NEW PREDICTIVE MODEL FOR AGRICOLTURE, BASED ON BOTH SOIL AND WEATHER CONDITIONS (CLIMATE CHANGE) FOR TO ENSURE THE SELECTION OF THE MOST SUITABLE-CROP

Autors: Sirio Cividino(*), Marco Colombo (*), Luigi Passariello (**), Angela Maiorana (*), Michele Passariello (**) ,

*CRSLaghi -Lake research and study centre
**INGV: National Institute di Geophisics and Volcanology, public scientific centre of research of Italian Ministry of Research
*** Ma.Pa.COM: ICT Company committed to the development of high added value solutions based on 4.0 and 5.0 technologies

Keywords: Wheater parameters, Climate change, Deep Learning, Convolutional networks, Soil parameters, crop parameters.

Abstract

Predictive analytics tools in agriculture are tools that use a variety of statistical methods including data mining, predictive modeling, and machine learning that analyze an array of current and historical agricultural, biological, climate, and hydrological data from various sources to make predictions about future outcomes on the farm. In this project, we developed a multi-class classification model using machine learning to predict the most suitable crop based on the provided soil measurements. Adding weather data into the model enhanced its predictive power.

INTRODUCTION

The impact of agriculture all over the world cannot be overemphasized as it remains the major source of food for mankind. Although the human population continue to increase while the land to cultivate remains static (Virk et al. 2020). As such, the knowledge about the compatibility of the key factors considered when planning or cultivating a farm such as a crop variety, soil characteristics, fertilizer and climate is highly required. These factors cannot be correctly selected using the traditional approach (human experience) due to their varied nature. The authors (Virk et al. 2020), asserted that the variability of these resources negatively afects agricultural output and crop yield in particular. Also, the authors (Mesgaran et al. 2017), observed that the ever-changing climate condition greatly reduces the output of agricultural products. Thus, farmers continue to experience low agricultural output due to the varying nature and poor management of the farming resources. The low output is alarming as the land for cultivation also degrades continuously due to persistent use without proper knowledge of the right soil chemical constituents (fertilizer or soil manure) to complement the lost nutrients. Hence, to meet the growing demand for food as the human population continues to increase. Predicting soil chemical properties, fertilizer and climate suitable for particular crops to excel is an important agricultural problem. More so, predicting the right combination of these resources is paramount to achieving a high crop yield (Kurşun and Dengiz 2020). Especially in underdeveloped countries, a poor understanding of the suitability of these factors to improve agricultural output has negatively impacted crop yield. Accurate and automatic information about the kind of soil and fertilizer compatible with crops is important. These pieces of information are greatly required in agriculture to curtail the impending food crisis (Feyisa et al. 2020) since low crop yield is by this means greatly reduced. Consequently, the advent of smart farming and precision agriculture came in as an innovation in the agricultural sector to automate farming processes towards achieving high yield quantity and quality for food sustainability. Smart farming particularly entails the application of technologies and Artifcial Intelligence (AI) in managing farms to increase the quantity and quality of crop yield (Adamides 2020).

The agricultural industry through proximity sensors and remote sensing is acquiring data at an unprecedented rate, ranging from agronomy to climate, logistics and market price volatility (Fenu et al., 2021). Multi-sensor and multi-source data flowing into large datasets of different volume, variety, velocity, veracity and variability. A massive wealth of data, from which to draw information and extract knowledge. Indeed, today's challenge lies in the analysis, understanding and representation of these heterogeneous data, in order to discern selective and systematic interventions to support agricultural decision-making processes. Artificial Intelligence (AI) (Fenu et al., 2020), and its Machine Learning (ML) (Fenu et al., 2019). and Deep Learning (DL) subfields (Fenu et al., 2021), are the most promising auxiliary analysis techniques for understanding this complex system, opening up new challenges and opportunities to the world of research. The main contribution is related to the analysis aimed at supporting decisions. Information access systems, such as Decision Sup port Systems (Fenu et al., 2021) and Recommendation Systems, can help filter the large amount of information collected by IoT sensors. The intersection of large amounts of data, obtained over long periods, leads in the medium-long term to create intelligent models that infer relationships, models and trends from the original data to suggest targeted and efficient actions based on actual crop needs and biochemical and physical characteristics of the soil. Operationally, the interventions are divided into several levels, such as monitoring, prevention, planning and treatment. Among the most relevant are: (i) prediction and monitoring of parasitic adversities, (ii) crop prediction, (iii) prediction of irrigation and nutritional needs, and (iv) monitoring of phenological phases. Therefore, the expected benefits are qualitative and quantitative in economic and environmental terms.

Predictive analytics tools in agriculture are tools that use a variety of statistical methods including data mining, predictive modeling, and machine learning that analyze an array of current and historical agricultural, biological, climate, and hydrological data from various sources to make predictions about future outcomes on the farm.
A good process for developing machine learning prediction models for smart farming, is aimed at discussing the essential stages of developing prediction models most times overlooked by researchers. This research was based on approach proposed In Kenneth.et al.,(2022) of following figure:

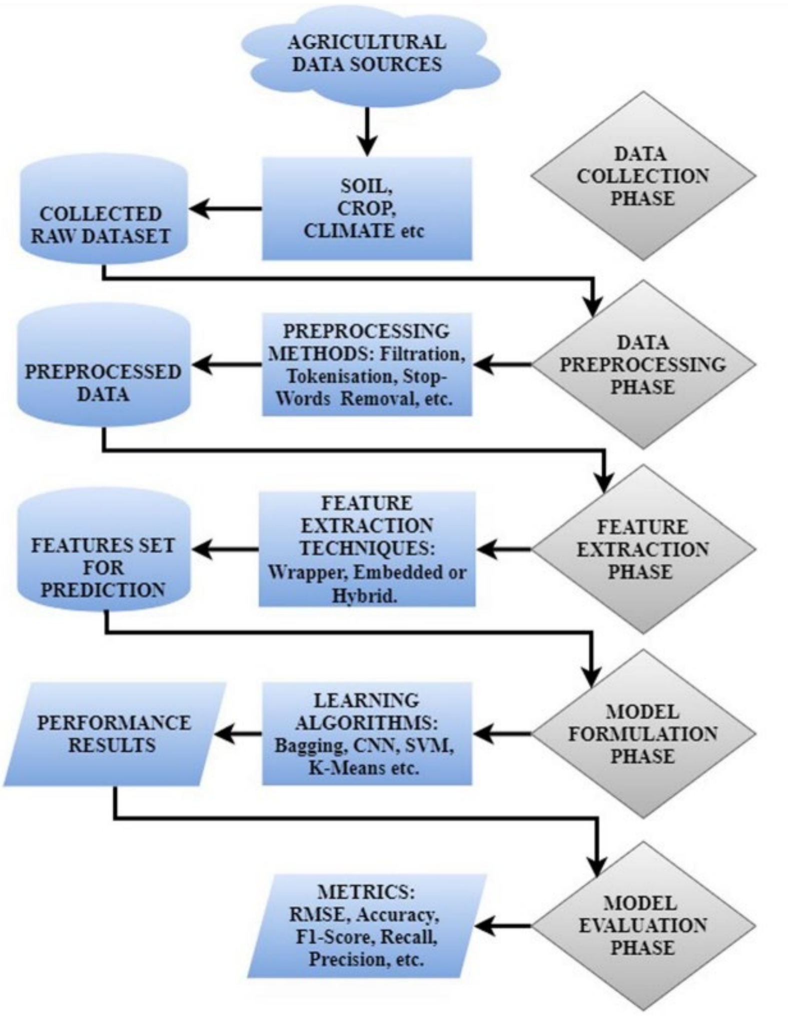


Figure 1 – Scientific approach adopteed

DISSERTATION CONTRIBUTION

Assessing soil condition is a critical aspect of farming as it significantly affects crop growth and yield. Essential soil metrics like nitrogen (N), phosphorous (P), potassium (K) levels, and pH value are pivotal in this assessment. However, obtaining these measurements can be costly and time-consuming, forcing farmers to selectively measure metrics based on their budget constraints.
Selecting the appropriate crop each season is a crucial decision for farmers, as their primary goal is to maximize crop yield. This decision involves considering various factors, including the soil condition. Different crops thrive in different soil conditions, so it is essential to assess the levels of nitrogen, phosphorous, potassium, and the pH value of the soil to ensure optimal growth and yield.

MATERIALS AND MODELS

Smart farming prediction models like other machine learning models depend on a dataset, and it is a fundamental stage in developing smart farming models. Hence, dataset gathering is the initial phase to train or test machine learning models for smart farming predictions. This has been a challenging task faced by several researchers in this area for many reasons including time and the variation of some parameters across geographic locations. Particularly, when modelling soil and or climatic conditions, a dataset from a diferent region cannot provide reliable results for another region due to weather discrepancies (Mesgaran et al. 2017) which afects the soil as well. For instance, the authors (Parikh et al. 2021; Pawar et al. 2021; Rajak et al. 2019), had implemented related works whose data source is publicly accessible. However, in agriculture, data from these public sites should not be used to carry out analysis in a diferent region for want of reliability. This is because, the soil chemical properties obtainable in a region with a diferent weather condition varies (Abhinav et al. 2021). Due to diferent geographic regions and their efect on these factors of farming, data must be collected on a regional basis for the proper implementation of models in smart agriculture predictions. For optimum crop yield analysis, datasets should encompass crops, soil, and fertilizer as well as weather situations. Particularly, useful features such as Nitrogen, Phosphorous, Potassium etc. and pH values of the soil are recommended for analyzing soil fertility (Parikh et al. 2021). Data containing humidity, temperature and rainfall as well as climatic parameters is required for crops and weather-related analysis (Pawar et al. 2021). The dataset should also consist of information about the Global Positioning System (GPS) locations of the test dataset samples (Bouighoulouden 2020)

Soil measures

The present dataset, soil_measures.csv, includes the following columns:

- **N:** Nitrogen content ratio in the soil
- **P:** Phosphorous content ratio in the soil
- **K:** Potassium content ratio in the soil
- **pH:** pH value (or acidity) of the soil
- **label:** label of the crop for that particular soil condition (target variable)

Weather measures

Moreover, weather conditions also play a vital role in crop growth and yield. Therefore, additional weather data, including:

- **temperature:** temperature (in degrees Celsius)
- **humidity:** relative humidity (in %)
- **rainfall:** rainfall (in mm)

Each row in the dataset represents different measurements of the soil from a specific field, and the corresponding "crop" column indicates the best crop for that field based on the measurements.

In this project,**the goal is to develop a multi-class classification model using machine learning to predict the most suitable crop** based on the provided soil measurements. The use of weather data into the model will enhance its predictive power. Additionally, it is essential to address multicollinearity, a phenomenon where two or more features are highly correlated, to ensure the model's accuracy and stability.

This comprehensive approach will ensure the selection of the most suitable crop, taking into account both soil and weather conditions.

The dataset has the following fields:

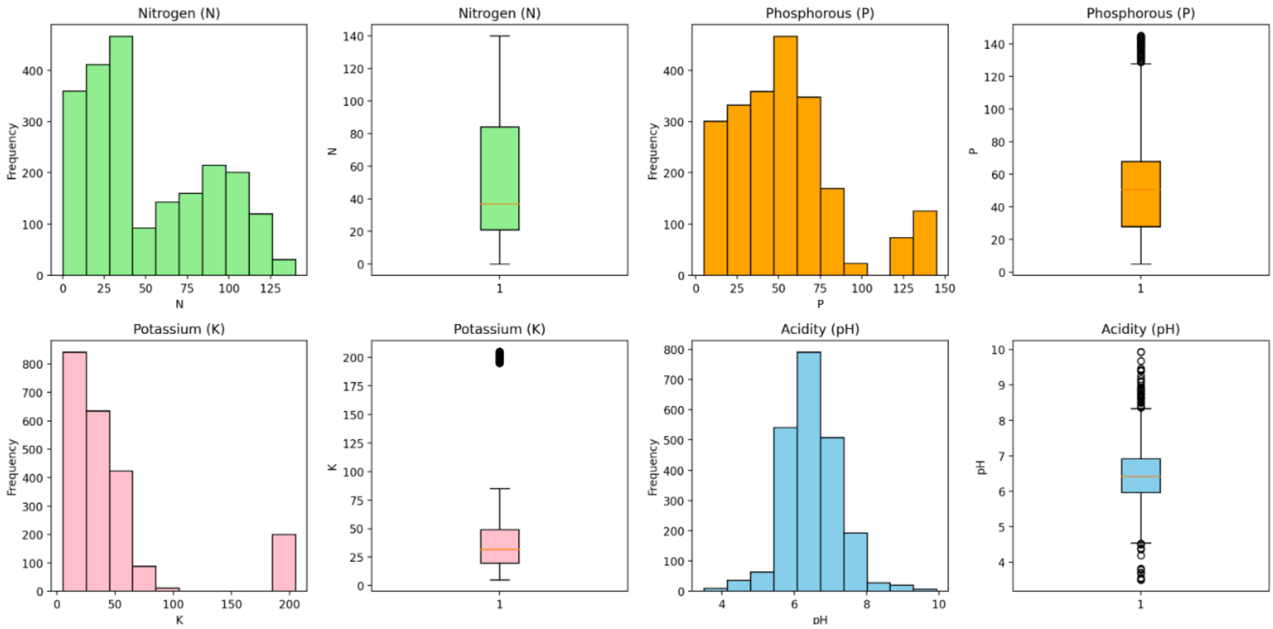
	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

DISCUSSION AND RESULTS

Data analysis shows data does'nt present missing values. Data statistics is done in the following table:

	N	P	K	temperature	humidity	ph	rainfall
count	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000
mean	50.551818	53.362727	48.149091	25.616244	71.481779	6.469480	103.463655
std	36.917334	32.985883	50.647931	5.063749	22.263812	0.773938	54.958389
min	0.000000	5.000000	5.000000	8.825675	14.258040	3.504752	20.211267
25%	21.000000	28.000000	20.000000	22.769375	60.261953	5.971693	64.551686
50%	37.000000	51.000000	32.000000	25.598693	80.473146	6.425045	94.867624
75%	84.250000	68.000000	49.000000	28.561654	89.948771	6.923643	124.267508
max	140.000000	145.000000	205.000000	43.675493	99.981876	9.935091	298.560117

Possiamo analizzare gli istogrammi e i Boxplot di tutti i parametri considerati per analizzare la loro distribuzione. Un grafico a box plot consente di identificare il 25° e 75° percentile meglio degli istogrammi, mentre un istogramma permette di visualizzare la forma complessiva dei dati meglio del box plot.



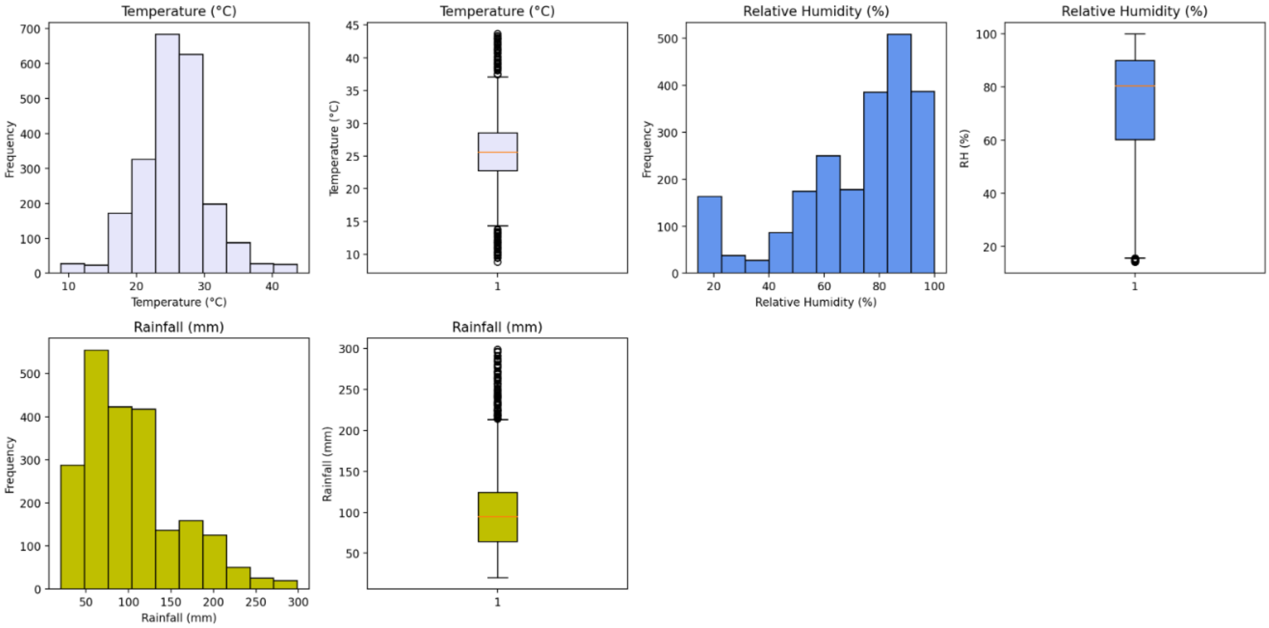


Figure 2 – Histograms and boxplots of dataset

Analizzando i boxplot e i valori della tabella precedente, notiamo che acidità e temperatura sembrano avere una distribuzione normale confermata dai valori di media e mediana molto vicini. Gli outlier influiscono su media, mediana e altri percentili. Poiché in un boxplot i punti estremi vengono evidenziati, i punti di dati da studiare sono di facile identificazione. Da un'analisi successiva, potrebbe emergere che gli outlier corrispondono a errori nella misurazione dei dati o nelle condizioni di misurazione. L'uso di metodi di misurazione manuale con metodi di misurazione tramite sistema IoT può generare un numero significativo di outliers.

Se analizziamo invece la Distribution of crop types it is uniform like showed in following figure:

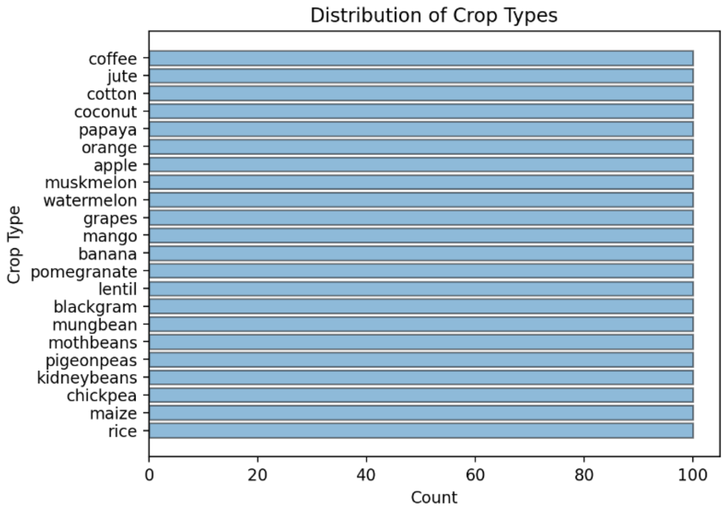


Figure 2 – Distribution of crop types

The data set suggests the use of a regressive model. So we use logistic regression and F1 Score metric to evaluate to Predict the crop using each feature, Following achieved results:

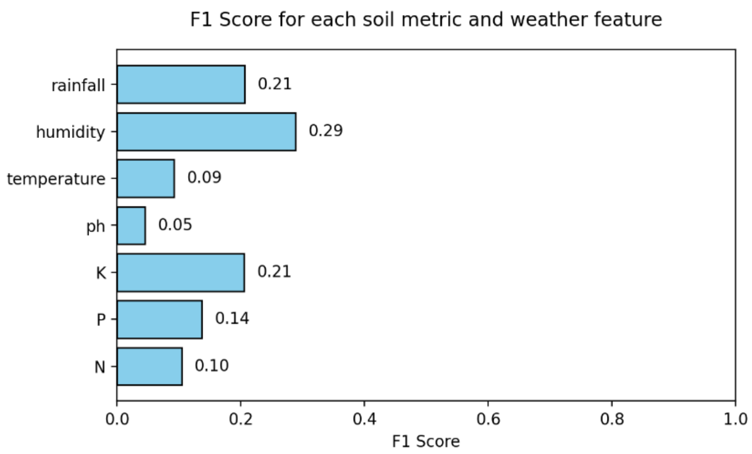


Figure 3 – Results using LR and F1-Score

The F1 scores for each feature are quite low, which indicates that each feature individually does not do a great job at predicting the crop type. The **highest F1 score** is for the **relative humidity** feature (0.29), which suggests that it is the most informative feature among those considered, although it still does not provide a very high level of predictive accuracy. Among the soil meatric feature, **Potassium (K)** (0.21) is the soil element that is better at predicting the crop type. We will test a combined approach to achive better results.

First we Estimate features correlation using the following heatmap:

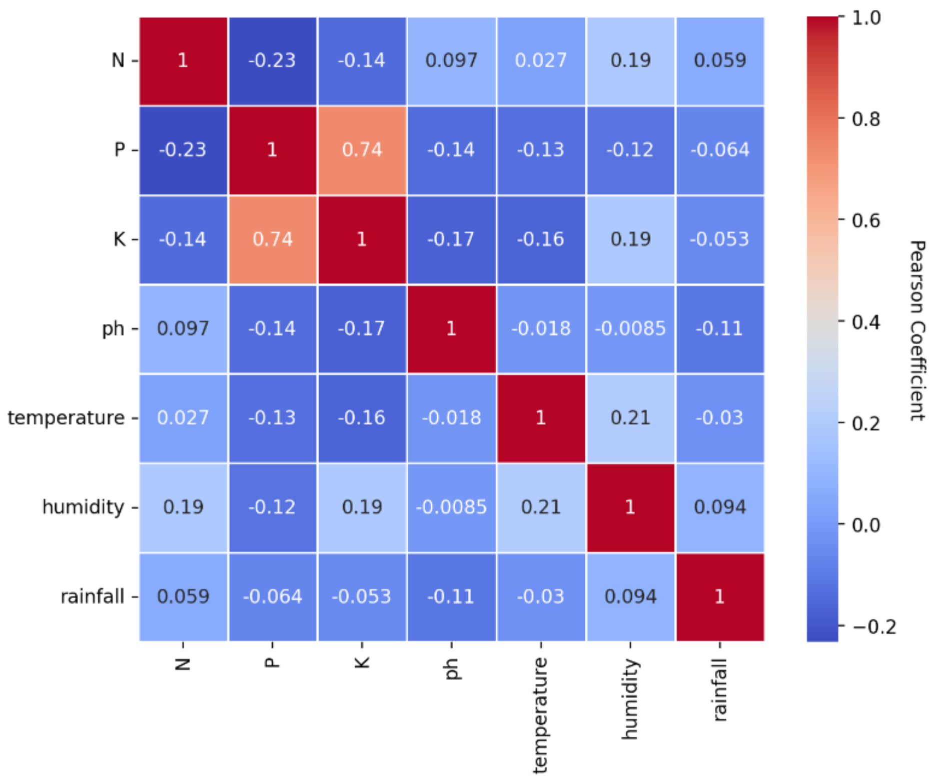


Figure 4– Soil metrics and weather data correlation,

The output shows the correlation between each pair of soil metric features. A high positive or negative value indicates a strong correlation. Ideally, we want to select features that are **NOT** highly correlated with each other to avoid multicollinearity.

In this case, the highest correlation is between P and K, with a correlation coefficient of 0.74. **This is relatively high, but not extreme**. The other correlations are all quite low.

Given this, we could reasonably use all four features (N, P, K, ph, humidity, temperature and rainfall) to build the final model, as none of them are extremely highly correlated. However, to be more conservative, we can exclude either P or K from the final model to reduce the risk of multicollinearity. This would be a judgement call based on how conservative we want to be.

If we decide to exclude one of the features, we should also consider the importance of each feature for predicting the crop type. If one of the features (P or K) is known to be more important for predicting the crop type, we should keep that feature in the model and exclude the other one.

Conclusion: Given the previous examination on **F1 score** conducted on individual soil metrics and weather feature, it might be beneficial to build a model using all the features, despite the correlation between P and K as this might lead to a better overall predictive performance

In the final part of this project, we are going first develop a model based on N, K, ph, humidity, temperature and rainfall features (P not included) (PART 1) and in the last part of the project, we are going to check with different combinations of features and choose the one that gives the best performance on a validation set.

Based on the previous analysis, the feature K had the **highest F1 score among the soil metric feature** (Relative humidity yielded the highest F1 score among all features) when used individually, and none of the features were extremely highly correlated. Therefore, it seems reasonable to build a final model using N, K and pH soil metric features combined with weather metric features.

We achieved the Final Model F1 Score: 0.95

That is a **much better F1 score than any of the individual features achieved**.

This indicates that, while the individual features did not provide a high level of predictive accuracy, combining them in a single model did lead to better predictive performance.

In order to check the obtained F1 score is the highest possible, let's create a list of all possible combinations of features and then loop through each combination, building a model and calculating the F1 score for each one. **The objective of this part is to examine whether there is a better combination of soil metric feature lead to a better performance.**

So doing we discover that The combination of features **Nitrogen (N), Phosphorus (P), Potassium (P) and weather metrics (T, RH and rainfall) gave the highest F1 score of 0.97**, which means that this combination of features gave the most accurate model according to the F1 score metric.

N° item	Combination of parameters	F1-score
1	N'	0.10
2	P'	0.14
3	K'	0.21
4	ph'	0.05
5	temperature'	0.09

6	humidity'	0.29
7	rainfall'	0.21
8	N', 'P'	0.31
9	N', 'K'	0.47
10	N', 'ph'	0.20
11	N', 'temperature'	0.25
12	['N', 'humidity'	0.54
13	N', 'rainfall'	0.50
14	P', 'K'	0.45
15	P', 'ph'	0.25
16	['P', 'temperature'	0.36
17	P', 'humidity'	0.63
18	P', 'rainfall'	0.60
19	K', 'ph'	0.40
20	K', 'temperature'	0.43
21	K', 'humidity'	0.63
22	K', 'rainfall'	0.61
23	ph', 'temperature'	0.23
24	ph', 'humidity'	0.38
25	ph', 'rainfall'	0.35
26	Features: ['temperature', 'humidity'	0.42
27	temperature', 'rainfall'], F1 Score: 0.42	0.42
28	humidity', 'rainfall'	0.65
29	N', 'P', 'K'	0.59
30	N', 'P', 'ph'	0.43
31	N', 'P', 'temperature'	0.56
32	['N', 'P', 'humidity'	0.75
33	N', 'P', 'rainfall'	0.72
34	N', 'K', 'ph'	0.56
35	N', 'K', 'temperature'	0.59
36	N', 'K', 'humidity'	0.77
37	N', 'K', 'rainfall'	0.76
38	N', 'ph', 'temperature'	0.37
39	N', 'ph', 'humidity'	0.56
40	N', 'ph', 'rainfall'	0.50
41	N', 'temperature', 'humidity'	0.59
42	N', 'temperature', 'rainfall'	0.54
43	N', 'humidity', 'rainfall'	0.73
44	P', 'K', 'ph'	0.54
45	P', 'K', 'temperature'	0.64
46	P', 'K', 'humidity'	0.82
47	P', 'K', 'rainfall'	0.75
48	P', 'ph', 'temperature'	0.49
49	P', 'ph', 'humidity'	0.63
50	P', 'ph', 'rainfall'	0.60
51	P', 'temperature', 'humidity'	0.71
52	P', 'temperature', 'rainfall'	0.69

53	P', 'humidity', 'rainfall'	0.85
54	K', 'ph', 'temperature'	0.54
55	K', 'ph', 'humidity'	0.68
56	K', 'ph', 'rainfall'	0.66
57	K', 'temperature', 'humidity'	0.75
58	K', 'temperature', 'rainfall'	0.72
59	['K', 'humidity', 'rainfall'	0.85
60	ph', 'temperature', 'humidity'	0.52
61	ph', 'temperature', 'rainfall'	0.51
62	ph', 'humidity', 'rainfall'	0.69
63	temperature', 'humidity', 'rainfall'	0.67
64	N', 'P', 'K', 'ph'	0.65
65	N', 'P', 'K', 'temperature'	0.73
66	N', 'P', 'K', 'humidity'	0.91
67	N', 'P', 'K', 'rainfall'	0.85
68	N', 'P', 'ph', 'temperature'	0.62
69	N', 'P', 'ph', 'humidity'	0.77
70	N', 'P', 'ph', 'rainfall'	0.74
71	N', 'P', 'temperature', 'humidity'	0.81
72	N', 'P', 'temperature', 'rainfall'	0.80
73	N', 'P', 'humidity', 'rainfall'	0.88
74	N', 'K', 'ph', 'temperature'	0.65
75	N', 'K', 'ph', 'humidity'	0.79
76	N', 'K', 'ph', 'rainfall'	0.77
77	N', 'K', 'temperature', 'humidity'	0.83
78	N', 'K', 'temperature', 'rainfall'	0.84
79	N', 'K', 'humidity', 'rainfall'	0.91
80	N', 'ph', 'temperature', 'humidity'	0.67
81	N', 'ph', 'temperature', 'rainfall'	0.60
82	N', 'ph', 'humidity', 'rainfall'	0.78
83	'N', 'temperature', 'humidity', 'rainfall'	0.76
84	P', 'K', 'ph', 'temperature'	0.69
85	P', 'K', 'ph', 'humidity'	0.79
86	P', 'K', 'ph', 'rainfall'	0.77
87	P', 'K', 'temperature', 'humidity'	0.89
88	P', 'K', 'temperature', 'rainfall'	0.84
89	P', 'K', 'humidity', 'rainfall'	0.90
90	P', 'ph', 'temperature', 'humidity'	0.71
91	P', 'ph', 'temperature', 'rainfall'	0.75
92	P', 'ph', 'humidity', 'rainfall'	0.84
93	P', 'temperature', 'humidity', 'rainfall'	0.88
94	K', 'ph', 'temperature', 'humidity'	0.78
95	K', 'ph', 'temperature', 'rainfall'	0.75
96	K', 'ph', 'humidity', 'rainfall'	0.84
97	K', 'temperature', 'humidity', 'rainfall'	0.89
98	ph', 'temperature', 'humidity', 'rainfall'	0.74
99	N', 'P', 'K', 'ph', 'temperature'	0.78

100	N', 'P', 'K', 'ph', 'humidity'	0.89
101	N', 'P', 'K', 'ph', 'rainfall'	0.64
102	N', 'P', 'K', 'temperature', 'humidity'	0.92
103	N', 'P', 'K', 'temperature', 'rainfall'	0.90
104	N', 'P', 'K', 'humidity', 'rainfall'	0.94
105	N', 'P', 'ph', 'temperature', 'humidity'	0.84
106	N', 'P', 'ph', 'temperature', 'rainfall'	0.80
107	N', 'P', 'ph', 'humidity', 'rainfall'	0.90
108	N', 'P', 'temperature', 'humidity', 'rainfall'	0.91
109	N', 'K', 'ph', 'temperature', 'humidity'	0.84
110	N', 'K', 'ph', 'temperature', 'rainfall'	0.83
111	N', 'K', 'ph', 'humidity', 'rainfall'	0.92
112	N', 'K', 'temperature', 'humidity', 'rainfall'	0.94
113	N', 'ph', 'temperature', 'humidity', 'rainfall'	0.82
114	P', 'K', 'ph', 'temperature', 'humidity'	0.90
115	P', 'K', 'ph', 'temperature', 'rainfall'	0.85
116	P', 'K', 'ph', 'humidity', 'rainfall'	0.90
117	P', 'K', 'temperature', 'humidity', 'rainfall'	0.96
118	P', 'ph', 'temperature', 'humidity', 'rainfall'	0.91
119	K', 'ph', 'temperature', 'humidity', 'rainfall'	0.91
120	N', 'P', 'K', 'ph', 'temperature', 'humidity'	0.92
121	N', 'P', 'K', 'ph', 'temperature', 'rainfall'	0.89
122	N', 'P', 'K', 'ph', 'humidity', 'rainfall'	0.94
123	**N', 'P', 'K', 'temperature', 'humidity', 'rainfall'**	**0.97**
124	N', 'P', 'ph', 'temperature', 'humidity', 'rainfall'	0.91
125	N', 'K', 'ph', 'temperature', 'humidity', 'rainfall'	0.95
126	P', 'K', 'ph', 'temperature', 'humidity', 'rainfall'	0.95
127	N', 'P', 'K', 'ph', 'temperature', 'humidity', 'rainfall'	0.96

On the other hand, the single *acidity (pH) feature* (pH) gave the lowest F1 score of 0.05, which means that using only the acidity as a feature results in a very inaccurate model according to the F1 score metric.

CONCLUSIONS AND FUTURE WORKS

In this project, we developed a multi-class classification model using machine learning to predict the most suitable crop based on the provided soil measurements. Adding weather data into the model enhanced its predictive power. Data Analysis showed like the the best data combination for our prediction is given by the combination of features Nitrogen (N), Phosphorus (P), Potassium (P) and weather metrics (T, pH and rainfall) gave the highest F1 score of 0.97.

Ph use doesn't provide useful information for goal of our analysis.

This comprehensive approach ensures the selection of the most suitable crop, taking into account both soil and weather conditions.

For the project, a database was considered resulting from the combination of different projects in different geographical areas of the world. The data set considered allowed us to verify the functional correctness of the application. As future work, the predictive capacity of the developed method will be tested on a dataset of agricultural products and surveys of the Italian agricultural system. Italy stands out not only for its multiple historical influences but also for the fortunate climatic variety and the rich diversity of its soils. From north to south, countless plants have found a way to evolve and adapt to very varied environmental conditions. This natural abundance has favored incredible biodiversity, making Italian soil a favorite habitat for various native species and varieties.

BIBLIOGRAPHY

[1] Virk AL, Noor MA, Fiaz S, Hussain S, Hussain HA, Rehman M, Ahsan M, Ma W (2020) Smart farming: an overview. In Smart Village Technology, pp 191–201

[2] Mesgaran MB, Madani K, Hashemi H, Azadi P (2017) Iran's land suitability for agriculture. Sci Rep 7(1):1–12. https://doi.org/10.1038/s41598-017-08066-y

[3] Kurşun G, Dengiz O (2020) Assessment of land suitability for the production of major crops in ayrancı district of Karaman province located at arid terrestrial ecosystem. Eurasian J Soil Sci 9(1):24–33. https://doi.org/10.18393/ejss.642200

[4] Feyisa GL, Palao LK, Nelson A, Gumma MK, Paliwal A, Win KT, Nge KH, Johnson DE (2020) Characterizing and mapping cropping patterns in a complex agro-ecosystem: an iterative participatory mapping procedure using machine learning algorithms and MODIS vegetation indices. Comput Electron Agric 175:105595. https://doi.org/10.1016/j.compag.2020.105595

[5] Adamides G (2020) A review of climate-smart agriculture applications in Cyprus. Atmosphere 11(9):1–15. https://doi.org/10.3390/ATMOS11090898

[6] Fenu G. and Maridina Malloci F., Forecasting plant and crop disease: An explorative study on current algorithms. Big Data and Cognitive Computing, 5(1):2, 2021.

[7] Fenu G. and Maridina Malloci F., Artificial intelligence technique in crop disease forecasting: a case study on potato late blight prediction. In International Conference on Intelligent Decision Technologies, pages 79–89. Springer, 2020.

[8] Fenu G. and Maridina Malloci F., An application of machine learning technique in forecasting crop disease. In Proceedings of the 2019 3rd International Conference on Big Data Research, pages 76–82, 2019. doi: https://doi.org/10.1145/3372454.3372474.

[9] Fenu G. and Maridina Malloci F., Using multi-output learning to diagnose plant disease and stress severity. Complexity, 2021(6663442): 18, 2021. doi: https://doi.org/10.1155/2020/6663442.

[10] Fenu G. and Maridina Malloci F., Lands dss: A decision support system for forecasting crop disease in southern sardinia. International Journal of Decision Support System Technology (IJDSST), 13(1):21–33, 2021. doi: https://doi.org/10.4018/IJDSST.2021010102.

[11] Parikh DP, Jain J, Gupta T, Dabhade RH (2021) Machine learning based crop recommendation system. Int J Adv Res Sci Commun Technol 6(1):891–897. https://doi.org/10.48175/ijarsct-1509 Pawar S, Dere S,

[12] Akangire A, Kamble H, Shrawne S (2021) Smart farming using machine learning. Smart Comput. https://doi.org/10.1201/9781003167488-11

[13] Rajak RK, Pawar A, Pendke M, Shinde P, Rathod S, Devare A (2019) Crop recommendation system to maximize crop yield in ramtek region using machine learning. Int Res J Eng Technol (IRJET) 4(12):485–489. https://doi.org/10.32628/ijsrst196172

[14] Abhinav S, Jain A, Gupta P, Chowdary V (2021) Machine learning applications for precision agriculture: a comprehensive review. IEEE Access 9:4843–4873. https://doi.org/10.1109/ACCESS.2020.3048415

[15]Bouighoulouden A (2020) Crop yield prediction using K-means clustering. School of Science and Engineering—Al Akhawayn University, pp 1–35
[16]Kenneth Kwaghtyo D., Ifeanyi Eke C., Smart farming prediction models for precision agriculture:a comprehensive survey, Artifcial Intelligence Review (2023) 56:5729–5772, https://doi.org/10.1007/s10462-022-10266-6