



Full Length Research Paper

Classification des apprenants en fonction des facteurs socio-économiques et environnementaux influençant leurs réussites dans un dispositif de formation à distance.

Ameth THIAM ^{1*}, Ndéye Massata NDIAYE ², Gaoussou CAMARA ¹¹Université Alioune Diop de Bambey (UADB), – Bambey, Sénégal²Université Numérique Cheikh Hamidou Kane (UNCHK) -- Dakar, Sénégal

Received September 2023 – Accepted November 2023

*Corresponding author. ameth.thiam@uadb.edu.sn

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License.

Résumé:

Fournir une formation de qualité et accessible à tous les diplômés du postsecondaire est un défi majeur pour tous les pays en voie de développement. Pour faire face à cette situation, beaucoup de pays font recours à la formation à distance. Cependant, cette nouvelle façon de dispenser des compétences et des connaissances est caractérisée par la rupture de l'unité de lieu contrairement à la formation en présentiel. À la différence des universités classiques où les étudiants partagent souvent le même environnement dans les campus sociaux, dans les dispositifs de formation à distance, les étudiants sont dans des endroits différents avec souvent des facteurs socio-économiques et environnementaux différents. Ces facteurs peuvent impacter sur les résultats des apprenants. L'objectif de notre article est de faire une classification non supervisée des apprenants d'un dispositif de formation à distance en fonction des facteurs socioéconomiques et environnementaux en utilisant la méthode K-means afin de déterminer des groupes d'étudiants identiques, complémentaires ou particuliers. Les données utilisées dans cette étude proviennent d'une enquête réalisée au cours de l'année académique 2021-2022 auprès de 236 étudiants de la licence 1 au master 2 de l'Université Numérique Cheikh Hamidou Kane (UNCHK). Les résultats de l'étude nous ont permis d'identifier cinq (5) groupes d'étudiants différents en fonction de leurs facteurs socio-économiques et environnementaux. Les facteurs les plus déterminants sont les facteurs financiers (avec ou sans source de revenu), les facteurs liés à l'habitation (nombre de personnes par chambre) et les facteurs liés à leurs disponibilités à suivre leurs enseignements (avec ou sans une autre formation en hors de l'UNCHK et avec ou sans des activités domestiques et ou professionnelles)..

Mots clés/Keyword: Formation à distance, classification, K-Means, facteurs socio-économiques, facteurs environnementaux, réussite.

Cite this article:

Ameth THIAM, Ndéye Massata NDIAYE, Gaoussou CAMARA. (2023). Classification des apprenants en fonction des facteurs socio-économiques et environnementaux influençant leurs réussites dans un dispositif de formation à distance. Revue RAMReS – Sci. Appl. & de l'Ing., Vol. 5(1), pp. 107-112. ISSN 2630-1164.

1. Introduction

Face à la demande d'accès à une formation de qualité postsecondaire qui ne cesse de croître d'année en année et le nombre très réduit de places qu'offre le système d'enseignement supérieur de la plupart des pays en voie de développement, l'orientation des bacheliers pose un réel problème. Pour faire face à cette situation, certains pays comme le Sénégal, la Côte d'Ivoire ou encore la Tunisie optent pour des formations à distance. Cependant, il faut noter qu'avec la formation à distance, l'interaction homme-homme est remplacée par une interaction homme-machine. Cette situation

peut constituer un facteur bloquant pour la réussite des apprenants. D'ailleurs, les auteurs de (Vidal, Grandbastien, et Mœglin 2010) considèrent la séparation des apprenants et des enseignants comme un handicap. De plus, les résultats (taux d'achèvement c'est-à-dire ceux qui parviennent à suivre leur formation jusqu'à la fin) obtenus avec une formation en présentiel sont meilleurs que ceux obtenus avec des formations à distance avec souvent des taux d'abandon très élevés selon (Dussarps 2015). En effet, contrairement aux universités classiques où les étudiants partagent le même environnement dans les

campus universitaires, dans les dispositifs de formation à distance, les apprenants se trouvent dans des endroits différents avec des facteurs socio-économiques et environnementaux très diversifiés. Ces facteurs peuvent impacter sur les résultats et les taux d'abandon des apprenants. Selon (Papi et al. 2022a), les attentes et le niveau de satisfaction ne dépendent pas seulement des interactions et rétroactions entre les apprenants et les enseignants mais aussi de la situation socio-économique, environnementale et professionnelle. Dans cet article, nous ferons dans un premier temps une présentation des données et la méthodologie utilisées dans le cadre de cette recherche. Ensuite, dans un second temps nous présentons les résultats obtenus avant de terminer par la conclusion et les perspectives.

2. Présentation des données et de la méthodologie.

2.1. Les données d'étude.

Nos travaux se sont réalisés à l'Université Numérique Cheikh Hamidou Kane (UNCHK) qui est le sixième établissement public d'enseignement supérieur du Sénégal avec plus de 60 000 apprenants et plus de 40 parcours de formations qui se déroulent essentiellement en ligne. Une enquête a été réalisée pour recueillir des données sur les facteurs socio-économiques et environnementaux des apprenants. Un questionnaire en ligne de 18 questions soumis à tous les étudiants de l'UNCHK (de la licence 1 au master 2) pendant un mois nous a permis de récolter 236 réponses. Notre dataset est constitué de 2 variables quantitatives et de 16 variables qualitatives. Après la phase de collecte des données, nous avons procédé à un prétraitement des données notamment l'encodage des données qualitatives, la standardisation, et la correction des données manquantes et ou aberrantes (52 individus ont été supprimés).

2.2. Méthodologie de recherche.

Pour pouvoir visualiser les résultats sur un plan à deux dimensions, nous avons réduit les variables de notre dataset à deux composantes. Pour gérer les problèmes de plusieurs variables, principalement deux approches existent dans la littérature statistique. L'une est la sélection de variable et l'autre la réduction de dimension selon (Ma et Zhu 2013). Dans cet article nous avons procédé pas une réduction de dimensions qui nous a permis de conserver toutes nos variables. Nous avons dans un premier temps fait une analyse à composantes principales linéaires (figure 1) et dans un second temps une analyse à composantes principales non linéaire (figure 2). Ensuite, nous avons choisi l'analyse en composantes principales linéaires car elle nous donne un meilleur résultat.

3. Résultats.

3.1. Réduction de dimension et classifications.

Pour pouvoir visualiser les résultats d'analyse de notre dataset multidimensionnel, nous avons utilisé l'analyse à composantes principale pour simplifier qui nous a

permis d'avoir le résultat de la figure 1 avec un Coefficient de corrélation = $4.85722573273506e-17$.

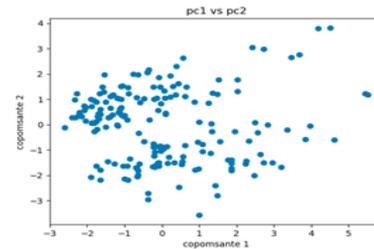


Figure 1 : Analyse de composantes principales linéaires

Nous avons également fait une analyse à composantes non linéaire qui nous a permis de faire une réduction de dimensions de façon non linéaire (figure 2) contrairement à l'analyse à composantes linéaire avec un Coefficient de corrélation $-4.753142324176451e-16$

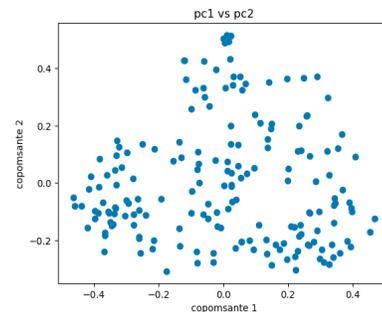


Figure 2 : Analyse de composantes principales non linéaires

La détermination du nombre de clusters est l'une des tâches les plus complexes quand on utilise la méthode K-means. Dans cette étude, nous avons utilisé la méthode Elbow (Syakur et al. 2018) pour déterminer le nombre de clusters. Cette méthode consiste à tracer l'évolution du coût de notre modèle en fonction du nombre de clusters et de déterminer dans ce graphique une zone de coude. Cette zone de coude permet de déterminer le nombre de clusters optimal c'est-à-dire déterminer le nombre de clusters qui nous permet de réduire le coût de notre modèle tout en conservant un nombre raisonnable de clusters.

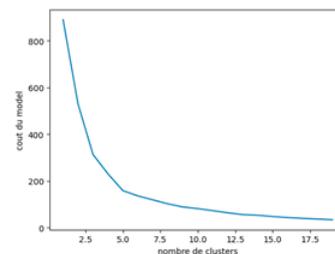


Figure 3 : Détermination du nombre de clusters avec la méthode Elbow

Après avoir déterminé le nombre optimal de clusters, nous avons appliqué une classification non supervisée

en utilisant la méthode des K-Means qui nous donne le résultat à la figure 4.

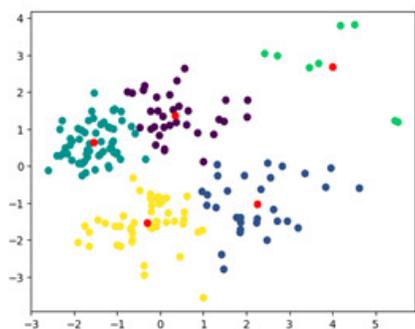


Figure 4 : Classification des apprenants avec la méthode des K-Means

3.2. Taux de représentativité de chaque groupe

Après avoir constitué les cinq groupes d'apprenants, nous constatons une représentativité inégale au niveau des groupes. Cette situation nous a amené à déterminer les taux de représentativité de chaque groupe (Tableau 1). Cette première analyse nous a permis de constater qu'une bonne partie des étudiants sont dans le groupe 1 soit 33,70%. Ensuite on a le groupe 5 avec 27,17%, le groupe 4 avec 18,48%, le groupe 2 avec 16,30% et le groupe 3 avec le plus petit effectif de 4,35%.

Tableau 1.: Répartition des apprenants dans les 5 groupes

Groupe	Effectif	Pourcentage
1	62	33,70%
2	30	16,30%
3	8	4,35%
4	34	18,48%
5	50	27,17%

3.3. Caractéristiques des individus de chaque groupe

Après avoir étudié le taux de représentativité (tableau 1), nous avons déterminé les caractéristiques des apprenants de chaque groupe. Chaque groupe étant représenté par 18 attributs, nous avons procédé par une segmentation des caractéristiques pour plus de clarté. Ainsi, on se retrouve avec quatre sous-groupes : un sous-groupe pour les facteurs économiques, un sous-groupe pour les facteurs liés à la disponibilité des apprenants à suivre les apprentissages, un sous-groupe pour les facteurs sociaux et un sous-groupe pour les facteurs environnementaux.

Tableau 2. – Les facteurs économiques

Variabes	Valeur	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Avez-vous une bourse d'études, aide ou subvention ?	Oui	0,00%	100,00%	25,00%	2,94%	100,00%
	Non	100,00%	0,00%	75,00%	97,06%	0,00%
Quel Type de Bourse, aide ou subvention avez-vous ?	Aide	0,00%	3,33%	12,50%	0,00%	10,00%
	Bourse entière	0,00%	63,33%	12,50%	2,94%	70,00%
	Bourse Excellence	0,00%	0,00%	0,00%	0,00%	2,00%
	Demi-bourse	0,00%	33,33%	0,00%	0,00%	18,00%
	Rien	100,00%	0,00%	75,00%	97,06%	0,00%
Avez-vous l'aide de vos parents ou proches ?	Oui	38,71%	43,33%	0,00%	35,29%	34,00%
	Non	61,29%	56,67%	100,00%	64,71%	66,00%
Avez-vous une autre source de revenu ?	Oui	6,45%	36,67%	75,00%	17,65%	6,00%
	Non	93,55%	63,33%	25,00%	82,35%	94,00%

Plus des 3/4 des étudiants des groupe 1, 3 et 4 sont des étudiants qui n'ont ni bourse d'étude ni aide. De plus, 61,29% des étudiants du groupe 1 n'ont pas d'aide de leurs parents et 93,55% d'entre eux n'ont pas d'autres sources de revenu. Les étudiants du groupe 4 sont dans les mêmes situations que les étudiants du groupe 1, 64,71% n'ont pas l'aide de leurs parents et 82,35% d'entre eux n'ont pas d'autres sources de revenu. Quant aux étudiants du groupe 3, même si aucun d'entre eux

n'a l'aide de leurs parents, 75,00% d'entre eux ont d'autres sources de revenu. Les étudiants des groupes 2 et 5 sont des apprenants qui ont tous au moins une aide. De plus, 94% des apprenants du groupe 5 ont en plus de leurs bourses ou aide d'autres sources de revenu et 34% ont l'aide de leurs parents. Les apprenants du groupe 2, en plus de leur bourse ou aide, 43,33% d'entre eux ont l'aide de leurs parents et 36,67% ont d'autres sources de revenu. En somme, on peut en

déduire que les apprenants des groupes 1 et 4, soit 52,18%, seraient dans des situations financières

difficiles contrairement aux apprenants des groupes 2, 3 et 5.

Tableau 3. – Les facteurs sociaux

Variabes	Valeurs	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Sexe	Homme	41,94%	56,67%	62,50%	73,53%	54,00%
	Femme	58,06%	43,33%	37,50%	26,47%	46,00%
Age	Minimum	19	20	25	19	20
	Moyenne	23	26	36	23	25
	Maximum	30	30	48	31	30
Situation matrimoniale	Célibataire	100,00%	80,00%	37,50%	76,47%	96,00%
	Marié	0,00%	20,00%	62,50%	20,59%	0,00%
	Divorcé	0,00%	0,00%	0,00%	2,94%	4,00%
Niveau d'étude de la femme ou mari	Aucun	Pas de Marié	0,00%	0,00%	14,29%	Pas de Marié
	Elémentaire		16,67%	0,00%	0,00%	
	Moyen		16,67%	20,00%	14,29%	
	Secondaire		33,33%	20,00%	14,29%	
	Supérieur		33,33%		60,00%	
Quelle est la taille de la famille	Famille moyenne (entre 6 et 4 personnes vous y compris)	11,29%	23,33%	25,00%	32,35%	22,00%
	Famille restreinte	0,00%	0,00%	12,50%	0,00%	0,00%
	Grande famille (avec plus de 6 personnes vous y compris)	88,71%	70,00%	0,00%	58,82%	74,00%
	Petite famille (avec moins de 4 personnes vous y compris)	0,00%	6,67%	37,50%	8,82%	4,00%
	Seul(e)	0,00%	0,00%	25,00%	0,00%	0,00%

La variable sexe n'est pas très déterminant. Dans l'échantillon, on a 46% de femme et 64% d'homme. Dans tous les groupes on a plus d'hommes que de femmes à l'exception les étudiants du groupe 1 avec 58% de femmes. Plus du ¼ des apprenants des groupes 1, 2, 4 et 5 sont des célibataires et sont dans la même tranche d'âge (19 à 30 ans). Cependant, 62,5% des apprenants du groupe 4 sont mariés et leurs âges varient entre 25 et 48 ans. Les étudiants mariés ont des partenaires instruits dont 88,89% ont au moins atteint le cycle moyen. En plus, on note également que plus de 58% des apprenants des groupes 1, 2, 4 et 5 vivent dans de grandes familles (avec plus de six personnes). Par contre, les apprenants du groupe 4 en grande majorité vivent dans de petites familles. 37,5% d'entre eux vivent dans des familles avec moins de quatre personnes et 25% vivent seul.

La majeure partie des apprenants vivent en milieu urbain. Néanmoins, il faut signaler que 48,39% des étudiants du groupe 1 et 42% du groupe 5 vivent en milieu rural. 63,33% des étudiants du groupe 2 et 75,00% du groupe 3 disposent d'une chambre individuelle et seulement la moitié des étudiants du groupe 4 ont des chambres individuelles. Par contre, pour les étudiants des groupes 1 et 5, ils vivent le plus

souvent dans leurs chambres avec d'autres personnes (90,32% du groupe 1 et 86,00% du groupe 5) avec des effectifs qui peuvent aller parfois jusqu'à 7 personnes par chambre.

Plus de la moitié des étudiants des groupes 1, 2, 3 et 5 ont des activités domestiques ou professionnelles familiales qui peuvent constituer un blocage pour la réussite. Cependant, seulement 38,24% des étudiants du groupe 4 ont des activités. En plus, au moins la moitié des apprenants des groupes 1, 2, 4 et 5 trouvent que l'appui en connexion Internet est insuffisant alors qu'une bonne partie des étudiants du groupe 3, 62,50% soutiennent le contraire. Les apprenants des groupes 1 et 5, en grande majorité (plus de 95%), suivent exclusivement leurs formations qu'à l'UNCHK alors que plus de 60% des apprenants des groupes 2 et 3, en plus de leurs formations à l'UNCHK, suivent d'autres formations. 77,78 % des formations que les étudiants suivent à l'extérieur sont liées à celles suivies à l'UNCHK. Plus de la moitié des apprenants ont des activités externes (autre formation et ou activités professionnelles) et soutiennent que l'appui à la connexion Internet est insuffisant

Tableau 4. – Les facteurs environnementaux

Variables	Valeurs	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Environnement de l'apprenant	Rural	48,39%	23,33%	0,00%	26,47%	42,00%
	Urbain	51,61%	76,67%	100,00%	73,53%	58,00%
Chambre individuelle	Oui	9,68%	63,33%	75,00%	50,00%	14,00%
	Non	90,32%	36,67%	25,00%	50,00%	86,00%
Combien êtes-vous dans votre chambre(en vous comptant) ?	Minimum	1	1	1	1	1
	Moyenne	3	2	2	2	3
	Maximum	6	7	4	6	7
Avez-vous de l'électricité permanente ?	Oui	61,29%	100,00%	100,00%	85,29%	78,00%
	Non	38,71%	0,00%	0,00%	14,71%	22,00%
Comment trouvez-vous l'environnement à ENO ?	A améliorer	17,74%	20,00%	0,00%	8,82%	26,00%
	Bon	27,42%	33,33%	62,50%	29,41%	34,00%
	Excellent	9,68%	10,00%	12,50%	17,65%	2,00%
	Moyen	19,35%	26,67%	0,00%	11,76%	28,00%
	Très bon	25,81%	10,00%	25,00%	32,35%	10,00%

Tableau 5. – Les facteurs liés à la disponibilité des apprenants à suivre les apprentissages

Variables	Valeur	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Avez-vous des activités domestiques ou professionnelles familiales?	Oui	58,06%	53,33%	50,00%	38,24%	52,00%
	Non	41,94%	46,67%	50,00%	61,76%	48,00%
Comment appréciez-vous l'accompagnement en connexion par mois ?	Pas du tout satisfaisant	16,13%	6,67%	12,50%	17,65%	26,00%
	Peu satisfaisant	58,06%	43,33%	25,00%	41,18%	44,00%
	Satisfaisant	24,19%	33,33%	25,00%	20,59%	22,00%
	Très satisfaisant	1,61%	16,67%	37,50%	20,59%	8,00%
Hors _UVS suivez-vous une autre formation ?	Oui	0,00%	60,00%	75,00%	32,35%	2,00%
	Non	100,00%	40,00%	25,00%	67,65%	98,00%
Degré de ressemblance entre la formation à et la formation hors de l'UVS	Liées	0,00%	16,67%	25,00%	8,82%	2,00%
	Pas du tout liées	0,00%	10,00%	25,00%	8,82%	0,00%
	Pas une autre formation	100,00%	40,00%	25,00%	67,65%	98,00%
	Très liées	0,00%	16,67%	12,50%	0,00%	0,00%
	Un peu liées	0,00%	16,67%	12,50%	14,71%	0,00%

4. Conclusion

Dans cet article, en utilisant des données collectées à travers une enquête, nous avons fait une classification non supervisée des apprenants d'un dispositif de formation à distance. Les résultats nous ont permis d'identifier 5 groupes d'apprenants en fonction des facteurs socio-économiques et environnementaux à des proportions différentes. Les facteurs financiers (avec ou sans source de revenu), les facteurs liés à l'habitation (nombre de personnes par chambre) et les facteurs liés à la disponibilité à suivre les enseignements (avec ou sans une autre formation hors de l'UNCHK et avec ou sans des activités domestiques et/ou professionnelles) sont les facteurs les plus déterminants dans la classification des apprenants. Par ailleurs, dans (Lee et Choi 2011), en faisant la revue de la littérature, les auteurs ont identifié les facteurs liés à l'environnement comme étant parmi un facteurs qui peut expliquer le taux d'abandon très élevé noté dans les dispositifs de formation à distance. Nous avons constaté que beaucoup de travaux de recherche avaient identifié ces mêmes facteurs comme étant des facteurs qui peuvent impacter sur la réussite des apprenants. La prochaine étape consiste à déterminer si ces facteurs impactent réellement sur la réussite des apprenants dans un dispositif de formation à distance.

REFERENCES

- [1] Dussarps, Clément. 2015. « L'abandon en formation à distance ». *Distances et médiations des savoirs. Distance and Mediation of Knowledge* 3 (10). <https://doi.org/10.4000/dms.1039>.
- [2] Ma, Yanyuan, et Liping Zhu. 2013. « A Review on Dimension Reduction: A Review on Dimension Reduction ». *International Statistical Review* 81 (1): 134-50. <https://doi.org/10.1111/j.1751-5823.2012.00182.x>.
- [3] Papi, Cathia, Louise Sauvé, Guillaume Desjardins, et Serge Gérin-Lajoie. 2022. « De la multiplicité des facteurs à prendre en compte pour mieux comprendre l'abandon en formation à distance ». *Distances et médiations des savoirs. Distance and Mediation of Knowledge* 37 (37). <https://doi.org/10.4000/dms.6904>. <https://doi.org/10.20448/journal.509.2018.51.37.42>.
- [4] Syakur, M A, B K Khotimah, E M S Rochman, et B D Satoto. 2018. « Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster ». *IOP Conference Series: Materials Science and Engineering* 336 (avril): 012017. <https://doi.org/10.1088/1757-899X/336/1/012017>.
- [5] Vidal, Martine, Monique Grandbastien, et Pierre Mœglin. 2010. « Éditorial ». *Distances et savoirs* 8 (2): 149-52. Zhang, Pingying, et Lakshmi Goel. 2011. « Is E-Learning for Everyone An Internal-External Framework of E-Learning Initiatives » 7 (2): 13
- [6] Lee, Youngju, et Jaeho Choi. 2011. « A Review of Online Course Dropout Research: Implications for Practice and Future Research ». *Educational Technology Research and Development* 59 (5): 593-618. <https://doi.org/10.1007/s11423-010-9177-y>.