

Research Article

ANALYSE COMPARATIVE DES ALGORITHMES DE CLASSIFICATIONS APPLIQUÉS À L'ÉVALUATION DES PERFORMANCES DES ENSEIGNANTS

***KADIATA MUKENGE Peter, MULUMBA MUAMBA Oscar, LUMU MUYA Steve, TSHIMANGA MBUYI Fortunat, NYUNGE KALUKANYI Chatty, KABAMBA KABAMBA Désiré**

Université pédagogique de Kananga, Democratic Republic of the Congo.

Received 20th May 2026; Accepted 21st June 2026; Published online 25th July 2026

RESUME

La baisse criante du niveau d'enseignement est due à la qualité des ressources humaines, c'est ainsi qu'une évaluation régulière de l'enseignant doit être une étape nécessaire pour chaque institution scolaire espérant avoir les bons résultats. Pour doter aux établissements scolaires un outil d'évaluation, nous avons récolté 1500 fiches d'inspections à la sous division Kananga 2, lesquelles nous ont permis de constituer un dataset de 56 variables. En effet trois algorithmes (Random Forest, Support à vecteurs machines et K-Nearest Neighbors) ont été appliqués sur les mêmes données pour voir lequel présenté une meilleure performance. C'est ainsi que Random Forest a été retenu comme meilleur algorithme car il a présenté un taux de performance de 99%.

Mots clés: Analyse comparative, Evaluation, Performances des enseignants.

INTRODUCTION

Aujourd'hui le système éducatif congolais est confronté à plusieurs difficultés entre autres la baisse du niveau de l'enseignement à tous les niveaux. En faisant une étude approfondie, nous avons découvert que l'enseignant fait partie de cette catastrophe. Il faudrait envisager une évaluation sérieuse et rigoureuse de ce dernier. Bien que cette évaluation se fasse déjà avec les méthodes traditionnelles qui ont montré leur limite.

Avec l'émergence de l'intelligence artificielle dans tous les domaines de la vie. Nous nous sommes dit, qu'il était temps d'appliquer l'intelligence artificielle dans le domaine éducatif à travers ses algorithmes. Dans notre article, nous avons fait appel à ces trois algorithmes: Random Forest, Support à vecteurs machines et K-Nearest Neighbors. Parmi les algorithmes ci-dessus lequel offre les meilleures performances pour l'évaluation des performances des enseignants des écoles secondaires à partir des données d'inspection? en effet, nous devons construire un jeu de données sur base des fiches d'inspections, entraîner les modèles Random Forest, SVM et KNN, faire une comparaison de leurs performances en se basant sur les métriques telles que: exactitude, précision, Rappel, F1-score et la matrice de confusion, pour identifier le modèle le plus performant.

REVUE DE LA LITTÉRATURE

La revue de littérature constitue une étape capitale dans la recherche scientifique, elle nous aide à analyser les recherches existantes similaires à notre sujet de recherche, c'est ainsi que nous avons eu à consulter les travaux ci-dessous:

- Pascal chatonnay, dans article intitulé: « Etude comparative d'algorithmes d'équilibrage de charge multicalculateurs », il présente les résultats obtenus en implantant trois algorithmes sur un multicalcateur à base de transputers, au-dessus du système Chorus.

***Auteur correspondant: KADIATA MUKENGE Peter,**

Assistant à l'université pédagogique de Kananga, Democratic Republic of the Congo.

- Oum-Hani Noulas (2023): dans son travail intitulé « Etude comparative des CNNs et de l'algorithme K-NN en mammographie », ce dernier met en œuvre un algorithme d'apprentissage profond (CNN) et l'algorithme d'apprentissage automatique (K-NN) pour détecter des anomalies dans l'image mammographie.
- Ma, X. (2024). : dans son article, il compare plusieurs algorithmes (régression logistique, arbre de décision, Random Forest, SVM et LSTM-RNN) pour l'évaluation des performances des étudiants. Les résultats montrent que le modèle LSTM-RNN+Random Forest obtient une meilleure précision de 98.5%.
- Ahmed, N. *et al.*, (2022): dans leur article, ils analysent les évaluations réalisées par les étudiants dans environnement d'enseignement en ligne après le covid-19. Les auteurs exploitent les données de 27 622 étudiants recueillies sur trois semestres afin d'identifier les facteurs qui influencent la qualité d'enseignement, pour cela ils utilisent l'arbre de décision, l'analyse des sentiments et les nuages mots. Et découvrent que la plus part des commentaires des étudiants sont positifs car le mot « Good » est le plus fréquent dans les commentaires.

Pour notre cas, nous avons utilisé séparément trois algorithmes sur les données de l'enseignement en vue de détecter l'algorithme ayant un meilleur taux de performance.

METHODES

Données

Notre étude repose sur échantillon de 1500 enseignants évalués sur base de 56 variables ci-dessous: présentation, Elocution, autorité, sens de responsabilité, assiduité, Tenue de la classe, Décoration, respect de l'autorité, Exactitude, dosage, formulation, adaptation au niveau, réponses aux questions, connaissance, conformité, progression, adaptation au milieu, décloisonnement, prérequis, motivation, inscription au JC des apprenants, Développement, Enchaînement, Synthèse, Application, Didactique générale, Didactique de discipline, imagination pédagogique, Facilité de transmission, Manuel de base, Matériel Didactique, utilisation du

tableau, questions de l'enseignant\ réponses des apprenants, Travail individuel, Travail en groupe, Discipline, Ambiance de travail, journal de classe, brouillon ou ardoise, cahiers de cours, Manuel de base, travaux scolaires, cahier de communication, Prévisions des matières, journal de classe, Fiche de leçon, fichier du cours, cahier de points, cahier de questions, livre du maître, Documentations autre, registre d'appel, Evaluation progressive, Evaluation, Atteinte de l'objectif, Evaluation inspectoriale.

Le Prétraitement

Les données ont été nettoyées en supprimant les colonnes qui ne contiennent pas d'informations et standardisées pour les mettre au même échelle en utilisant l'instruction **StandardScaler()**.

Algorithmes et logiciels utilisés

Nous avons utilisé trois algorithmes cités ci-haut à travers le langage python, et avons déterminé les métriques telles que :

- Exactitude
- Rappel
- F1-score
- Matrice de confusion

Nous allons démontrer notre démarche à travers un algorithme et pour les deux autres, les résultats seront dans le tableau synthétique.

Importation de la base de données

```
BDD=pd.read_excel('EVALUATION.xlsx')
BDD.shape
```

suppression de colonnes qui ne contiennent pas d'informations

```
df=BDD.loc[:,~BDD.columns.str.contains('^Unnamed')]
```

Les variables

```
x=df.drop("OBSERVATION",axis=1)
y=df["OBSERVATION"]
```

Séparation des données

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.30,random_state=42,stratify=y)
```

Normalisation

```
from sklearn.preprocessing import StandardScaler
scaler=StandardScaler()
x_train=scaler.fit_transform(x_train)
x_test=scaler.fit_transform(x_test)
```

Usage de Random Forest

```
from sklearn.ensemble import RandomForestClassifier
rf=RandomForestClassifier(n_estimators=100,random_state=42)
rf.fit(x_train,y_train)
```

les predictions

```
y_pred=rf.predict(x_test)
# Rapport de classifications
print("nRapport de classification")
print(classification_report(y_test,y_pred,target_names=["Elite","TB","B","AB","M"]))
```

Classification Report Data

Class / Metric	Precision	Recall	F1-Score	Support
Elite	0.97	1.00	0.98	90
TB	1.00	0.98	0.99	90
B	1.00	1.00	1.00	90
AB	1.00	0.99	0.99	90
M	1.00	1.00	1.00	90
accuracy			0.99	450
macro avg	0.99	0.99	0.99	450
weighted avg	0.99	0.99	0.99	450

La matrice de confusion

```
import matplotlib.pyplot as plt
from sklearn.metrics import ConfusionMatrixDisplay
disp=ConfusionMatrixDisplay(confusion_matrix=cm,display_labels=["Elite","TB","B","AB","M"])
disp.plot(cmap="Blues",xticks_rotation=45)
plt.title("Matrice pour l'algo Random Forest")
plt.show()
```

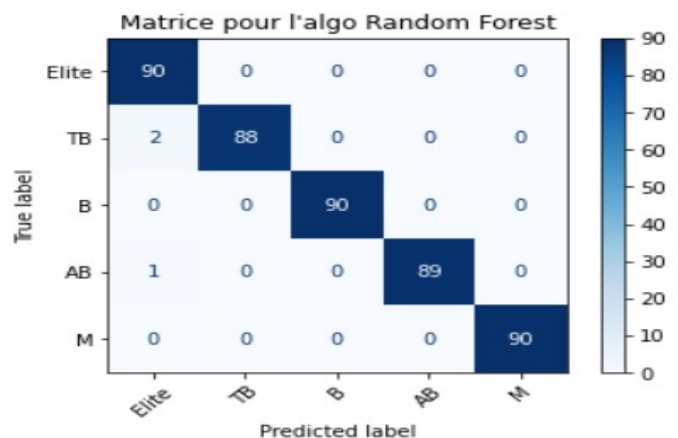
RESULTATS ET DISCUSSIONS

TABLEAU SYNTHETIQUES DES METRIQUES

	Exactitude	Précision	Rappel	F1-Score
Random Forest	99	99	99	99
SVM	96	97	96	96
KNN	91	92	91	91

LES MATRICES DE CONFUSION

RANDOM FOREST



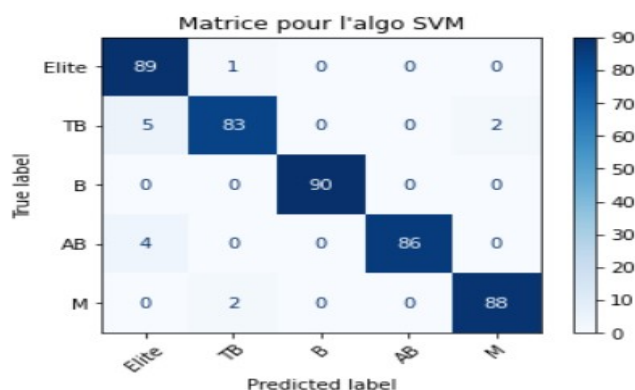
Les vrais positifs : les valeurs de diagonale (90,88,90,89,90) présentent le nombre de bonnes classifications.

- 90 enseignants correctement classés comme Elite
- 88 enseignants correctement classés comme TB
- 90 Enseignants correctement classés comme B
- 89 Enseignants correctement classés comme AB
- 90 Enseignants correctement comme M

Les Faux Positifs : les valeurs en dehors de la diagonale sur la même colonnes indiquent les erreurs où d'autres classes ont été prédites à la place.

- 2 enseignants TB et 1 Enseignant AB ont été classés comme Elite

SVM



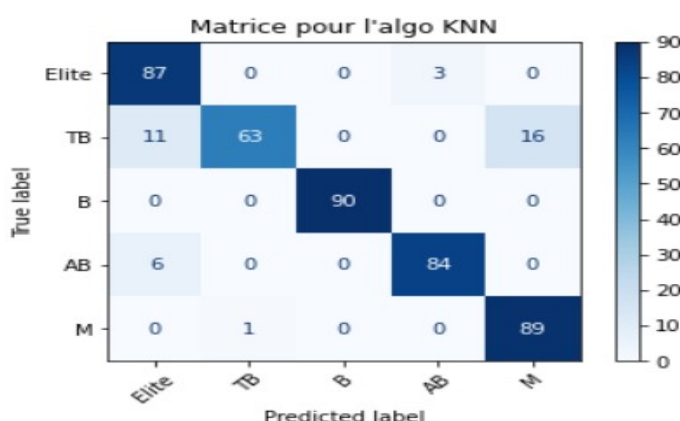
Les vrais positifs : les valeurs de diagonale (89,83,90,86,88) présentent le nombre de bonnes classifications.

- 89 enseignants correctement classés comme Elite
- 83 enseignants correctement classés comme TB
- 90 Enseignants correctement classés comme B
- 86 Enseignants correctement classés comme AB
- 88 Enseignants correctement comme M

Les Faux Positifs : les valeurs en dehors de la diagonale sur la même colonnes indiquent les erreurs où d'autres classes ont été prédites à la place.

- 5 enseignants TB et 4 Enseignant AB ont été classés comme Elite
- 2 Enseignants M et 1 Enseignant Elite ont été classés comme TB
- 2 Enseignants TB ont été classés comme M.

KNN



Les vrais positifs : les valeurs de diagonale (87,63,90,84,89) présentent le nombre de bonnes classifications.

- 87 enseignants correctement classés comme Elite
- 63 enseignants correctement classés comme TB
- 90 Enseignants correctement classés comme B
- 84 Enseignants correctement classés comme AB
- 89 Enseignants correctement comme M

Les Faux Positifs : les valeurs en dehors de la diagonale sur la même colonnes indiquent les erreurs où d'autres classes ont été prédites à la place.

- 11 enseignants TB et 6 Enseignant AB ont été classés comme Elite
- 1 Enseignant M a été classé comme TB
- 3 Enseignants Elite ont été classés comme AB
- 16 Enseignants TB ont été classés comme M.

Conclusion : En comparant les pourcentages des différentes métriques et les erreurs commises au niveau de chacune des matrices de confusion, nous pouvons affirmer que c'est le modèle développé avec l'algorithme Random Forest qui présente une meilleure performance et qui est l'unique modèle que nous pouvons utiliser dans l'évaluation des performances des enseignants des écoles secondaires.

CONCLUSION GENERALE

Notre recherche avait pour objectif de comparer les performances de trois algorithmes de classification qui sont: support à vecteurs machines, Random Forest et K-Nearest Neighbors dans l'évaluation des performances des enseignants des écoles secondaires. Tout est parti d'un jeu de données formé de 1500 fiches d'inspection avec 56 variables chacune. Les trois modèles ont été entraînés et évalués des métriques comme exactitude, précision, rappel, F1-score ainsi que les matrices de confusion.

Les résultats obtenus montrent que l'algorithme Random Forest (Forêts aléatoires) dépasse les autres algorithmes avec une exactitude de 99%. Les matrices de confusion montrent également que ce modèle commet moins d'erreurs de classification, ce qui prouve que sa capacité de généraliser correctement les données sur l'évaluation des enseignants. Cette étude présente des limites car les données proviennent d'une seule sous division scolaire. Dans prochaines recherches, nous pensons étendre sur l'espace grand Kasai et incorporer d'autres algorithmes de classification tels que XGBoost ou les réseaux de neurones.

REFERENCES

- Ahmed,N.,Nandi,D. & Zaman,A.G.M(2022). Analyzing student evaluations of teaching in a completely online environment. International Journal of modern Education and Computer Science, 14(6), 13-24.
- Chatonnay, P. Étude comparative d'algorithmes d'équilibrage de charge sur multicalculateurs.
- Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media.
- Gutiérrez,G.et al(2018). Mining students comments about teacher performance assesment using machine learning algorithms. International Journal of Combinatorial optimization problems and informatics, 9(3), 26-40
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques (3rd ed). Morgan Kaufmann.
- Ma,X.(2024)., Artificial intelligence-driven education evaluation and scoring : Comparative exploration of machine learning algorithms. Journal of intelligent Systems, 33(1).
- Oum-Hani Noukas. (2023). Étude comparative des CNNs et de l'algorithme K-NN en mammographie.
- Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
- UNESCO. (2023). Global Education Monitoring Report 2023.
- Yun,G. et al (2023). Analysis of the teaching quality on deep learning-based innovative ideological political education platform. Progress in Artificial Intelligence, 12(2), 175-186.