

Research Article

LE TRAITEMENT DES DONNÉES PAR SCORING ET SEGMENTATION POUR LES PROJETS DE DÉVELOPPEMENT HUMAINITAIRE

***Steve LUMU MUYA, Jully MBOMBO KABEYA, Charles MAPETOMATHUNDA**

Assistant à l'Institut Supérieur d'Etudes Sociales de Kananga, République Démocratique du Congo.

Received 01st July 2026; Accepted 02nd August 2026; Published online 11th September 2026

RESUMÉ

L'étude porte sur l'évaluation de l'éligibilité des Organisations Non Gouvernementales (ONG) dans le cadre des interventions humanitaires. Elle propose une approche analytique basée sur un système de scoring utilisant des critères clés tels que le statut juridique, l'expérience et la capacité opérationnelle, afin d'évaluer chaque ONG et de lui permettre de soumissionner aux appels à manifestation d'intérêt qui seront lancés par les bailleurs de fonds en vue d'améliorer les conditions de vie des populations. Les résultats sont ensuite enrichis par une technique de clustering issue de l'apprentissage non supervisé, permettant de regrouper les ONG selon des caractéristiques similaires. Cette approche facilite l'identification des organisations éligibles et favorise la mutualisation des services dans des secteurs clés tels que la sécurité alimentaire, l'éducation, le WASH et la lutte contre les violences basées sur le genre (VBG). Dans l'ensemble, le modèle proposé améliore la transparence du processus de sélection des ONG et soutient une meilleure prise de décision dans la coordination des actions humanitaires.

Mots Clés: ONG, Éligibilité, Scoring, Clustering, K-means, Apprentissage non supervisé.

INTRODUCTION

À l'ère de la transformation numérique, la croissance rapide des volumes de données représente à la fois un défi majeur et une opportunité pour les organisations engagées dans le développement humanitaire. L'exploitation efficace de ces données est devenue indispensable pour améliorer la compréhension des besoins des populations, optimiser l'allocation des ressources et renforcer l'impact des interventions. Dans ce contexte, l'utilisation conjointe des techniques de scoring et de segmentation apparaît comme une approche pertinente pour soutenir les processus décisionnels.

En République Démocratique du Congo, où les besoins humanitaires demeurent importants malgré l'augmentation des financements, la sélection des organisations chargées de mettre en œuvre les projets reste confrontée à des difficultés liées à l'hétérogénéité des capacités organisationnelles, opérationnelles et financières des ONG. Cette situation souligne la nécessité de disposer d'outils analytiques objectifs permettant d'évaluer les performances des organisations, de réduire les risques liés à l'allocation des ressources et d'améliorer la transparence des décisions.

Le choix de cette recherche s'inscrit dans la volonté de proposer une solution informatique innovante destinée à renforcer la gestion des financements des projets de développement. L'étude s'appuie sur la constitution d'un entrepôt de données et sur l'exploitation des technologies décisionnelles afin d'intégrer un système de scoring capable de mesurer les performances des ONG à partir d'indicateurs objectifs, puis de les regrouper en catégories homogènes grâce aux techniques de segmentation.

Cette démarche présente un intérêt scientifique, technique et opérationnel. Elle contribue à l'évolution des systèmes décisionnels

en adaptant des méthodes de fouille de données, traditionnellement utilisées dans le secteur bancaire, au domaine humanitaire. Elle vise également à fournir aux gestionnaires des projets un outil d'aide à la décision favorisant une allocation plus rationnelle des financements, un meilleur alignement des interventions sur les besoins réels des populations et une amélioration des conditions de vie des bénéficiaires.

L'analyse de la littérature met en évidence plusieurs travaux portant sur le scoring et la segmentation dans des domaines tels que les services publicset les méthodes d'apprentissage statistique. Toutefois, ces recherches ne répondent pas directement aux exigences de l'évaluation des ONG dans le contexte humanitaire. La présente étude se distingue par son objectif de combiner les approches de scoring et de segmentation afin de classer les organisations selon leur niveau d'éligibilité et leurs capacités d'intervention, en vue d'améliorer les mécanismes de sélection et de financement des projets de développement.

La recherche s'articule autour de la problématique de l'évaluation objective des capacités des ONG face à la multiplication des acteurs humanitaires, à la complexité croissante des données et à la nécessité de garantir une gestion transparente des ressources.

Elle cherche ainsi à déterminer comment un système de scoring associé à des techniques de segmentation peut permettre de regrouper les organisations en catégories homogènes, de renforcer la transparence dans l'allocation des financements et de concilier les exigences de performance avec les contraintes propres aux organisations non gouvernementales.

L'hypothèse principale postule que l'utilisation d'indicateurs mesurables combinée à des techniques de classification, notamment l'algorithme K-means, permet d'améliorer l'objectivité des décisions, de réduire la subjectivité dans les processus de sélection et d'optimiser l'utilisation des ressources destinées au développement. Pour vérifier cette hypothèse, la recherche poursuit plusieurs objectifs consistant à évaluer la conformité administrative et juridique des

***Auteur correspondant: Doctorant Steve LUMU MUYA,**

Assistant à l'Institut Supérieur d'Etudes Sociales de Kananga, République Démocratique du Congo.

ONG, leurs capacités techniques, organisationnelles et financières, leurs mécanismes de suivi-évaluation ainsi que leur alignement avec les priorités nationales de développement.

La démarche méthodologique repose sur une approche analytique intégrant les techniques de scoring et de clustering, complétées par l'utilisation de l'algorithme K-means, des arbres de décision et d'outils informatiques tels que Python, MySQL, Excel et Power BI.

L'ensemble de cette approche vise à proposer un système décisionnel capable d'améliorer l'efficacité, la transparence et la pertinence des décisions relatives au financement des projets humanitaires et de développement.

REVUE DE LITTÉRATURE

- ❑ **Benamar Houmadi (2007)** : Sa thèse soutenue au Québec ("*Etude exploratoire d'outils pour le data mining*") a mené une étude exploratoire comparative d'outils de data mining, se concentrant sur les réseaux de neurones et les arbres de décision. Il a démontré la haute capacité de classification de ces outils, atteignant une précision allant jusqu'à 99,55%.
- ❑ **SAFA El. Ayeb (2019)** : Sa thèse soutenue en France (Polyclinique Monterial), l'auteur a travaillé sur le « *développement d'un outil de segmentation des comportements d'achat des clients en se basant sur leurs données morphologiques* », dans le but d'améliorer les stratégies marketing et la personnalisation des offres.
- ❑ **Marine DEPECKER (2010)** : thèse soutenue à l'Université Télécom Paris Tech, a abordé les *Méthodes d'apprentissage statistique pour le Scoring*, l'auteur a fait l'étude théorique et le développement algorithmique d'une méthode non-paramétrique pour l'apprentissage supervisé de règles d'ordonnement à partir des données étiquetées de façon binaire.
- ❑ **PLISSON Manuel et ROSSARD Benjamin (2014)** : « *Arbitrage entre mutualisation et segmentation : cas d'assurance et collectivités* » Leur mémoire soutenu en France vise à développer un modèle de tarification qui équilibre les principes de mutualisation et de segmentation afin d'optimiser la gestion des risques et la compétitivité des offres d'assurances pour les collectivités.

La particularité de ce travail se trouve dans l'application de l'approche scoring et la méthode de segmentation dans un système décisionnel pour la bonne compréhension des comportements des divers groupes homogènes (d'individus). Constituant une population nous appliquerons la méthode des scores en vue de mesurer le risque ou la probabilité de défaut de choix du décideur. Enfin, nous présenterons un recueil des résultats possibles partant de fait liés aux segments et scores.

Hormis l'introduction et la conclusion générale, notre démarche dans cette réflexion gravite au tour de trois grands points:

- ❑ **L'approche scoring pour l'évaluation des ONG**: Ici, nous verrons également les notions générales des scores, quelques types et méthodes de scoring.
- ❑ **La méthode de segmentation de données** où nous développerons les techniques de segmentation en partant des concepts de base, les types, les critères, jusqu'aux stratégies d'une bonne segmentation.
- ❑ **Le dernier point** sera consacré à la nouvelle approche de l'outil d'aide à la décision en partant de sa conception jusqu'à sa matérialisation.

APPROCHE DE SCORING POUR L'ÉVALUATION DES ONG

Dans le cadre des politiques de développement orientées vers la justice sociale et la croissance économique durable, l'octroi des fonds constitue un enjeu stratégique majeur. Toutefois, la diversité des agences locales d'exécution, tant en matière de capacités institutionnelles que de performance opérationnelle, complexifie considérablement le processus de prise de décision. Selon **Mosse**, l'efficacité des projets de développement dépend largement de la capacité réelle des acteurs chargés de leur mise en œuvre.¹

Face à cette problématique, le recours à une approche de scoring apparaît comme une solution méthodologique pertinente permettant d'objectiver l'évaluation des agences locales d'exécution.

Le scoring constitue la phase analytique cruciale permettant de traduire des réalités de terrain complexes en données comparables. Pour que la croissance économique soit durable et la justice sociale effective, l'allocation des ressources ne peut plus reposer sur une simple conformité administrative.

Le présent point analyse ainsi les fondements, les critères et la mise en œuvre du scoring comme outil d'aide à la décision dans l'octroi des fonds de développement.

Il propose un modèle de scoring rigoureux destiné à mesurer la capacité réelle des agences avec comme objectif fournir aux décideurs un instrument de mesure capable d'identifier les structures qui présentent le meilleur équilibre entre performance opérationnelle, intégrité éthique et impact social.

Définition

Le scoring est défini comme une méthode d'évaluation quantitative consistant à attribuer un score numérique à une entité (organisation, projet ou individu) sur base d'un ensemble de critères prédéfinis et pondérés, afin de mesurer sa performance, sa capacité ou son niveau de risque. Cette approche vise à transformer des informations qualitatives et quantitatives en un score synthétique facilitant la comparaison et l'aide à la prise de décision².

Rôle du scoring dans l'octroi des fonds

L'utilisation du scoring dans l'octroi des fonds de développement vise à renforcer l'équité et la rationalité du processus décisionnel. En l'absence d'un outil structuré d'évaluation, l'allocation des ressources peut être influencée par des considérations subjectives ou politiques, au détriment de l'efficacité et de l'impact social.

Dans cette perspective, on souligne que le développement doit être conçu comme un processus d'élargissement des libertés réelles des individus³. Le scoring permet d'identifier les agences les plus à même de traduire les ressources financières en résultats sociaux tangibles, contribuant ainsi à la réduction des inégalités et à la promotion de la justice sociale.

¹Mosse, D., (2005) : *Cultivating Development: An Ethnography of Aid Policy and Practice*, Pluto Press, Londres, p. 101.

²Kaplan, R. S. & Norton, D. P., (1996) : **The Balanced Scorecard: Translating Strategy into Action**, Harvard Business School Press, Boston, 1996, pp. 72-78.

³Sen, A., (1999) : **Development as Freedom**, Oxford University Press, Oxford, p.42

Définition des critères de scoring

Capacité institutionnelle et organisationnelle (20%)

La capacité institutionnelle constitue un facteur déterminant de la réussite des projets de développement. Elle renvoie à la structuration de l'organisation, aux compétences du personnel et à la qualité des mécanismes de gouvernance interne. Comme le souligne Mintzberg, une organisation efficacement structurée est mieux à même de coordonner ses activités et d'atteindre ses objectifs.⁴

Dans le cadre de ce projet, les agences disposant de procédures formalisées et d'un personnel qualifié obtiennent des scores plus élevés, traduisant une meilleure capacité d'exécution.

Elle analyse la pérennité de l'agence au-delà du cycle de financement.

- Capacité de Mobilisation de Ressources (10 pts): Diversification des sources de revenus (autofinancement, dons locaux).
- Efficience des Coûts (10 pts) : Part des frais de fonctionnement par rapport au budget total alloué aux projets.

Performance historique dans l'exécution des projets (20%)

La performance passée constitue un indicateur essentiel de la capacité future d'une agence. Elle prend en compte le respect des délais, la maîtrise des budgets et l'atteinte des résultats escomptés. Selon Kerzner, l'historique de performance est l'un des meilleurs prédicteurs du succès des projets ultérieurs.⁵

Le scoring permet ainsi d'intégrer de manière systématique les enseignements tirés des expériences antérieures dans l'évaluation des demandes de financement.

Elle mesure l'aptitude à transformer les fonds en résultats tangibles.

- Expertise Qualifiée (10 pts): Ratio de personnel technique par rapport au personnel administratif.
- Suivi et Évaluation (10 pts): Capacité à produire des rapports d'impact basés sur des indicateurs précis.

Gestion financière et transparence (30%)

La gestion financière est un élément central de la crédibilité des agences locales d'exécution. Elle comprend la tenue régulière de la comptabilité, la réalisation d'audits financiers et le respect des exigences des bailleurs de fonds. D'après la Banque mondiale⁶, les défaillances en matière de gouvernance financière constituent une cause majeure d'échec des projets de développement.

Dans le modèle de scoring proposé, les agences démontrant une gestion financière rigoureuse et transparente sont donc valorisées. Cette section évalue la solidité institutionnelle pour prévenir les risques de corruption.

- Contrôle Interne (15 pts): Existence de manuels de procédures et d'audits financiers certifiés.

- Démocratie Organisationnelle (15 pts): Renouvellement des mandats et transparence dans le choix des dirigeants.

Impact social des interventions

L'impact social constitue la finalité ultime des fonds de développement. Il s'agit d'évaluer la capacité des agences à produire des changements positifs et durables dans les conditions de vie des populations bénéficiaires. Comme le rappelle Sen, le développement n'a de sens que s'il améliore concrètement le bien-être humain⁷.

Les projets ciblant prioritairement les groupes vulnérables et générant des effets sociaux mesurables obtiennent ainsi des scores élevés. Ce pilier est le cœur de la validation éthique du projet.

- Équité de Genre et Diversité (15 pts) : Pourcentage de femmes et de minorités dans les instances de décision.
- Redevabilité envers les Bénéficiaires (15 pts) : Mécanismes permettant aux populations locales de contester ou d'influencer les actions de l'agence.

Ancrage territorial et connaissance du contexte local

La connaissance du contexte local et la proximité avec les bénéficiaires favorisent l'efficacité des interventions. Chambers met en évidence l'importance des approches participatives et de l'ancrage territorial dans la réussite des projets de développement⁸.

Les agences disposant d'une présence effective sur le terrain et de partenariats communautaires solides sont donc mieux notées dans le scoring.

Pondération des critères et calcul du score global

Les critères de scoring ne présentent pas tous le même niveau d'importance. Une pondération est appliquée afin de refléter les priorités stratégiques du programme de financement. Le score global obtenu par chaque agence résulte de l'agrégation pondérée des scores partiels, permettant ainsi une évaluation synthétique et objective des capacités institutionnelles⁹.

Apport du scoring à la prise de décision

Le scoring constitue un outil d'aide à la décision permettant aux bailleurs de fonds d'apprécier de manière rigoureuse la capacité des agences locales d'exécution. Il favorise une allocation plus efficace des ressources, limite les risques financiers et renforce la transparence du processus décisionnel. Comme le souligne Mosse, la clarté des critères d'évaluation renforce la légitimité et l'acceptabilité des décisions prises¹⁰.

Le point suivant sera consacré à la segmentation des données, où nous montrerons comment les ONG peuvent être classées en K groupes homogènes sur la base de leurs scores.

⁴Mintzberg H., (1982) : *Structure et dynamique des organisations*, Paris, éd. Organisation. ISBN, p. 60

⁵Kerzner H., (2017) : *Project Management: Q systems Approach to planning, scheduling and controlling*, 12^e éd., Chicago, p.110

⁶World Bank, *the world Bank annual report 2018*, DC: international Bank for reconstruction and developpement, Washington(USA)

⁷Sen A., *op. cit.*, p. 50

⁸Chambers R., (1997) : *Whose Reality Counts? Putting the first Last*, London, p.58

⁹Kaplan & Norton, (1996) : *The Balanced scorecard : Translating Strategy into action*. 1^{ère} Ed. Boston MA : Havard Business School Press, p. 83

¹⁰Mosse D., (2005) : *Cultivating Development : An Ethnography of Aid Policy and Practice*, 1^{ère} éd., London Pluto Press, p.112.

LA SEGMENTATION DE DONNEES

La segmentation de données est un processus fondamental dans l'analyse de données, le marketing, la gestion de la relation client (CRM) et de nombreuses autres disciplines. Elle consiste à diviser un ensemble de données hétérogène en sous-groupes plus homogènes, appelés segments ou clusters, en fonction de caractéristiques ou d'attributs partagés¹¹. Chaque segment représente une partie cohérente de la population étudiée ou du système à modéliser.

L'objectif principal est de regrouper des éléments similaires pour mieux comprendre, mieux décider et agir plus efficacement.

L'idée centrale du machine learning est d'apprendre à partir de données. Pour cela, il existe trois grandes méthodes d'apprentissage:

Apprentissage Non Supervisées

L'apprentissage non supervisé est une branche très connue du machine learning dans laquelle un modèle analyse des données sans étiquettes ni réponses prédéfinies contrairement à l'apprentissage supervisé où on ne dit pas à l'algorithme ce qu'il doit trouver.

Les méthodes non supervisées visent à découvrir des structures cachées dans les données sans avoir de connaissance préalable des groupes. Elles sont particulièrement utiles pour l'exploration de données et la découverte de segments initiaux.

Le clustering est donc une méthode intéressante lorsqu'on souhaite laisser à la machine le pouvoir de proposer sa propre solution, et ainsi découvrir une approche différente de la nôtre.

Il a comme principe d'analyser les différentes variables X qui caractérisent nos données, afin de regrouper les points en clusters, sans pour autant connaître la nature de ces points.

a) Algorithme K-Means:

Un algorithme itératif qui partitionne les données en k clusters en minimisant la somme des distances au carré entre chaque point de données et le centroïde de son cluster¹².

Le nombre de clusters (k) doit être spécifié à l'avance.

Cet algorithme vise à partitionner un ensemble de données en K clusters en minimisant la distance entre chaque point de données et le centre (centroïde) de son cluster. Il est couramment utilisé pour la segmentation de clients, la classification d'images et la compression de données.

1. Avantages:

- **k-means** est très populaire en raison de sa simplicité de compréhension et de mise en œuvre;
- Il est rapide et efficace, même pour de grands volumes de données;
- Il est applicable à des jeux de données de grande taille et peut s'adapter à différents types de données (y compris textuelles), à condition de choisir une mesure de distance appropriée.

2. Inconvénients:

- Le nombre de clusters K doit être défini à l'avance;
- Les résultats peuvent varier selon l'initialisation des centres;
- Les clusters sont construits autour de centroïdes artificiels, qui ne correspondent pas toujours à des objets réels ;
- L'algorithme est sensible aux valeurs aberrantes (outliers) et aux données mal normalisées.

3. Algorithme classique de k-means

Étape 1: Choisir aléatoirement K centres initiaux (centroïdes).

Étape 2: Affecter chaque objet au cluster dont le centre est le plus proche (selon la distance choisie).

Étape 3: Recalculer le centre de gravité de chaque cluster à partir des objets qui lui sont assignés.

Étape 4: Répéter les étapes 2 et 3 jusqu'à ce que:

- les centres ne changent plus, ou
- les objets ne changent plus de cluster, ou
- un nombre maximal d'itérations soit atteint.

Algorithme classique

Cet algorithme permet de partitionner un ensemble de données X en K groupes homogènes. Il commence par initialiser aléatoirement K centres, puis affecte chaque donnée au groupe dont le centre est le plus proche. Ensuite, les centres des groupes sont recalculés en prenant la moyenne des éléments de chaque groupe. Ce processus d'affectation et de mise à jour des centres est répété jusqu'à ce que les groupes deviennent stables, ce qui signifie que l'algorithme a convergé vers une partition finale des données.

Entrée

Ensemble de N données, noté par X

Nombre de groupes souhaité, noté par K

Sortie

Une partition de K groupes ($C_1, C_1, \dots, C_k$)

Debut

1) initialisation aléatoire de centres CK ; répéter

2) affectation : générer une nouvelle partition en assignant chaque objet au groupe dont le centre est le plus proche :

$X_i \in C_k \text{ si } \forall j | x_i - \mu_k | < | x_i - \mu_j |$

Avec μ_k le centre de la classe K ;

3) Représentation : calculer le centre associé à la nouvelle partition

$$\mu_k = \frac{1}{N} \sum_{x \in C_k} x$$

jusqu'à convergence de l'algorithme vers une partition stable
Fin

3. Critère de convergence

L'algorithme K-means se termine lorsque l'un des critères suivants est satisfait :

- **Les centres des clusters ne changent plus** : les positions des centres des clusters restent constantes entre les itérations.

¹¹ IBM Cloud Education, What is Data Segmentation?, 2023.

<https://www.ibm.com/cloud/learn/data-segmentation>, consulté le 03 juin 2026 à 13h 17

¹² MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability Vol. 11, pp. 281-297

- **Les affectations des points de données ne changent plus :** les points de données restent dans le même cluster entre les itérations.
- **Un nombre maximal d'itérations est atteint.**

4. Formule de la distance

La distance euclidienne entre un point x et un centre de cluster μ est calculé de cette manière :

$$\text{Distance} : (x, \mu) = \sqrt{\sum_{i=1}^d (x_i - \mu_i)^2}$$

Où d est la dimension des données, x_i est la i ème caractéristique du point x_i et μ_i est la i ème caractéristique du centre du cluster μ

5. Choix du nombre de clusters

Le choix du nombre de clusters K peut être délicat. Voici quelques méthodes courantes pour le déterminer :

-  **Méthode du coude :** tracez la somme des carrés des distances intra-clusters (inertia) en fonction du nombre de clusters et cherchez le « coude » dans le graphique où la diminution de la variance intra-cluster commence à ralentir.
-  Cette méthode consiste à tester différentes valeurs de K pour l'algorithme du K-Means Clustering, en mesurant à chaque fois la variance (écart des points par rapport à leur moyenne) finale de chaque clustering, puis en traçant sur un graphique l'évolution de cette variance en fonction de K pour identifier une forme de coude dans le graphique.
-  **Critère de silhouette :** évaluez la qualité du clustering en mesurant la cohésion et la séparation des clusters. Les scores de silhouette proches de +1 indiquent un bon clustering
-  **Validation croisée :** divisez les données en plusieurs ensembles, effectuez le clustering pour différents nombres de clusters, et évaluez la stabilité des clusters.

Nous signalons aussi que cet algorithme converge au bout d'un nombre fini d'itérations.

Pour créer une segmentation respectant les deux critères indiqués plus haut, il faut maximiser I_B et minimiser en même temps I_w . Donc, d'après le théorème de Huygens, il suffit de respecter l'un des deux critères : le second sera respecté en même temps.

Segmentation hiérarchique¹³

En principe, il existe deux classes de méthodes de segmentation hiérarchique :

- **Segmentation hiérarchique ascendante :** qui part des N points comme N groupes à partir desquels on construit une partition à $N-1$ classes par fusion de deux groupes, puis $N-2$ groupes par fusion de deux groupes, jusqu'à avoir rassemblé tous les points dans un seul groupe ;
- **Segmentation hiérarchique descendante :** qui rassemble initialement les N points dans un seul groupe à partir duquel on construit 2 groupes, puis 3, ... puis N .

Observation : les méthodes hiérarchiques ne sont pas beaucoup utilisées.

Algorithme de la segmentation hiérarchique ascendante

Cet algorithme vise à regrouper les individus, en commençant par le bas (les deux plus proches) et en construisant progressivement un arbre, ou un dendrogramme, regroupant finalement tous les individus en une seule classe à la racine.

Nécessité : un paramètre : le jeu de données x

Initialiser les N groupes à raison d'une donnée par groupe :

$G_i \leftarrow x_i$

Marquer tous les G_i comme "prenable"

Pour d de $N+1$ à $2N+1$ faire

Chercher les deux groupes prenables à fusionner :

G_i et G_j

Les fusionner : $G_d \leftarrow G_i \cup G_j$

Marquer G_i et G_j comme "prenable"

Marquer G_d comme prenable "prenable"

Fin pour

Algorithme de Classification hiérarchique ascendante (agglomérative)

Soit $X = \{x_1, x_2, \dots, x_N\}$ un jeu de données.

Initialisation :

Chaque observation constitue un cluster initial :

$C_i \leftarrow \{x_i\}, i=1, \dots, N$

Tous les clusters C_i sont marqués comme clusters actifs.

Itération :

Tant qu'il existe plus d'un cluster actif :

1. **Identifier les deux clusters actifs C_i et C_j présentant la plus faible dissimilarité**
2. **Fusionner ces deux clusters pour former un nouveau cluster**
 $C_d \leftarrow C_i \cup C_j$
3. **Marquer C_i et C_j comme clusters inactifs**
4. **Marquer C_d comme cluster actif**

Le processus se termine lorsqu'un seul cluster actif subsiste, formant la racine du dendrogramme.

Dans le cadre de l'octroi de financement, chaque ONG est considérée comme une observation décrite par des indicateurs de performance, de gouvernance et de capacité opérationnelle. La classification hiérarchique ascendante permet de regrouper les ONG en clusters homogènes, facilitant l'identification de profils distincts d'éligibilité au financement.

Début

Créer un groupe pour chaque ONG

Marquer tous les groupes comme PRENABLES

Tant qu'il reste plus d'un groupe prenable :

Trouver les deux groupes prenables les plus proches

Les fusionner en un nouveau groupe

Marquer les anciens groupes comme NON PRENABLES

Marquer le nouveau groupe comme PRENABLE

Fin

Principe : créer à chaque étape une partition obtenue en agrégeant deux à deux les éléments (individus ou groupe d'individus) les plus proches.

¹³Bourbaki, N., (2007) : *théorie des ensembles*, Springer Science & Business Media, Paris, 2007, pp. 1-40.

L'algorithme fournit une hiérarchique de partition à l'arbre contenant l'historique de la classification et permettant de retrouver N-1 partition.

Nécessité: il faut avoir une métrique (distance euclidienne), c'est-à-dire fixer à priori une règle pour agréger un individu et un groupe d'individus. Bref, il faut définir un critère d'agrégation.

Clustering Hiérarchique: Construit une hiérarchie de clusters, soit de manière agglomérative (en commençant par des points individuels et en les fusionnant progressivement) soit de manière divisive (en commençant par un seul cluster et en le divisant). Le résultat est un dendrogramme qui représente la structure des clusters à différents niveaux de granularité¹.

DBSCAN (Density-Based Spatial Clustering of Applications with Noise): Un algorithme basé sur la densité qui regroupe les points de données proches les uns des autres et marque comme bruit les points isolés dans des régions de faible densité¹.

Il ne nécessite pas de spécifier le nombre de clusters à l'avance.

Modèles de MélangeGaussiens (GMM): Supposent que les données sont générées à partir d'un mélange de plusieurs distributions gaussiennes. L'algorithme estime les paramètres de chaque gaussienne (moyenne, covariance) et attribue chaque point de données à la gaussienne la plus probable¹. La segmentation de données est un outil puissant pour comprendre la diversité au sein des ensembles de données et pour adapter les stratégies en conséquence. En utilisant des méthodes et des techniques appropriées, les organisations peuvent identifier des groupes distincts, comprendre leurs besoins et leurs comportements, et personnaliser leurs interactions pour obtenir de meilleurs résultats.

Cependant, il est crucial d'aborder la segmentation avec rigueur, en tenant compte de la qualité des données, des objectifs de l'analyse et des considérations éthiques pour éviter les biais et garantir une utilisation responsable de cette technique. La segmentation de données continue d'évoluer avec les avancées technologiques et l'augmentation de la disponibilité des données, offrant des opportunités toujours plus grandes pour une prise de décision éclairée et une personnalisation efficace.

Dans le point suivant, nous parlerons de la nouvelle approche constituant notre contribution à cette étude, en expliquant comment les ONG peuvent être classées en **K groupes** à partir de leurs scores.

NOUVELLE APPROCHE

Afin d'objectiver le processus d'allocation de financement aux ONG locales, cette étude a recours à des méthodes d'apprentissage non supervisé. Contrairement aux approches supervisées, ces méthodes ne nécessitent pas de variable cible prédéfinie et permettent d'identifier des structures latentes au sein des données.

Les ONG ont été décrites à l'aide d'indicateurs relatifs à leur expérience institutionnelle, leur capacité opérationnelle, leur couverture géographique, leur gouvernance financière et leurs résultats opérationnels. Après une phase de normalisation des données, une Analyse en Composantes Principales (ACP) a été appliquée afin de réduire la dimensionnalité et faciliter l'interprétation.

La segmentation a été réalisée à l'aide de l'algorithme K-means, qui a permis d'identifier trois groupes homogènes d'ONG. Le premier groupe regroupe des ONG matures et bien structurées, le second des ONG intermédiaires présentant un potentiel de développement, et le troisième des ONG fragiles nécessitant un accompagnement institutionnel.

Cette segmentation a ensuite été combinée à un système de scoring multicritère afin de soutenir la prise de décision en matière d'octroi de financement. L'approche proposée renforce ainsi la transparence, l'équité et l'efficacité des décisions de financement dans les projetshumanitaires.

ID	Acronyme	Nom Partenaire d'implémentation	Type d'organisation	Cluster	Existence juridique	Opérationnalité	Adresse physique	Capacité de gestion	Compte bancaire	Personnel clé	Expérience	Note (0-100)	Pourcentage (% ID (clusters))
1	AACDS	Aide et Action à la Coopération au développement et la Solidarité	ONG Nationale	Sécurité alimentaire	1	1	1	1	0	0,5	0	0,675	67,5 Cluster4
2	ABCOM	Action pour le Bien-Être Communautaire	ONG Nationale	Protection de l'enfant	1	1	1	1	0	0,5	0	0,675	67,5 Cluster4
3	ACDI	Actions Civiliennes pour le Développement Intégral	ONG Nationale	Protection de l'enfant	1	1	1	1	0,5	0	0,775	77,5 Cluster1	
4	ACDR	Assistance Communautaire pour le Dialogue Réconciliation	ONG Internationale	Sécurité alimentaire	1	1	1	1	0	0,5	0	0,675	67,5 Cluster4
5	ACE	Appel au Cœur l'Enfant Africain	ONG Nationale	Hygiène et assainisseme	1	1	1	1	0,5	0	0,775	77,5 Cluster1	
6	ACDPE	Actions Concrètes pour la Protection de l'Enfance	ONG Nationale	Protection de l'enfant	1	1	1	1	0,5	0	0,775	77,5 Cluster1	
7	ACDOI	Actions de Communautés Paysannes pour le Développement Intégré	ONG Internationale	Sécurité alimentaire	1	1	1	1	1	0,5	0	0,775	77,5 Cluster1
8	ACTED	ACTED-Agence d'aide à la coopération technique et au développement	ONG Internationale	Protection	1	1	1	1	0,5	1	0,925	92,5 Cluster2	
9	AKCTO	Agence d'Aide à la Coopération Technique et au Développement	ONG Internationale	Articles ménagers essen	1	1	1	1	1	0,5	1	0,925	92,5 Cluster2
10	APROBES	Actions pour le Développement du Bien Être Social	ONG Nationale	Sécurité alimentaire	1	1	1	1	1	0,5	1	0,925	92,5 Cluster2
11	SDC	SDC	ONG Nationale	Wauash	1	0	1	1	1	0,5	1	0,725	72,5 Cluster1
12	CARITAS KANI CARITAS KANANGA		ONG Nationale	Sécurité alimentaire	1	0	1	0	0	0,5	0	0,275	27,5 Cluster3
13	ADDP	Action pour le développement durable des populations	ONG Nationale	Abiris	1	0	1	1	1	0,5	1	0,725	72,5 Cluster1
14	ASDECO	ASSOCIATION POUR LE DEVELOPPEMENT COMMUNAUTAIRE	ONG Nationale	Protection de l'enfant	1	0	1	1	1	0,5	1	0,725	72,5 Cluster1
15	ADRES	Agence de Développement Economique et Social	ONG Internationale	Nutrition	1	0	1	1	1	0,5	1	0,725	72,5 Cluster1
16	ADRC	Actions pour le Développement Intégral par la Conservation Communautaire	ONG Nationale	Protection Lutte anti-mi	1	0	1	0	1	0,5	1	0,525	52,5 Cluster4
17	ADIRA	Adventist development relief and agency	ONG Internationale	Santé	1	0	1	0	0	0,5	1	0,425	42,5 Cluster4
18	LUMIERE DU « LUMIERE DU KASAI		ONG Nationale	Sécurité alimentaire	1	0	1	1	1	1	1	0,8	80 Cluster1
19	AIS	Armée Du Salut	ONG Nationale	Éducation	1	0	1	1	1	1	1	0,8	80 Cluster1
20	ADISSE	Association pour le Développement Social et la Sauvegarde de l'Environnement	ONG Nationale	Sécurité alimentaire	1	0	1	1	1	1	1	0,8	80 Cluster1
21	AFPD	Action pour l'encadrement de filles mères désecurisées	ONG Internationale	Santé	1	0	1	1	1	1	1	0,8	80 Cluster1
22	ACDOE	ACDOE	ONG Nationale	Santé	1	0	1	1	1	1	1	0,8	80 Cluster1
23	AROCNA	Assistance aux Enfants abandonnés et Orphelins,Centres Nutritionnels et Aliment	ONG Nationale	Nutrition	1	0	1	1	1	1	1	0,8	80 Cluster1
24	AFYFE	Association des Femmes pour la Promotion et le Développement Endogène	ONG Nationale	Nutrition	1	0	1	0,5	1	1	1	0,7	70 Cluster1
25	APADI	Action Humanitaire pour l'Assistance et le Développement Intégré	ONG Nationale	Nutrition	1	0	1	0,5	1	1	1	0,7	70 Cluster1
26	AMUSADEC	Action Humanitaire pour la santé et le développement communautaire	ONG Nationale	Éducation	1	0	1	0,5	1	1	1	0,7	70 Cluster1
27	AIDES	Actions et interventions pour le Développement et l'Enseignement Social	ONG Nationale	Abiris	1	0	1	0,5	1	1	1	0,7	70 Cluster1
28	AIDE	Actions et Interventions pour le Développement et l'Enseignement	ONG Nationale	Sécurité alimentaire	1	0	1	0,5	1	1	1	0,7	70 Cluster1
29	AIRD	African Initiative for Relief and Development	ONG Internationale	Abiris	1	0	1	0,5	1	1	1	0,7	70 Cluster1
30	AIEDEC	Association des Jeunes pour le Développement Communautaire	ONG Nationale	Protection de l'enfant	1	0	1	0,5	1	1	1	0,7	70 Cluster1
31	AJD	Association des Jeunes Islamiques pour le développement	ONG Nationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
32	ALDI	Association Locale pour le Développement Intégral	ONG Nationale	Sécurité alimentaire	1	0	1	0,5	1	1	1	0,7	70 Cluster1
33	ALIMA	The Alliance for International Medical Action	ONG Nationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
34	AMAIN	Action pour la mobilisation et l'atténuation de l'impact négatif	ONG Nationale	Protection de l'enfant	1	0	1	0,5	1	1	1	0,7	70 Cluster1
35	AMOR	Association des Amis du Monde Rural	ONG Nationale	Sécurité alimentaire	1	0	1	0,5	1	1	1	0,7	70 Cluster1
36	AMUD	Association multisectoriel pour les urgences et le développement durable	ONG Nationale	Protection de l'enfant	1	0	1	0,5	1	1	1	0,7	70 Cluster1
37	ADF	Action Of the Future	ONG Nationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
38	AP	Ami des Personnes-ONG	ONG Nationale	Nutrition	1	0	1	0,5	1	1	1	0,7	70 Cluster1
39	APANWIP	Action de Promotion et d'Assistance pour l'Amélioration du Niveau de Vie des P	ONG Internationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
40	APDR	Association Paysanne pour le Développement Rural	ONG Nationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
41	AFPEI	Amis des Personnes en Détresse	ONG Nationale	Nutrition	1	0	1	0,5	1	1	1	0,7	70 Cluster1
42	APFEFFVU	Action pour la protection et l'encadrement des filles et femmes vulnérables	ONG Nationale	Santé	1	0	1	0,5	1	1	1	0,7	70 Cluster1
43	APES	Action pour la Promotion de l'Environnement Social	ONG Nationale	Protection de l'enfant	1	0	1	0,5	1	1	1	0,7	70 Cluster1
44	APROSHAV	Action pour la Protection de la Santé Humaine, Animale et Végétale	ONG Nationale	Sécurité alimentaire	1	0	1	1	1	1	1	0,8	80 Cluster1
45	APROSHAC	Action pour la Protection de la Santé Humaine, Animale et Culture	ONG Nationale	Abiris	1	0	1	1	1	1	1	0,8	80 Cluster1
46	APFSA	ACTION POUR LA PROMOTION SOCIALE DE LA FEMME ET DE L'ENFANT	ONG Nationale	Protection de l'enfant	1	0	1	1	1	1	1	0,8	80 Cluster1
47	APYTDV	Association des Personnes Victimes des Troubles, Déplacées et Vulnérables de l'E	ONG Nationale	Protection de l'enfant	1	0	1	0,5	1	1	1	0,7	70 Cluster1
48	AIRD-ROC	Association de Réveil pour le Développement Endogène	ONG Nationale	Hygiène et assainisseme	1	0	1	0,5	1	1	1	0,7	70 Cluster1
49	AS	Amika Sasa	ONG Nationale	Nutrition	1	0	1	0,5	1	1	1	0,7	70 Cluster1
50	ASDC	ACTION SOCIALE POUR LE DEVELOPPEMENT COMMUNAUTAIRE	ONG Nationale	Protection de l'enfant	1	1	1	0,5	0	1	1	0,9	90 Cluster2
51	ASDI	Actions et Soutiens au Développement Intégral	ONG Nationale	Nutrition	1	1	1	0,5	0	1	1	0,8	80 Cluster1
52	AVIEO	Association des Vulnérables pour la Réintégration des Enfants Orphelins Abando	ONG Internationale	Protection de l'enfant	1	0	1	0,5	0	1	1	0,8	80 Cluster1
53	AVSI	Associazione Volontari per il Servizio Internazionale	ONG Internationale	Éducation	1	0	1	0	0	1	1	0,5	50 Cluster4
54	AVS	Associazione Volontari per il Servizio	ONG Internationale	Nutrition	1	0	1	0	1	1	1	0,6	60 Cluster4
55	ASBEKA	Association pour le Bien Être Social	ONG Nationale	Sécurité alimentaire	1	0	1	0	1	1	1	0,6	60 Cluster4
56	BADU	Bureau des Actions de Développement et des Urgences	ONG Nationale	Protection Lutte anti-mi	1	0	1	0	1	1	1	0,6	60 Cluster4
57	BODOM	Bureau D'opératins des secours Médicales	ONG Nationale	Nutrition	1	1	1	1	1	1	1	1	100 Cluster2
58	CBK	Communauté Baptiste au Kivu	ONG Nationale	Nutrition	1	1	1	1	1	1	1	1	100 Cluster2
59	CACSG	CACSG	ONG Nationale	Violences basées sur le s	1	1	1	1	1	1	1	1	100 Cluster2
60	CADERGA	Centre d'Appui au Développement Rural et à la Sécurité Alimentaire	ONG Nationale	Nutrition	1	0	1	1	1	1	1	0,8	80 Cluster1
61	CADMIR	CADMIR	ONG Internationale	Sécurité alimentaire	1	0	1	1	1	1	1	0,8	80 Cluster1
62	CARE	Care International	ONG Internationale	Nutrition	1	0	1	1	1	1	1	0,8	80 Cluster1
63	CARITAS	CARITAS	ONG Nationale	Sécurité alimentaire	1	0	1	1	1	1	1	0,8	80 Cluster1
64	CARITAS BUNI CARITAS BUNGA		ONG Nationale	Santé	1	0	1	1	1	1	1	0,8	80 Cluster1
65	CARITAS CONI CARITAS CONGO Ajaï		ONG Nationale	Protection de l'enfant	1	0	1	1	1	1	1	0,8	80 Cluster1
66	Caritas Dévelup Caritas Développement		ONG Nationale	Santé	1	0	1	1	1	1	1	0,8	80 Cluster1

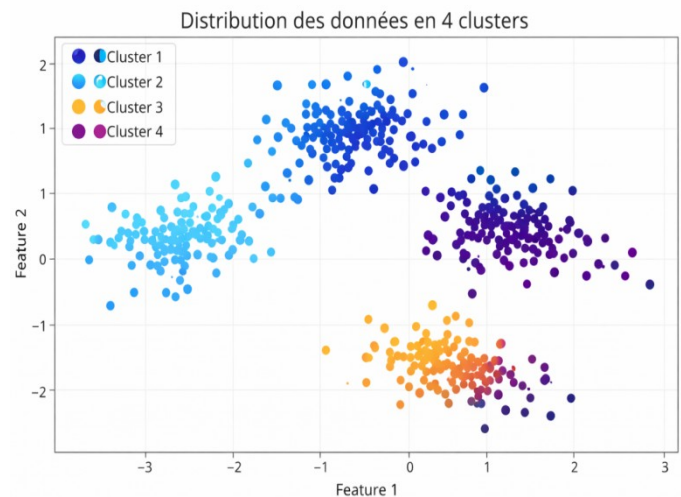
ALGORITHME DE SCORING**DEBUT***Pour chaque ONG dans la liste :**Lire :**existence_juridique**operationnalite**adresse_physique**capacite_gestion**compte_bancaire**personnel_cle**experience_anterieur**Calculer SCORE = (**existence_juridique +**operationnalite +**adresse_physique +**capacite_gestion +**compte_bancaire +**personnel_cle +**experience_anterieur**)**Calculer POURCENTAGE = (SCORE / SCORE_MAX) * 100**SI POURCENTAGE >= 75**DECISION = "ELIGIBLE"**SINON SI POURCENTAGE >= 60**DECISION = "CONDITIONNEL"**SINON SI POURCENTAGE >= 40**DECISION = "NON ELIGIBLE"**SINON**DECISION = "REJETE"**FIN POUR**FIN*

L'algorithme analyse chaque ONG individuellement en quatre étapes:

1. Collecte des Données (Variables) : Attribution d'une note précise pour chaque capacité testée :
 - 1 : Validé (Oui)
 - 0,5 : Partiellement validé
 - 0 : Non validé
2. Calcul du Score Global: Addition de l'ensemble des points obtenus. Ce total définit le niveau de robustesse de l'organisation.
3. Conversion en Pourcentage: Transformation du score brut en valeur relative pour faciliter l'interprétation et la comparaison.
4. Classification : Application de seuils critiques pour déterminer la catégorie de l'ONG :
 - Excellent (> 75%) **Éligible** (Validation immédiate).
 - Correct (60-74%) **Conditionnel** (À surveiller/renforcer).
 - Insuffisant (40-59%) **Non Éligible** (Échec technique).
 - Critique (< 40%) **Rejeté** (Dossier écarté).

ALGORITHME DE SEGMENTATION (K-MEANS)

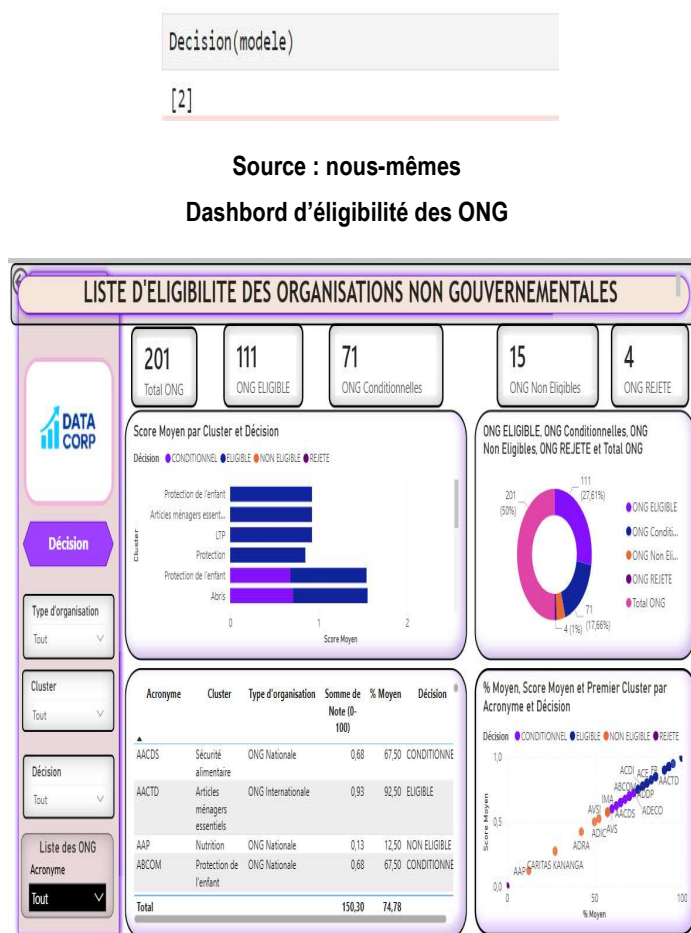
Cet algorithme permet de regrouper automatiquement les ONG similaires et ici il n'y a pas de décision au départ.

DEBUT*// Choisir le nombre de groupes**Lire K**Normaliser ONG dans ONG_normalise**Initialiser aléatoirement les K centres**REPETER**ANCIENS_CENTRES ← CENTRES**// Affectation**Pour chaque ONG i**Affecter i au centre le plus proche**FinPour**// Mise à jour des centres**Pour chaque groupe k**CENTRES[k] ← moyenne des ONG du groupe k**FinPour**JUSQU'A CENTRES = ANCIENS_CENTRES**Afficher GROUPE**FIN***Source: nous mêmes****2. Arbre de décision (Classification)****DEBUT***// Initialisation du traitement pour l'ensemble des partenaires**POUR CHAQUE ONG DANS LA LISTE :**// --- 1. LECTURE DES DONNÉES ---**Lire : existence_juridique, operationnalite, adresse_physique, capacite_gestion, compte_bancaire, personnel_cle, experience_anterieur**// --- 2. CALCUL DU SCORE GLOBAL ---**// L'algorithme additionne les points de chaque critère technique*

SCORE <- (existence_juridique + operationnalite +
adresse_physique +
capacite_gestion + compte_bancaire + personnel_cle +
experience_anterieur)

*// Conversion du score en pourcentage pour la normalisation**POURCENTAGE <- (SCORE / SCORE_MAX) * 100**// --- 3. CRITERE DE CLASSIFICATION ---**// Attribution de la décision selon les seuils métiers définis**SI (POURCENTAGE >= 75) ALORS**DECISION <- "ELIGIBLE"**SINON SI POURCENTAGE >= 60 ALORS**DECISION <- "CONDITIONNEL"**SINON SI POURCENTAGE >= 40 ALORS**DECISION <- "NON ELIGIBLE"**SINON**DECISION <- "REJETE"**FIN SI**// Affichage du résultat pour l'ONG en cours**Afficher "ONG : ", Acronyme, " | Score : ", POURCENTAGE, "%»**| Décision : ", DECISION**FIN POUR // Fin de la boucle de parcours de la liste**FIN*

Tableau 6. : Affichage de résultat de la prédiction



Source: nous mêmes

CONCLUSION

Au terme de cette étude, il ressort que l'évaluation des capacités des Organisations Non Gouvernementales (ONG) constitue un enjeu important pour améliorer l'efficacité des programmes humanitaires et de développement. Face à la diversité des ONG et à leurs différents niveaux de capacités organisationnelles et opérationnelles, cette recherche a proposé une approche analytique fondée sur plusieurs critères liés notamment à la gouvernance, aux ressources humaines, aux capacités financières et à l'expérience opérationnelle.

L'attribution de scores a permis d'évaluer le niveau de performance et d'éligibilité des ONG, tandis que l'algorithme K-means a permis de les regrouper en classes homogènes. Cette approche contribue ainsi à rendre le processus de sélection de partenaires plus objectif, transparent et rationnel, tout en constituant un outil d'aide à la décision pour les institutions publiques, les agences d'exécution et les partenaires techniques et financiers.

En perspective, cette démarche pourrait être renforcée par la mise en place d'un système intégré de suivi et d'évaluation des ONG et des projets, reposant sur une infrastructure de Cloud computing. La centralisation et le partage des données opérationnelles, financières et de performance permettraient un suivi en temps réel et faciliteraient la prise de décision. À terme, cette approche pourrait contribuer au développement d'un système numérique intégré favorisant la transparence, la redevabilité et l'efficacité des interventions de développement.

REFERENCES

1. Ouvrages

- Borja SOTOMAYOR et al., (2009): "Virtual infrastructure management in private and hybrid clouds". In : Internet computing, IEEE 13.5.
- Bourbaki, N., (2007): théorie des ensembles, Springer Science & Business Media, Paris.
- Breiman, L. (2001). Random forests. Machine learning, 45(1),
- Chambers R., (1997): Whose Reality Counts? Putting the first Last, London.
- Christian B., (2011), cloud computing: Web-based Dynamic IT Services, 2ème Ed. US, livre.
- Cortes, C., & Vapnik, V. (1995): Support-vector networks. Machine learning, 20(3)
- Darcy, J., & Hofmann, (2003): C. A. According to need? Needs assessment and decision- making in the humanitarian sector. Overseas DevelopmentInstitute. (Souligne le rôle crucial de l'information partagée).
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977) : Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal 4 Statistical Society. Serie (Methodological), 5 39(1),
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996): A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the 2 second international conference on knowledge discovery and data mining.
- Fisher, R. A. (1936): The use of multiple measurements in taxonomic problems. Annals of eugenics, 7(2),
- Jain A. K. (2010): Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8)
- Kaplan & Norton, (1996): The Balanced scorecard: Translating Strategy into action. 1ère Ed. Boston MA: Havard Business School Press.
- Kaplan, R. S. & Norton, D. P., (1996): The Balanced Scorecard: Translating Strategy into Action, Harvard Business School Press, Boston,
- Kerzner H., (2017): Project Management: Q systems Approach to planning, scheduling and controlling, 12è éd., chicago.
- KRUTZ, Ronald L. et VINES Russell D.,(2010) : Cloud Security: A Comprehensive Guide to Secure Cloud Computing. Indianapolis: Wiley Publishing.
- MacQueen, J. (1967): Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability Vol. 1 (1)
- MAGOULES F., (2009): Fundamentals of Grid Computing: Theory, Algorithms and Technologies, Boca Raton: CRC Press.
- MARSHALL D.; MCFERRIN Wade A., et WILLIAMS K. (2010):Advanced Server Virtualization: VMware and Microsoft Platform in the Virtual Data Center. Boca Raton: CRC Press.
- McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426. 8
- Mintzberg H., (1982): Structure et dynamique des organisations, Paris, éd. Organisation. ISBN.
- Mosse D., (2005): Cultivating Development : An Ethnography of Aid Policy and Practice, 1ère éd., London Pluto Press.
- Mosse, D., (2005): Cultivating Development: An Ethnography of Aid Policy and Practice, Pluto Press, Londres.

- Murtagh, F. (1983): A survey of recent advances in hierarchical clustering algorithms. The Computer Journal, 26(4),
- Ortiz Jr., S. (2011). The problem with cloud-computing standardization. Computer, Los Alamitos, Vol., 44, n°7.
- Rutkowska, J. (2006). Introducing the Blue Pill (présentation à la conférence Black Hat Briefings, Las Vegas, 3 août 2006). Invisible Things Lab
- Sen, A., (1999) : Development as Freedom, Oxford University Press, Oxford.
- Trend M., (2011): Cloud Security Survey - Global Executive Summary. Trend Micro Inc.report.
- Van Wijk G., & Klijn, E. H.(2010): Network management in the public sector: Past achievements and future challenges. Public Management Review, vol.12, n°1,
- Viktors B., (2002):Fundamentals of Grid Computing. IBM Redbook
- Spearman, C. (1904):. "General intelligence," objectively determined and measured. The American Journal of Psychology, 15(2)
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. Journal of Machine Learning Research, 9(Nov).

2. Thèses

- Pierre M. (2014): l'impact des pratiques de mutualisation sur la performance des services départementaux d'incendie et de secours. Thèse soutenue en France (Université de Pau),
- Benamar H. (2007): ("Etude exploratoire d'outils pour le data mining", thèse soutenue au Québec
- SAFA El. Ayeb (2019): « développement d'un outil de segmentation des comportements d'achat des clients en se basant sur leurs données morphologiques », thèse soutenue en France (Polyclinique Monterial).
- Marine D. (2010): Méthodes d'apprentissage statistique pour le Scoring, thèse soutenue à l'Université Télécom Paris Tech.

3. Source internet

- https://fr.wikiversity.org/wiki/Cloud_computing_et_entreprise/ , consulté 10 avril 2024 à 15h 30'
- <http://blog.3li.com/les-caracteristiques-essentielles-du-cloud/>, 10 avril 2024 à 16 h 35'
- <https://www.cloud-council.org/deliverables/CSCC-Practical-Guide-to-Cloud-Service-Level-Agreements.pdf>, consulté 12 mai à 11h 20'
- <http://aws.amazon.com/importexport/>, consulté le 15 mai 2024 à 16H 10'
- Stroud F., cloudstack, Webodia, 17 Avril 2024 disponible sur <https://www.webodia.com/definition/cloudstack>, consulté le 18 Mai 2014 à 16h 30'
- <http://archives.opennebula.org/detail/documentation:rel1.2:one-architecture>, consulté le 22 mai 2026 à 11h 20
- <https://www.ibm.com/cloud/learn/data-segmentation>, consulté le 03 juin 2026 à 13h 17

4. Rapport

- World Bank, the world Bank annuel report 2018, DC: international Bank for reconstruction and developpement, Washington(USA)

5. Journals

- Wells, W. D. (1975). Psychographics: A critical review. Journal of Marketing Research, 12(3), 196-213.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 62(11)