

Développement d'un modèle prédictif d'orientation académique intégré dans une application web intuitive

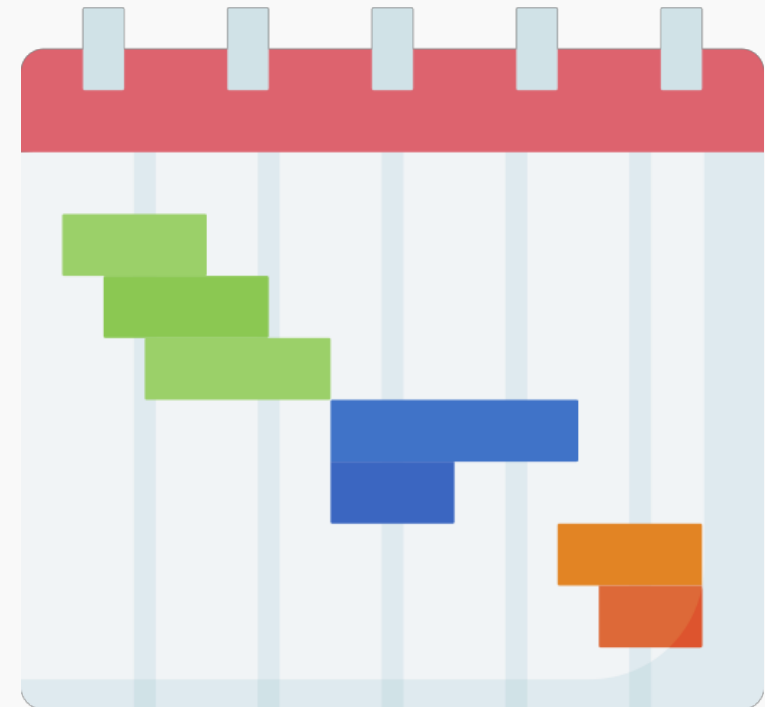
Presenté par :

Keroudine
Belladjo

Encadré par : Camille SABBAH

Plan

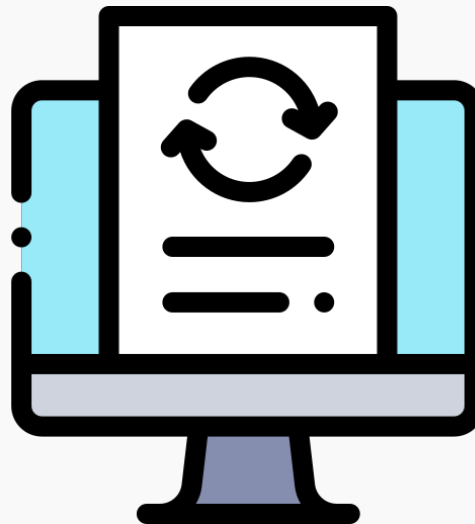
- ☐ **Introduction du projet**
- ☐ **Processus de collecte de données**
 - ✓ Élaboration du questionnaire
 - ✓ Critères de sélection des données
- ☐ **Traitement des données**
 - ✓ Nettoyage et prétraitement
 - ✓ Structuration des réponses
- ☐ **Modèle prédictif d'orientation académique**
 - ✓ Algorithmes d'apprentissage employés
 - ✓ Évaluation des performances des modèles
- ☐ **Intégration du modèle dans une application**
 - ✓ Conception de l'Interface utilisateur



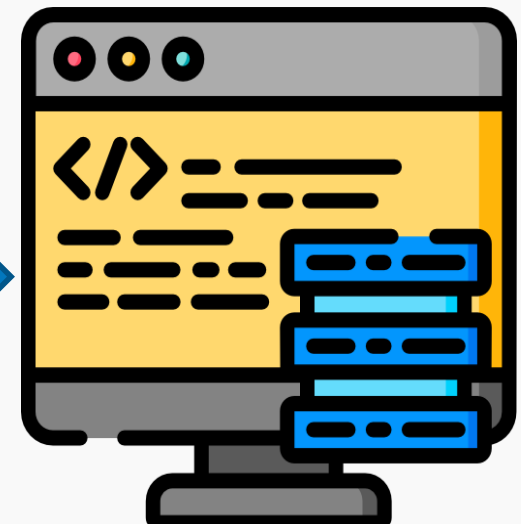
Introduction du projet



Collection des données



Traitement des données



Création d'une application

Processus de collecte de données : Élaboration du questionnaire

Questions

Réponses

219

Paramètres

Total des points : 0

Formulaire d'Orientation Académique

Dans le cadre d'un projet universitaire visant à développer un modèle d'orientation pour les bacheliers et les étudiants, nous aimerions solliciter votre précieuse participation en répondant à quelques questions. Votre contribution est très importante, et soyez assuré(e) que vos réponses seront traitées de manière anonyme. Nous vous remercions sincèrement de bien vouloir partager ce questionnaire avec vos contacts. Votre soutien est grandement apprécié.

10 Question à choix unique
3 Question à choix multiple
9 Questions ouvertes dont 5 sont optionnelles



Keroudine Belladjo • 1er

En recherche d'un stage de fin d'étude dans le domaine de la data (data ...
3 sem. •

Dans le cadre d'un projet universitaire visant à développer un modèle d'orientation pour les bacheliers et les étudiants, nous aimerions solliciter votre précieuse participation en répondant à quelques questions. Votre contribution est très importante, et soyez assuré(e) que vos réponses seront traitées de manière anonyme. Nous vous remercions sincèrement de bien vouloir partager ce questionnaire avec vos contacts. Votre soutien est grandement apprécié.

https://lnkd.in/eNXu_D5d.



Vous et 19 autres personnes

6 replications



r/etudiants · Posted by u/Chocolatine00 4 months ago



Projet de recherche pour créer un modèle d'orientation pour les bacheliers et les étudiants

BONJOUR ! Dans le cadre d'un projet universitaire visant à développer un modèle d'orientation pour les bacheliers et les étudiants, nous aimerions solliciter votre précieuse participation en répondant à quelques questions. Votre contribution est très importante, et soyez assuré(e) que vos réponses seront traitées de manière anonyme.

https://docs.google.com/forms/d/e/1FAIpQLSfFauHjKjKN00ws2v0i1GPNWgvcWVjdKKa44zkog9sF2C7e8w/viewform?usp=sf_link



1 Comment



Share



Save



...

[Lien vers le questionnaire Google form](https://lnkd.in/eNXu_D5d)

Processus de collecte de données : Critères de sélection des données

Complétude des réponses :

	V
1	Question optionnelle: Pouvez-vous formuler une remarque ou avez-vous quelque chose à ajouter ?
2	vu que j'ai répondu à votre questionnaire , je veux une compensation
3	RAS
4	
5	
6	
7	Non
8	

Suppression des colonnes supplémentaires :

Horodateur

Score

Adresse e-mail

Question 1 - 22

Anonymat des réponses et considérations éthiques :

Question optionnelle:
Pouvez-vous formuler une remarque ou av

33758107369

#helshyr

Traitement des données : Nettoyage et prétraitement

Modification des noms des colonnes :

Question 1 - 22

bac	mention	specialite	matière_prefere	filière	type_etu	domaine_etu	type_lieu_etu
pays_bac	tranche_age	genre	qualité	hobbies	domaine_activite_pro	satisfact_formation	
type_orientation	centre_interet	pays_residence	type_cours	type_orientation	remarque		
Niveau_etu	annee_bac	etablissement					

Suppression des espaces :

```
# Suppression des espaces au début et à la fin à tous les éléments textuels des colonnes  
data = data.applymap(lambda x: x.strip() if isinstance(x, str) else x)
```

Remplacement des valeurs manquantes :

```
#remplacer les valeurs manquantes par non renseigné  
data.fillna("non renseigné", inplace=True)
```

Mise en majuscules :

```
# Mettre en majuscule tous les éléments textuels des colonnes  
data = data.applymap(lambda x: x.upper() if isinstance(x, str) else x)
```

Traitement des données : Structuration des réponses

Regroupement des valeurs redondantes

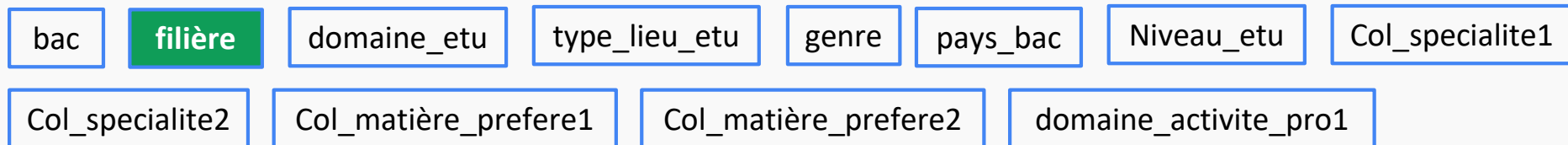
Traitement des réponses à choix multiple :

Traitement des réponses optionnelles :

Modèle prédictif d'orientation académique

Algorithmes d'apprentissage employés

Variable retenu :



Algorithmes d'apprentissage employés :

Random Forest

Un algorithme d'apprentissage supervisé utilisé pour la classification et la régression, construisant plusieurs arbres de décision et combinant leurs prédictions

Régression Logistique

Un algorithme d'apprentissage supervisé utilisé pour la classification, modélisant la probabilité qu'une observation appartienne à une classe particulière en fonction de ses caractéristiques.

K plus proches voisins

Un algorithme d'apprentissage supervisé utilisé pour la classification, où chaque observation est classée en fonction de la classe majoritaire parmi ses k voisins les plus proches dans l'espace des caractéristiques.

Évaluation des performances des modèles

	Précision (%)	Erreur Empirique	Précision en Généralisation (%)	Erreur en Généralisation
Random Forest	65.85	2.52	74.39	2.68
Régression Logistique	83.17	1.87	81.95	1.35
KNN	58.78	3.11	62.73	3.47

Intégration du modèle dans une application

Dans le cadre de notre projet, nous avons mis en œuvre une application interactive intégrant notre modèle d'orientation.

The application, titled "Application pour s'orienter", contains the following fields and their selected values in the right-hand screenshot:

- Bac: SCIENTIFIQUE (ÉQUIVALENT)
- Domaine d'études: SCIENCES ET TECHNOLOGIE
- Type de lieu d'études: UNIVERSITÉ
- Genre: MASCULIN
- Niveau d'études: MASTER
- Spécialité 1: MATHÉMATIQUES
- Spécialité 2: PAS DE SPÉ
- Matière préférée 2: SCIENCES NUMÉRIQUES ET T
- Matière préférée 1: MATHÉMATIQUES
- Domaine d'activité professionnel: DOMAINE DE LA DATA

The result displayed at the bottom of the right-hand application is: "Vous êtes fait pour les filières de MATHÉMATIQUES ET INFORMATIQUE".