

Mineralogical Characterization of Subsurface Soils Using Machine Learning: Application of Support Vector Machines

Ahmed Babacar Sarr*, Mapathe Ndiaye, Sabou Sarr, Ndiouga Camara

Soil Mechanics and Modelling Laboratory, Faculty of Engineering Sciences, Iba Der Thiam University of Thies, Thies, Senegal
Email: *ababacar.sarr@univ-thies.sn

How to cite this paper: Sarr, A. B., Ndiaye, M., Sarr, S., & Camara, N. (2025). Mineralogical Characterization of Subsurface Soils Using Machine Learning: Application of Support Vector Machines. *Journal of Geoscience and Environment Protection*, 13, 121-134. <https://doi.org/10.4236/gep.2025.1310009>

Received: September 19, 2025

Accepted: October 25, 2025

Published: October 28, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Support vector machines are recognized as a powerful tool for supervised analysis and classification in different fields, particularly geophysics. In summary, SVMs are binary classifiers. Thus, for the multiclass study, the problem is divided into a series of binary classifications. At the end, all the results obtained from these binary classifications are combined into a one-to-one or one-to-all comparison. In this article, the strategy consists of classifying soils using their chemical composition as characteristics specific to a soil. The prediction consists of 5 classes, which are: White Clay, Red Clay, Black Clay, Granite, and Sand. The dataset is composed of basic oxides, which contribute to increasing soil salinity, acidic oxides such as silica, which do not influence soil fertility, and amphoteric oxides. These data are divided into training, test, and validation data. The one-vs-all strategy was used. The results obtained showed the strength of the one-vs-all associated with SVM on all classification metrics. The selection of the kernel as well as hyperparameters also played an important role in the prediction score. From the results obtained, the one-vs-all associated with SVM can be used for classification problems. For further studies, geo-location can be introduced to have knowledge of the evolution according to the different sectors of the same region.

Keywords

Classification, Support Vector Machine, Kernel Function, Soil, Raw Materials

1. Introduction

Raw material exploration is a complex estimation process (Wittenberg & de Oliveira, 2023; Koutsopoulou et al., 2021) whose goal is to find a material of sufficient qual-

ity and quantity, in the estimation (Nie et al., 2024; Wang et al., 2018) for exploitation. For this, various processes are implemented to achieve these objectives. Among these data, geographical and hydraulic (Camara et al., 2024), hydrogeological, and geophysical (Ndiaye et al., 2012) are a set of data that are collected, analyzed, and interpreted. This allows us to construct a raw material distribution map. Raw material mapping is a task based on a set of criteria, particularly volume estimation models from core drilling. There are several raw material mapping models. There are data-based models and knowledge-based models (Goldstein et al., 2025).

Data-driven models use previous data to develop the spatio-temporal variability between the different characteristics of a terrain. This relationship makes it possible to build a prediction model in the face of new data from said terrain. Empirical models do not actually build a decision model. This is the case for algorithms such as: k-Nearest Neighbors (KNN) (Halder et al., 2024), Local Weighted Regression (LWR), Memory-based learning or Case-Based learning (CBR), as well as neural networks (Miller, Kaminsky, & Rana, 1995). These algorithms, called machine learning, are divided into two categories: supervised and unsupervised. The existence of data from a site composed of quantitative data that describes data, whether qualitative (classification) or quantitative (regression), allows for the construction of supervised models. As soon as our data is unlabeled or unstructured, the algorithm learns to group the data on its own. This is called unsupervised learning.

We chose the SVM algorithm with RBF kernel because the data show non-linear separation between classes. This model offers good generalization capacity while being robust to local variations.

Each field of engineering approaches classification according to these objectives. Geophysicists use them to characterize the physical properties of soils. Chemical composition is an interdisciplinary point of convergence. Machine learning models allow for a globalist approach while integrating the specificities of each discipline in advance.

Support vector machines are classification algorithms introduced by Vapnik in 1995 (Zuo & Carranza, 2010) and are now widely used in geological classification problems.

2. Materials

2.1. Data

The data come from ten of the fourteen regions of Senegal and The Gambia. Senegal, located in West Africa, and the Gambia at its center, as shown in **Figure 1**. These countries share the same geological heritage. Indeed, Gambia is an enclave in the center of Senegal, where geological continuity ensures the presence of the same basin. Soil samples are taken and tested in the lab to determine their chemical element content.

The samples taken from different regions depend on the diversity of colors and

especially the types of materials, most often clayey, found in these different regions.

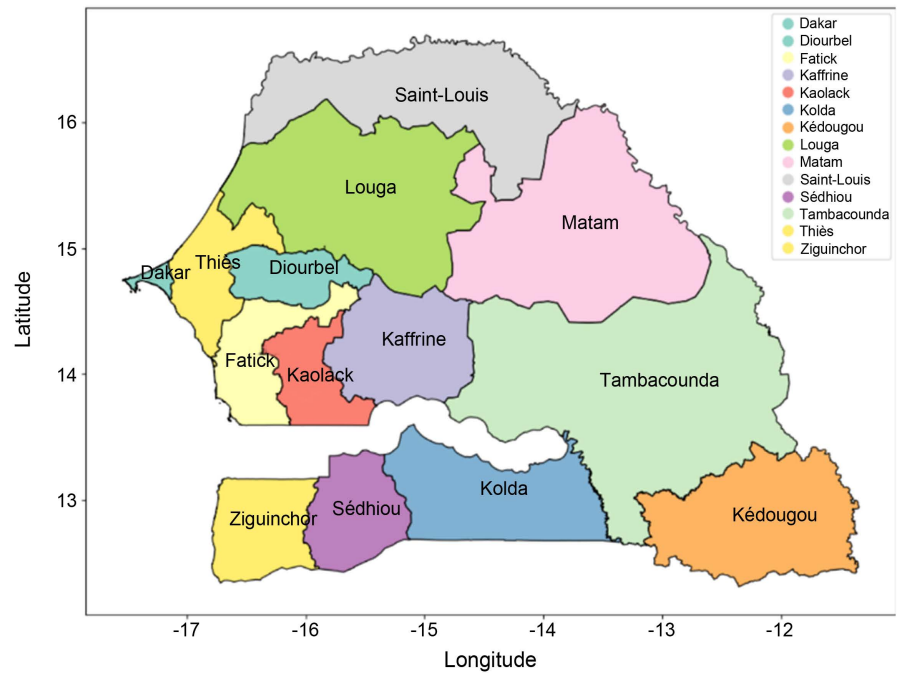


Figure 1. Administrative division of Senegal.

Hence, this great variability of samples depends on the region and needs. **Figure 2** shows the quantities sampled by region.

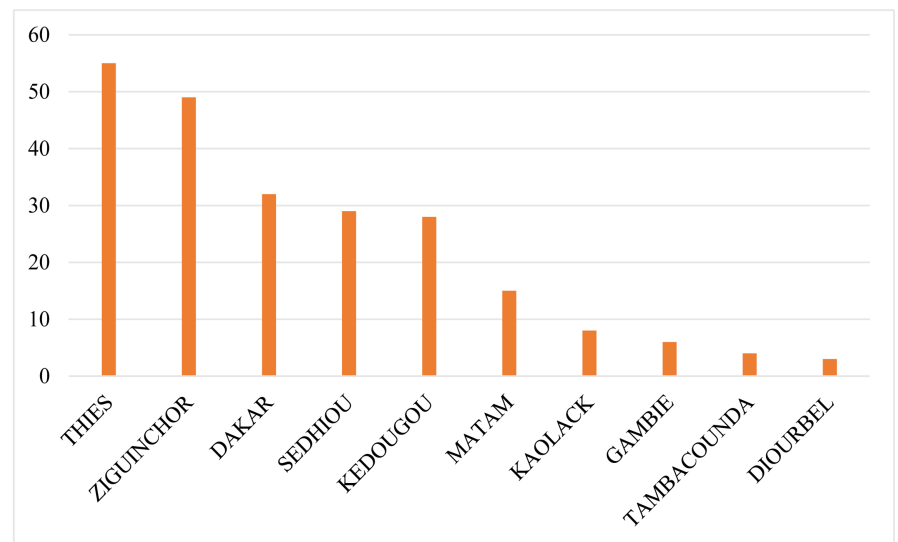


Figure 2. Geographical distribution of samples.

Thus, depending on the soil type, we note the considerable contribution of data on clays, particularly black, to other types of soils. We show in **Figure 3** the quantities of raw material in our samples according to the soil type.

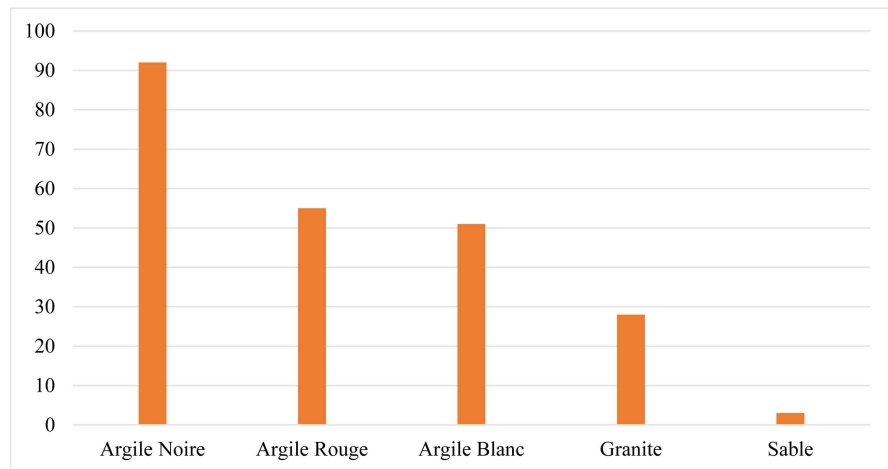


Figure 3. Occurrence of samples.

This difference can be explained by the fact that the large quantity of data collected, especially in the south, is of lake origin. This environment is strongly marked not only by a dancing hydrographic network. But also, by a strong vegetation decomposition of living organisms, making these lands very rich in organic matter.

This distribution, **Table 1**, broadly represents the data used to carry out the studies in this article and is distributed as follows:

- The samples come from two regions (Dakar and Thies). These regions represent approximately 4% of the country's total area.
- 86% represent clays which cover all regions of Senegal except the outcrops noted in eastern Senegal.

In this table:

- Occ = Occurrence
- Per = Percentage

Table 1. Representativeness of samples.

REGION	Occ	Per (%)	
THIES	55	24.02	%
ZIGUINCHOR	49	21.4	%
DAKAR	32	13.97	%
SEDHIOU	29	12.66	%
KEDOUGOU	28	12.23	%
MATAM	15	6.55	%
KAOLACK	8	3.49	%
GAMBIE	6	2.62	%

Continued

TAMBACOUNDA	4	1.75	%
DIOURBEL	3	1.31	%
Type of Soil			
Black Clay	92	40.17	%
Red Clay	55	24.02	%
White Clay	51	22.27	%
Granite	28	12.23	%
Sand	3	1.31	%

This data allows us to see the different ranges of variation by soil type and their mineral content. Sands have a high silica content and low alkaline and ferric mineral content, unlike clays. The geographical diversity of soils in Senegal highlights their mineralogical variability depending on their nature. It shows that sandy soils, which are widespread in certain regions, are richer in silica, while clay soils, which are more alkaline, contain about half as much basic oxides. This chemical distinction between soil types is essential for understanding their mechanical, physical, and biochemical behavior.

The data contains the content of silica (SiO_2), alumina (Al_2O_3), titanium monoxide (TiO), quicklime (CaO), sodium oxide (Na_2O), hematite (Fe_2O_3), magnesia (MgO), and potassium oxide (K_2O). Using this data, we calculated: the silica saturation index (IndSatSilice), soil alkalinity (indAlcalin), the proportion of iron relative to other oxides (IndFer), basic potential (SomOxyBasik), and total composition (Compotot). An initial classification is based on regions. The second classification is based on soil type, without considering the region of origin.

Senegal is a flat country whose altitude rarely exceeds 40 m (Maignien, 1965) in the sedimentary basin. This flat terrain sometimes has certain consequences, such as flooding in combination with land use. This basin covers a large part of the country, unlike the Precambrian basement. It is an outcrop of hard rock in eastern Senegal, but especially in its eastern part. In this area, hills rise to a height of around 400m. The sedimentary basin provides a wealth of knowledge about the pedology of Senegal. It extends from Mauritania to Guinea-Bissau.

2.2. Acquisition

As part of this study, data were collected from field samples taken in several targeted regions. Sampling sites were chosen along riverbanks, areas that are particularly representative of alluvial deposits and local hydrogeological dynamics. The collection operations were carried out using a pickup truck, which allowed access to rural areas that are difficult to reach. The tools used included a GPS, a shovel, a pickaxe, an auger, and bags for packaging the materials. Each sample was collected according to a standardized protocol to ensure the representativeness and

traceability of the samples. They were collected at different depths, taking into account their spatial and temporal variability.

2.3. Preprocessing

After collection, the samples underwent a series of laboratory tests to determine their chemical composition (calcium, titanium, silica, etc.). These results were compiled in Excel, allowing for structured organization of the data. These files were then processed using Python scripts, including data cleaning for management purposes and anomaly detection. As part of this study, min-max normalization was used to control deviations, standardize variable scales, and improve algorithm convergence. Where applicable, exploratory statistical analyses reveal the influences and dominance of variables.

3. Methodology

3.1. The Support Vector Machine Method

Support vector machines (SVMs) are supervised learning algorithms used for classification, regression, and anomaly detection problems. They are increasingly used in geoscience, particularly in geophysics, for the classification of geological formations. Determining the estimator parameters is a convex optimization problem. Any local solution to such a problem will be globally optimal (Agrawal, Barratt, & Boyd, 2021). It is an estimator that uses labels. These labels are characterized by a set of features.

Each feature is a data vector used as a training set to construct a hyperplane separating the classes of a dataset. SVM is a binary classifier. For complex cases, SVM transforms the data into a higher-dimensional space to make it linearly separable. To understand the basics of SVM (Kim et al., 2003), we will illustrate the approach using a binary classification problem. Suppose we have a training set $\{(x_i, y_i)\}_{i=1,2}$ in a two-dimensional plane, where the x_i are descriptors of a class y_i ; $y_i \in \{-1, +1\}$. If the set is linearly separable, then there exists a group of linear separators called separation hyperplanes. This is valid for N -dimensions or N descriptors (Kovacevic et al., 2009). Thus, these hyperplanes are modeled by the function:

$\omega^T x_i + b = 0$ where x is an input vector, w is a weight vector, and b is the threshold. Overall, we can write:

$$\omega^T x_i + b \geq 0 \text{ pour } y_i = +1$$

$$\omega^T x_i + b < 0 \text{ pour } y_i = -1$$

The space between the separating hyperplane and the support points is called the geometric margin. The objective of SVMs is to maximize this margin, hence their name “wide margin separator.” If w_0 and b_0 are the optimal weights and bias, respectively, then the optimal hyperplane is defined by:

$$w_0^T x + b_0 = 0$$

The relationship between x_i and $f(x_i)$ is defined by:

$$f(x) = w^T x + b$$

If we call $f(x)$ the functional separation margin, then the distance between the decision function and the support points can be written as follows:

$$r = \frac{|f(x)|}{\|w_0\|}$$

This relationship assumes that all training points are classified on or behind the hyperplanes. The distance between the two support vectors is defined by d :

$$d = 2r = \frac{2}{\|w\|}$$

The objective function can be represented by:

$$\max \frac{2}{\|w\|} \quad y_i (w^T x_i + b) \gg 1, i = 1, \dots, n$$

The solution to such an optimization problem is given by the Lagrange function.

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum \alpha_i [y_i (w^T x_i + b)]$$

where α_i are Lagrange multipliers. This function is minimized by the values of w and b and maximized by the positivity of the multiplier coefficients. The multiplier coefficients are determined by the following optimization function:

$$\alpha_i - \frac{1}{2} \alpha_i \alpha_j y_i y_j (x_i x_j) \text{ avec } \alpha_i \gg 0, i = 1, \dots, l \text{ et } \alpha_i y_i = 0$$

It relies on a kernel to determine a separating hyperplane. This hyperplane can be linear if no mixing of data is observed. It can be flexible if a certain degree of tortuosity is observed in the separating hyperplane. It will be associated with another data separation model called “One versus All”, of which another variable, “One versus One”, exists.

The “one versus all” approach results in one classifier per class, for a total of five classifiers. For class i , it will consider the labels of that class as positive and the others as negative.

Whereas “one versus one” forms a separate classifier for each pair of different labels. This leads to $5(5-1) / 2 = 10$, or 10 classifiers. However, despite the high number of models, this method is much less sensitive to data set imbalance issues and is much more computationally expensive.

3.2. Kernel Selection

In soil studies, data cannot generally be separated in a linear fashion, as decision boundaries are often complex and non-trivial. Indeed, the characteristic properties of a soil type may appear locally in the form of inclusions, lenses or dome-shaped structures within the same geological layer, making classification more difficult. This spatial heterogeneity requires the use of models capable of capturing

the non-linear relationships between variables.

Thus, these learning data are projected into an N-dimensional space corresponding to the size of the descriptors (Liu & Xu, 2013). Depending on the flexibility of the kernel and the hardware and software resources chosen, they can be separated as best as possible. The following kernels are used:

- Linear kernel:

$$K(x_i, x_j) = x_i^T x_j$$

- Polynomial kernel:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$$

- Radial Basis Function (RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$$

- Sigmoid kernel:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$$

The values γ , r , and d are parameters of the kernel in question. In geophysics, the RBF kernel is the most commonly used (Abedi, Norouzi, & Bahroudi, 2011):

- It projects data into a higher dimension.
- It has fewer hyperparameters than the polynomial kernel.
- It has lower spatial and temporal complexity than the polynomial kernel.

3.3. Parameter Selection

SVMs have proven successful in classification problems. This is largely due to parameter selection (Agrawal, Barratt, & Boyd, 2021). For a linear model, defining the regularization parameter C is sufficient for good classification. Unfortunately, data are not always linearly separable, particularly in structural geology, where layers are not horizontal. The parameters d for a polynomial kernel or γ for the RBF are added to the regularization parameter. In order to find the most optimal values for these parameters, the grid search method was introduced. It allows us to find the best combination of parameters for our problem.

3.4. Data Partitioning

The data was divided according to the following percentages in both cases: training: 80%, testing: 15%, and 5% set aside to verify the model (Agrawal, Barratt, & Boyd, 2021; Kovacevic et al., 2009). In the first case, raw values are used, and in the second, the data is normalized. To improve the performance of our different models, we varied the search criteria for training. The following performance metrics: overall precision or accuracy, true positive rate, recall, and F-1 score are used to evaluate the performance of our models (Ruano et al., 2013).

4. Results and Discussion

4.1. Analysis and Interpretation of Results

The results below were recorded in this study. **Table 2** shows the mean and stand-

ard deviation of the numerical variables in our data. There is a significant difference in the ranges of variation of the different vectors in our dataset. This led to normalization in the second part of the study. Based on these variables, new variables were created. They are indicators of overall chemical composition, soil fertility, and the processes of alteration and mineral formation in the soil. These indicators are shown in **Table 2**.

Table 2. Variable statistics.

	SiO ₂		Al ₂ O ₃		TiO		CaO		MgO		Fe ₂ O ₃		K ₂ O		Na ₂ O	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Argile Blanche	72.191	18.254	12.316	7.144	0.947	0.369	2.309	7.716	0.471	1.297	5.908	7.136	0.266	0.402	0.681	0.679
Argile Noire	59.816	8.866	22.293	5.104	1.291	0.383	0.290	0.113	0.253	0.066	2.046	1.068	0.325	0.100	0.611	0.142
Argile Rouge	57.580	27.270	10.805	5.704	0.707	0.317	10.155	17.971	1.710	2.463	7.521	8.182	0.480	0.475	0.735	0.482
Granite	69.837	11.449	17.564	5.203	0.543	0.699	0.391	0.371	0.298	0.186	2.982	6.637	2.823	2.125	2.149	1.484
Sable	78.743	0.612	9.900	0.128	0.507	0.061	2.047	0.225	0.463	0.035	2.803	0.035	0.300	0.035	0.470	0.056

This table provides additional information on the chemical behavior of the soils studied. Like the previous table, it details the average concentrations and standard deviations of several oxides, revealing contrasts between clays, granite, and sand in terms of composition. This quantitative approach makes it possible to identify overall trends and mineralogical specificities unique to the soil. **Table 3** provides a summary in the form of chemical indices (IndSatSilice, IndAlcalin, IndFer, etc.). These indices facilitate the comparison of soils according to their degree of silica saturation, their richness in alkaline oxides or iron, and their total composition. This table presents data that has been normalized according to the extreme values observed, i.e., based on the minimum and maximum values measured.

Table 3. Chemical profile of soils.

Soil Types	IndSatSilice		IndAlcalin		IndFer		CompoTot		SomOxyBasik	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Argile Blanche	0.8	0.14	2.25	2.34	0.07	0.08	95.09	7.2	3.73	8.98
Argile Noire	0.71	0.07	1.9	0.84	0.02	0.01	86.93	4.48	1.48	0.32
Argile Rouge	0.76	0.12	1.04	0.79	0.09	0.09	89.69	13.83	13.08	19.39
Granite	0.77	0.11	8.05	7.63	0.03	0.07	96.59	4.39	5.66	3.73
Sable	0.86	0	0.3	0.02	0.03	0	95.23	0.41	3.28	0.32

Figure 4 shows the variability around the mean by region according to chemical component, in dark lines. Each line shows the variability of measurements in the same region.

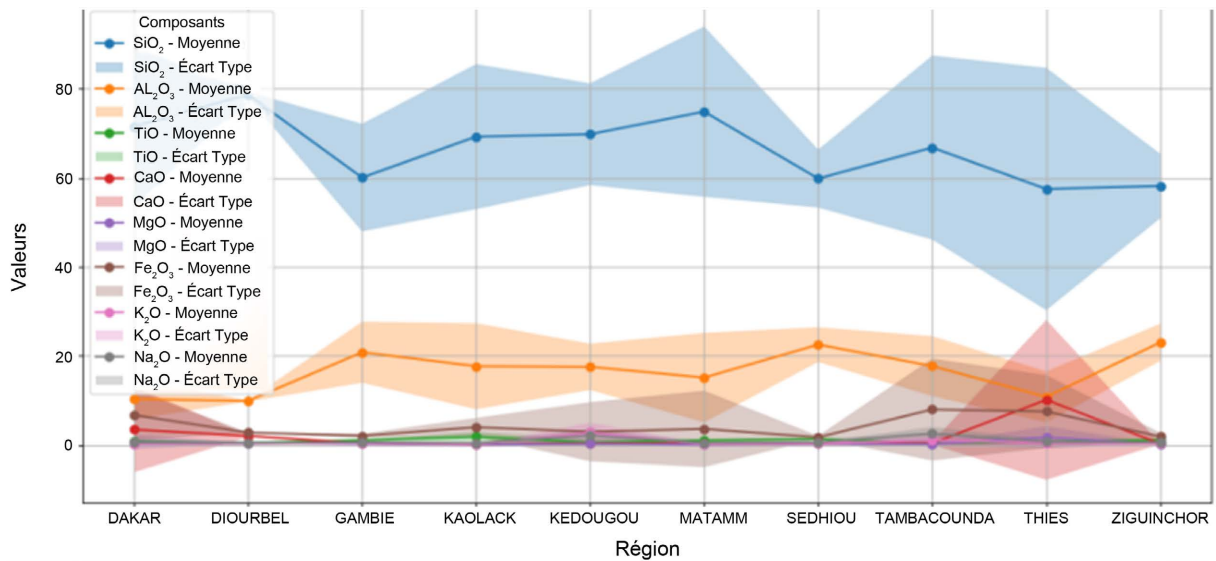


Figure 4. Average and standard deviation of ores.

The graph compares the average oxide content in the following regions: Dakar, Diourbel, Gambia, Kaolack, Kédougou, Matam, Sedhiou, Tambacounda, Thies, and Ziguinchor. The values measured are: basic oxide content (CaO, MgO, TiO, K₂O, Na₂O: increases pH and improves soil fertility), acid oxide content (Si: acidifies the soil), and amphoteric oxide content (alumina: behaves in a complex manner). The lines show the different average values of oxides in each region. The shaded areas around the lines indicate the variability of measurements by region.

The graph in Figure 5 compares the average saturation indices of five soil types: black clay, red clay, granite, and sand. These initial observations enable a more in-depth mineralogical analysis, particularly by examining the dispersion of data around the mean values.

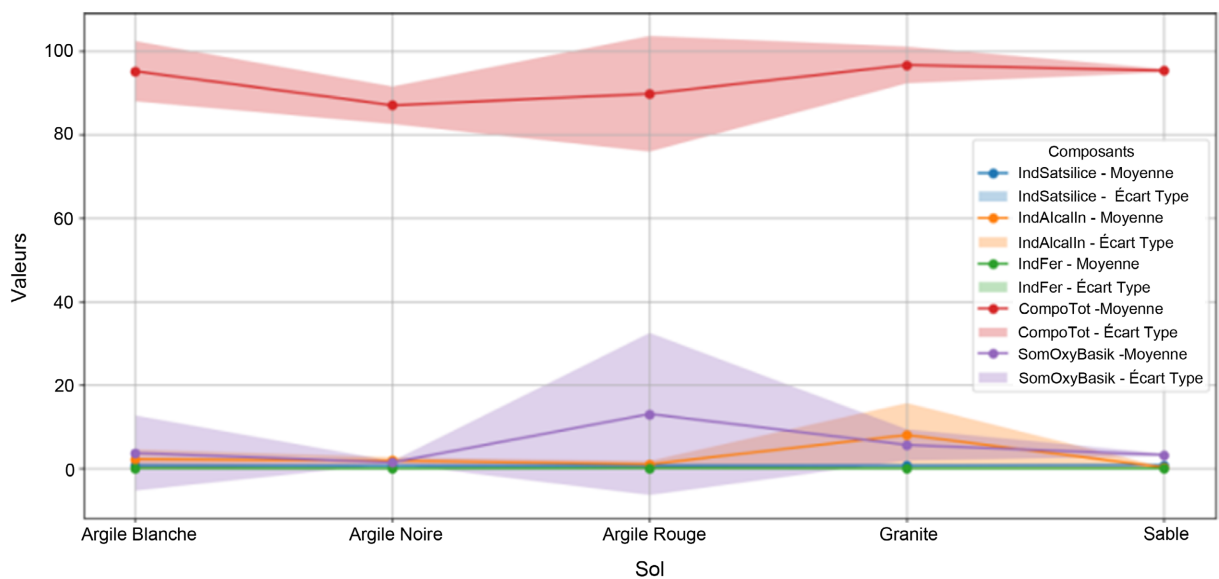


Figure 5. Mean and standard deviation per compound.

The components measured are: saturation index, alkaline ratio, iron index, and sum of basic oxides. The lines show the different average values of the components in each soil type. The shaded areas around the lines indicate the variability of the measurements. **Figure 6** shows the richness of Senegal's subsurface soils in alumina, titanium oxide, iron oxide, lime, and, above all, silica. But it also shows the low presence of potassium oxide. The following numbers are used to replace the names of minerals in **Figure 7**.

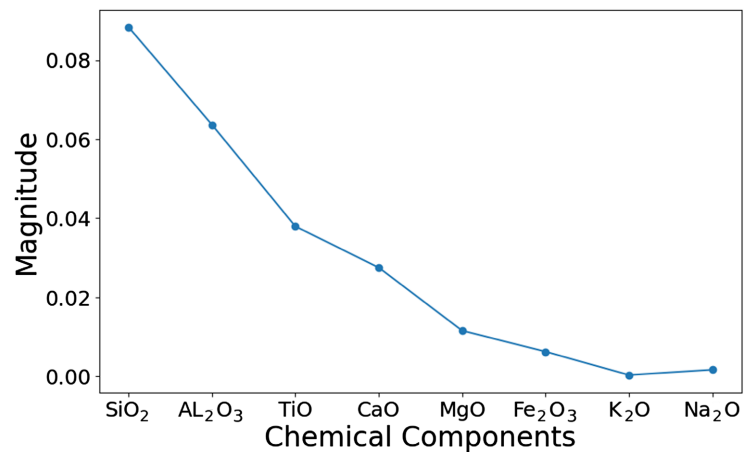


Figure 6. Average mineral composition of soil in Senegal.

1	2	3	4	5	6	7	8
SiO ₂	Al ₂ O ₃	TiO	CaO	MgO	Fe ₂ O ₃	K ₂ O	Na ₂ O

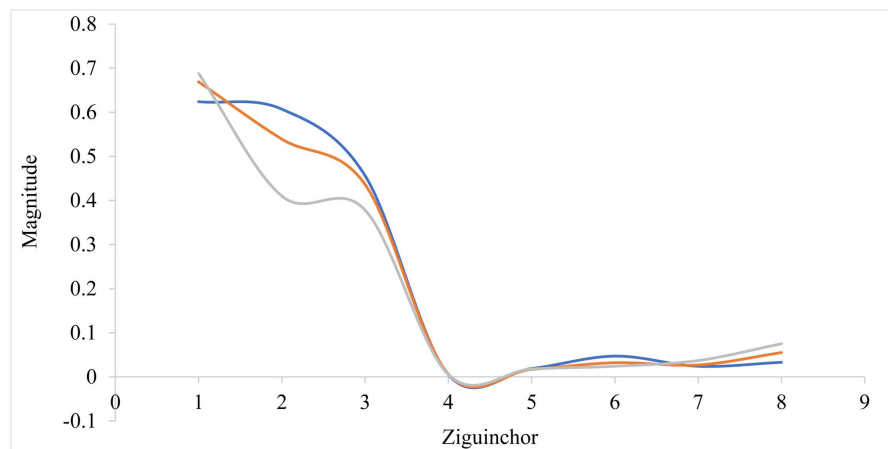


Figure 7. Magnitude according to the mineral in Ziguinchor.

Table 4 shows a comparison of the performance of different machine learning models. Here, the table mainly shows the non-separability and linearity of the data.

The SVC, NuSVC, and one-versus-all estimators perform better than LinearSVC. The RBF kernel of the SVM offers better classification performance due to its flexibility and simplicity compared to the linear kernel (Saha, 2023). This reflects the

reality of field data. Most of the time, they are not linearly separable. Unlike the linear model, the others perform well but relatively poorly, as for SVC and NuSVC, barely 1 in 2 predictions are correct. And for the one-versus-all model, 6 out of 10 predictions are correct. After normalization and application of the estimator, **Table 5** shows the results obtained, which are significantly improved.

Table 4. Performance before normalization.

Model	k	Acc	P	R	F1
SVC	rbf	57	54	57	54
NuSVC	rbf	57	56	57	54
LinearSVC	rbf, SVC	34	31	34	28
OneVsRest	rbf, SVC	62	59	62	56

k = kernel; Acc = Accuracy; P = Precision; R = Recall.

Table 5. Performance after normalization.

Model	k	Acc	P	R	F1
SVC	rbf	80	83	80	79
OneVsRest	SVM, rbf	91	92	91	92

The one-versus-all approach combined with the SVC model outperforms the SVC model in all performance metrics. The one-versus-all model is more accurate, with a better ability to identify true positives. It takes class imbalance into account but maintains a good balance between precision and recall. The model has an overall accuracy of 91%, indicating more than 9 out of 10 correct predictions most of the time, with more details provided in **Table 6**. With an overall accuracy of 92%, the model is very reliable for predicting positive classes and has a recall of 91% for predicting true positives. It also indicates a good balance between prediction accuracy and the ability to predict true positives. The one-versus-all model combined with SVC and an RBF kernel is overall better. It has high scores in all metrics.

Table 6. Classification performance.

	SiO ₂	Al ₂ O ₃	TiO	CaO	MgO	Fa ₂ O ₃	K ₂ O	Na ₂ O
SiO ₂	1	-0.19	-0.09	-0.7	-0.52	-0.16	0.19	0.17
Al ₂ O ₃	0	1	0.53	-0.49	-0.39	0.04	-0.03	-0.04
TiO	0	0	1	-0.25	-0.19	-0.05	-0.37	-0.39
CaO	0	0	0	1	0.7	-0.15	-0.1	-0.07
MgO	0	0	0	0	1	-0.09	-0.01	0.04
Fa ₂ O ₃	0	0	0	0	0	1	-0.16	-0.16
K ₂ O	0	0	0	0	0	0	1	0.91
Na ₂ O	0	0	0	0	0	0	0	1

This is the confusion matrix, **Table 6**, showing the performance of our classification model. It makes very few classification errors. Yes, indeed, the diagonal of the matrix concentrates almost all the prediction values. Only a few rare mispredictions were made. Classes 3 and 4 are the only classes where the model has difficulty making predictions.

4.2. Limitations and Prospects

This study is based on a dataset limited to specific geographical regions and well-targeted sectors within those regions. This may restrict the scope and generalization of the results to other geological contexts. Indeed, the physical and chemical characteristics of soils vary considerably depending on local climatic, geological and anthropogenic conditions. For example, a model trained on lateritic clays and black clays in tropical areas may not be directly applicable to marl-limestone or sandy soils in temperate environments, due to marked structural and behavioral differences.

5. Conclusion

In this work, a comprehensive framework based on SVM multi-class classification is proposed. Soil classification based on mineralogical composition is thus performed using a search grid. Comparative analysis in this study shows the effectiveness of the method used for supervised classification compared to other variants of the algorithm. The performance of a counter-all association with different classifiers and hyperparameters was also studied. The various stages, from data cleaning to the selection of study vectors, are important for achieving overall performance. This work demonstrates the importance of information processing through machine learning in solving engineering tasks, thus opening up interdisciplinary perspectives.

This work can be expanded by adding new data, such as the geolocation of samples. This will enable the production of a map showing mineralogical variation by region.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Abedi, M., Norouzi, G., & Bahroudi, A. (2011). Support Vector Machine for Multi-Classification of Mineral Prospectivity Areas. *Computers & Geosciences*, *46*, 272-283. <https://doi.org/10.1016/j.cageo.2011.12.014>
- Agrawal, A., Barratt, S., & Boyd, S. (2021). Learning Convex Optimization Models. *IEEE/CAA Journal of Automatica Sinica*, *8*, 1355-1364. <https://doi.org/10.1109/jas.2021.1004075>
- Camara, N., Diatta, I., Badiatte, L. A. D. L., & Tamba, S. (2024). Diagnosis and Solution of the Drinking Water Supply Network of the Municipality of Diembéring. *Computational Water, Energy, and Environmental Engineering*, *13*, 133-141. <https://doi.org/10.4236/cweee.2024.134008>

- Goldstein, D., Aldrich, C., Shao, Q., & O'Connor, L. (2025). A Machine Learning Classification Approach to Geotechnical Characterisation Using Measure-While-Drilling Data. *Geosciences*, *15*, Article No. 93.
- Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-Nearest Neighbor Algorithm: A Comprehensive Review and Performance Analysis of Modifications. *Journal of Big Data*, *11*, 1-50.
<https://doi.org/10.1186/s40537-024-00973-y>
- Kim, H.-C., Pang, S., Je, H. M., Kim, D., & Bang, S. Y. (2003). Constructing Support Vector Machine Ensemble. *Pattern Recognition*, *36*, 2757-2767.
- Koutsopoulou, E., Servou, A., Aggelopoulos, G., & Laskaridis, K. (2021). Research and Innovation in Exploration and Mining of Raw Materials: The ROBOMINERS Project. *Materials Proceedings*, *5*, Article No. 59. <https://doi.org/10.3390/materproc2021005059>
- Kovacevic, M., Bajat, B., Trivic, B., & Pavlovic, R. (2009). Geological Units Classification of Multispectral Images by Using Support Vector Machines. In *2009 International Conference on Intelligent Networking and Collaborative Systems* (pp. 267-272). IEEE.
<https://doi.org/10.1109/incos.2009.44>
- Liu, Z., & Xu, H. (2013). Kernel Parameter Selection for Support Vector Machine Classification. *Journal of Algorithms & Computational Technology*, *8*, 163-177.
<https://doi.org/10.1260/1748-3018.8.2.163>
- Maignien, R. (1965). *Carte Pédologique du Sénégal*.
<https://gifex.com/fr/fichier/carte-pedologique-du-senegal/>
- Miller, D. M., Kaminsky, E. J., & Rana, S. (1995). Neural Network Classification of Remote-Sensing Data. *Computers & Geosciences*, *21*, 377-386.
[https://doi.org/10.1016/0098-3004\(94\)00082-6](https://doi.org/10.1016/0098-3004(94)00082-6)
- Ndiaye, M., Davaud, E., Ariztegui, D., & Fall, M. (2012). A Semi Automated Method for Laminated Sediments Analysis. *International Journal of Geosciences*, *3*, 206-210.
<https://doi.org/10.4236/ijg.2012.31023>
- Nie, D., Li, W., Xie, L., Deng, M., Ding, H., & Liu, K. (2024). Effects of Fineness and Morphology of Quartz in Siliceous Limestone on the Calcination Process and Quality of Cement Clinker. *Materials*, *17*, Article No. 3601. <https://doi.org/10.3390/ma17143601>
- Ruano, A. E., Madureira, G., Barros, O., Khosravani, H. R., Ruano, M. G., & Ferreira, P. M. (2013). A Support Vector Machine Seismic Detector for Early-Warning Applications. *IFAC Proceedings Volumes*, *46*, 405-410.
<https://doi.org/10.3182/20130902-3-cn-3020.00082>
- Saha, A. (2023). An Empirical Comparison of Linear and Non-linear Classification Using Support Vector Machines. *International Journal of Computer Sciences and Engineering*, *11*, 123-215.
- Wang, C., Huang, Y., Cheng, I., & Nguyen, V. T. (2018). A Multi-Criteria Decision-Making (MCDM) Approach Using Hybrid SCOR Metrics, AHP, and TOPSIS for Supplier Evaluation and Selection in the Gas and Oil Industry. *Processes*, *6*, Article No. 252.
<https://doi.org/10.3390/pr6120252>
- Wittenberg, A., & de Oliveira, D. (2023). Critical Raw Material Resource Potentials in Europe. *Materials Proceedings*, *15*, Article No. 24.
<https://doi.org/10.3390/materproc2023015024>
- Zuo, R., & Carranza, E. J. M. (2010). Support Vector Machine: A Tool for Mapping Mineral Prospectivity. *Computers & Geosciences*, *37*, 1967-1975.
<https://doi.org/10.1016/j.cageo.2010.09.014>