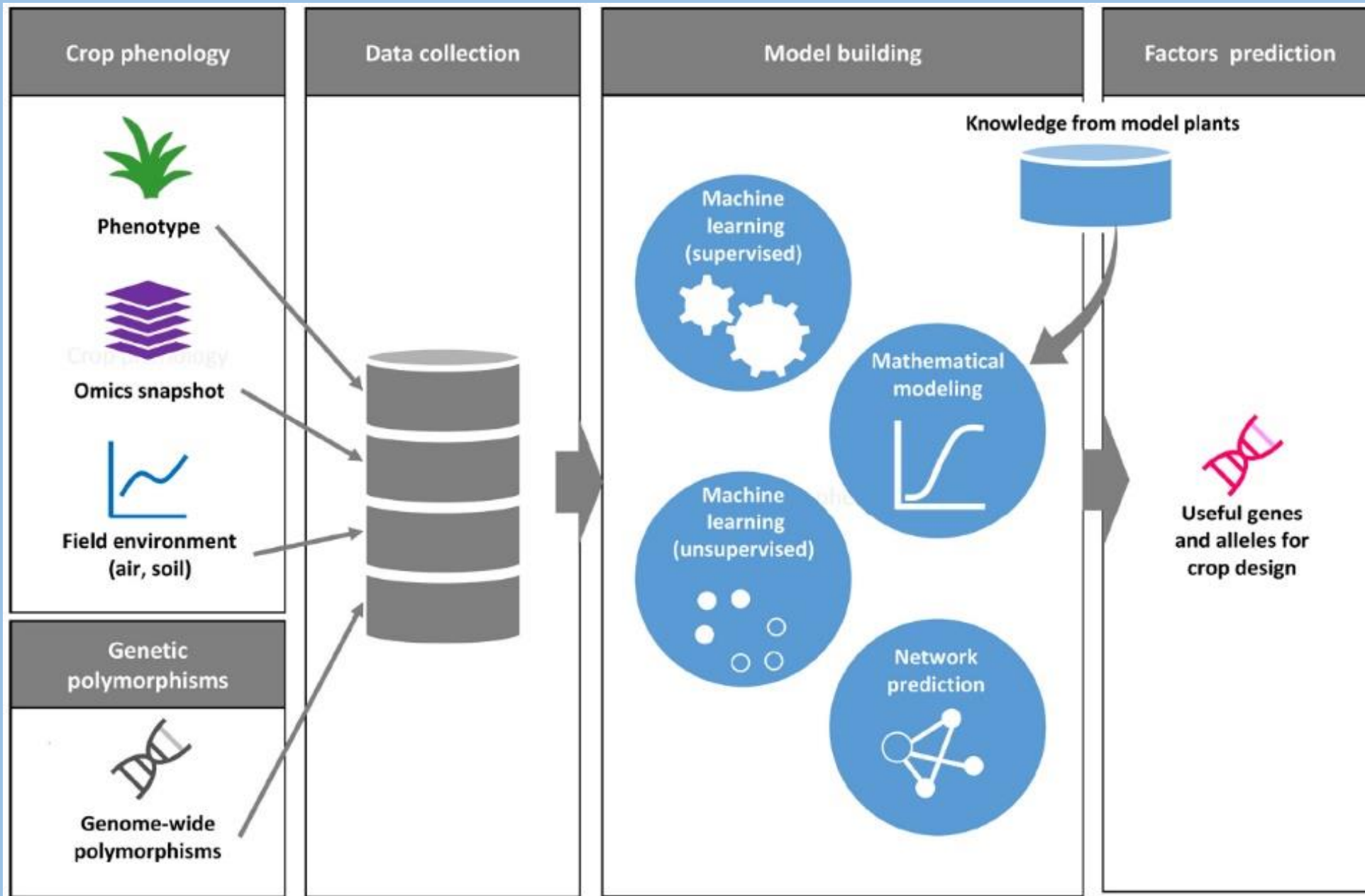
# Machine Learning in Plant Breeding



# High Throughput Phenotyping







# Artificial Neural Network

*Protoplasma*. 2017 Jan;254(1):335-341. doi: 10.1007/s00709-016-0953-3. Epub 2016 Apr 11.

## Artificial neural network-based model for the prediction of optimal growth and culture conditions for maximum biomass accumulation in multiple shoot cultures of *Centella asiatica*.

Prasad A<sup>1</sup>, Prakash O<sup>2</sup>, Mehrotra S<sup>1</sup>, Khan F<sup>2</sup>, Mathur AK<sup>1</sup>, Mathur A<sup>3</sup>.

⊕ Author information

Journal of Plant Physiology 167 (2010) 1226–1231



Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: [www.elsevier.de/jplph](http://www.elsevier.de/jplph)



## Artificial neural networks modeling the *in vitro* rhizogenesis and acclimatization of *Vitis vinifera* L.

Jorge Gago<sup>a</sup>, Mariana Landín<sup>b</sup>, Pedro Pablo Gallego<sup>a,\*</sup>

<sup>a</sup> Applied Plant and Soil Biology, Faculty of Biology, University of Vigo, 36310 Vigo, Spain

<sup>b</sup> Department of Pharmacy and Pharmaceutical Technology, Faculty of Pharmacy, University of Santiago, 15782 Santiago de Compostela, Spain

### ARTICLE INFO

#### Article history:

Received 20 October 2009  
Received in revised form 6 April 2010  
Accepted 7 April 2010

#### Keywords:

ANNs  
Artificial intelligence  
Grapevine

### ABSTRACT

This study employs artificial neural networks (ANNs) to create a model to identify relationships between variables affecting the *in vitro* rhizogenesis and acclimatization of two cultivars of *Vitis vinifera* L. Albariño and Mencía. The effects of three factors (inputs), the type of cultivar, concentration and exposure time to indolebutyric acid (IBA), on the success of *in vitro* rhizogenesis and acclimatization were evaluated. The developed model, using ANNs software, was assessed using a separate set of validation data and was in good agreement with the observed results. Exposure time to IBA was found to have the dominant role in influencing the height of acclimatized plantlets. ANNs can be a useful tool for modeling different complex processes and data sets, in plant tissue cultures or, more generally, in plant biology.

...d modelling approach is used to determine the synergistic effect of five major components of growth (se) on improved *in vitro* biomass yield in multiple shoot cultures of *Centella asiatica*. The back is employed to concentrations 30, 40, 60 ml forming three d is of the model ral concentrati )W/culture) in ca herb that ca

REDES NEURAS ARTIFICIAIS E *WARD-MLM* APLICADOS À ANÁLISE DA DIVERGÊNCIA GENÉTICA EM GOIABA (*Psidium guajava* L.)

1),  
%  
6

...lling; *Centella a*

...33-3

**BIANCA MACHADO CAMPOS**

Dissertação apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Genética e Melhoramento de Plantas.

# Support Vector Machine

Proceedings of the 4th GEOBIA, May 7-9, 2012 - Rio de Janeiro - Brazil. p.035

## COMPARISON OF MACHINE LEARNING ALGORITHMS RANDOM FOREST, ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE MAXIMUM LIKELIHOOD FOR SUPERVISED CROP TYPE CLASSIFICATION

I. Nitze<sup>a</sup>, U. Schulthess<sup>b</sup>, H. Asche<sup>c</sup>

a,c Universität Potsdam, Institut für Geographie, 14476 Potsdam – ingmarnitze@gmail.com, gislab@uni-potsdam.de  
b 4DMaps, 10405 Berlin - usschulthess@4dmaps.de

**KEY WORDS:** Crop Classification, Machine Learning Algorithms, Support Vector Machine, RapidEye

### ABSTRACT:

The classification and recognition of agricultural crop types is an important application of remote sensing. New machine learning algorithms have emerged in the last years, but so far, few studies only have compared their performance and usability. We compared three different state-of-the-art machine learning classifiers, namely Support Vector Machine (SVM), Artificial Neural Network (ANN) and Random Forest (RF) as well as the traditional classification method Maximum Likelihood (ML). For this purpose we classified a dataset of more than 500 crop fields located in the Canadian Prairies with a randomized sampling approach. Up to four multi-spectral RapidEye images from the 2009 growing season were used. We compared the mean overall classification accuracies as well as standard deviations. Furthermore, the classification accuracy of each classifier was analysed. Support Vector Machine classifiers using radial basis function or polynomial kernels exhibited superior results compared to ANN and RF in terms of overall accuracy and robustness, while ML exhibited inferior accuracies and higher variability. ML exhibited the best results for early-season mono-temporal analysis. With a multi-temporal approach, the highest accuracy was achieved for Rapeseed and Field Peas. Other crops, such as Wheat, Flax and Lentils were also successfully classified. Producer's accuracies were higher than 85 %.

## Support vector machines-based identification of alternative splicing in *Arabidopsis thaliana* from whole-genome tiling arrays

Johannes Eichner<sup>1,3\*</sup>, Georg Zeller<sup>1,2,4</sup>, Sascha Laubinger<sup>2,5</sup>, Gunnar Rättsch<sup>1</sup>

### Abstract

**Background:** Alternative splicing (AS) is a process which generates several distinct mRNA isoforms from the same gene by splicing different portions out of the precursor transcript. Due to the (patho-)physiological importance of AS, a complete inventory of AS is of great interest. While this is in reach for human and mammalian model organisms, our knowledge of AS in plants has remained more incomplete. Experimental approaches for monitoring AS are either based on transcript sequencing or rely on hybridization to DNA microarrays. Among the microarray platforms facilitating the discovery of AS events, tiling arrays are well-suited for identifying intron retention, the most prevalent type of AS in plants. However, analyzing tiling array data is challenging, because of high noise levels and limited probe coverage.

**Results:** In this work, we present a novel method to detect intron retentions (IR) and exon skips (ES) from tiling arrays. While statistical tests have typically been proposed for this purpose, our method instead utilizes support vector machines (SVMs) which are appreciated for their accuracy and robustness to noise. Existing EST and cDNA sequences served for supervised training and evaluation. Analyzing a large collection of publicly available microarray and sequence data for the model plant *A. thaliana*, we demonstrated that our method is more accurate than existing approaches. The method was applied in a genome-wide screen which resulted in the discovery of 1,355 IR events. A comparison of these IR events to the TAIR annotation and a large set of short-read RNA-seq data showed that 830 of the predicted IR events are novel and that 525 events (39%) overlap with either the TAIR annotation or the IR events inferred from the RNA-seq data.

**Conclusions:** The method developed in this work expands the scarce repertoire of analysis tools for the identification of alternative mRNA splicing from whole-genome tiling arrays. Our predictions are highly enriched with known AS events and complement the *A. thaliana* genome annotation with respect to AS. Since all predicted AS events can be precisely attributed to experimental conditions, our work provides a basis for follow-up studies focused on the elucidation of the regulatory mechanisms underlying tissue-specific and stress-dependent AS in plants.

## Strengths of artificial neural networks in modelling complex plant processes

---

Jorge Gago,<sup>1</sup> Mariana Landín<sup>2</sup> and Pedro Pablo Gallego<sup>1,\*</sup>

<sup>1</sup>Applied Plant and Soil Biology; Faculty of Biology; University of Vigo; Vigo, Spain; <sup>2</sup>Department of Pharmacy and Pharmaceutical Technology; Faculty of Pharmacy; University of Santiago; Santiago de Compostela, Spain

Commonly, simple mathematical models can not be used to describe exactly the biological processes due to their higher complexity. In fact, most biological interactions cannot be elucidated by a simple stepwise algorithm or a precise formula, particularly when the data are complex or noisy. ANNs allows an accurate description of those kind of biological processes in plant science, offering new advantages over tra-

and binary data should to be analyzed by Poisson regression and binary logistic regression, respectively. For this purpose, plant researchers need a high level statistical background therefore, in many cases, have to be assisted by statisticians in spite of it the outputs being difficult to understand.<sup>3</sup> Furthermore, techniques to model the whole process or to obtain optimized values taking in account all the factors which influence the study are hard to use

# Hidden Markov Model

Article

## Efficiency of neural network-based combinatorial model predicting optimal culture conditions for maximum biomass yield in hairy root cultures

November 2012 · Plant Cell Reports 32(2)

DOI: 10.1007/s00299-012-1364-3

Source · [PubMed](#)

● Shakti Mehrotra · ● Om Prakash · ● Feroz Khan · A K Kukreja

[Overview](#) [Stats](#) [Comments](#) [Citations \(20\)](#) [References \(29\)](#) [Related res](#)



Recommend this research to let the authors know you liked their work

[Not now](#) [Recommend](#)

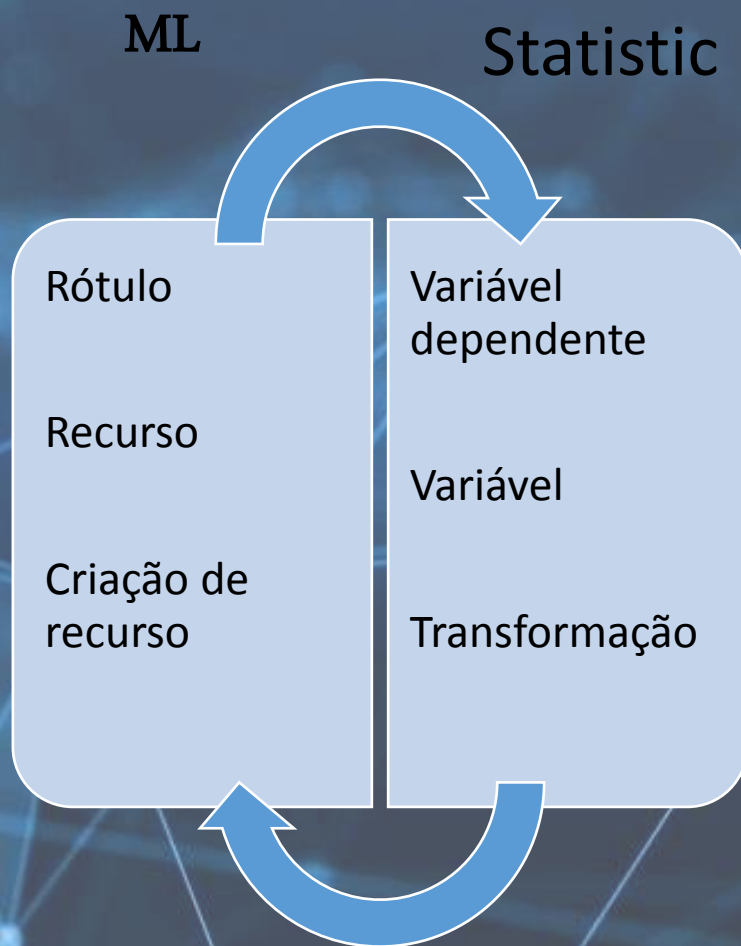
### Abstract

Key message ANN-based combinatorial model is proposed and its efficiency is assessed for the prediction of optimal culture conditions to achieve maximum productivity in a bioprocess in terms of high biomass. Abstract A neural network approach is utilized in combination with Hidden Markov concept to assess the optimal values of different environmental factors that result in maximum biomass productivity of cultured tissues after definite culture duration. Five hidden Markov models (HMMs) were derived for five test culture conditions, i.e. pH of liquid growth medium, volume of medium per culture vessel, sucrose concentration (%w/v) in growth medium, nitrate

# What is necessary?

- Data preparation capability;
- Algorithms - basic and advanced;
- Automated and iterative processes;
- Scalability;
- Conjoint modeling.

# Did You Know?



# How to make?

- Gathering data
- Preparing that data
- Choosing a model
- Training
- Evaluation
- Hyperparameter tuning
- Prediction



# Truths

- Machine learning is not magic
- Data must have quality
- Non-directed data does not contribute at all
- Generalization
  - overfitting

# MACHINE LEARNING



**THANK YOU!**

# References

- Links:

- [https://www.sas.com/pt\\_br/insights/analytics/machine-learning.html](https://www.sas.com/pt_br/insights/analytics/machine-learning.html)
- <https://tecnoblog.net/247820/machine-learning-ia-o-que-e/>
- [https://pt.wikipedia.org/wiki/Aprendizado\\_de\\_m%C3%A1quina](https://pt.wikipedia.org/wiki/Aprendizado_de_m%C3%A1quina)
- <https://medium.com/data-science-brigade/a-diferen%C3%A7a-entre-intelig%C3%Aancia-artificial-machine-learning-e-deep-learning-930b5cc2aa42>
- <https://www.datascienceacademy.com.br/course?courseid=machine-learning-engineer>

- Artigos:

- Gago et al. Artificial neural networks modeling the in vitro rhizogenesis and acclimatization of *Vitis vinífera* L. Journal of plant physiology, 2010.
- Prasad A. et al. Artificial neural network-base model for the prediction of optimal growth and culture conditions for maximum accumulation in multiple shoot cultures of *Centelha asiática*. Protoplasma, 2017.
- Campos. Redes neurais artificiais e Ward –mlm aplicados a análise da divergência genética em goiaba (*Psidium guajava* L.), 2012.
- Nitze et al. Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. Proceedings of the 4th GEOBIA, May 7-9, 2012 - Rio de Janeiro - Brazil. p.035.
- Eichner et al., Support vector machines-based identification of alternative splicing in *Arabidopsis thaliana* from whole-genome tiling arrays, BMC Bioinformatics, 2011.
- Gago et al. Strengths of artificial neural networks in modeling complex plant processes, plant Signaling & Behavior, 2010.