Improvement of the Automated Orientation Criteria for New Graduates

N'GUESSAN B. Gerard†., ACHIEPO Odilon Y M†, KONE Tiémoman†, BAMBA Sita†
behou.nguessan@uvci.edu.ci, yapo.achiepo@uvci.edu.ci, dg@uvci.edu.ci, sita.bamba@uvci.edu.ci
†Research and Digital Expertise Unit (UREN),
Virtual University of Ivory Coast

Summary
Education can be defined as the action of instilling in individuals a set of knowledge, moral and intellectual values, considered necessary to reach the desired level of culture. In Côte d'Ivoire, for several years, the orientation of new high school graduation as remained a major challenge. Indeed, the plethoric number of students in Ivorian universities the insufficiency of the classrooms, the insufficiency of the teaching staff and trainers, etc. make orientation a very selective mechanism. Today, these assignments are made almost automatically. It is therefore imperative that these automatic assignment devices become increasingly sophisticated in order to minimize academic failure. It is in this context that this article is written situate the objective of which is to propose an automatic classification of new hig school graduates in order to identify homogeneous profiles. The interest of this work lies in the understanding of the criteria necessary for an effective personalization of the automatic assignment system for new graduates.

Keywords:
Educational system, Automatic assignment, Automatic classification.

1. Introduction

The Ivorian educational system has been facing difficulties for years. It is crossed by a series of crises and no level escapes it (primary, secondary, and higher) [1]. According to Seydou OUATTARA, the high unemployment rate explains the presence of less passionate and motivated teachers in the educational system [2]. An important element of this crisis is the phenomenon of university dropouts and complaints of bad orientation, especially at the level of baccalaureate holders. Given the very high number of baccalaureate holders to be assigned each year in relation to the reception capacities, it is not obvious that these assignments are in line with the cognitive abilities of future students. In such a situation, assignments are made by digital platforms that automatically assign students based on static criteria.
Indeed, since 2016 Côte d'Ivoire has started a phase of integration of ICTs in its schools in order to make its educational system one of the most competitive on the continent [3], [4]. Despite all these efforts, we are witnessing the case of failures, even drop-outs in some Universities. One of the arguments commonly put forward to explain this situation is the inadequate orientation of high school graduates in the courses. Indeed, each year, many graduates present themselves at the doors of the Universities. Faced with this plethoric number of candidates, orientation becomes digital. The problem therefore arises from the performance of automated orientation criteria in order to limit or even avoid university failures. In view of all the above, one wonders how to improve the school orientation system for Ivorian baccalaureate holders.

The objective of this article is to use classification methods in order to determine the relevant indicators of orientation and success of the baccalaureate holders. It will be about using, in a complementary way, multivariate statistical analysis methods Multiple Correspondence Analysis (MCA), Principal Component Analysis (PCA) to end up with an Ascending Hierarchical Classification (HAC).
The first step is to present a literature review of guidance methods. Then present the methodological approach and the results of the classification method. Finally, we will end up end with a conclusion.

State of the art

The orientation process is characterized by several factors and differs from one country to another. The works of Marion Erouart of the University of Lyon stated that this orientation process must be integrated upstream and then into higher education training. Indeed, it suggests that the first two years constitute a core curriculum then the students specialize when they enter the third year in one of the nine specialties offered by Erouart and al.. Hubert, and al. proposes in his work a recommendation system for orientation towards higher education [4]. This system is based on the operation of an ontology in order to specify the application domain of the student in a set of classes (the training) linked together by relations. GRENET and al., and IEHLE and al. offer a critical analysis of assignment procedures such as the SIGEM software and

Manuscript received November 5, 2022
Manuscript revised November 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.11.79

registrations used in France in French secondary and higher education [6], [7]. Combe and al. proposes an assignment procedure. This procedure not only makes it possible to quantify the impact that its adoption could have compared to the current system. But also, is based on the specificities of the least attractive academies to increase the rate of movement of tenured teachers up to 30% [8]. Omar and al. offers an orientation system based on the neural network algorithm. This system is able to classify the academic orientations according to the professional category (Activities, Aptitudes, Abilities, Personality) [9]. Sakers and al. is interested in finding a match between the data characterizing candidates for higher education courses (candidate and class marks, high school of origin, etc.) and the marks obtained by these same candidates during their first year at university [10].

Material and method

As part of the study, we used an analytical and statistical processing mechanism for a database containing a mixture of secondary school leavers from 2000 to 2021. To implement our typology, we made the choice of a multivariate analysis. The data for baccalaureate holders is recorded in Table 1 below. This table includes 303 high school diploma holders and 32 variables describing information on them.

Table 1 : List of information about bachelors

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
1	date_naissar	genre	matricule	serie_bac	note_mat	moy_class_n	note_phy_ch	m	n	m	n	m	n	m	n	m	n	moy_class_a	note_allema	moy_class_a	note_espagn	moy_class_e	date_bac	filiere	categorie_filiere	choix	categorie_fil	choix_respec	diplome
2	8/24/2009	Male	686-45-5488	A2	10	11	0	0	0	0	#	#	#	#	#	#	#	14	14	15	11	12	02/07/2000	Allemand	Litterature	Anglais	Litterature	0	1
3	08/07/2007	Female	671-11-6066	C	13	14	14	#	5	8	7	#	#	9	8	#	#	18	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Maths-Infor	MATHS-INFC	1	1
4	8/21/1999	Female	535-27-6330	A2	6	11	0	0	0	0	#	#	#	#	#	8	#	15	13	14	12	14	02/07/2001	Anglais	Litterature	Anglais	Litterature	1	1
5	11/19/2005	Male	631-09-0269	D	15	14	15	#	#	#	5	8	#	#	9	8	#	9	0	0	0	0	02/07/2010	Maths-Informatiq	MATHS-INFO	Maths-Infor	MATHS-INFC	1	0
6	06/02/2001	Female	412-72-7361	D	12	13	16	#	#	#	#	#	#	#	#	#	#	15	0	0	0	0	02/07/2001	Sociologie	Scs Sociologie	Sociologie	Scs Sociologi	1	0
7	11/01/2001	Male	827-71-4365	C	14	15	16	#	#	#	#	#	#	#	#	#	9	7	0	0	0	0	02/07/2001	Sciences Economi	Scs Sociologie	Lettres mod	Litterature	0	0
8	3/18/2005	Male	734-83-2627	A2	10	11	0	0	#	#	#	#	#	#	#	#	#	14	14	15	12	13	02/07/2001	Psychologie	Scs Sociologie	Psychologie	Scs Sociologi	1	1
9	07/03/2003	Female	693-20-3439	D	13	14	16	#	#	#	#	#	#	#	#	#	#	12	0	0	0	0	02/07/2001	Psychologie	Scs Sociologie	Psychologie	Scs Sociologi	1	1
10	06/05/2002	Male	808-77-7670	C	13	15	14	#	#	#	#	#	#	#	#	#	#	11	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	0
11	10/27/2001	Female	720-79-2174	C	13	14	15	#	#	#	#	#	#	#	#	#	#	15	0	0	0	0	02/07/2001	Sciences Economi	Scs Sociologie	Sciences Eco	Scs Sociologi	1	1
12	1/31/1998	Female	850-96-6434	D	11	13	14	#	#	#	#	#	#	#	#	#	#	15	0	0	0	0	02/07/2006	Maths-Informatiq	MATHS-INFO	Maths-Infor	MATHS-INFC	1	0
13	05/10/2002	Male	614-81-9785	C	9	14	14	#	#	#	#	#	#	#	#	#	#	14	0	0	0	0	02/07/2006	Biologie	Biologie	Biologie	Biologie	1	1
14	10/21/2000	Male	596-84-4902	A2	7	10	0	0	0	0	#	#	#	#	#	#	#	15	12	11	14	13	02/07/2010	Espagnol	Litterature	Anglais	Litterature	0	0
15	11/17/2003	Female	140-80-2563	A2	8	10	0	0	0	0	#	#	#	#	#	#	#	12	12	13	14	13	02/07/2001	Espagnol	Litterature	Espagnol	Litterature	1	1
16	3/30/2002	Male	394-38-6123	A2	9	11	0	0	0	0	#	#	#	#	#	#	#	9	13	12	14	13	02/07/2001	Lettres modernes	Litterature	Lettres mod	Litterature	1	1
17	8/24/2008	Female	743-73-7240	A2	9	11	0	0	0	0	#	#	#	#	#	#	#	13	11	12	13	13	02/07/2001	Lettres modernes	Litterature	Anglais	Litterature	0	0
18	2/26/2003	Male	560-28-2791	D	14	15	15	#	#	#	#	#	#	#	#	#	#	14	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Maths-Infor	MATHS-INFC	1	0
19	3/25/2015	Male	459-34-4072	C	12	12	14	#	#	#	#	#	#	#	#	8	#	15	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Physique	Scs Physique	0	0
20	6/25/2008	Male	735-71-8349	C	12	14	14	#	#	#	#	#	#	#	#	#	#	12	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Physique	Scs Physique	0	1
21	1/16/2000	Female	371-81-0229	D	13	12	14	#	#	#	#	#	#	#	#	#	#	14	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Maths-Infor	MATHS-INFC	1	1
22	09/01/2001	Female	615-45-4956	A2	10	11	0	0	0	0	#	#	#	#	#	#	#	12	14	15	12	13	02/07/2001	Allemand	Litterature	Allemand	Litterature	1	1
23	1/26/2002	Female	652-25-9615	A2	7	12	0	0	0	0	#	#	9	#	9	7	#	14	14	13	12	13	02/07/2001	Allemand	Litterature	Allemand	Litterature	1	1
24	02/02/2012	Male	693-68-7331	A2	9	11	0	0	0	0	#	#	#	#	7	6	#	15	12	11	13	14	02/07/2021	Anglais	Litterature	Lettres mod	Litterature	0	0
25	6/23/2003	Male	391-93-6036	A2	7	10	0	0	0	0	#	#	#	#	#	#	#	16	13	12	12	13	02/07/2001	Anglais	Litterature	Anglais	Litterature	1	1
26	5/23/2001	Male	815-95-6387	C	13	14	15	#	#	#	#	#	9	6	#	#	8	9	0	0	0	0	02/07/2021	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	1
27	05/09/2002	Male	710-23-6160	C	12	13	14	#	#	#	#	#	7	#	#	#	#	7	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	1
28	04/04/2003	Male	816-96-3267	D	14	15	15	#	#	#	#	#	#	7	#	#	17	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	0
29	06/12/1999	Male	809-98-0047	D	13	14	14	#	#	#	#	#	#	8	#	5	#	11	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	0
30	01/08/2005	Male	443-64-3282	D	12	13	14	#	#	#	#	#	7	#	#	#	#	14	0	0	0	0	02/07/2001	Physique	Scs Physique	Physique	Scs Physique	1	1
31	1/23/2000	Female	260-07-3007	C	12	14	14	#	#	#	#	#	9	#	#	#	#	14	0	0	0	0	02/07/2001	Physique	Scs Physique	Physique	Scs Physique	1	1
32	09/05/1998	Male	885-65-8344	A2	9	10	0	0	0	0	#	#	#	#	#	#	#	16	12	13	11	12	02/07/2001	Anglais	Litterature	Anglais	Litterature	1	1
33	10/20/2002	Female	572-74-4659	A2	7	9	0	0	0	0	#	#	#	#	#	#	#	14	11	10	14	15	02/07/2001	Anglais	Litterature	Anglais	Litterature	1	1
34	4/25/2004	Female	331-61-6761	A2	5	9	0	0	0	0	#	#	#	#	#	#	#	15	12	13	10	11	02/07/2001	Anglais	Litterature	Lettres mod	Litterature	0	0
35	12/13/2004	Male	567-46-6802	C	12	13	15	#	#	#	#	#	#	#	#	#	#	13	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	1
36	02/10/2000	Male	444-50-5852	C	13	14	12	#	#	#	#	#	#	#	#	#	#	12	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	1
37	7/20/2004	Male	357-33-0239	D	13	14	15	#	#	#	#	#	#	8	#	#	#	13	0	0	0	0	02/07/2001	Maths-Informatiq	MATHS-INFO	Sciences Eco	Scs Sociologi	0	0
38	8/27/2005	Male	188-77-6979	D	12	13	14	#	#	#	#	#	#	#	#	#	#	12	0	0	0	0	02/07/2001	Physique	Scs Physique	Physique	Scs Physique	1	1
39	02/12/2005	Female	189-32-2304	D	13	14	13	#	#	#	#	#	#	9	#	#	#	13	0	0	0	0	02/07/2001	Physique	Scs Physique	Physique	Scs Physique	1	0
40	12/26/2000	Female	802-78-0668	C	12	13	14	#	#	#	#	#	#	#	#	#	#	13	0	0	0	0	02/07/2001	Physique	Scs Physique	Maths-Infor	MATHS-INFC	0	0
41	10/18/2000	Male	164-84-4934	C	13	13	12	#	#	#	#	#	#	#	#	#	#	15	0	0	0	0	02/07/2001	Sciences Economi	Scs Sociologie	Sciences Eco	Scs Sociologi	1	1

As part of the classification, the R 4.2.0 software was used for the data analyses with the FactoMineR packages for the data analysis and factoextra for the data visualization. One- and two-dimensional descriptive analyses were performed. In order to highlight the linear links between the individuals of our sample, a multidimensional analysis, the Principal Component Analysis (PCA) was carried out. Its objective is not only to reduce the number of variables assigned to individuals and facilitate observations while saving as much student information as possible but also to determine homogeneous groups of observations or atypical observations. Among the variables available to our sample, some are qualitative, others quantitative. Their joint analysis uses an MCA. This analysis makes it possible to create clusters (groupings or distances) of objects. Subsequently, the information from this analysis was used to carry out a CAH. The Ascending Hierarchical Classification (AHC) made it possible not to have to choose a priori provisional class centers or the number of classes. This method makes it possible to reveal the "true" classes if they exist since they are determined automatically and do not depend on an a priori choice as in the case of the Classification method around Mobile Centers (CCM) [11]. Kaiser's criterion was used to determine the axes to be retained for PCA and MCA. According to this Kaiser criterion, it is advisable to retain the axes whose relevant variables are clearly > 1 [12], [13].

Results

– Results of the first PCA

A first Principal Component Analysis (PCA) was performed to determine the relevant variables as shown in the following figure

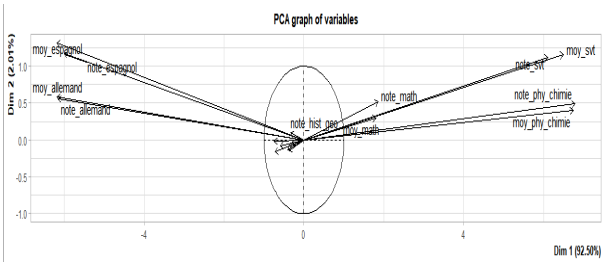


Fig 1 : Graph of the variables of the first PCA

This Principal Component Analysis allowed us to observe that the variables "grade_allemand", "mean_german", "grade_spanish" and "mean_spanish', are well correlated with each other on the second dimension. These variables are negatively correlated. While the variables "notes_svt", "avg_svt", avg_phy_chemistry", "note_phy_chemistry", are correlated with each other on the first dimension; the correlation is positive.

The fact that the variables note_español and note_allemand are correlated and the variables note_phy_chimie and note_svt are correlated, shows that the pupils of the literary series do not have the subjects physics chemistry and svt in their school program and the pupils of the scientific series do not have the subjects Spanish and German in their school curriculum.

At the end of this analysis, we conclude that class averages are not necessary in the orientation of baccalaureate holders; therefore, we will remove the averages from our data.

– Results of the second PCA

A second presentation analysis was carried out as shown in the following figure 2:

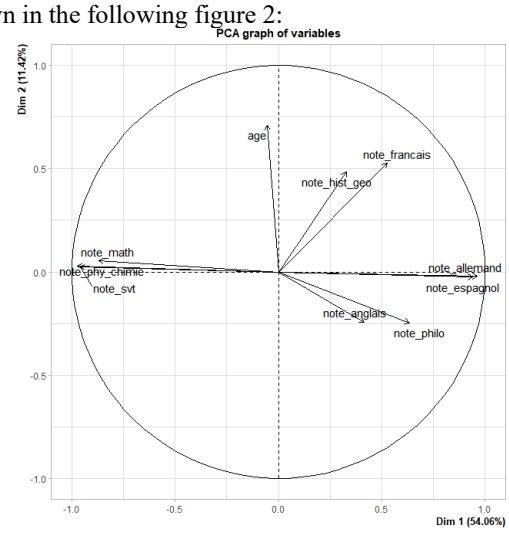


Fig 2 : Graph of the variables of the second PCA

This analysis shows that the variables "note_svt", note_phy_chimie", "note_math" are correlated with each other, but the correlation is negative on the first dimension; it also shows that the variables "note_french", note_hist_geo", "note_allemand", "note_aespañol", "note_philo", are correlated with each other and the correlation is positive on the first dimension.

Subsequently, a CAH was carried out as shown in the following figures 3.4 and 5:

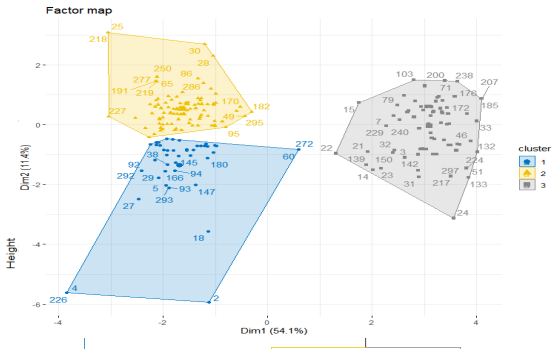


Fig 3 : Visualization of the individuals of each group

Fig 4 : Dendrogram of the classification on the result of the PCA carried out

```
$`1`
                    v.test Mean in category overall mean sd in category
note_phy_chimie   5.763254         14.3461538          9.347682       0.7568275
note_svt          5.660722         13.1730769          8.685430       1.6257679
note_math         5.617443         12.9615385         11.380795       0.7585801
note_espagnol    -5.064723          0.4615385          4.410596       2.3076923
note_allemand    -5.109715          0.5384615          4.579470       2.6923077
note_hist_geo    -5.192397         12.5576923         13.440397       1.5861596
note_francais    -5.298705         12.5769231         13.483444       1.9547393
age             -10.790064         20.3269231         21.619205       0.5084348
                  Overall sd      p.value
note_phy_chimie  6.8625323 8.250750e-09
note_svt         6.2728037 1.507371e-08
note_math        2.2265768 1.938036e-08
note_espagnol    6.1695380 4.089936e-07
note_allemand    6.2576027 3.226445e-07
note_hist_geo    1.3451251 2.076041e-07
note_francais    1.3537014 1.166268e-07
age              0.9476506 3.835224e-27

$`2`
                    v.test Mean in category overall mean sd in category
note_phy_chimie  12.090748         14.32414          9.347682       0.9386914
note_svt         11.779774         13.11724          8.685430       0.9210447
note_math         9.646132         12.66897         11.380795       1.1145135
age               8.762539         22.11724         21.619205       0.5446093
note_francais    -3.151322         13.22759         13.483444       0.9153468
note_anglais     -4.128334         14.26207         14.519868       1.1080937
note_philo       -8.077474         11.61379         12.241722       1.0251069
note_espagnol    11.919610          0.00000          4.410596       0.0000000
```

Fig 5 : Quantitative variables describing the classes

This classification gave three classes, namely:

Class1:

For the first class, the most significant variables are "note_svt", "note_phy_chimie", "note_math". For the "note_svt" variable, the individuals of class 1 take values that are significantly different from zero and above the average. They are nonzero because the test value (6.355075) is greater than 2; and they are above average because the test value is positive. Class 1 individuals have an average of 13,300,000 in SVT. For the variable "note_phy_chimie", the individuals of class 1 also take values significantly different from zero and higher than the average. The test value is 6.339025. Individuals in class 1

have an average of 14.3833333 in Physics Chemistry. For the "note_math" variable, the individuals of class 1 also take values different from zero and higher than the average. The test value is 5.506286. Individuals in class 1 have an average of 12.8000000 in Mathematics.

Class 2:

For the second class, the most important variables are "note_phy_chimie", "note_svt", "note_math". For the "note_phy_chime" variable, the individuals of class 2 take values that are significantly different from zero and above the average. They are nonzero because the test value (12.090748) is greater than 2; and they are above average because the test value is positive. Individuals in class 2 have an average of 14.32414 in Physics and Chemistry. In this class, the test and average values of the "note_phy_chimie" variable are higher than those of class 1. With regard to the "note_svt" variable, the individuals of class 2 also take significantly different values zero and above average. The test value is 11.779774. Class 2 individuals have an average of 13.11724 in SVT. In this class, the test value of the "note_svt" variable is greater than that of class 1; but the average value is lower than that of class 1 (13.11724 <13.3000000).
For the "note_math" variable, the individuals of class 2 also take values different from zero and higher than the average. The test value is 9.646132. Individuals in class 2 have an average of 12.66897 in Mathematics. We then observe that the test value of the "note_math" variable is higher than that of class 1, while the average is lower than that of class 1. We note that in class 2, the variable ' 'age" has the positive test value, which is not the case in class 1 where it has a negative value. This variable is thus a significant variable in class 2, which is not the case in class 1.

Class 3

Only the variables of the literature category characterize the third class; in particular "note_allemand", "note_spanish", "note_philo", "note_french", "note_english", "note_hist_geo". For the variable "note_allemand", the individuals of class 3 take specific values different from zero and higher than the average. The test value is 16.851410. Individuals in class 3 have an average of 12.9047619 in German. With regard to the variable "note_español", the individuals of class 3 also take significant values different from zero and higher than the average. The test value is 16.519682. Individuals in class 3 have an average of 12.4571429 in Spanish. For the variable "note_philo", the individuals of class 3 also take values different from zero and higher than the average. The test value is 9.550628. Individuals in class 3 have an average of 13.2190476 in Philosophy.
As for the "note_francais" variable, the individuals of class 3 also take significant values different from zero and above the average. The test value is 7.506592. Individuals in class 3 have an average of 14.2857143 in French.
As for the variable "note_anglais", the individuals of class 3 also take values different from zero and higher than the average. The test value is 5.493361. Individuals in class 3 have an average of 14.9714286 in English. For the variable "note_hist_geo", the individuals of class 3 have values different from zero and higher than the average. The test value is 4.193256. Individuals in class 3 have an average of 13.8857143 in History-Geography.

- Results of the second ACM

The figure below represents the visualization of the biplot of individuals and variables. In this figure 6, individuals are in blue and variables are in red

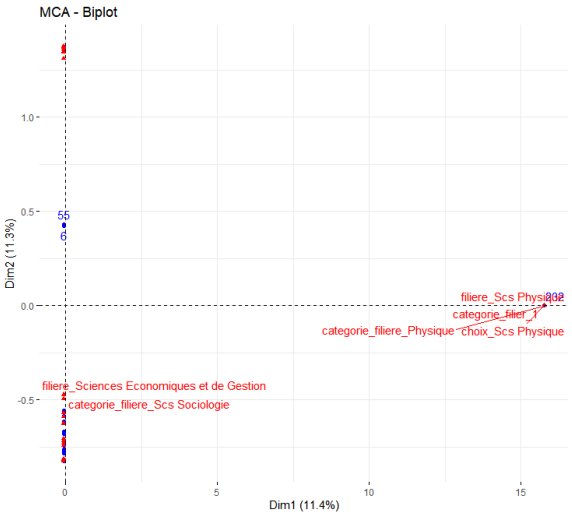


Fig 1: Biplot of individuals and variables

The variables filiere_Scs Physique, categorie_filiere_Physique and choice_Scs Physique, are close on the graph; therefore these variables have a similar profile. The same is true for the variables filiere_Economics and Management Sciences, categorie_filiere_Scs Sociologie, which are close on the graph.
Let us now proceed to the verification of the existence of correlation between the variables and the main axes of the MCA. Figure 7 below specifies the existence of correlation between the variables and the main axes of the MCA.

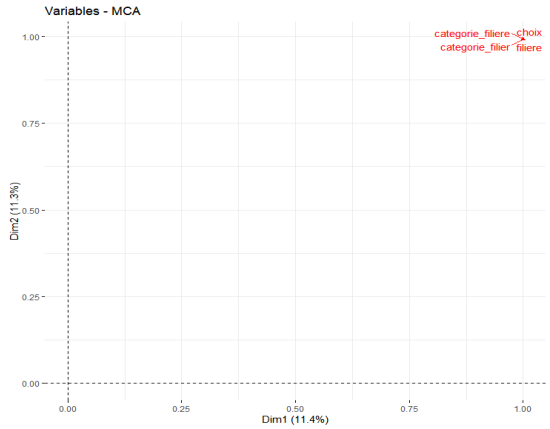


Fig 2: Correlation between variables

The graphic representation above indicates that the variables "categorie_filiere", "choice", "categorie_filier" and "filiere" are correlated with each other. And this correlation is positive on the first axis and negative on the second axis.

At the end of this multiple correspondence analysis, we obtained significant results; it will now be necessary to apply the ascending hierarchical classification to these results.

– Hierarchical ascending classification on MCA results

The classification on the results of the ACM gives 4 classes, namely: class1, class 2, class 3 and class 4. The first axis is associated with class 1. As for classes 2 and 3, they are associated with the axis 2 and axis 3. While class 4 is associated with axis 1. Let's visualize the dendrogram highlighting the four classes. Figures 8 and 9 below present the dendrogram and the 3D representation generated by the classification:

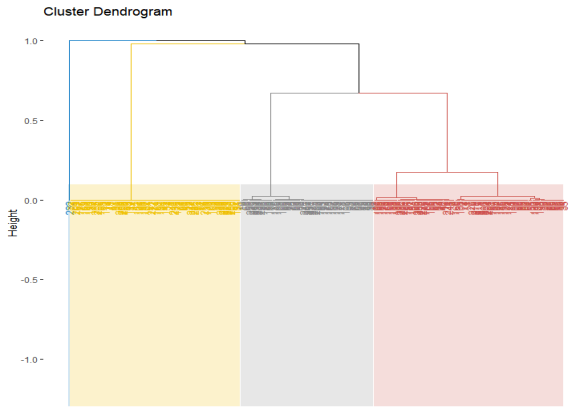


Fig 3: Dendrogram of the classification carried out on the results of the MCA

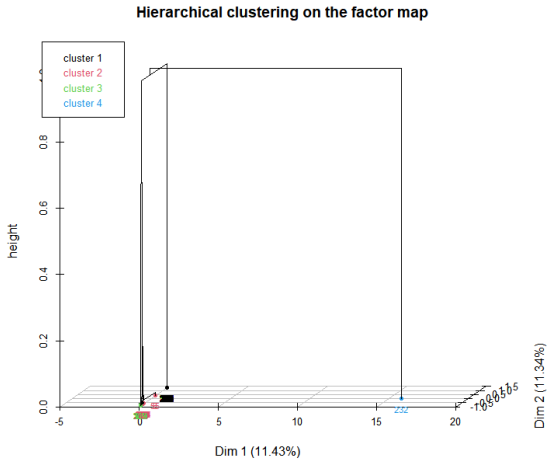


Fig 4: 3D graph combining hierarchical classification and factor plan

Conclusion

The work carried out in this article made it possible to make the classification of baccalaureate holders from data comprising the marks obtained in the Baccalaureate examination and information concerning their orientations. To do this, a Principal Component Analysis (PCA) was first used to determine the relevant variables in the orientation process. It made it possible to understand that the marks in the Baccalaureate examination were very strongly correlated with the averages obtained in class. This result proves that it is not necessary to take class averages into account in the orientation of new baccalaureate holders. Then, a Multiple Correspondence Analysis (MCA) was carried out to be able to take into account the qualitative variables, in particular the chosen sector and the sector of assignment of the high school graduate, in automatic classification. Finally, a Hierarchical Ascending Classification (HAC) was performed. It made it possible to have three (3) groups of graduates with different characteristics. Therefore, if one wishes to avoid failures, the specificities of each group should be taken into account in the automated criteria used in the online orientations of new graduates.

References

[1]. Algiers Jean-Francis EkOUNGOUN. The current model of school management will put society at risk if it remains unchanged, 2019.

[2]. Seydou OUATTARA. THE IVORIAN SCHOOL: SUCCESS RATE OR QUALITY TRAINING? 2020.

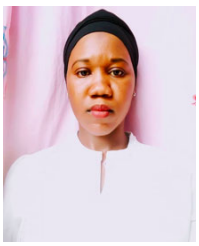
[3]. SEHI, Bi, MIAN, Antoine, and ABIDJAN, ENS “The practical training of teachers in Côte d'Ivoire” Formation et Profession, 2015, vol. 23, No. 3, p. 203-210.

[4]. HUBERT, Nicolas, BRUN, Armelle, and MONTICOLO, Davy. Towards an explainable recommender system for school guidance. In: Atelier EXPLAIN'AI-Conference EGC Blois 2022. 2022.
[5]. EROUART, Marion. Psychosocial analyzes of gender orientation processes towards and during engineering training. Educational and vocational guidance, 2022, no 51/1, p. 171–177.
[6]. GRENET, Julien. Allocation algorithms in the French education system. Simioni M., Steiner P., ed., Comment ça matche, 2022, p. 21-59.
[7]. IEHLÉ, Vincent and JACQMIN, Julien. SIGEM: Analysis of the Assignment Procedure in the Grandes Ecoles of Management. 2021.
[8]. COMBE, Julien, TERCIEUX, Olivier, and TERRIER, Camille. Improve teacher mobility. Education & training, 2016, no 92, p. 57-75.
[9]. Omar Zahour, EL Habib Benlahmar, A. Eddaoui. "Automatic Classification of Educational and Career Guidance Questions Using Machine Learning Algorithms: A Comparative Study", 2020.
[10]. Saker Amine, Christel Dartigues-Pallez, Rey Gaetan. Supervised Classification of Pedagogical Data for Success in Higher Education. [Research Report] I3s, University Côte D'azur. 2020. Hal-02486729
[11]. LESUR-IRICHABEAU, Gabrielle and POINT, Patrick. Classification methods applied to censuses: creation of a typology of oyster farms using an Ascending Hierarchical Classification. In: 41st Statistics Days, SFdS, Bordeaux. 2009.
[12]. S. CHAMPELY, "Introduction to multivariate (factorial) analysis under R", (2005),https://docplayer.fr/1127216-Introduction-al-analyse-multivariee-factorielle-sous-r-stephane-champely.html
[13]. J. BRAEKEN & MAL M VAN ASSEN, Psychological Methods, 74(2016),http://dx.doi.org/10.1037/met0000074

Behou Gerard N'Guessan has a doctorate in computer engineering. He holds a master's degree in media engineering from the University of May 08, 1945 in Guelma, Algeria. He obtained his PhD at the University Nangui Abrogoua, Abidjan-Côte d'Ivoire in the Faculty of Applied Basic Sciences. He is a member of the Research Laboratory in Computer Science and Telecommunications of the Institut National Polytechnique Houphouët Boigny (INP-HB), Abidjan, Côte d'Ivoire, member of the Laboratory of Data Engeering and Artificial Intelligence and associate member of the Research Unit and Digital Expertise of the Virtual University of Côte d'Ivoire. His research interests include mathematical modeling, media engineering, traditional medicine and application inventor. His work focuses on their method of research and training in traditional medicine. He is currently working as a Master Assistant at the Virtual University of Côte d'Ivoire in Abidjan (Côte d'Ivoire)

After graduating in electronics engineering at the end of the 1980s from the National Higher Institute of Technical Education (INSET) in Yamoussoukro – now a polytechnic – **Koné Tiémoman** set sail for France. He will successively obtain his DEA (Diploma of Advanced Study) in automation, option control and management of industrial processes (1990) and his Doctorate in computer science, Very honorable mention, also in the said university (1993) at the University of Paris 12 Val de Marne in France. This work enthusiast has a rich professional career to his credit both in Côte d'Ivoire in the Ivorian administration and through numerous consultancies for private institutes recognized worldwide. Since 2016, he has been Director General of the Virtual University of Côte d'Ivoire (UVCI)

Odilon Yapo M. ACHIEPO is a Doctor in Computer Science. He holds a degree in Statistician-Economist Engineer at ENSEA in Abidjan (Côte d'Ivoire) and a Master's degree in Computer Science with a specialization in Knowledge Extraction from Data obtained at Polytechnic School of Nante's University (France). He is also a consultant, Teacher-researcher at Péléforo Gon Coulibaly University of Korhogo (Ivory Coast). His areas of expertise and interests include Statistical Engineering, Artificial Intelligence, Big Data, Internet of Things, Data Science, Fast Data and Ethical Hacking. He is also a member of the Laboratory of Data Engineering and Artificial Intelligence of Abidjan, and associate member of the Unit of Research and Digital Expertise of the Virtual University of Cote d'Ivoire. He is also the author-creator of Resiliometrics, a modeling discipline that consists of developing and applying computational models for measurement, analysis and simulation of social resilience process.



Bamba Sita is a Master's student Big Data Analytics of the Virtual University of Côte d'Ivoire. she is to Analysis, Information and Decision (AIDE) research team of the Digital Research and Expertise Unit.