

Analyse des traces d'apprentissage : Modélisation d'un profil d'apprenant dans Moodle à l'aide du Data Mining

Kouassi Rodrigue Ahoussou

Département d'informatique et de génie logiciel

Laboratoire ERICAE



Contexte de Recherche

- Adoption croissante des cours mixtes (présentiel et distantiel)
 - Traces d'apprentissage disponible
 - Défis liés à l'apprentissage en ligne et mixte
 - Modélisation des profils des apprenants
-
- Étudiants de premières années universités française.

Recherches Antérieures

Predicting student dropouts with machine learning: An empirical study in Finnish higher education (Matti Vaarma, 2024)

Remarque:

- a) Détection des étudiants à risque à la fin des semestres.
- b) Une seule étude réalisée à trois moments différents dans un semestre, mais sans données de LMS.
- c) Analyser comment les modèles et l'importance des fonctionnalités varient au fil du temps.

Recherches Antérieures

Analysis and Prediction of Student Performance Based on Moodle Log Data using Machine Learning Techniques (Chayapon Kaensar,2023)

Remarque:

- a) Prédiction des performances des étudiants à 25 %, 50 %, 75 % et 100 % du cours
- b) Arbre de décision (81,10 %) à 100 % du cours, SVM (86,90 %) à 25 % du cours et la régression linéaire (80 %) à 50 % et 75 % du cours
- c) Le choix de l'algorithme peut dépendre du moment de la prédiction
- d) Déterminer les étudiants à risque à chaque étape en raison de la variation des performances
- e) Nécessité d'une bonne interaction entre les étudiants et la plateforme pour une bonne prédiction

Résumé des recherches antérieures.

Ces différentes études nous montrent qu'il est crucial de déterminer les étudiants à risque dans un délai convenable afin d'agir à temps. Elles mettent également en évidence la complexité de trouver le meilleur algorithme pour la prédiction. Ces études soulignent également que les étudiants changent de comportement tout au long de leur apprentissage, ce qui pourrait créer des biais dans les modèles de prédiction.

Problématique

- Détecter les étudiants à risque
- Prédire des étudiants en échec avec peu de donnée et le plus tôt possible

Objectifs

- Analyser la performance des étudiants et leurs comportements en fonction du temps
- Détecter les étudiants à risque ainsi que leurs changements d'état.
- Déterminer les données nécessaires pour prédire les étudiants en situation d'échec
- Déterminer le temps minimum requis pour qu'un étudiant soit considéré en échec

Présentation du cours (Excel)

période: 23 janvier-19 mars 2024

- Le cours se déroule en 6 séances et est réparti en 6 groupes différents. Chaque groupe a son heure et ses jours de cours spécifiques.
- Le cours est dispensé en présentiel et dure deux heures. La première partie consiste à dispenser le cours, tandis que la deuxième partie nécessite que les étudiants se connectent à Moodle pour effectuer des activités. Ils doivent soumettre leurs travaux pratiques avant la fin de la séance.
- Les étudiants ont aussi à réaliser les exercices en distantiel pour préparer la séance suivante.

• Séance 1 - Tri, Sous Totaux, Filtrage et barre d'état 23 29 janvier

• Présentiel

- TriFiltre Fichier
- Dépôt du TP1 Tri Sous-Totaux Filtres Devoir

• Distantiel

- Préparation TP2 : PDF Les fonctions Fichier
- Préparation TP2 : FICHER EXCEL SI
- Préparation TP2 : DEPOT fichier SI Devoir
- Préparation TP2 : QCM Fonction SI Test

Séance 2 - Fonctions imbriquées 5 6 février

- Présentiel
 - Les Fonctions imbriquées Fichier
 - Exercices sur les fonctions imbriquées Fichier
 - Déposez du TP2 Fonctions Imbriquées Devoir
- Distantiel
 - Préparation TP3 : PDF Fonctions conditionnelles. Fichier
 - Préparation TP3 : QCM Fonctions conditionnelles

Séances des cours

		8h-10h	10h-12h	14h-16h	16h-18h
23-Jan	Séance 1	4	6	5	
29-Jan	Séance 1		2	3	1
5-Feb	Séance 2		2	1	6
6-Feb	Séance 2		5	3	4
26-Feb	Séance 3		4	5	6
27-Feb	Séance 3		2	3	1
4-Mar	Séance 4		1	3	6
5-Mar	Séance 4		2	5	4
11-Mar	Séance 5		6	5	1
12-Mar	Séance 5	4	3	2	
18-Mar	Séance 6		3	5	6
19-Mar	Séance 6		2	4	1

Sources des données

- Listing étudiants (groupe de TP, bac, Rang de recrutement)
- Notes semestre précédent
- Traces de Moodle
- Emploi du temps des étudiants

Choix des données

- Données statiques des étudiants
- Données Dynamiques(fichiers Logs)
- Données Dynamiques (Évaluation des Quiz)

Défi du traitement de données

Informations liées aux différents dépôts des étudiants.

Impossible de faire directement des tris sur Object_name, Object_ID et le verbe action Submitted dans le fichier Excel.

- Lorsqu'on fait un tri sur Submitted , on constate que tous les Object_Name sont nuls.
- Impossible de faire un double filtre sur Object_Name et Submitted
- Impossible de faire un double filtre sur Object_ID et Submitted

Date	Hour	ACTION_VERB	CourseID	CourseName	OBJECT_ID	OBJECT_NAME	OBJECT_TYPE
23/01/2024	8:00:12 AM	updated	1580	S2 Ressources et culture numeriques	10705	Depot du TP1 Tri Sous-Totaux Filtres	grade_items
23/01/2024	8:54:43 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:54:44 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:54:46 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:55:09 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:55:09 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:56:05 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	8:56:07 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:34:17 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:34:19 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:34:42 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:34:42 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:00 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:03 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:05 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:07 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:16 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:25 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:27 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:33 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:41 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:40:42 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:42:21 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:42:40 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:42:41 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:42:45 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:42:47 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:43:00 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:43:01 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign

Date	Hour	ACTION_VERB	CourseID	CourseName	OBJECT_ID	OBJECT_NAME	OBJECT_TYPE
23/01/2024	8:55:09 AM	submitted	1580	S2 Ressources et culture numeriques	1093636	null	assign_submission
23/01/2024	9:34:42 AM	submitted	1580	S2 Ressources et culture numeriques	1093663	null	assign_submission
23/01/2024	9:40:27 AM	submitted	1580	S2 Ressources et culture numeriques	1093667	null	assign_submission
23/01/2024	9:40:42 AM	submitted	1580	S2 Ressources et culture numeriques	1093668	null	assign_submission
23/01/2024	9:42:41 AM	submitted	1580	S2 Ressources et culture numeriques	1093667	null	assign_submission
23/01/2024	9:43:16 AM	submitted	1580	S2 Ressources et culture numeriques	1093673	null	assign_submission
23/01/2024	9:43:24 AM	submitted	1580	S2 Ressources et culture numeriques	1093674	null	assign_submission
23/01/2024	9:43:39 AM	submitted	1580	S2 Ressources et culture numeriques	1093675	null	assign_submission
23/01/2024	9:45:04 AM	submitted	1580	S2 Ressources et culture numeriques	1093666	null	assign_submission
23/01/2024	9:45:10 AM	submitted	1580	S2 Ressources et culture numeriques	1093676	null	assign_submission
23/01/2024	9:45:17 AM	submitted	1580	S2 Ressources et culture numeriques	1093679	null	assign_submission
23/01/2024	9:45:20 AM	submitted	1580	S2 Ressources et culture numeriques	1093680	null	assign_submission
23/01/2024	9:45:22 AM	submitted	1580	S2 Ressources et culture numeriques	1093677	null	assign_submission
23/01/2024	9:45:25 AM	submitted	1580	S2 Ressources et culture numeriques	1093678	null	assign_submission
23/01/2024	9:45:36 AM	submitted	1580	S2 Ressources et culture numeriques	1093682	null	assign_submission
23/01/2024	9:45:54 AM	submitted	1580	S2 Ressources et culture numeriques	1093681	null	assign_submission
23/01/2024	9:46:06 AM	submitted	1580	S2 Ressources et culture numeriques	1093683	null	assign_submission
23/01/2024	9:46:12 AM	submitted	1580	S2 Ressources et culture numeriques	1093685	null	assign_submission
23/01/2024	9:46:21 AM	submitted	1580	S2 Ressources et culture numeriques	1093686	null	assign_submission
23/01/2024	9:46:24 AM	submitted	1580	S2 Ressources et culture numeriques	1093687	null	assign_submission
23/01/2024	9:46:34 AM	submitted	1580	S2 Ressources et culture numeriques	1093689	null	assign_submission
23/01/2024	9:46:38 AM	submitted	1580	S2 Ressources et culture numeriques	1093688	null	assign_submission
23/01/2024	9:47:32 AM	submitted	1580	S2 Ressources et culture numeriques	1093690	null	assign_submission
23/01/2024	9:47:56 AM	submitted	1580	S2 Ressources et culture numeriques	1093636	null	assign_submission
23/01/2024	9:48:20 AM	submitted	1580	S2 Ressources et culture numeriques	1093692	null	assign_submission
23/01/2024	9:48:26 AM	submitted	1580	S2 Ressources et culture numeriques	1093684	null	assign_submission
23/01/2024	9:48:36 AM	submitted	1580	S2 Ressources et culture numeriques	1093691	null	assign_submission
23/01/2024	10:42:00 AM	submitted	1580	S2 Ressources et culture numeriques	1093692	null	assign_submission
23/01/2024	11:28:38 AM	submitted	1580	S2 Ressources et culture numeriques	1093778	null	assign_submission
23/01/2024	11:32:02 AM	submitted	1580	S2 Ressources et culture numeriques	1093783	null	assign_submission
23/01/2024	11:32:16 AM	submitted	1580	S2 Ressources et culture numeriques	1093784	null	assign_submission
23/01/2024	11:32:29 AM	submitted	1580	S2 Ressources et culture numeriques	1093785	null	assign_submission

Date	Hour	ACTION_VERB	CourseID	CourseName	OBJECT_ID	OBJECT_NAME	OBJECT_TYPE
23/01/2024	9:45:10 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:10 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:17 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:17 AM	uploaded	1580	S2 Ressources et culture numeriques	1093679	null	assign_submission
23/01/2024	9:45:17 AM	created	1580	S2 Ressources et culture numeriques	653174	null	assignsubmission_file
23/01/2024	9:45:17 AM	submitted	1580	S2 Ressources et culture numeriques	1093679	null	assign_submission
23/01/2024	9:45:17 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:18 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:20 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:20 AM	uploaded	1580	S2 Ressources et culture numeriques	1093680	null	assign_submission
23/01/2024	9:45:20 AM	created	1580	S2 Ressources et culture numeriques	653175	null	assignsubmission_file
23/01/2024	9:45:20 AM	submitted	1580	S2 Ressources et culture numeriques	1093680	null	assign_submission
23/01/2024	9:45:20 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:20 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:22 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:22 AM	uploaded	1580	S2 Ressources et culture numeriques	1093677	null	assign_submission
23/01/2024	9:45:22 AM	created	1580	S2 Ressources et culture numeriques	653176	null	assignsubmission_file
23/01/2024	9:45:22 AM	submitted	1580	S2 Ressources et culture numeriques	1093677	null	assign_submission
23/01/2024	9:45:22 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:23 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:24 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:24 AM	uploaded	1580	S2 Ressources et culture numeriques	1093678	null	assign_submission
23/01/2024	9:45:24 AM	created	1580	S2 Ressources et culture numeriques	653177	null	assignsubmission_file
23/01/2024	9:45:25 AM	submitted	1580	S2 Ressources et culture numeriques	1093678	null	assign_submission
23/01/2024	9:45:25 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:25 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:36 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:36 AM	uploaded	1580	S2 Ressources et culture numeriques	1093682	null	assign_submission
23/01/2024	9:45:36 AM	created	1580	S2 Ressources et culture numeriques	653178	null	assignsubmission_file
23/01/2024	9:45:36 AM	submitted	1580	S2 Ressources et culture numeriques	1093682	null	assign_submission
23/01/2024	9:45:37 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:37 AM	viewed	1580	S2 Ressources et culture numeriques	null		null
23/01/2024	9:45:41 AM	viewed	1580	S2 Ressources et culture numeriques	4112	Depot du TP1 Tri Sous-Totaux Filtres	assign
23/01/2024	9:45:41 AM	viewed	1580	S2 Ressources et culture numeriques	null		null



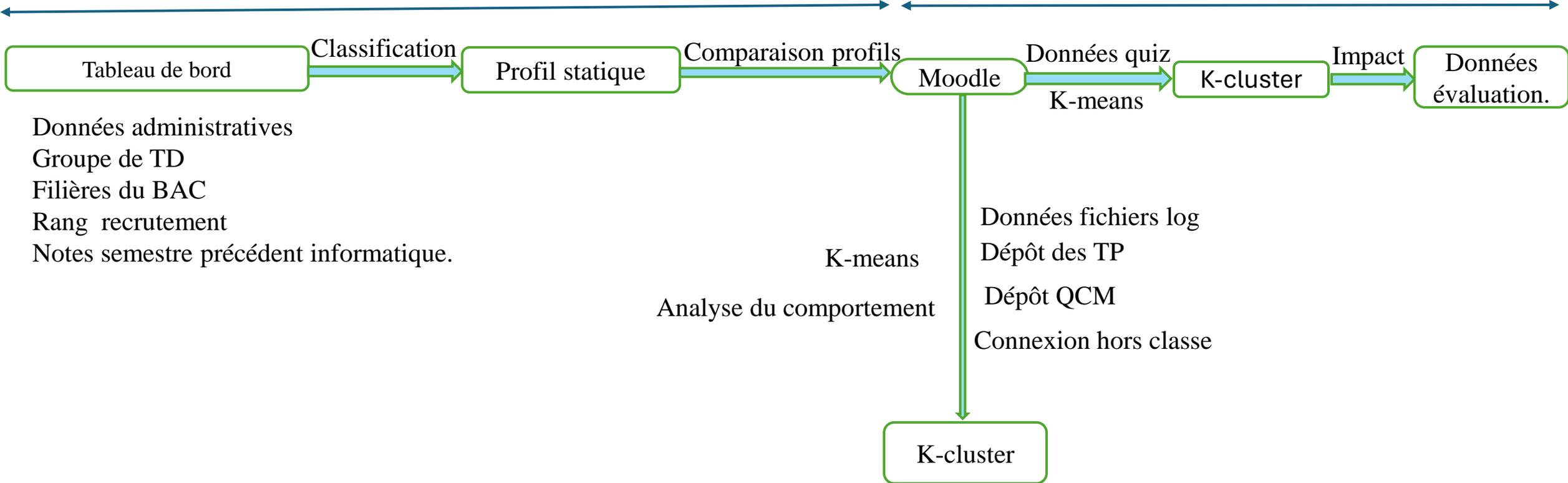
Méthodologie

- Modélisation d'un profil statique des étudiants à l'aide d'un Tableau de Bord dynamique
- Modélisation d'un profil dynamique des étudiants à l'aide des données d'apprentissage.
- Voir les variations des différents indicateurs de performance dans chaque cluster.
- Comparaison du profil statique des étudiants et des profils dynamiques des étudiants après l'algorithme de K-means
- Faire une analyse des comportements des étudiants en classe et hors classe.
- Faire une analyse générale de la performance des étudiants et leur comportement en fonction du temps.

Analyse de la performance des étudiants et leur comportement en fonction du temps

Modélisation d'un profil statique

Modélisation d'un profil dynamique



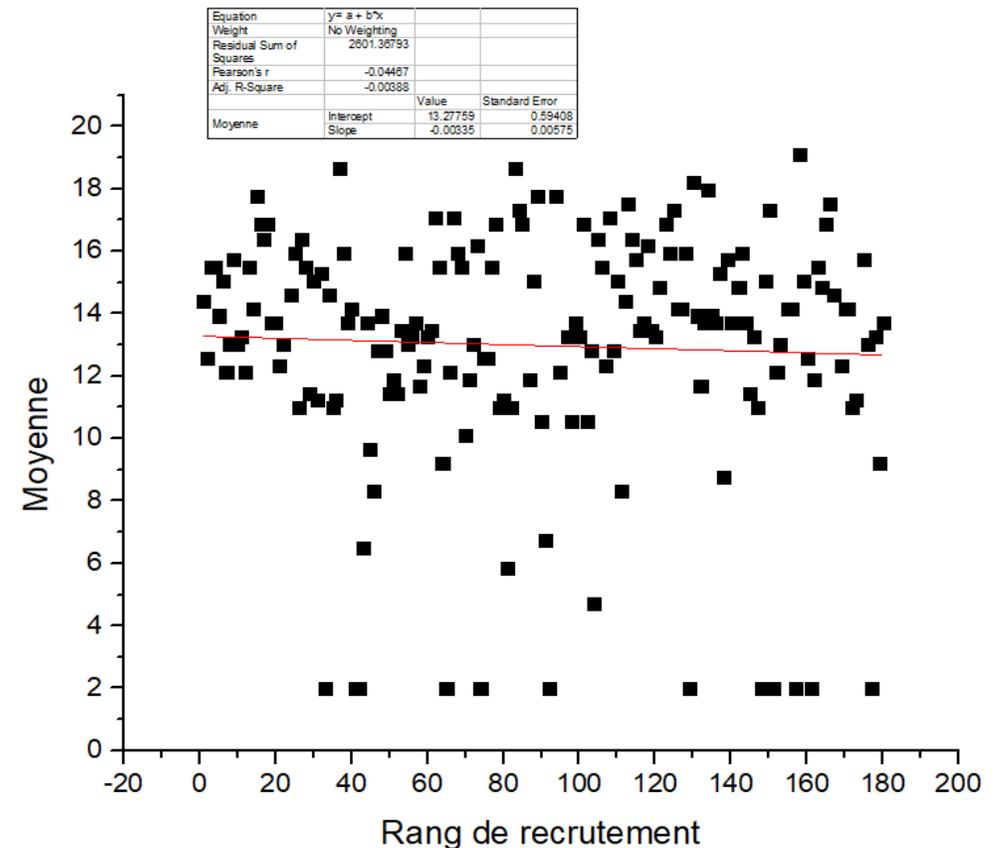
Données administratives
Groupe de TD
Filières du BAC
Rang recrutement
Notes semestre précédent informatique.

K-means
Analyse du comportement

Données fichiers log
Dépôt des TP
Dépôt QCM
Connexion hors classe

Corrélation: Moyenne et le rang de Recrutement

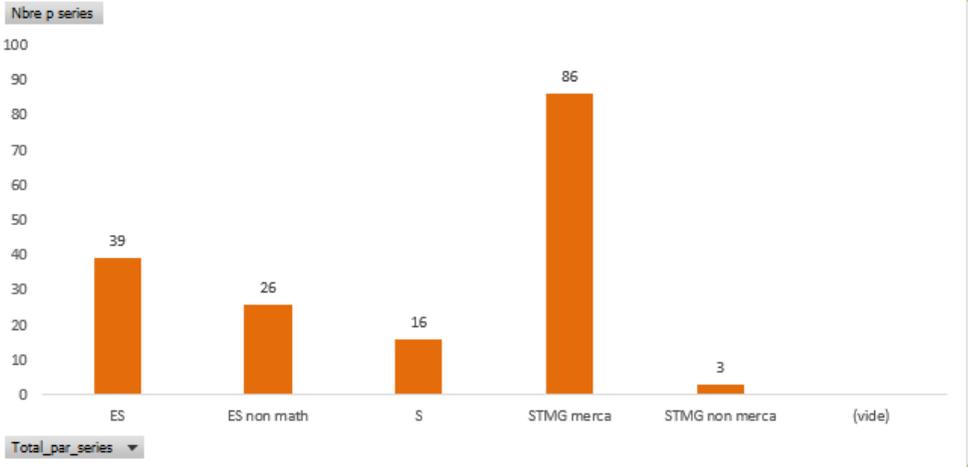
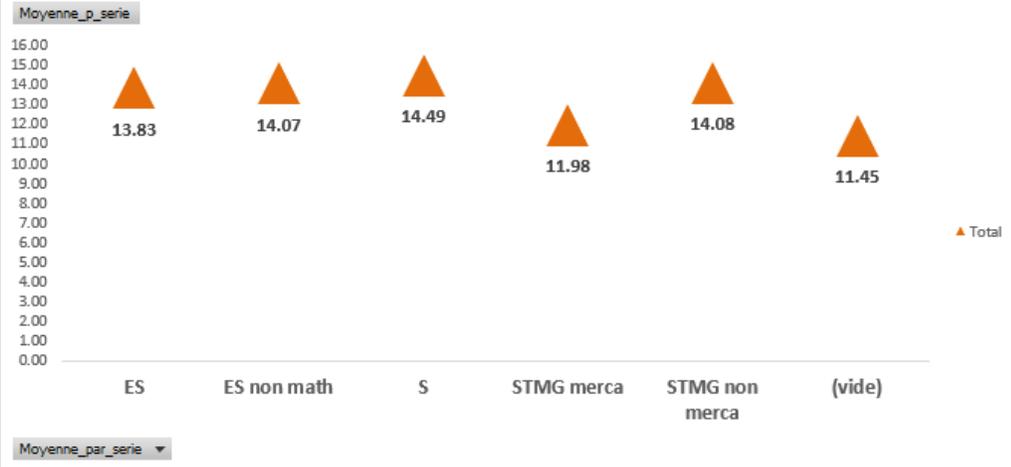
Le coefficient de corrélation de Pearson de -0,0446 entre les données de recrutement et les moyennes respectives suggère une faible corrélation linéaire entre ces variables. Cela indique que les rangs de recrutement des étudiants n'ont pas un impact significatif sur les moyennes obtenues.



Relation entre les séries et les moyennes

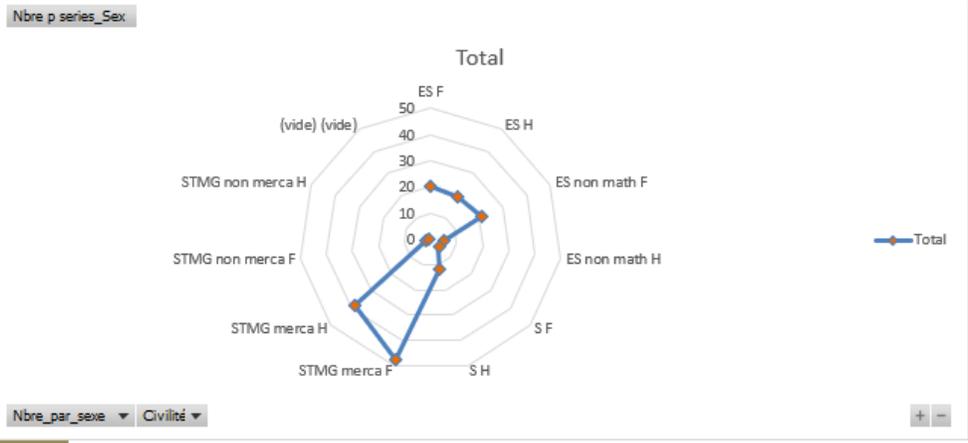
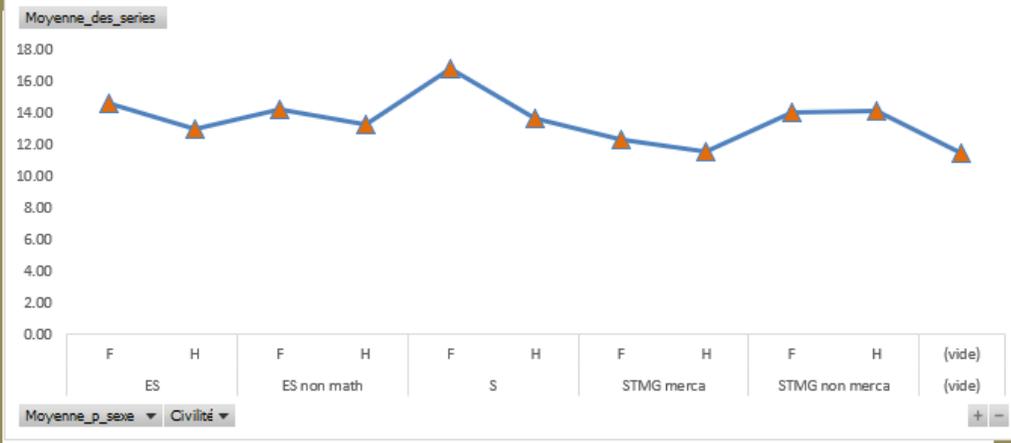
Dans la continuité de nos analyses, nous avons cherché à déterminer s'il existe une relation entre les différentes séries et les moyennes des étudiants. Pour ce faire, nous avons utilisé des tableaux croisés dynamiques pour examiner les moyennes en fonction de chaque série..

TABLEAU DE BORD DE VISUALISATION



Total_par_serie

- ES
- ES non math
- S
- STMG merca
- STMG non merca
- (vide)



Civilité

- F
- H
- (vide)

Remarque

Il est observé que les étudiants en série S ont des moyennes plus élevées que les autres. Les étudiants des filières ES (non mathématiques) et STMG non mercatique présentent des moyennes sensiblement similaires. En revanche, les étudiants de la filière STMG mercatique affichent généralement des moyennes plus basses par rapport aux autres filières. De plus, une tendance se dégage où les filles ont généralement des moyennes plus élevées que les garçons dans les différentes filières.

MERCI!

Bibliographie

Kaensar, C., & Wongnin, W. (2023). Analysis and Prediction of Student Performance Based on Moodle Log Data using Machine Learning Techniques. *International Journal of Emerging Technologies in Learning*, 18(10), 184-203.

Ünal, F. (2020). Data mining for student performance prediction in education. *Data Mining-Methods, Applications and Systems*, 28, 423-432.

Vaarma, M., & Li, H. (2024). Predicting student dropouts with machine learning: An empirical study in Finnish higher education. *Technology in Society*, 102474.