

# Digital modeling of soil-borne fusarium's nutritional: investigating microbiological and microecological dynamics in Moroccan agroecosystems

Youssef El Hilali Alaoui <sup>1,5\*</sup>, Said Bouda <sup>1</sup>, Samira Chabaa <sup>2</sup>, Mohammed Elouali Alami <sup>3</sup>, Zakaria Khoudi <sup>4</sup>, and Adil Essarioui <sup>5</sup>

<sup>1</sup> Laboratory of Agro-Industrial and Medical Biotechnology, Faculty of Sciences and Technics, Sultan Moulay Slimane University, B.P. 523, Beni Mellal, Morocco

<sup>2</sup> LISAD Laboratory, Industrial Engineering Dept. ENSA, Ibn Zohr University, Agadir, Morocco

<sup>3</sup> Intelligent System and Security of System, Mohamed V university, Rabat, Morocco

<sup>4</sup> Equipe des Mathematiques et Interactions, Faculty of Sciences and Technics, Sultan Moulay Slimane University, Beni Mellal, Morocco

<sup>5</sup> Oasis System Research Unit, Regional Center of Agricultural Research of Errachidia, National Institute of Agricultural Research, PO. Box 415, Rabat 10090, Morocco

**Abstract.** The areas planted with date palm trees within the Moroccan oases cover more than 48,000 hectares and play a key role in both the environment and socioeconomic stability. Unfortunately, these ecosystems are threatened by the Bayoud disease caused by *Fusarium oxysporum* f. sp. *albedinis*, a vascular wilt pathogen that has already devastated millions of date palms in Morocco and Algeria since the 19th century. Any similar outbreak over time poses a serious threat to the long-term sustainability of these oases. This study aimed to elucidate the biological mechanisms associated with Bayoud decline in suppressive soils. To achieve this, soil samples were collected from the Ziz and Draa Valleys, where date palms are infected by Bayoud, as well as from the palm grove of Marrakech, which is considered a suppressive zone for this disease. In other words, the samples were taken from two disease-conducive zones and one suppressive zone for Bayoud. A total number of eighteen samples were removed from various depths to compare two conducive soils and one suppressive soil. Ninety *Fusarium* strains were isolated using this approach and tested for their antagonistic or competitive properties against the Bayoud pathogen. The bacterial and fungal communities were characterized using ITS1 and 16S amplicon sequencing, respectively, with growth tests conducted on Biolog SF-P2 plates. Using soil samples from the three research regions, we investigated three machine learning techniques to determine the feeding patterns of *Fusarium* communities: Decision tree models, k-nearest neighbors, and Logistic regression. The performance scores of the models were as follows: the k-nearest neighbors model achieved 80%, the logistic regression model scored 77.78%, and the decision tree classifier obtained a score of 68%. These results highlight the potential of machine learning approaches in understanding the nutritional behavior of *Fusarium* communities. Our research provides a foundation for modeling efforts aimed at generating forecasts to mitigate the damages caused by Bayoud on Morocco's vital date palm ecosystems.

## 1 Introduction

Over 48,000 hectares and more than 4.8 million trees are supported by the date palm (*Phoenix dactylifera*), which is essential to the socioeconomic and environmental stability of Morocco's oasis [1]. But since the late 19th century, *Fusarium oxysporum* f. sp. *albedinis*, the pathogenic fungus that causes Bayoud disease, has devastated approximately 10 million trees in Morocco and Algeria, putting these distinctive ecosystems in jeopardy [2]. Effective, large-scale control strategies are still elusive despite decades of research, in part because this soil-borne disease is so persistent [3].

In this work, we propose a novel model of the nutritional behavior of soil-borne *Fusarium* populations as a means of explaining Bayoud illness. According to our bibliographic research, this is considered the first model of its kind that thoroughly focuses on both conducive and suppressive soils in Morocco's oases. Our model is a potent tool to replicate the environmental variables that promote or inhibit *Fusarium* growth, as it combines nutrient competition experiments with microbial community studies. Furthermore, the development of early detection kits and simulation tools made possible by this approach could revolutionize the management of bayoud in date palm fields. It also has major practical applications.

Our study is one of the first studies to predict the nutritional interactions between *Fusarium* and various soil types in the context of vascular wilt disease in date palms

\*Corresponding author: [hilalialaoui.79@gmail.com](mailto:hilalialaoui.79@gmail.com)

within Moroccan and Algerian oases. In addition to enhancing our understanding of soil suppressiveness, this model could serve as a basis for the development of diagnostic kits and simulation tools designed for early detection and prevention of Bayoud disease. Such tools would be valuable for reducing the financial and environmental damage caused by this pathogen in Morocco's date palm ecosystems.

Within the materials and methods section, we go over how to sample soil, isolate *Fusarium* strains from various soils, perform microbial competition assays, and employ sequencing approaches to examine bacterial and fungus populations. This section also explains the modeling methods used to replicate the nutritional behavior of *Fusarium* under various circumstances, including logistic regression, the k-neighbors model, and decision trees. The outputs of the modeling studies, emphasizing each model's prediction accuracy, are then presented in the results section along with information on the microbial community composition in conducive and suppressive soils and the outcomes of nutrient competition assays. Lastly, the discussion section provides an interpretation of these findings within the framework of managing Bayoud disease, outlining the consequences for the production of suppressive soils and pathogen early detection kits. We also talk about the study's shortcomings and suggest areas for further investigation, such as building on the model to include more edaphic and genetic components.

## 2 Materials and Methods

### 2.1 Soil Sampling

To examine the relationship between soil microorganisms and the pathogen *Fusarium*, three types of soils were chosen for this study, comprising two conducive soils and one suppressive soil. A total of eighteen soil samples were taken from six replicate trees at two different depths, 0–20 cm and 20–40 cm. The sampling was conducted following the method described by Essarioui et al. (2018), [1].

### 2.2 Isolation and Nutrient Testing of *Fusarium*

*Fusarium* isolates were taken out of the soil between 0 and 20 cm deep. Six palm duplicates and five isolates per plot throughout three plots were collected to yield ninety isolates. Inhibition tests were conducted to assess the isolates' ability to suppress pathogens. To study *Fusarium* development and edaphic characteristics that can affect nutrient competition among native *Fusarium* communities, PDA (Potato Dextrose Agar) and Biolog SP-F2 plates were utilized [5].

The Biolog SP-F2 microplates are made especially for measuring how well fungi use carbon sources. Different carbon sources are present in each well, and over time, the

optical density (OD) of each well is measured to determine how much of each isolate has consumed carbon sources. In this work, we measured the use of eight different groups of carbon sources at intervals ranging from twenty-four to one hundred and sixty hours. The selected carbon sources are important constituents of soil organic matter, and the way in which they are utilized by the *Fusarium* isolates demonstrates their metabolic adaptability and ability to compete under diverse soil conditions. To ascertain the metabolic profiles of the isolates and their capacity to compete for nutrients under various conditions, the OD measurements were collected and examined [6].

### 2.3 Numerical Data Normalization and Model Testing

Data normalization was an essential step before modeling to guarantee accurate and trustworthy statistical analysis. The min-max normalization method was used to normalize the dataset, which had 68,400 numerical values and was made up of 720 rows and 95 columns (Table 1). In order to prevent factors with higher values from disproportionately influencing the outcomes, this strategy was adopted to bring all variables to the same scale. Using the following formula, the min-max normalization converts each variable's values to a range between 0 and 1:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

**Where:**  $x$  is the original value,  $x_{min}$  and  $x_{max}$  are the minimum and maximum values of the variable, respectively [7]. In order to categorize the soil samples into favorable or suppressive soils based on microbial interactions and edaphic variables, three predictive models were explored.

#### 2.3.1 Logistic Regression Model

Typically employed in binary classification, logistic regression models the likelihood of a binary result (in this example, conducive or suppressive soils). The logistic regression model uses a number of characteristics, including nutrient competition and the makeup of the microbial community, to calculate the likelihood that a soil is suppressive. Any real-valued number may be mapped by the model using the sigmoid function into a probability between 0 and 1, which can be expressed as follows:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}$$

**Where:**

- **P(Y=1|X):** Probability that the outcome is 1 (e.g., suppressive soil) given the characteristics X.
- **$\beta_0$ :** Intercept term representing the baseline value when all  $X_i$  are zero.

- $\beta_i$ : Coefficient of the  $i$ -th variable  $X_i$ , indicating its influence on the probability of  $Y=1$ .
- $X_i$ : The  $i$ -th explanatory variable (Total growth in Carbone source)
- $e$ : The base of the natural logarithm, approximately equal to 2.71828.

Because it provides a clear understanding of how many factors affect the possibility that soil is suppressive, this model was selected for its simplicity and interpretability [8].

### 2.3.2 K-nearest Neighbors (KNN) Model

A sample is classified using the KNN algorithm, a non-parametric technique, based on the majority class of its closest neighbors in the feature space. The Euclidean distance formula is commonly utilized to compute the distance between samples:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Where:

- $d(x,y)$ : Euclidean distance between two samples  $x$  and  $y$ .
- $x_i$ : The  $i$ -th feature of sample  $x$ .
- $y_i$ : The  $i$ -th feature of sample  $y$ .
- $n$ : The total number of features.
- $\sum_{i=1}^n$ : Summation notation, indicating the sum of the squared differences across all features.

### 2.3.3 Decision Tree Model

Using feature values to maximize the separation between classes, decision trees are a well-liked machine learning approach for classification tasks. The decision tree algorithm determines the Gini impurity at each split to assess how effectively a certain feature divides the classes:

$$Gini = 1 - \sum_{i=1}^c p_i^2$$

where  $P_i$  is the proportion of samples belonging to class  $i$  at a particular node.

### 2.3.4 Data Splitting and Performance Measures

The dataset has been divided into training and testing datasets, with an 90% to 10% split, respectively. Once the training is complete, the model is tested with the test dataset and the various performance metrics are recorded. The performance of the proposed model was evaluated using the most popular metrics: Accuracy, Precision, Recall, and F1-score, defined as follows:

**Table 1.** Normalized data on the nutritional behavior of 90 isolates of saprophytic Fusarium in different media and across 8 time points using the SF-P2 Biolog plate: preparation for numerical modeling

Region	Hour	Total nich>0.005	Total Growth	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
M1	24	5.851	5.869	0.160228	0.104577	0.078993	0.033810	0.003208	0.058673	0.042420	0.031250	0.056075	0.004785	0.000000
M2	24	4.850	4.878	0.059921	0.086799	0.044580	0.020806	0.076183	0.021259	0.150398	0.006250	0.068925	0.122010	0.271382
M3	24	4.062	4.071	0.129556	0.081171	0.025340	0.000000	0.000000	0.011054	0.047722	0.001563	0.037383	0.019139	0.027176
M4	24	3.814	3.832	0.129124	0.078700	0.024871	0.000000	0.000000	0.021259	0.018318	0.010937	0.058411	0.000000	0.005067
M5	24	5.364	5.394	0.159833	0.103689	0.048334	0.032510	0.094627	0.034864	0.059050	0.018750	0.056075	0.035885	0.027176
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Z26	120	39.651	39.656	0.988365	1.000000	1.000000	0.350540	0.721000	1.000000	1.000000	0.432371	0.859438	0.105245	0.034815
Z27	120	32.836	32.846	0.868686	0.973006	0.854787	0.124880	0.447000	0.634000	0.995883	0.271741	0.606190	0.000000	0.000000
Z28	120	34.928	34.935	1.000000	1.000000	1.000000	0.212515	0.541000	0.828000	0.881737	0.265703	0.980905	0.054785	0.005356
Z29	120	31.638	31.639	1.000000	0.975821	1.000000	0.223469	0.516000	0.600000	1.000000	0.334544	0.695276	0.010092	0.000000
Z30	120	30.302	30.307	1.000000	1.000000	1.000000	0.273859	0.366000	0.470000	1.000000	0.283819	0.425118	0.000000	0.000000

The KNN model was used because it can capture similarities and local correlations between soil samples, which makes it especially useful for complicated datasets with non-linear variable interactions. Based on its proximity to other known samples, KNN was used in our investigation to identify whether a soil sample is conducive or suppressive [9].

**Table 2.** Formulas for Key Performance Metrics Used to Evaluate the Model

Metric	Formula
<b>Accuracy</b>	$(TP + TN) / (TP + TN + FP + FN)$
<b>Precision</b>	$TP / (TP + FP)$
<b>Recall</b>	$TP / (TP + FN)$
<b>F1-score</b>	$2 \times (Precision \times Recall) / (Precision + Recall)$

Where:

- **TP** = True Positive
- **TN** = True Negative
- **FP** = False Positive
- **FN** = False Negative

According to Breiman et al. (2017) [10], decision trees are especially helpful for pinpointing the key variables affecting soil suppressiveness and make it simple to visualize the decision-making process. This model was chosen because of its interpretability and capacity to manage intricate, non-linear interactions between features.

Through the integration of these models, our goal is to offer a thorough comprehension of the elements that contribute to Fusarium suppression in various soil settings, as well as useful instruments for the creation of early detection kits for Bayoud disease.

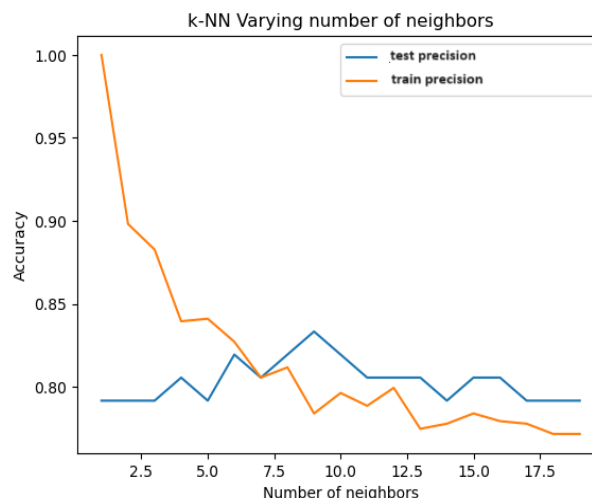
### 3 Results

The results of the model testing for the prediction of *Fusarium* behavior in soil samples yielded distinct performances across the three models.

On the test dataset, the **Logistic Regression model** achieved an accuracy of 77.78%. This model, frequently employed for binary classification, successfully predicted the presence or absence of *Fusarium* based on various environmental and nutritional factors. Logistic regression is well-suited for this type of biological prediction task because it can evaluate the probability of an outcome based on input features and explain the relationship between these variables and the binary response. However, the comparatively lower performance of this model, in contrast to the k-nearest neighbors model, suggests that the complex interactions within the soil's nutritional composition affecting *Fusarium* growth may not be fully captured by its linear assumptions [11,12].

The **K-nearest neighbors model (KNN)** performed better, with an accuracy of 80%. This model groups data points based on their proximity to known samples, allowing a more sophisticated approach for identifying soil samples with similar microbial and nutritional compositions (figure1). KNN's strength lies in its ability

However, its computational cost increases with the data size, which could limit its use with larger datasets. Additionally, the number of neighbors (k) selected significantly impacts system performance, requiring careful tuning to avoid underfitting or overfitting [13]



**Fig 1.** k-Neighbors model applied to the study of the nutritional behavior of *Fusarium* saprophytic isolates: variation in the number of neighbors and accuracy of testing and training.

Finally, the **Decision Tree model** achieved an accuracy of 68%, lower than both the logistic regression and k-nearest neighbors' models. Decision trees are known for their interpretability, providing clear insights into how various factors affect the classification result. In this study, the decision tree likely identified important variables such as specific nutrient levels or environmental conditions influencing *Fusarium* growth (figure 2). However, the model's lower accuracy could indicate poor generalization to the test dataset or overfitting to the training data—common issues with decision trees, especially when noise is present or when there are many ways for variables to split the data.

Ultimately, the Decision Tree Model's accuracy of 68% was less than that of the k-neighbors and logistic regression models. Decision trees are renowned for being easily interpreted since they offer distinct perspectives on the ways in which various factors influence the

Table 3: Performance Comparison of Different Models Across Precision, Recall, F1-Score, and Accuracy Metrics

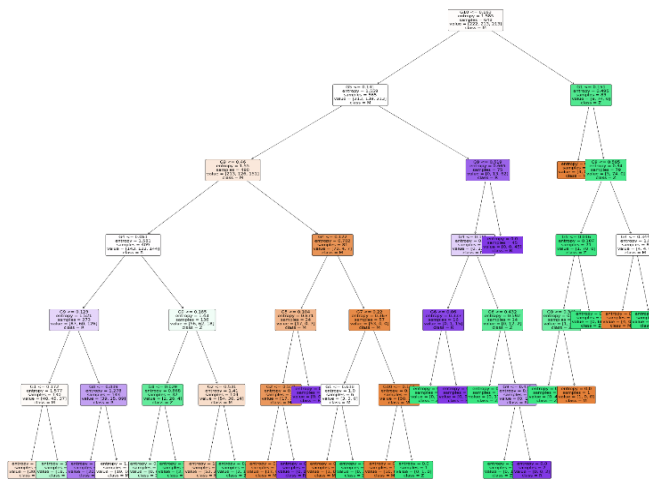
Model	Logistic Regression				K-nearest Neighbors	Decision Tree	
Metrics	Precision	Recall	F1-Score	Accuracy	Accuracy	Accuracy	
Region	R	0.95	0.74	0.83	0.78	0.79	0.68
	M	0.65	0.72	0.68			
	Z	0.74	0.85	0.79			

to capture non-linear relationships in the data, which makes it ideal for datasets with complex nutritional interactions, as demonstrated in several biological studies.

classification result. The decision tree in this study probably identified important variables, like particular nutrient levels or environmental circumstances, that affect



*Fusarium* growth. The model might not have generalized well to the test dataset or may have overfit the training set, based on the comparatively lower accuracy. Decision trees are vulnerable to these kinds of problems, particularly when there is noise in the data or when there are numerous ways for the variables to split [14].



**Fig 2.** Decision Tree model applied to the study of the nutritional behavior of *Fusarium* saprophytic isolates using normalized data from the SF-P2 Biolog plate.

## 4 Discussion

The present study offers a comprehensive and innovative approach to understanding the nutritional behavior of *Fusarium* populations in different soil types found in Moroccan agroecosystems. Through the integration of machine learning models, including logistic regression, k-nearest neighbors (KNN), and decision trees, the study sheds light on the microbiological and microecological dynamics that contribute to soil suppressiveness or conduciveness in relation to *Fusarium* development. This work not only provides valuable insights into the interactions between microbial communities and soil pathogens but also lays the groundwork for future advancements in disease prediction and management, particularly concerning the Bayoud pathogen (*Fusarium oxysporum* f. sp. *albedinis*) (Table 3).

The modeling results indicate that the KNN model outperforms both logistic regression and decision trees, achieving an accuracy of 80%. This is a significant finding, as it highlights the importance of non-linear relationships in the dataset, which are likely a result of the complex interactions between soil nutrients, microbial communities, and environmental factors. The ability of KNN to capture local correlations makes it an ideal choice for datasets where non-linear variable interactions are predominant. In comparison, the logistic regression model, which is inherently linear, achieved an accuracy of 77.78% [12]. This suggests that while logistic regression can explain some of the simpler, linear relationships between variables, it falls short in capturing the

complexity of nutrient competition and microbial interactions present in suppressive and conducive soils. The decision tree model, with an accuracy of 68%, provided some interpretative value by identifying key variables influencing soil suppressiveness, but its lower accuracy may be attributed to overfitting and the inherent noise in the dataset, a common issue with decision trees in biological systems [15].

The results of this study have several important implications for our understanding of soil-borne pathogens and microbial ecology. First, the use of BIOLOG SP-F2 microplates to assess the metabolic profiles of *Fusarium* isolates provides critical insights into their ability to utilize various carbon sources, which in turn reflects their adaptability to different soil environments. This adaptability, as measured by optical density changes over time, highlights the competitive nature of *Fusarium* in both conducive and suppressive soils. The ability of suppressive soils to inhibit *Fusarium* growth, as observed in this study, likely stems from the complex interplay between nutrient availability and microbial competition. These findings are consistent with prior research, which has shown that microbial communities in suppressive soils often possess a higher level of functional diversity and competition for nutrients, thereby limiting the growth of soil-borne pathogens like *Fusarium* [4].

The practical applications of this research are manifold. By developing a predictive model of *Fusarium* nutritional behavior, we can improve early detection methods for Bayoud disease, a devastating pathogen affecting date palms in Morocco [15]. The ability to classify soils as suppressive or conducive using machine learning techniques could pave the way for the development of soil management practices aimed at enhancing soil suppressiveness, thereby reducing the incidence of Bayoud disease in vulnerable agroecosystems. This is particularly relevant given the socioeconomic importance of date palm cultivation in Moroccan oases [16]. Additionally, the insights gained from this study could inform the design of diagnostic kits that use microbial and nutrient profiles to predict disease outbreaks before they manifest, allowing for timely interventions and more sustainable agricultural practices.

Moving forward, there are several avenues for further research that could build on the foundation established by this study. One promising direction is the incorporation of genetic and edaphic factors into the existing model. For example, by integrating data on soil pH, moisture content, and microbial gene expression, future models could provide even more accurate predictions of soil suppressiveness. Furthermore, expanding the dataset to include a broader range of soil types and environmental conditions would increase the generalizability of the model and improve its applicability to other regions affected by *Fusarium*-related diseases [17].

While this study provides valuable insights into the nutritional behavior of *Fusarium* populations, there are

some limitations that should be acknowledged. One potential limitation is the relatively small sample size, which may have contributed to the overfitting observed in the decision tree model. Additionally, the reliance on optical density measurements as the sole indicator of nutrient utilization could be complemented by other biochemical assays to provide a more comprehensive understanding of *Fusarium* metabolic activity [18]. Future studies should aim to address these limitations by increasing the sample size, incorporating additional biochemical markers, and employing advanced machine learning techniques such as random forests or neural networks to further improve model accuracy.

This study is the first to fully explore and model the nutritional interactions between *Fusarium* and different soil types in Moroccan agroecosystems, offering a novel perspective on the factors that contribute to soil suppressiveness. By integrating nutrient competition experiments with microbial community analysis and machine learning, this research provides a multifaceted approach to understanding and managing Bayoud disease. The development of early detection kits and simulation tools based on this model has the potential to revolutionize the management of soil-borne pathogens in date palm cultivation, reducing both the economic and environmental impact of Bayoud disease. This research is therefore not only scientifically valuable but also of significant practical relevance for the agricultural sector in Morocco and beyond.

## 5 Conclusion and Perspectives

This study presents a groundbreaking approach to modeling the nutritional behavior of *Fusarium* in various soil environments. The findings underscore the importance of non-linear interactions in microbial communities and highlight the potential of machine learning models to improve disease prediction and management strategies. Future research should aim to refine the model by incorporating additional environmental and genetic variables, as well as expanding its applicability to other regions and soil types. Ultimately, this research sets the stage for more effective and sustainable management of soil-borne pathogens in Moroccan agroecosystems.

## References

1. Essarioui, A., Ben-Amar, H., Khoulassa, S., Meziani, R., Amamou, A., & Mokriani, F., 2018, "Gestion du Bayoud du palmier dattier dans les oasis marocaines," *Rev. Mar. Sci. Agron. Vét.*, 6, no. 4, pp. 537-543.
2. El Hilali Alaoui Y, Bouda S, Amar HB, Bouhlali EDT, Khoulassa S, Boutagayout A, et al. Enhancing Date Palm (*Phoenix dactylifera* L.) resilience: Strategies against Bayoud spread in Moroccan oases. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing; 2024. p. 012001.
3. Sedra, M. H., 2000, "Biological and genetic characteristics of bayoud resistant Moroccan date palm cultivars and strains," *Agric Water*, 20, pp. 55–66.
4. El Bouhssini, M., & Faleiro, J. R., 2018, *Date Palm Pests and Diseases: Integrated Management Guide*, Beirut, Lebanon: International Center for Agricultural Research in the Dry Areas (ICARDA).
5. Essarioui, A., Meziani, R., & Mokriani, F., 2020, "Resource competitive interactions as mechanism of date palm Bayoud disease suppression," *Moroccan Journal of Agricultural Sciences*, 1, no. 2. (Numéro de page manquant).
6. LeBlanc, N., Essarioui, A., Kinkel, L., & Kistler, H. C., 2017, "Phylogeny, plant species, and plant diversity influence carbon use phenotypes among *Fusarium* populations in the rhizosphere microbiome," *Phytobiomes*, 1, no. 3, pp. 150-157. <https://doi.org/10.1094/PBIOMES-06-17-0028-R>.
7. Muhammad, Ali P. J., 2022, "Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements," *ARO-The Scientific Journal of Koya University*, 10, no. 1, pp. 85-91. <https://doi.org/10.14500/aro.10955>.
8. Kumar, S., 2024, "Logistic Regression," *Python for Accounting and Finance*, Cham: Palgrave Macmillan, pp. 99–115. [https://doi.org/10.1007/978-3-031-54680-8\\_19](https://doi.org/10.1007/978-3-031-54680-8_19).
9. Acito, F., 2023, "K Nearest Neighbors," *Predictive Analytics with KNIME*, Cham: Springer, pp. 189–204. [https://doi.org/10.1007/978-3-031-45630-5\\_10](https://doi.org/10.1007/978-3-031-45630-5_10).
10. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J., 2017, *Classification and Regression Trees*, Taylor & Francis. (Ville de publication manquante).
11. Jauk, S., Kramer, D., Sumerauer, S., Veeranki, S. P. K., Schrempf, M., & Puchwein, P., 2024, "Machine learning-based delirium prediction in surgical inpatients: a prospective validation study," *JAMIA Open*, 7, no. 3, ooae091. <https://doi.org/10.1093/jamiaopen/ooae091>.
12. Wang, P., Chen, S., & Yang, S., 2022, "Recent Advances on Penalized Regression Models for Biological Data," *Mathematics*, 10, no. 19, p. 3695. <https://doi.org/10.3390/math10193695>.
13. Uddin, S., Haque, I., Lu, H., & Moni, M. A., 2022, "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Scientific Reports*, 12, no. 1, pp. 1-12. (DOI manquant).
14. Guidotti, R., Monreale, A., Setzu, M., & Volpi, G., 2024, "Generative Model for Decision Trees," *Proceedings of the AAAI Conference on Artificial Intelligence*, 38, no. 19, pp. 21116-21124. (DOI manquant).
15. Han, J., Kamber, M., & Pei, J., 2011, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, ISBN: 978-0-12-381479-1. (Ville de publication manquante).

16. Elbouazaoui, A., Douira, A., Ilyass, M., & Ahmed, S., 2022, "Integrating Sowing Date with Chickpea Genotypes in Managing Fusarium Wilt in Morocco," *Agriculture*, 12, no. 6, p. 773. <https://doi.org/10.3390/agriculture12060773>.
17. Segura-Mena, R. A., Stoorvogel, J. J., García-Bastidas, F., et al., 2021, "Evaluating the Potential of Soil Management to Reduce the Effect of *Fusarium oxysporum* f. sp. *ubense* in Banana (Musa AAA)," *European Journal of Plant Pathology*, 160, no. 3, pp. 441-455. <https://doi.org/10.1007/s10658-021-02255-2>.
18. Lal, D., Dev, D., & Kumari, S., 2024, "Fusarium wilt pandemic: current understanding and molecular perspectives," *Funct Integr Genomics*, 24, p. 41. <https://doi.org/10.1007/s10142-024-01319-w>.