



École Supérieure de Technologie de Salé  
Université Mohammed V de Rabat

# PROJET DE FIN D'ETUDE

Plateforme de Détection du harcèlement  
scolaire via l'IA et Data mining

**DEPARTEMENT:** INFORMATIQUE

**FILIERE :** GENIE INFORMATIQUE

**SOUTENU LE:** 17/04/2025

**JURRY:** P.LEFDAOUI/P.ELHAZITI

**ENCADRÉ PAR:** P.LEFDAOUI

**RÉALISÉ PAR:**

YASSINE EL ABBAOUI/ ZINEB MABROUK





École Supérieure de Technologie de Salé  
Université Mohammed V de Rabat

# PROJET DE FIN D'ETUDE

Plateforme de Détection du harcèlement  
scolaire via l'IA et Data mining

## Remerciements

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce projet de fin d'études.

Notre reconnaissance s'adresse tout particulièrement au Professeur LEFADOUI, notre encadrant, dont l'expertise, la disponibilité et la bienveillance ont été déterminantes. Ses conseils avisés et son accompagnement rigoureux nous ont guidés avec justesse à chaque étape de notre démarche.

Nous adressons également nos sincères remerciements à l'ensemble du corps enseignant et administratif de l'École Supérieure de Technologie de Salé, en particulier ceux du Département Informatique, pour la qualité de l'enseignement dispensé et le cadre favorable à notre développement personnel et professionnel.

Nos remerciements vont aussi à nos familles qui ont été notre pilier durant ces années d'études. Leur soutien indéfectible, leurs sacrifices et leur compréhension nous ont donné la force et la persévérance nécessaires pour atteindre nos objectifs.

À nos amis et camarades de promotion, merci pour cette solidarité précieuse, ces moments d'entraide et ces instants de complicité qui ont rendu ce parcours inoubliable.

Enfin, nous remercions toutes les personnes qui ont, de près ou de loin, contribué à la réussite de ce projet.

## Sommaire

|   |           |
|---|-----------|
| <b>Remerciements.....</b>                       |           |
| <b>Sommaire.....</b>                            |           |
| <b>Introduction .....</b>                       | <b>1</b>  |
| o Contexte et Justification .....               | 1         |
| o Objectifs du projet .....                     | 1-2       |
| o Méthodologies adoptée .....                   | 2 -3      |
| <b>Génération et analyse des données.....</b>   | <b>4</b>  |
| o Génération des données .....                  | 4         |
| o Analyse des données(Data mining) .....        | 5-6       |
| <b>Conception de l'Application .....</b>        | <b>7</b>  |
| o Spécifications Fonctionnelles.....            | 7         |
| o Spécifications Techniques .....               | 9         |
| <b>Développement de l'Application .....</b>     | <b>12</b> |
| o Description de l'Implémentation... ..         | 12        |
| o Défis Rencontrés et Solutions Apportées ..... | 20        |
| <b>Perspectives Futures .....</b>               | <b>21</b> |
| <b>Conclusion Générale .....</b>                | <b>22</b> |
| <b>Annexe .....</b>                             | <b>23</b> |
| o Diagrammes.....                               | 23        |
| o Captures d'Écran.....                         | 28        |
| <b>Table de Matières.....</b>                   | <b>38</b> |

## Table de figures

**Figure 1 :** Diagramme de package

**Figure 2 :** Diagramme de cas d'utilisation

**Figure 3 :** Diagramme de classe

**Figure 4 :** Diagramme de flux

**Figure 5 :** Diagramme de gantt

**Figure 6 :** Page de login

**Figure 7 :** Dashboard admin

**Figure 8 :** Dashboard admin

**Figure 9 :** Dashboard admin

**Figure 10 :** Dashboard admin

**Figure 11 :** Dashboard admin

**Figure 12 :** Page d'analyse

**Figure 13 :** Page d'analyse

**Figure 14 :** Page de gestion des élèves

**Figure 15 :** Page de gestion des enseignants

**Figure 16 :** Page de gestion des parents

**Figure 17 :** Page de vue de retours

**Figure 18 :** Page d'accueil pour les parents

**Figure 19 :** Page de Recommendation de l'AI pour les parents

**Figure 20 :** Page de Feedback pour les parents

**Figure 21 :** Page d'accueil pour les enseignants

**Figure 22 :** Page de Recommendation de l'AI pour les enseignants

**Figure 23 :** Page de feedback pour les enseignants

## *I-Introduction*

### *A. Contexte et Justification*

Le **harcèlement scolaire** est souvent difficile à repérer à cause de sa nature insidieuse. Les victimes peuvent hésiter à en parler, et les comportements de harcèlement peuvent parfois passer inaperçus des adultes. La technologie et les nouvelles approches analytiques offrent pourtant des possibilités intéressantes pour identifier des schémas de comportements caractéristiques du harcèlement. L'utilisation des **techniques de fouille de données** permet de traiter des ensembles de données complexes et d'extraire des informations utiles, telles que des indicateurs de comportements anormaux ou des relations sociales tendues entre élèves, facilitant ainsi la détection des cas de harcèlement.

Dans ce cadre, ce projet de fin d'études explore l'utilisation de ces techniques pour analyser des données scolaires simulées, dans le but d'identifier des indices de harcèlement et de proposer des solutions adaptées pour prévenir ce phénomène.

### *B. Objectif du projet*

#### **1. Analyser des données simulées représentant des situations scolaires de harcèlement pour identifier les comportements à risque :**

- Comprendre les profils typiques d'élèves exposés au harcèlement en analysant les variables telles que l'âge, le genre, le niveau scolaire, la fréquence des incidents et leur nature (verbale, physique, cyber).
- Déceler les signaux faibles précurseurs d'une situation de harcèlement (ex : dégradation des performances académiques, isolement social, augmentation des absences).

## **2 .Appliquer des techniques de fouille de données pour extraire des schémas récurrents de harcèlement et anomalies comportementales :**

- . Utilisation de la classification pour évaluer la gravité des cas selon des indicateurs mesurables.
- . Emploi du clustering pour segmenter les élèves selon leur exposition au risque.
- . Mise en place de détection d'anomalies pour identifier les cas extrêmes ou ignorés.

## **3. Proposer des stratégies de prévention basées sur les résultats de l'analyse, en recommandant des actions concrètes pour intervenir de manière précoce :**

- . Suggestions personnalisées pour chaque acteur (parents, enseignants, élèves) en fonction des données analysées.
- . Mise en place d'un moteur de recommandation intelligent pour orienter les interventions pédagogiques ou psychologiques.

## **4. Collecter et analyser les feedbacks des parents et des enseignants sur les suggestions proposées, afin d'ajuster et d'optimiser l'approche pédagogique et émotionnelle :**

### **C. Méthodologie adoptée**

**1. Planification et Analyse :** Nous avons commencé par une phase approfondie de planification et d'analyse pour comprendre les besoins et les exigences du projet. Cela comprenait la définition des fonctionnalités clés, l'identification des technologies à utiliser, et l'estimation des délais et des ressources nécessaires.

- **Préparation des données :** Les données utilisées dans ce projet ont été générées artificiellement pour simuler des situations de harcèlement scolaire dans un environnement éducatif. Un travail de prétraitement a été effectué pour préparer les données à l'analyse.
- **Analyse exploratoire :** Nous avons exploré les données afin d'identifier les tendances et les comportements récurrents associés au harcèlement scolaire.
- **Fouille de données :** Des algorithmes de classification et de clustering ont été appliqués pour détecter des groupes à risque et identifier des motifs comportementaux caractéristiques du harcèlement.
- **Proposition de solutions :** Sur la base des résultats de l'analyse, nous avons développé des suggestions d'actions préventives et des recommandations de soutien adaptées aux élèves concernés.



**2. Choix des Technologies :** L'application repose sur une architecture moderne et performante. MySQL est utilisé pour stocker les données structurées (utilisateurs, élèves, feedbacks, suggestions), avec Sequelize ou TypeORM assurant un accès fluide et sécurisé. Le front-end, conçu avec React.js et stylisé via Tailwind CSS, offre une interface réactive et ergonomique, intégrant React Router pour la navigation, Redux ou Context API pour la gestion d'état, et Axios ou Fetch API pour la communication avec le serveur. Le back-end, développé avec Node.js et Express.js, gère les requêtes API et l'authentification via JWT pour sécuriser les échanges. Enfin, les analyses comportementales sont visualisées à l'aide de Chart.js ou Recharts, permettant une interprétation claire et interactive des données.

**3. Test et Validation :** Tout au long du processus de développement, nous avons effectué des tests rigoureux pour garantir le bon fonctionnement de l'application. Cela comprenait des tests d'unité, des tests d'intégration et des tests de convivialité pour s'assurer que l'application répondait aux normes de qualité élevées.

## ***II-Génération et analyse des données***

### **A. Génération des données**

La génération de données est une phase cruciale dans tout projet d'analyse et d'optimisation des systèmes éducatifs. En effet, disposer de données variées, précises et représentatives est essentiel pour permettre une analyse fine du comportement des élèves, l'engagement des parents, ainsi que l'utilisation de la plateforme. Ces données constituent la base sur laquelle les algorithmes d'intelligence artificielle peuvent établir des modèles de prédiction, générer des recommandations personnalisées et contribuer à une meilleure compréhension des dynamiques éducatives. En synthétisant des informations issues de diverses sources, le système peut ainsi proposer des stratégies d'intervention adaptées et améliorer globalement le suivi pédagogique.

- ***Données démographiques et comportementales :***

Ces données offrent une vue d'ensemble détaillée des caractéristiques des élèves, telles que l'âge, le genre, et d'autres indicateurs socio-économiques ou comportementaux. Elles permettent de segmenter les élèves par groupe et de détecter des tendances ou des écarts dans leurs comportements au quotidien.

- **Engagement parental :**

Cette catégorie regroupe des indicateurs quantitatifs et qualitatifs mesurant la participation des parents dans le suivi scolaire. Elles permettent d'évaluer le niveau d'implication, l'assiduité aux réunions ou toute forme de communication directe entre parents et école, contribuant ainsi à une meilleure compréhension de l'impact de l'engagement familial sur la réussite des élèves.

- **Comportement sur la plateforme :**

Ces informations analysent la manière dont les utilisateurs interagissent avec le système, en observant notamment les temps de connexion, les parcours de navigation et les clics effectués sur différentes sections. Ces données sont essentielles pour optimiser l'ergonomie de la plateforme, identifier les points de friction et adapter l'expérience utilisateur.

- **Changements liés aux médias sociaux :**

Cette catégorie mesure et analyse l'évolution des comportements digitaux des élèves sur les réseaux sociaux. Elle permet d'identifier les tendances, d'évaluer l'impact des plateformes sociales sur la vie scolaire et de détecter d'éventuels signaux faibles de dérives comportementales ou émotionnelles, afin de mettre en place des actions de prévention adaptées.

## **B. Analyse des données (data mining)**

### **1. Classification avec Random Forest**

La classification a été utilisée dans ce projet pour prédire le profil comportemental de chaque élève à partir de plusieurs variables clés : le temps passé sur les réseaux sociaux, le niveau de stress déclaré, le nombre d'absences mensuelles et le nombre d'amis. Cette approche permet de regrouper les élèves en catégories distinctes, facilitant l'identification des individus à risque de harcèlement scolaire.

#### **1-1 :Définition des catégories d'élèves**

Trois catégories ont été définies selon des critères spécifiques :

- **Utilisateurs actifs des réseaux sociaux** : ce groupe comprend les élèves qui passent plus de 5 heures par semaine sur les réseaux sociaux tout en affichant un niveau de stress modéré (inférieur ou égal à 5). Ces élèves sont considérés comme socialement engagés, sans signes apparents de détresse émotionnelle.
- **Élèves à risque potentiel** : ces élèves cumulent plusieurs facteurs de vulnérabilité. Ils présentent un niveau de stress élevé (supérieur à 7), s'absentent fréquemment (plus de 5 absences par mois) et ont peu d'amis (moins de 3). Ce profil est fortement corrélé à un risque accru d'isolement et de harcèlement.
- **Élèves peu engagés** : cette catégorie regroupe les élèves qui ne remplissent aucune des conditions précédentes. Ils ne présentent ni sur-engagement social ni signes de vulnérabilité évidente, mais leur implication générale dans la vie scolaire semble limitée.

#### **1-2 :Prédiction par l'algorithme Random Forest**

Pour prédire ces catégories, un modèle de classification de type Random Forest a été entraîné sur les données simulées. Ce modèle, reconnu pour sa robustesse et sa capacité à gérer des interactions complexes entre variables, a permis de générer une prédiction de catégorie pour chaque élève, facilitant ainsi l'identification des profils à risque de harcèlement scolaire.

### **1-3 :Visualisation et évaluation des résultats**

Les prédictions du modèle ont été présentées sous forme d'un graphique en barres, qui illustre la répartition des élèves dans chacune des catégories définies. Pour valider la performance du modèle, un rapport de classification a été généré, incluant des métriques telles que la précision, le rappel et la F-mesure. Ces indicateurs permettent d'évaluer avec précision l'efficacité du modèle dans la détection et la classification des élèves susceptibles d'être exposés au harcèlement scolaire.

## **2. Analyse des données (Data Mining)**

Le clustering permet de regrouper les élèves en fonction de similarités comportementales, sans avoir à définir d'avance des catégories fixes. Dans notre projet, nous avons exploité l'algorithme KMeans pour segmenter les élèves en trois groupes de risque, en nous basant sur des critères combinant l'utilisation des réseaux sociaux, le niveau de stress, le nombre d'absences et le nombre d'amis.

- **Groupe à risque faible** : Ce cluster regroupe les élèves ayant un faible niveau de stress, peu d'absences et un nombre élevé d'amis. Ces élèves ne présentent pas de signes de vulnérabilité marqués et sont généralement moins susceptibles d'être exposés au harcèlement scolaire.
- **Groupe à risque modéré** : Ce groupe comprend les élèves affichant des indices de vulnérabilité, sans pour autant adopter des comportements extrêmes. Par exemple, un élève avec un niveau de stress modéré, quelques absences et un nombre raisonnable d'amis sera classé dans ce groupe. Ces élèves pourraient bénéficier d'un suivi plus attentif afin d'éviter une détérioration de leur situation.
- **Groupe à risque élevé** : Ce cluster regroupe les élèves qui présentent des signes évidents de vulnérabilité, tels qu'un niveau de stress élevé, des absences fréquentes et un faible nombre d'amis. Ces profils sont les plus exposés au risque de harcèlement scolaire et nécessitent une intervention immédiate.

L'algorithme KMeans génère ainsi une répartition claire des élèves en clusters de risque, permettant d'identifier rapidement les profils les plus vulnérables. Les résultats de cette segmentation sont visualisés par un graphique en barres, qui indique le nombre d'élèves dans chaque groupe de risque, et complétés par un tableau détaillé affichant l'ID des élèves et leur niveau de risque correspondant.

## III-Conception de l'application

### A. Spécifications fonctionnelles

L'application intègre trois rôles – Administrateur, Professeur et Parent – mais seul le rôle d'Administrateur dispose d'un tableau de bord personnalisé, offrant une vue centralisée pour la gestion complète du système, tandis que les rôles de Professeur et de Parent bénéficient d'un ensemble de fonctionnalités adaptées à leurs besoins spécifiques.

Voici une description détaillée des spécifications fonctionnelles pour chaque rôle.

#### 1. Rôle Administrateur

Tableau de Bord de l'Administrateur

- **Visualisations des données générées :**

L'administrateur peut accéder à une série de graphiques interactifs issus des analyses de classification et de clustering. Ces visualisations illustrent la répartition des élèves selon différents niveaux de risque et fournissent des informations détaillées sur leurs comportements, facilitant ainsi la compréhension des tendances et des points sensibles.

- **Vue d'ensemble des analyses :**

Un tableau de bord centralisé permet à l'administrateur de consulter l'ensemble des résultats analytiques en un coup d'œil. Il peut ainsi suivre l'évolution des profils des élèves au fil du temps, identifier rapidement les changements notables et évaluer l'efficacité des mesures d'intervention.

- **Gestion des élèves :**

L'administrateur dispose d'outils pour ajouter, modifier ou supprimer des informations relatives aux élèves. Il peut consulter la liste complète des élèves et accéder à leurs profils détaillés, qui incluent non seulement leurs données personnelles, mais également les suggestions générées par l'IA et les catégories de risque attribuées.

- **Gestion des professeurs :**

Le système permet à l'administrateur de gérer le personnel enseignant : il peut créer, modifier ou supprimer des comptes de professeurs. De plus, l'interface offre la possibilité de consulter les feedbacks fournis par ces derniers et de leur transmettre des recommandations pour améliorer leur prise en charge des élèves.

- **Gestion des parents :**

L'administrateur a également la responsabilité de gérer les comptes des parents. Cette fonctionnalité comprend l'ajout, la modification et la suppression des informations relatives aux parents et à leurs enfants. Par ailleurs, il peut consulter les retours des parents concernant la situation de leurs enfants ainsi que les suggestions de l'IA.

- **Consultation des feedbacks :**

Un espace dédié permet à l'administrateur de visualiser l'ensemble des feedbacks provenant des parents et des professeurs. Cela offre une vue globale de l'impact des actions entreprises et sert de base pour identifier des axes d'amélioration et ajuster les stratégies d'intervention.

## 2. Rôle Parent

Le rôle du parent dans l'application consiste à consulter les Consultation des suggestions de l'IA : suggestions générées par l'intelligence artificielle, ainsi qu'à une fonctionnalité lui permettant de soumettre des feedbacks pour contribuer au suivi et à l'amélioration de l'accompagnement proposé.

- **Consultation des suggestions de l'IA :**

Le parent a la possibilité de consulter les suggestions personnalisées générées par l'IA, qui peuvent inclure des conseils pédagogiques, émotionnels ou sociaux pour améliorer l'accompagnement de l'élève.

- **Soumission de feedback :**

Enfin, le parent peut transmettre un retour d'expérience concernant son enfant, incluant des observations sur son état émotionnel, comportemental ou social. Ces feedbacks sont essentiels pour enrichir le système et adapter les recommandations en fonction de la réalité vécue par les familles.

## 3. Rôle Professeur

Le rôle du professeur dans l'application est centré sur le suivi pédagogique et comportemental de ses propres élèves. Il dispose d'un espace dédié lui permettant de consulter les données clés de ses élèves, d'examiner les recommandations générées par l'IA et de fournir des retours constructifs pour améliorer l'accompagnement.

- **Consultation des suggestions de l'IA :**

Pour chaque élève, l'enseignant peut accéder aux suggestions proposées par l'IA, qu'elles soient d'ordre pédagogique, comportemental ou émotionnel. Ces recommandations permettent d'adapter l'accompagnement aux besoins réels de l'élève.

- **Soumission de feedback :**

Le professeur a la possibilité de fournir des feedbacks sur le comportement, l'engagement et l'état émotionnel des élèves. Ces retours enrichissent les données du système et contribuent à affiner les recommandations de l'IA, pour un suivi plus personnalisé et efficace.

## B. Spécifications techniques

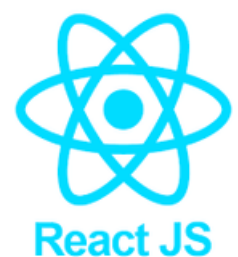
### 1. Base de données (SGBD) : MySQL

L'application utilise MySQL, un système de gestion de base de données relationnelle, pour stocker et organiser les informations des utilisateurs, élèves, feedbacks et suggestions. La base de données est structurée en plusieurs tables : Users, Students, Feedbacks, Suggestions, et Risk\_Analysis. Sequelize ou TypeORM sera utilisé pour gérer l'accès à la base de données.



### 2. Front-End :

- **React.js** : Le front-end sera développé avec React.js, permettant de créer une interface dynamique et réactive. Les composants seront utilisés pour structurer l'application, avec React Router pour la navigation entre les différentes pages. Redux ou Context API gèrera l'état global de l'application, et Axios ou Fetch API pour la communication avec le back-end.
- **Tailwind CSS** : L'interface sera stylisée avec Tailwind CSS, un framework utilitaire moderne qui permet de concevoir rapidement des interfaces responsive et personnalisées. Grâce à ses classes directement intégrées dans le JSX, il offre un contrôle précis sur le design (espacements, couleurs, typographie) sans écrire de CSS manuel. Tailwind s'intègre parfaitement à React, garantissant une interface cohérente, modulable et facile à maintenir.



### 3. Back-End : Node.js

Le back-end sera construit avec Node.js et Express.js pour gérer les requêtes HTTP. L'application interagira avec la base de données MySQL via un ORM tel que Sequelize. JWT sera utilisé pour gérer l'authentification et assurer la sécurité des échanges de données entre le front-end et le back-end.



### 4. Visualisation des Données

Les données seront visualisées à l'aide de bibliothèques comme Chart.js ou Recharts pour afficher des graphiques interactifs représentant les analyses de risque et le comportement des élèves. Ces visualisations permettront aux utilisateurs de mieux comprendre les résultats et d'agir en conséquence.



### 5. Outils Utilisés pour la Classification et le Clustering

- **pandas (pd)**: Bibliothèque essentielle pour la manipulation et l'analyse de données structurées en Python. Elle fournit des structures de données flexibles comme les DataFrames qui permettent le traitement efficace d'ensembles de données tabulaires.
- **matplotlib.pyplot (plt)**: Bibliothèque de visualisation qui permet de créer des graphiques statistiques de haute qualité. Elle offre une interface similaire à MATLAB pour générer des diagrammes à barres, des histogrammes, des nuages de points et bien d'autres représentations visuelles.
- **sklearn.metrics** : Module fournissant des fonctions pour évaluer quantitativement les performances des modèles prédictifs. Le `classification_report` génère un rapport détaillé incluant précision, rappel et score F1 pour chaque classe, permettant d'analyser finement l'efficacité du modèle de classification des étudiants.





- **Python** : Langage de programmation polyvalent et puissant, largement utilisé dans le domaine de la science des données et de l'intelligence artificielle. Dans notre projet, Python a été employé pour le traitement des données, l'application d'algorithmes de classification (Random Forest) et de clustering (KMeans), ainsi que pour la visualisation des résultats. Il a permis une analyse efficace et automatisée des comportements à risque liés au harcèlement scolaire.



## 6.

**Hugging Face** : Plateforme d'intelligence artificielle spécialisée dans le traitement du langage naturel. Grâce à son API accessible via Python, elle permet d'exploiter des modèles NLP avancés (comme BERT ou GPT) pour générer du texte de manière intelligente. Dans notre projet, Hugging Face a été utilisé pour analyser les feedbacks textuels et produire des recommandations personnalisées adaptées aux profils des élèves.



## ***IV-Développement de l'application***

### **A. Description de l'Implémentation**

L'implémentation de notre application s'est déroulée en plusieurs étapes, chacune visant à concrétiser les fonctionnalités définies dans la phase de conception. Voici un aperçu des principales étapes d'implémentation :

#### **1. Développement de l'interface de login**

Mise en place d'une interface de connexion simple et sécurisée, comprenant les champs suivants :

- Implémentation d'une interface de consultation et de suivi des feedbacks soumis par les parents et les professeurs, permettant d'ajuster les recommandations proposées et de renforcer l'accompagnement des élèves.
  - Adresse email ou nom d'utilisateur.
  - Mot de passe.
  - Rôle de l'utilisateur (administrateur, professeur ou parent), sélectionné via un menu déroulant
- Un bouton « Se connecter » permet d'authentifier l'utilisateur et de le rediriger vers l'espace correspondant à son rôle (dashboard admin, espace professeur ou espace parent).

#### **2. Développement de l'interface administrateur**

- Mise en place d'un tableau de bord administrateur centralisant la gestion des élèves, des professeurs, des parents et des analyses de risque générées par le système.
- Intégration des fonctionnalités d'ajout, de modification et de suppression des données liées aux utilisateurs (élèves, parents, professeurs).
- Gestion des analyses de risque issues des algorithmes de classification et de clustering, avec visualisation des résultats sous forme de graphiques interactifs (répartition par catégorie, évolution des cas, etc.).
- Implémentation d'une interface de consultation et de suivi des feedbacks soumis par les parents et les professeurs, permettant d'ajuster les recommandations proposées et de renforcer l'accompagnement des élèves.

### 3. Développement de l'interface de suggestions par l'IA

Mise en place d'une interface dédiée à l'affichage des recommandations personnalisées générées par l'IA, à destination des administrateurs et enseignants, dans le but d'améliorer le suivi et l'accompagnement des élèves identifiés comme à risque.

#### 3.1 Intégration de l'API Hugging

##### 3.1.1 Objectif de l'intégration

L'un des objectifs clés de ce projet est de fournir un accompagnement ciblé et personnalisé aux élèves en difficulté, en s'appuyant sur l'analyse de leur profil individuel. Pour cela, nous avons intégré l'**API Hugging Face**, qui permet d'interagir avec des modèles de traitement du langage naturel (NLP) avancés, capables de générer du texte cohérent, contextualisé et pertinent.

Cette intégration vise à produire automatiquement deux types de recommandations :

- **Pour les parents** : conseils pratiques, activités à domicile, stratégies d'encadrement émotionnel et académique.
- **Pour les enseignants** : interventions pédagogiques ciblées, aménagements de classe, stratégies différenciées.

##### 3.1.2 Architecture de la solution

L'architecture suit les étapes suivantes :

**1. Collecte des données de l'élève** : informations de base (âge, sexe, classe) et observations spécifiques liées à son comportement, ses performances académiques et son bien-être socio-émotionnel.

**2. Prétraitement** : génération d'un texte d'observation synthétique à partir des données brutes, via une fonction JavaScript personnalisée (`constructObservation()`).

**3. Génération du prompt** : création d'un prompt structuré en langage naturel, contenant à la fois les métadonnées de l'élève et la description synthétique de son profil.

Appel à l'API Hugging Face :

**4. Appel à l'API Hugging Face** :

- Clé d'authentification API utilisée pour sécuriser l'appel.
- Modèle NLP appelé pour générer le texte de sortie à partir du prompt.

### 4.1 Exemple de prompt envoyé à l'API

```
const API_KEY = "hf_EvyJVJULOUTS2RbGBxShr0mdThxygxkEAO"; // Replace with your actual key

// Construct an observation from student data (excluding basic fields)
const observationText = constructObservation(student);

const prompt = `
En tant que spécialiste de l'enfant et de l'éducation axé sur l'orientation parentale, veuillez analyser les données du profil de l'élève suivant et créer un plan de remédiation complet adapté aux parents.
Utilisez les informations fournies sur l'élève pour générer un plan comprenant des recommandations claires et réalisables ainsi que des activités de remédiation que les parents peuvent mettre en œuvre à la
maison pour soutenir le développement académique, émotionnel et social de leur enfant.

ID de l'élève: ${student.ID}
Âge: ${student.Age}
Genre: ${student.Gender}
Classe: ${student.Grade}

Données du profil de l'élève:
${observationText}

Votre réponse doit inclure:
1. Un plan détaillé pour l'implication parentale et la remédiation.
2. Des activités ou interventions spécifiques que les parents peuvent utiliser pour soutenir l'apprentissage et le bien-être émotionnel de l'enfant.
3. Des stratégies pratiques pour améliorer les performances académiques et les compétences sociales de l'enfant.
4. Un calendrier suggéré ou une approche par étapes pour la mise en œuvre de ces stratégies.

Assurez-vous que les recommandations sont claires, pratiques et adaptées spécifiquement au rôle des parents dans le soutien de leur enfant.
`;
```

### 4.2 Modèle NLP utilisé

Pour la génération des recommandations personnalisées à destination des **parents** et des **enseignants**, nous utilisons l'**API de Hugging Face** qui donne accès à plusieurs modèles de **Text Generation**. Ces modèles sont conçus pour générer du texte de manière cohérente à partir d'un prompt fourni, en suivant une consigne structurée (Instruction-following).

Dans le cadre de ce projet, plusieurs modèles ont été testés ou considérés selon les critères suivants : **qualité du langage généré en français**, **capacité à suivre des instructions complexes**, et **disponibilité via l'API Hugging Face**. Les modèles principalement utilisés ou compatibles sont :

- **GPT-2** (gpt2) : un modèle de génération de texte classique, plus léger, utilisé pour des tests rapides ou des prototypes.
- **GPT-Neo** (EleutherAI/gpt-neo-1.3B) : un modèle open-source de génération plus puissant que GPT-2, capable de produire des réponses plus longues et plus nuancées.
- **LLaMA** (Meta's **Large Language Model Meta AI**) : via les modèles llama-2 disponibles sur Hugging Face en version fine-tunée pour l'instruction (llama-2-7b-chat, etc.). Ces modèles offrent une meilleure cohérence, surtout pour des consignes en français.

### 4.3 Avantages de l'approche

- **Personnalisation automatisée** : génération de contenus adaptés au profil de chaque élève.
- **Gain de temps** : réduction de la charge de travail des professionnels.
- **Accessibilité** : les recommandations sont rédigées dans un langage clair pour les parents non spécialistes.
- **Cohérence pédagogique** : coordination entre les actions des parents et celles des enseignants.

### 4.4 Limites et perspectives

Bien que la génération soit efficace, certaines limites subsistent :

- Qualité variable selon les données fournies.
- Nécessité de relecture humaine pour éviter les recommandations génériques.
- Sensibilité aux biais du modèle.

**5. Post-traitement** : formatage des réponses en deux blocs distincts : un pour les parents, un pour les enseignants.

## 4. Développement de phase du data mining

### 2-1. Classification avec Random Forest

- **Préparation des données** :

✓ Chargement d'un ensemble de données démographiques et comportementales.

✓ Création d'une nouvelle colonne "Student Category" qui définit trois catégories selon des critères spécifiques :

- "Active Social Media User" (utilisateur actif des réseaux sociaux)
- "Potential at-Risk Student" (étudiant potentiellement à risque)
- "Low Engagement Student" (étudiant peu engagé)

- **Modélisation**

✓ Utilisation des caractéristiques comme l'usage des réseaux sociaux, les absences, le niveau de stress et le nombre d'amis.

✓ Application d'un algorithme Random Forest avec 100 estimateurs pour classifier les étudiants.

- **Évaluation et visualisation**

✓ Génération d'un rapport de classification pour évaluer les performances du modèle.

✓ Création d'un graphique à barres pour visualiser la distribution des catégories prédites.

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report

# Load the dataset (use the demographics data with new labels)
df = pd.read_csv("demographics_behavior_100.csv")

# Create a new column for categories based on conditions
def categorize_student(row):
    if row['Social Media Usage'] > 5 and row['Stress (1-10)'] <= 5:
        return 'Active Social Media User'
    elif row['Stress (1-10)'] > 7 and row['Absences/Month'] > 5 and row['Friends'] < 3:
        return 'Potential at-Risk Student'
    else:
        return 'Low Engagement Student'

df['Student Category'] = df.apply(categorize_student, axis=1)

# Features and target
X = df[['Social Media Usage', 'Absences/Month', 'Stress (1-10)', 'Friends']]
y = df['Student Category']

# Split the data into training and test sets (80% for training, 20% for testing)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create and train the Random Forest model
clf = RandomForestClassifier(n_estimators=100, random_state=42)
clf.fit(X_train, y_train)

# Predict on test data
y_pred = clf.predict(X_test)

# Visualization of predicted categories (Bar chart)
category_counts = pd.Series(y_pred).value_counts()
category_counts.plot(kind='bar', color=['blue', 'green', 'orange'])
plt.title('Predicted Student Categories')
plt.xlabel('Category')
plt.ylabel('Number of Students')
plt.xticks(rotation=0)
plt.show()

# Print classification report
print("\nClassification Report:\n")
print(classification_report(y_test, y_pred))
```

## 2-2. Clustering avec K-Means

- **Préparation des données :**

- ✓ Utilisation des mêmes caractéristiques que pour la classification.
- ✓ Pas de catégories prédéfinies car il s'agit d'un apprentissage non supervisé.

- **Modélisation:**

- ✓ Application de l'algorithme K-Means avec 3 clusters.
- ✓ Définition des étiquettes pour les clusters identifiés : "Low Risk", "Moderate Risk" et "High Risk".

- **Visualisation et interprétation**

- ✓ Création d'un graphique à barres pour montrer la distribution des étudiants selon les niveaux de risque.
- ✓ Affichage du dataframe mis à jour incluant les identifiants des étudiants et leurs niveaux de risque.

```
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import pandas as pd

# Load the data (features for clustering)
df = pd.read_csv("demographics_behavior_100.csv")
X = df[['Social Media Usage', 'Absences/Month', 'Stress (1-10)', 'Friends']]

# Apply KMeans clustering with 3 clusters (Low, Moderate, High Risk)
kmeans = KMeans(n_clusters=3, random_state=42)
df['Cluster'] = kmeans.fit_predict(X)

# Map clusters to risk levels (based on feature importance or average cluster values)
cluster_labels = {0: 'Low Risk', 1: 'Moderate Risk', 2: 'High Risk'}
df['Risk Level (Cluster)'] = df['Cluster'].map(cluster_labels)

# Visualization: Bar Chart showing the distribution of students across risk levels
risk_level_counts = df['Risk Level (Cluster)'].value_counts()
risk_level_counts.plot(kind='bar', color=['green', 'orange', 'red'])
plt.title('Distribution of Students Across Risk Levels')
plt.xlabel('Risk Level')
plt.ylabel('Number of Students')
plt.xticks(rotation=0) # Keeps x-axis labels horizontal
plt.show()

# Display the updated dataframe with risk levels from clustering
print(df[['Student ID', 'Risk Level (Cluster)']])
```

## **5. Développement de l'interface du parent**

Création d'une interface dédiée permettant aux parents de se connecter et d'accéder aux recommandations personnalisées générées par l'IA concernant leur enfant. Cette interface a pour objectif de favoriser la communication entre l'école et les familles, en impliquant les parents dans le suivi et l'accompagnement de leur enfant.

- **Affichage des recommandations de l'IA :**

- + Le parent peut consulter les suggestions issues des analyses automatisées basées sur les comportements observés de l'élève (ex. : signes d'isolement, baisse de participation, interactions sociales limitées).

- **Mécanisme de feedback :**

- + Le parent peut donner son avis ou partager ses observations concernant l'efficacité des recommandations ou les évolutions constatées à la maison.

- + Ces retours sont pris en compte par le système d'IA pour ajuster dynamiquement les futures suggestions.

- **Suivi collaboratif :**

- + L'interface favorise une collaboration tripartite entre parents, enseignants et administrateurs, afin de garantir un accompagnement individualisé et réactif des élèves à risque.

## **6. Développement de l'interface du Professeur**

Mise en place d'une interface dédiée aux enseignants, permettant une connexion sécurisée afin de consulter et d'évaluer les recommandations personnalisées générées par l'IA pour l'ensemble des élèves, notamment ceux identifiés comme présentant des signaux de risque. Cette interface vise à renforcer la collaboration entre les enseignants, les parents et les administrateurs pour un suivi éducatif plus réactif et coordonné.



- **Affichage des recommandations de l'IA :**

- + Une fois connecté, l'enseignant peut visualiser les suggestions issues des analyses de risque de l'IA (ex. : signes d'isolement, harcèlement verbal, baisse d'interactions) pour chaque élève ou groupe d'élèves.

- **Mécanisme de feedback :**

- + L'enseignant peut fournir des retours sur l'efficacité des recommandations et signaler des observations sur le comportement ou les progrès des élèves.

- + Ces feedbacks permettent à l'IA d'affiner continuellement ses analyses et propositions.

- **Suivi collaboratif :**

- +L'interface facilite la communication entre enseignants, parents et administrateurs, permettant ainsi une prise en charge coordonnée et proactive des élèves à risque.

## **B. Défis Rencontrés et Solutions Apportées**

Au cours du développement de l'application, plusieurs défis techniques et fonctionnels ont été rencontrés. L'un des principaux défis a été la simulation de données réalistes pour entraîner les modèles de classification et de clustering, en l'absence de données réelles liées au harcèlement scolaire. Ce problème a été résolu par la génération d'un jeu de données synthétique structuré à l'aide de bibliothèques Python comme Pandas et NumPy, en s'appuyant sur des cas types et des études existantes. Un autre défi important concernait l'interprétation des résultats algorithmiques pour les rendre compréhensibles aux utilisateurs finaux (administrateurs, professeurs, parents). Pour y répondre, des visualisations interactives ont été intégrées via Chart.js, permettant une lecture simple et intuitive des niveaux de risque et des catégories d'élèves.

## ***V-Perspectives Futures***

Dans cette section, nous présentons les opportunités d'évolution fonctionnelle de l'application ainsi que les axes stratégiques envisagés pour son déploiement à plus grande échelle au sein d'établissements scolaires et institutions partenaires.

### **A. Intégration de Nouvelles Fonctionnalités**

- Ajout d'un module de suivi psychopédagogique, permettant aux professionnels d'évaluer les progrès et de proposer un accompagnement personnalisé.
- Développement d'une application mobile multiplateforme (Android/iOS) pour une accessibilité facilitée à tous les utilisateurs.

### **B. Déploiement de l'Application au Sein des Établissements Scolaires**

- Stratégie de déploiement ciblé auprès d'établissements scolaires pilotes (collèges, lycées), pour une première phase de test en conditions réelles.
- Mise en avant des bénéfices pour les établissements partenaires, notamment en termes de prévention, de bien-être scolaire et de conformité aux politiques anti-harcèlement.
- Plan de communication et de sensibilisation à destination des acteurs éducatifs pour encourager l'adoption de l'outil (ateliers, démonstrations, supports pédagogiques).

## ***VI-Conclusion générale***

La conception de cette application dédiée à la détection et à la prévention du harcèlement scolaire constitue une réponse innovante et pertinente aux enjeux actuels du milieu éducatif. Face à la montée des cas de harcèlement, notamment dans les environnements numériques, il devient essentiel de doter les établissements scolaires d'outils intelligents, capables d'identifier les signaux faibles et de proposer des interventions ciblées.

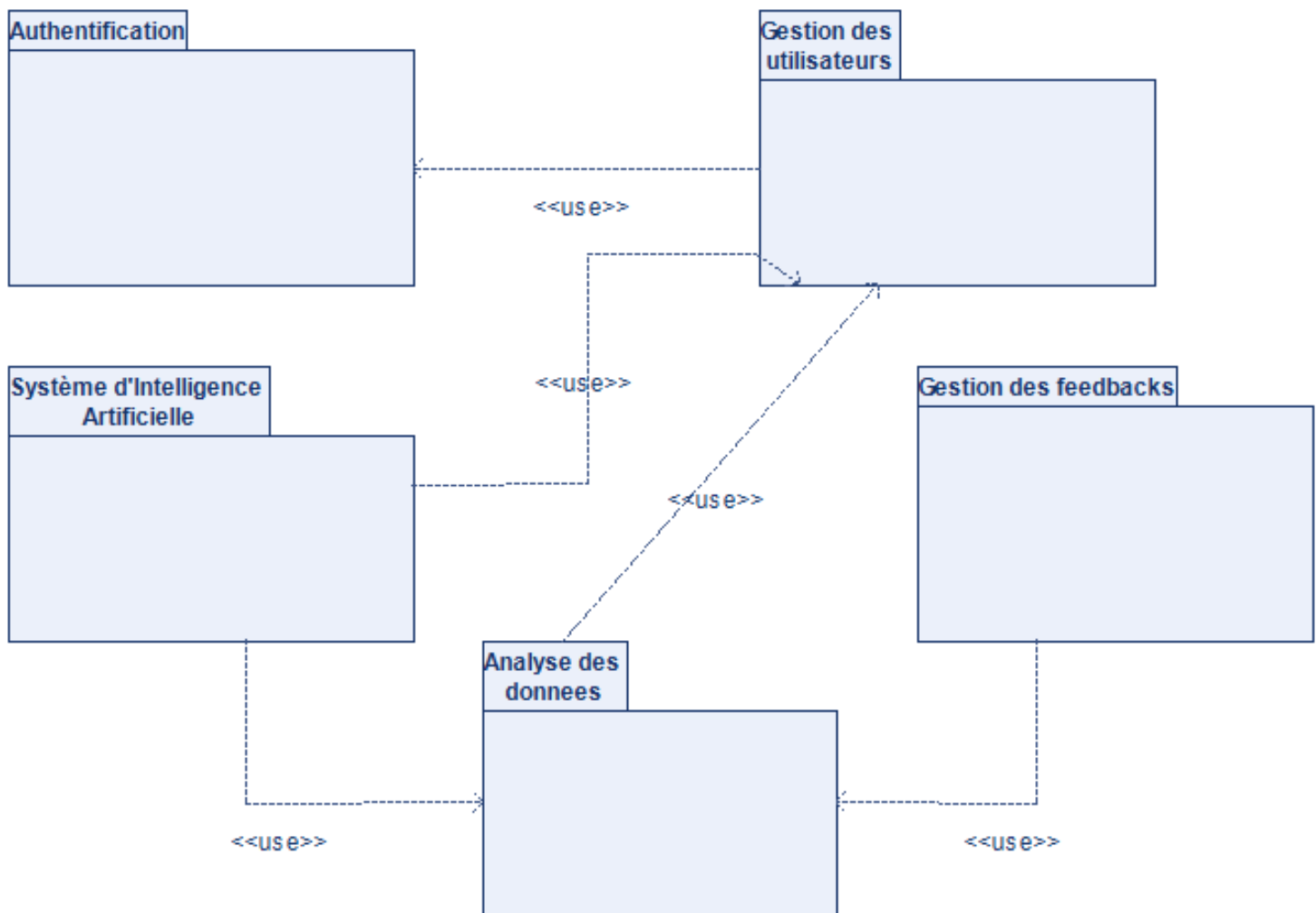
À travers ce projet, nous avons mis en place une plateforme fonctionnelle, intégrant un tableau de bord administratif complet, des interfaces personnalisées pour les parents et les professeurs, ainsi qu'un système d'analyse de données fondé sur des algorithmes de classification et de clustering. Ces algorithmes permettent de catégoriser les élèves en fonction de leur niveau de risque et de générer des recommandations adaptées, appuyées par l'intelligence artificielle.

Grâce aux fonctionnalités implémentées – telles que la consultation des profils élèves, les visualisations des risques, la gestion des feedbacks et les suggestions de l'IA – nous avons posé les bases d'un outil d'aide à la décision orienté vers le bien-être scolaire.

En somme, ce projet représente non seulement une réalisation académique aboutie, mais également une contribution concrète au monde éducatif, en exploitant les technologies de l'intelligence artificielle au service de la prévention et du suivi personnalisé des élèves. Nous espérons qu'à travers cette solution, il sera possible de renforcer la détection précoce des situations à risque et de promouvoir un climat scolaire plus sain et plus sécurisé.

## VII-Annexe

### A. Diagrammes



**Figure 1 : Diagramme de package**

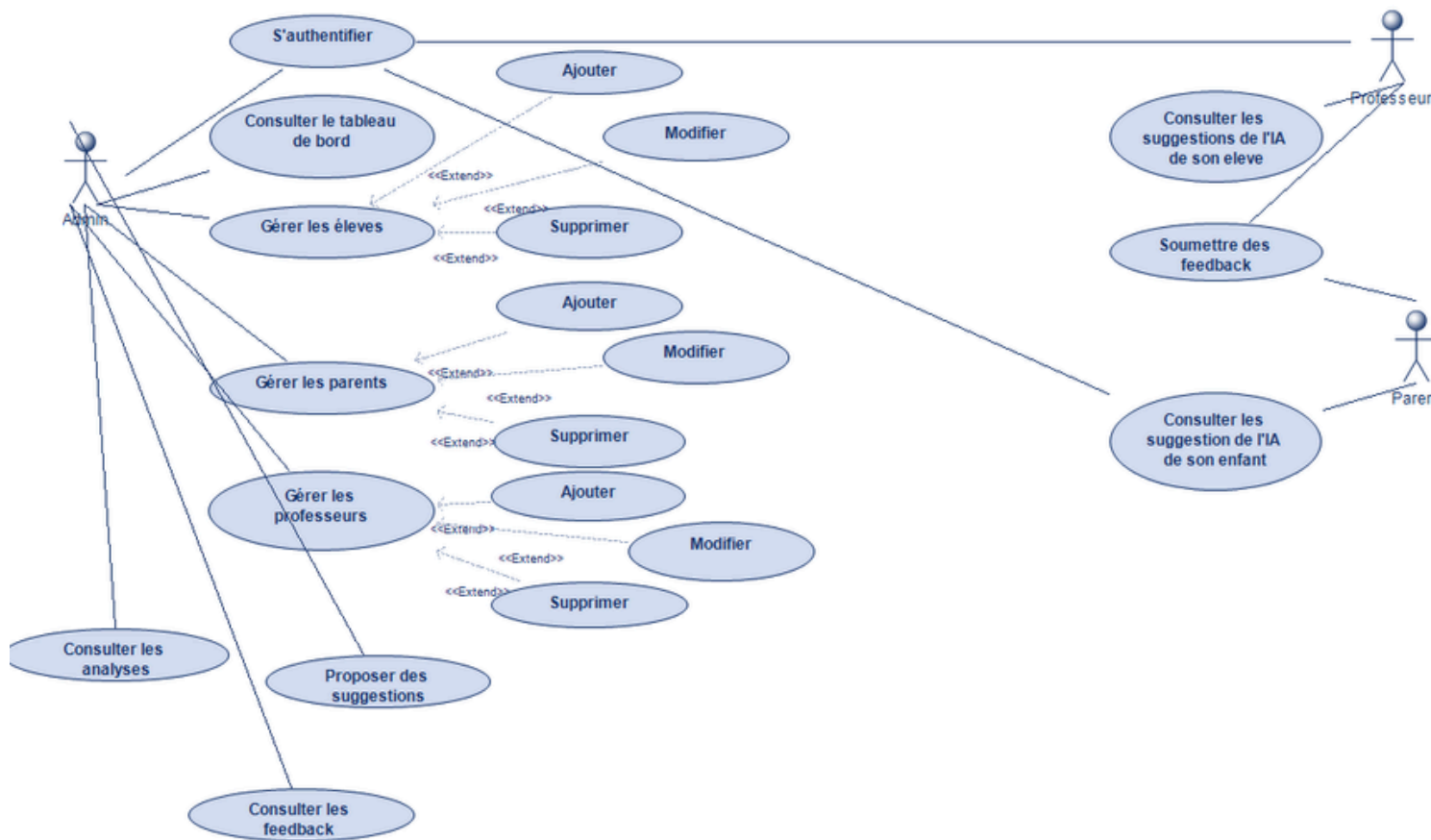


Figure 2 : Diagramme de cas d'utilisation

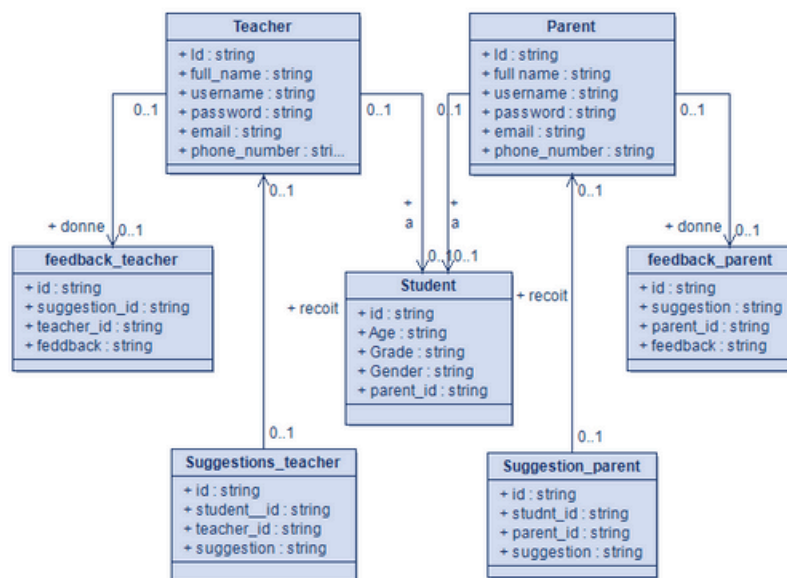


Figure 3 : Diagramme de classe

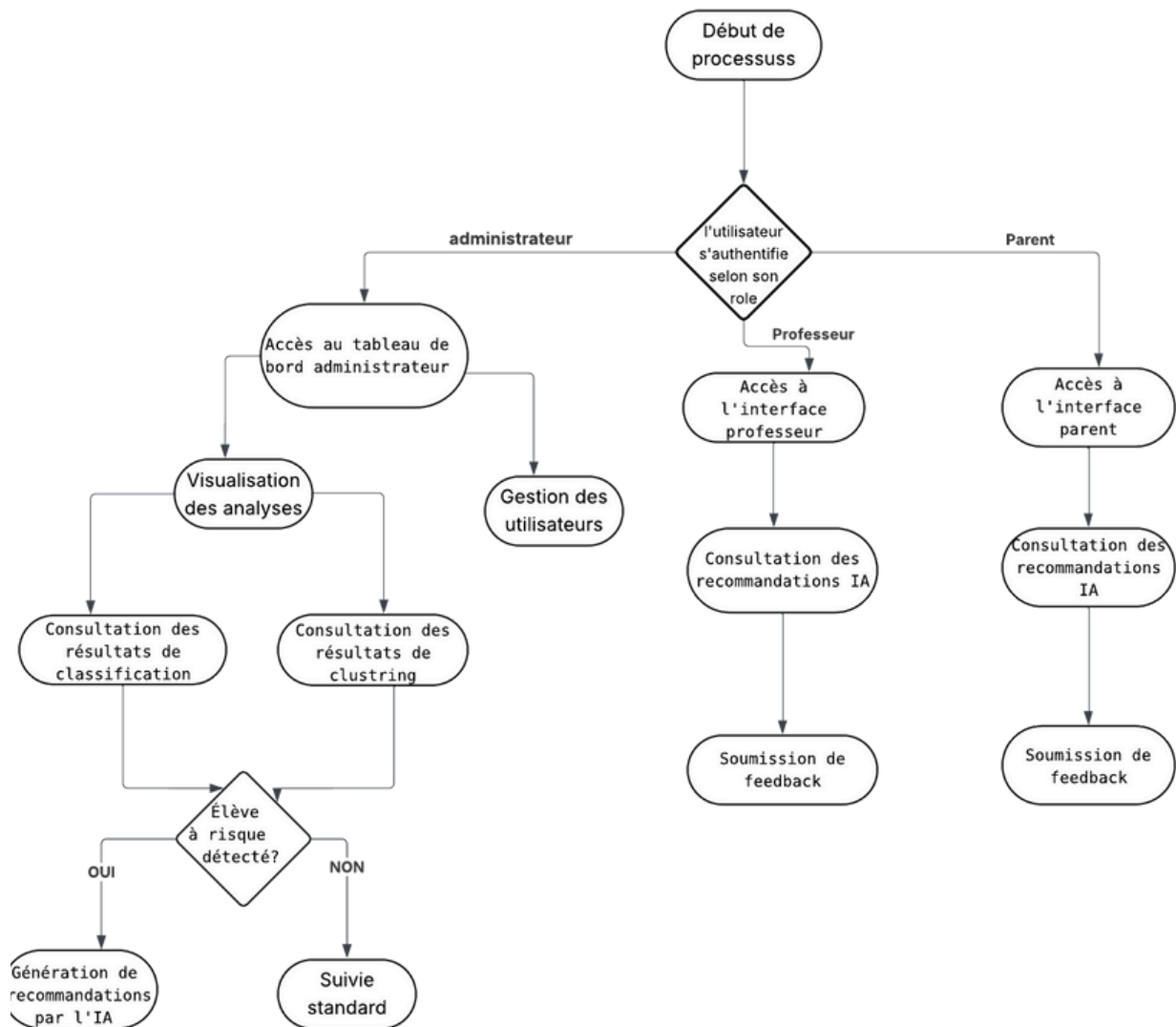
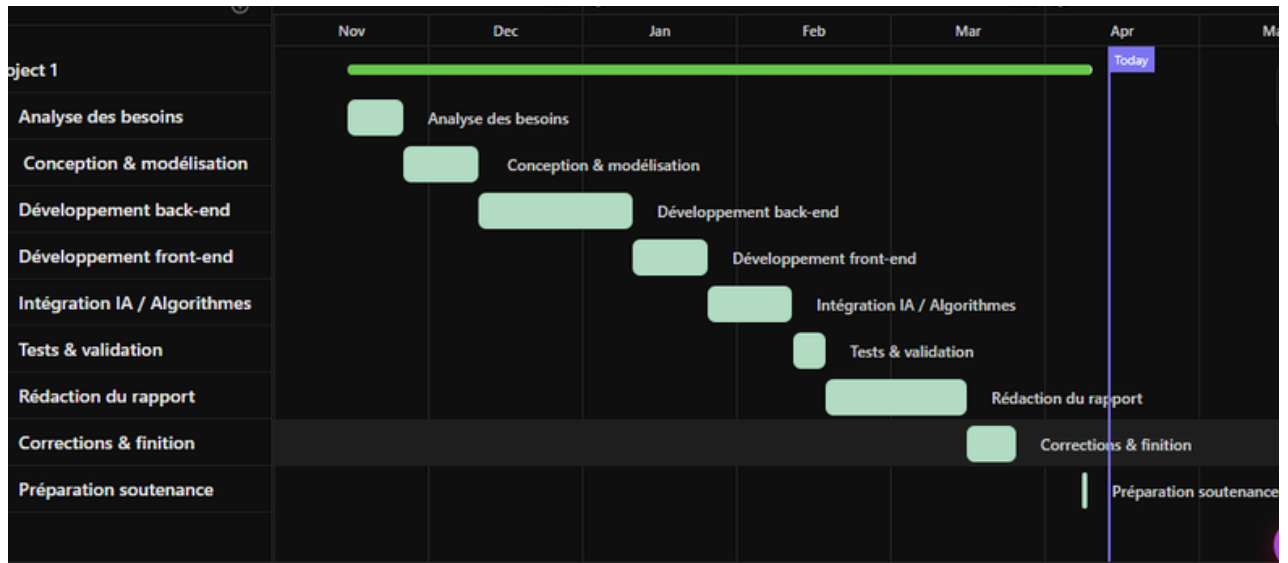


Figure 4 : Diagramme de flux





| Name                           | Assignee | Due date   | Priority | Status     | Comments |  |
|--------------------------------|----------|------------|----------|------------|----------|--|
| ✓ Analyse des besoins          | 👤        | 11/25/24   | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Conception & modélisation    | 👤        | 12/10/24   | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Développement back-end       | 👤        | Jan 10     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Développement front-end      | 👤        | Jan 25     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Intégration IA / Algorithmes | 👤        | Feb 10     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Tests & validation           | 👤        | Feb 16     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Rédaction du rapport         | 👤        | Mar 15     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Corrections & finition       | 👤        | Mar 25     | 🚩        | ✓ COMPLETE | 💬        |  |
| ✓ Préparation soutenance       | 👤        | 5 days ago | 🚩        | ✓ COMPLETE | 💬        |  |

Figure 5 : Diagramme de gantt

## B. Captures d'Écran

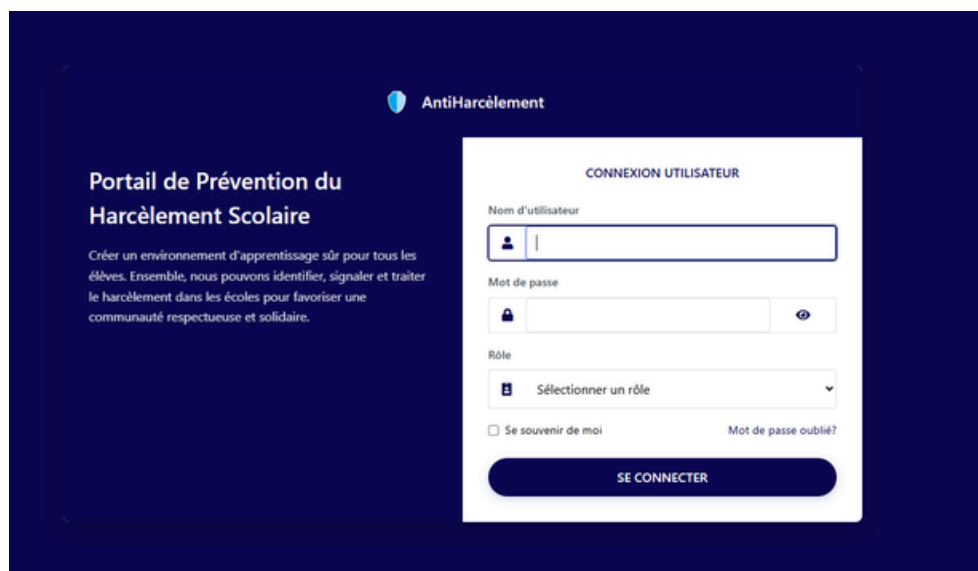


Figure 6 : Page de login

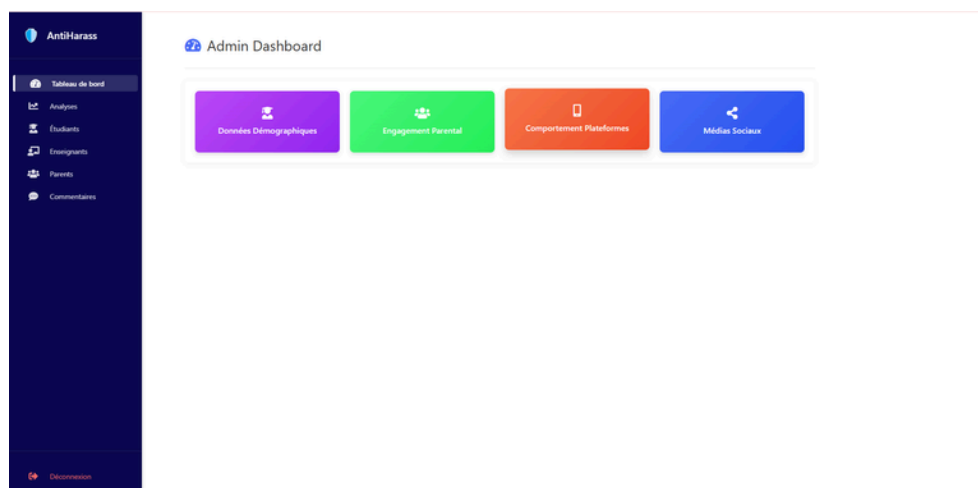


Figure 7 : Dashboard Admin

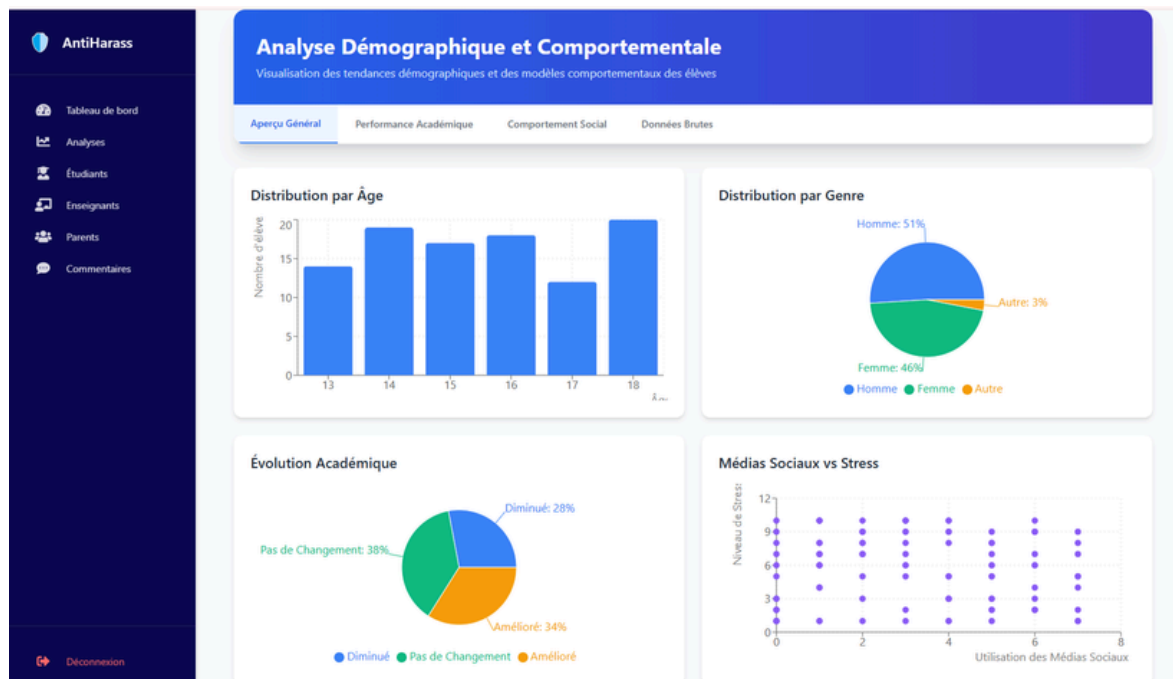


Figure 8 : Dashboard Admin

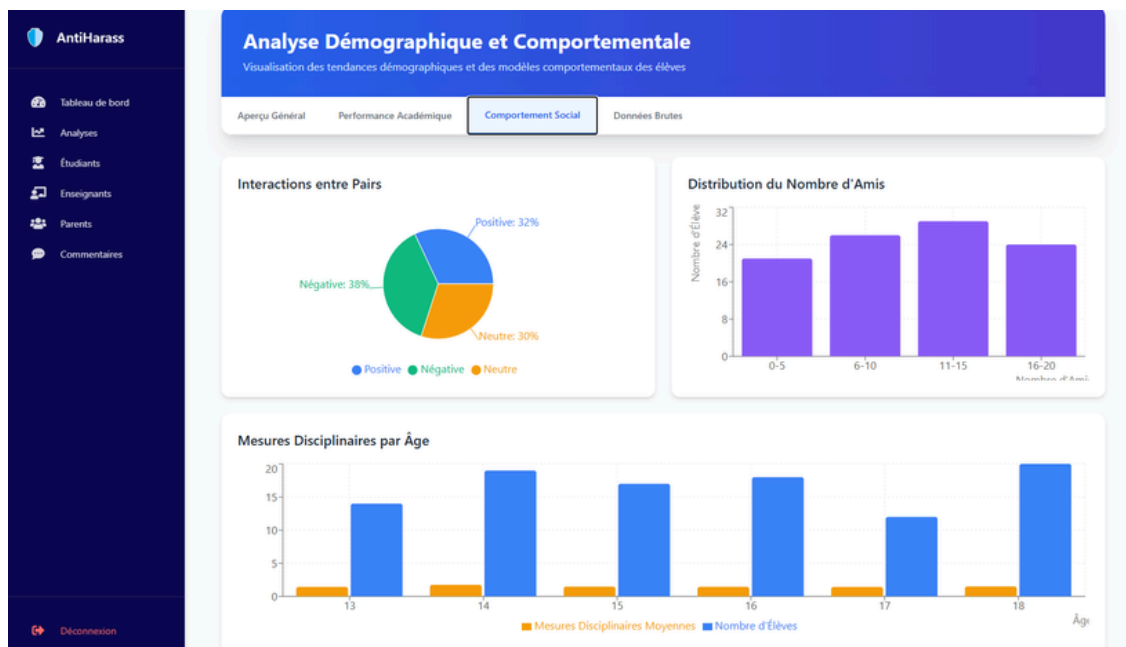


Figure 9 : Dashboard Admin

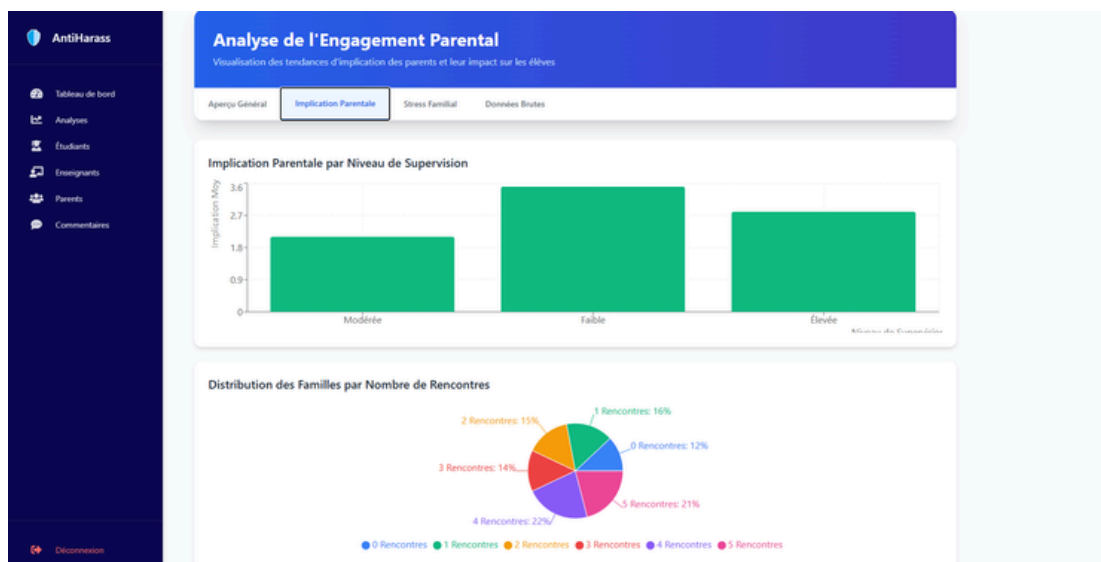


Figure 10 : Dashboard Admin

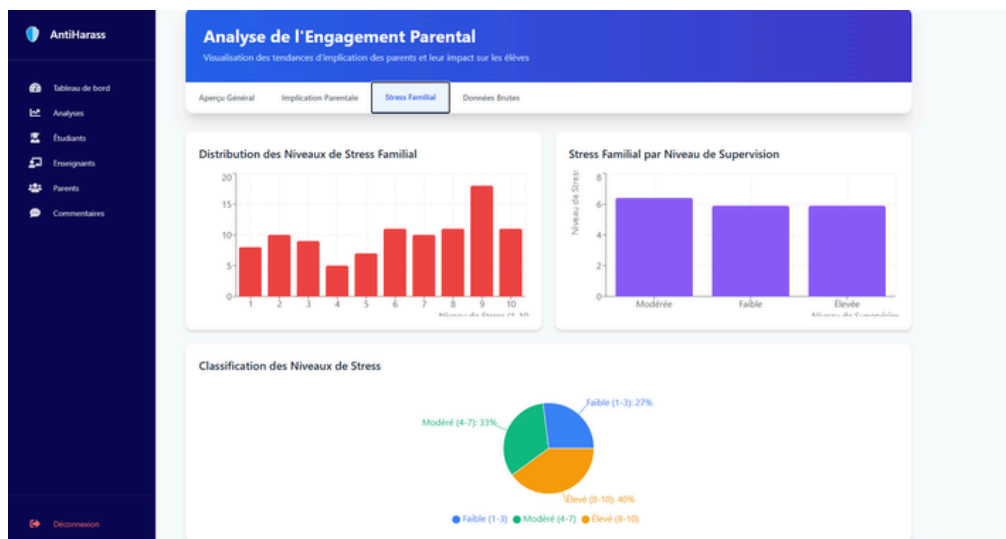


Figure 11 : Dashboard Admin

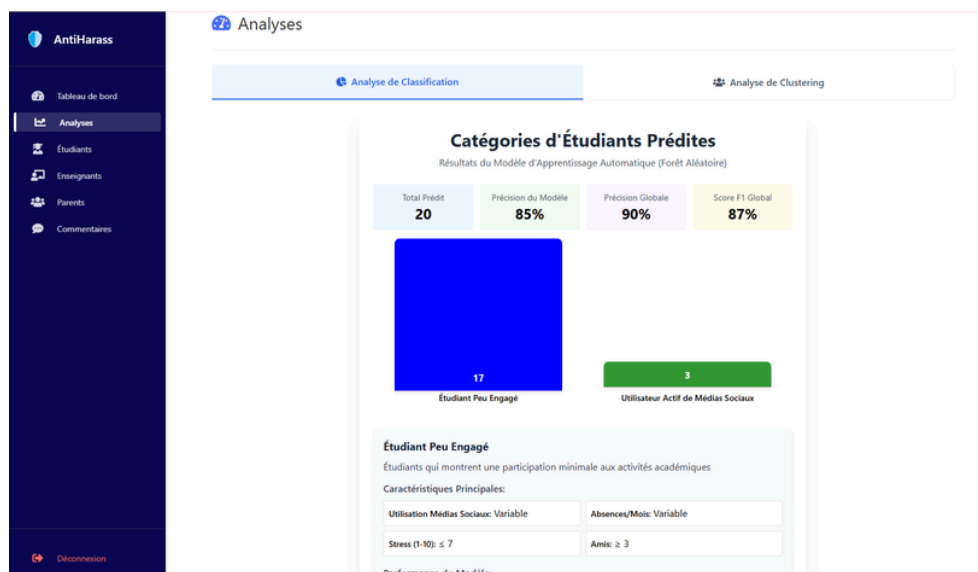


Figure 12 : Page d'analyse

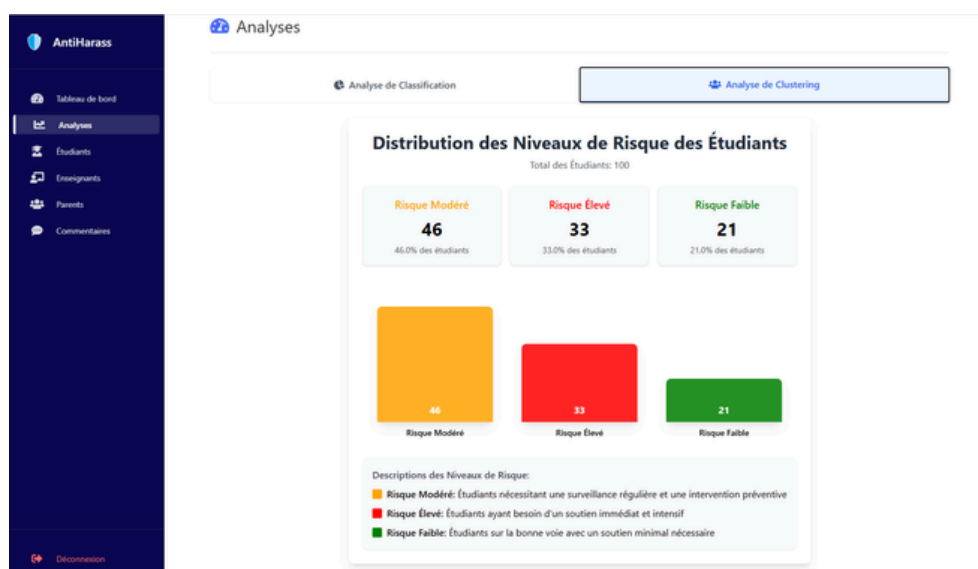


Figure 13 : Page d'analyse

**AntiHarass**

Tableau de bord  
Analyses  
**Étudiants**  
Enseignants  
Parents  
Commentaires

Déconnexion  
sachin@2000-student

### Gestion des élèves

Gérer les profils et les affectations des élèves

Rechercher par ID ou niveau... Rechercher

| ID | Âge | Sexe   | Niveau | Statut  | Actions |
|----|-----|--------|--------|---------|---------|
| 1  | 18  | Male   | D      | Inactif |         |
| 2  | 18  | Female | C      | Inactif |         |
| 3  | 15  | Male   | C      | Inactif |         |
| 4  | 13  | Female | C      | Inactif |         |
| 5  | 15  | Male   | C      | Inactif |         |
| 6  | 16  | Female | B      | Inactif |         |
| 7  | 13  | Female | C      | Inactif |         |
| 8  | 14  | Male   | C      | Inactif |         |
| 9  | 13  | Male   | A      | Inactif |         |
| 10 | 16  | Male   | B      | Inactif |         |

Page 1 sur 10

Figure 14: Page de gestion des eleves

**AntiHarass**

Tableau de bord  
Analyses  
Étudiants  
**Enseignants**  
Parents  
Commentaires

Déconnexion  
sachin@2000-teacher

### Gestion des enseignants

Gérer les comptes et informations des enseignants

Rechercher par nom ou email... Rechercher + Ajouter un enseignant

| ID | Nom complet | Nom d'utilisateur | Email             | Téléphone | Actions |
|----|-------------|-------------------|-------------------|-----------|---------|
| 1  | S SSSS      | ss                | ss@gmail.com      | 05        |         |
| 2  | S SS        | ssss              | s@gmail.com       | 85        |         |
| 3  | P parent1   | parent1           | parent1@gmail.com | 06596826  |         |
| 4  | T teacher   | teacher1          | yy@gmail.com      | 58        |         |
| 5  | S SS        | s                 | s@gmail.vom       | 5         |         |

Page 1 sur 1

Figure 15: Page de gestion des enseignants

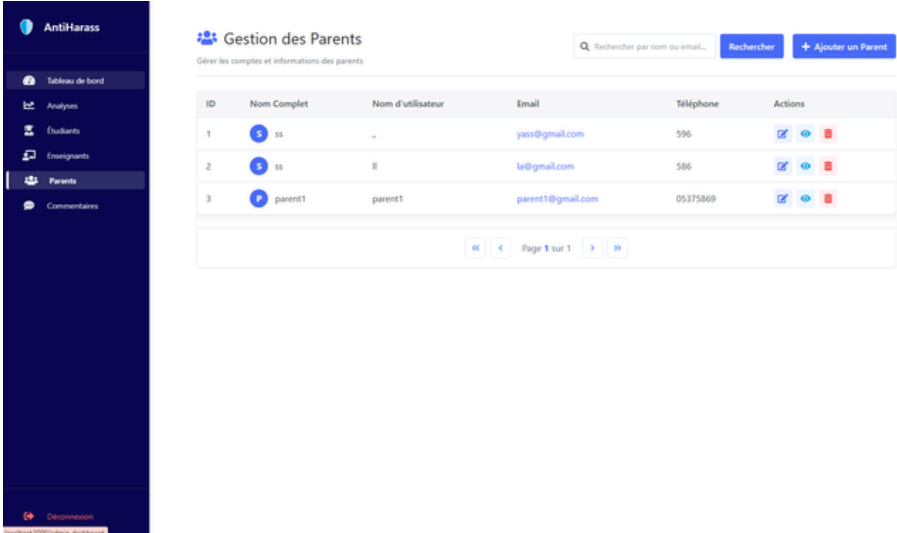


Figure 16: Page de gestion des parents



Figure 17: Page de vue de retours

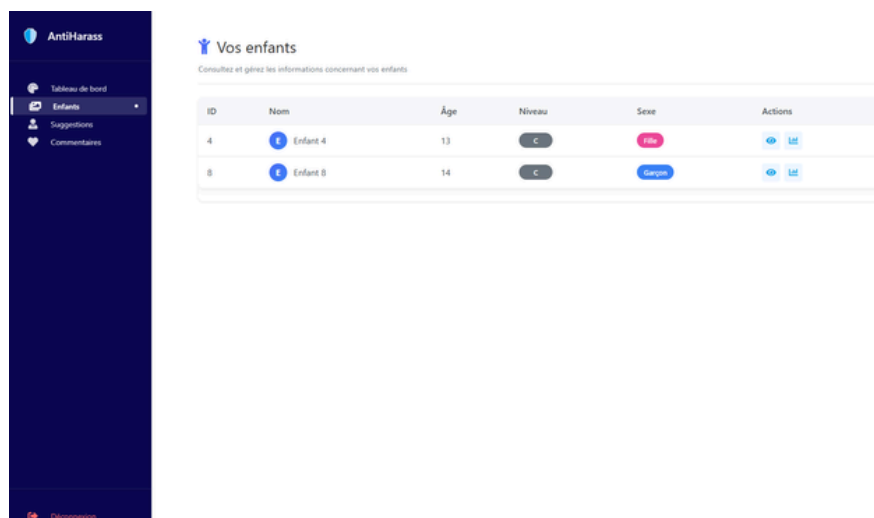


Figure 18: Page d'accueil des parents

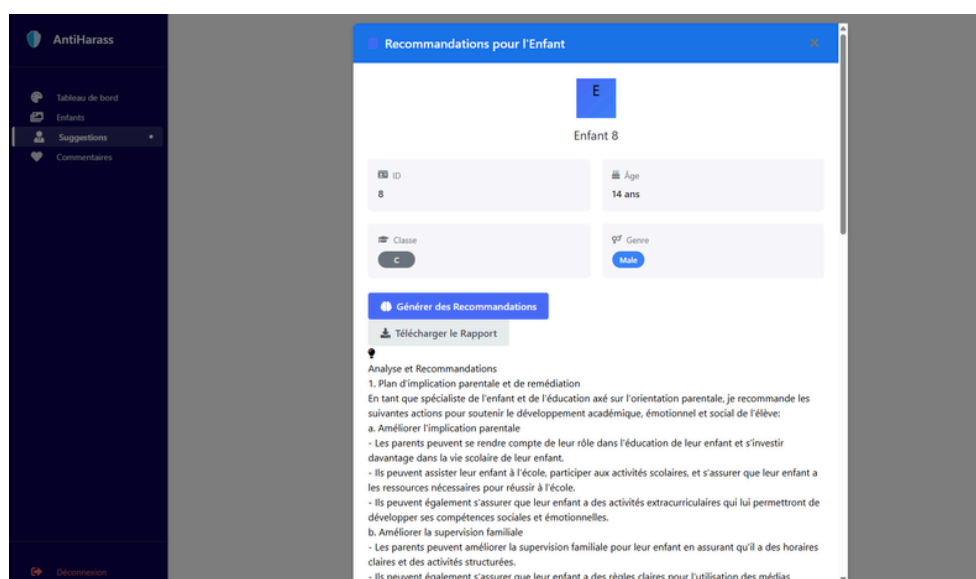


Figure 19: Page de Recommendation de l'AI pour les parents





**Figure 20: Page de Feedback pour les parents**

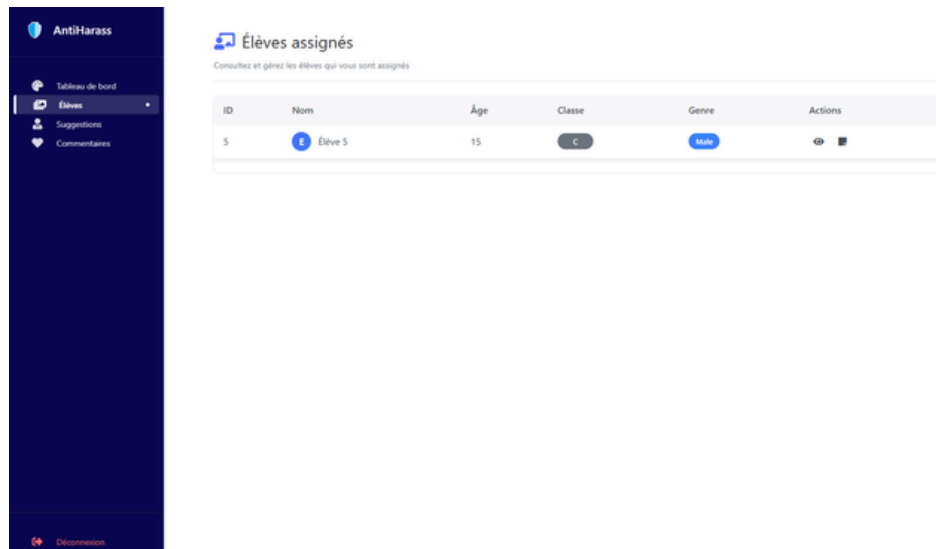


Figure 21: Page d'accueil des enseignants

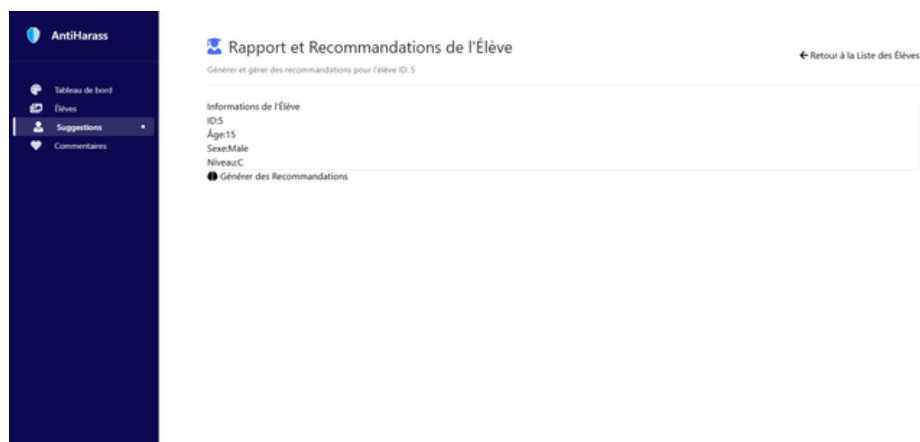


Figure 22: Page de Recommandation de l'AI pour les enseignants

**AntiHarass**

- Tableau de bord
- Élèves
- Suggestions
- Commentaires

### Commentaires des enseignants sur les recommandations

Aidez-nous à améliorer nos suggestions d'IA en apportant votre expertise pédagogique

Sélectionnez une recommandation :

Étudiant 5 - 1. Teacher Involvement and Remediation Plan: The ...

Votre commentaire :

pas mal

Envoyer le commentaire

**Figure 23: Page de Feedback pour les enseignants**

## Table de Matières

|   |           |
|---|-----------|
| <b>Remerciements.....</b>                       |           |
| <b>Sommaire.....</b>                            |           |
| <b>Introduction .....</b>                       | <b>1</b>  |
| o Contexte et Justification .....               | 1         |
| o Objectifs du projet .....                     | 1-2       |
| o Méthodologies adoptée .....                   | 2 -3      |
| <b>Génération et analyse des données.....</b>   | <b>4</b>  |
| o Génération des données .....                  | 4         |
| o Analyse des données(Data mining) .....        | 5         |
| <b>Conception de l'Application .....</b>        | <b>7</b>  |
| o Spécifications Fonctionnelles.....            | 7         |
| o Spécifications Techniques .....               | 9         |
| <b>Développement de l'Application .....</b>     | <b>12</b> |
| o Description de l'Implémentation... ..         | 12        |
| o Défis Rencontrés et Solutions Apportées ..... | 20        |
| <b>Perspectives Futures .....</b>               | <b>21</b> |
| <b>Conclusion Générale .....</b>                | <b>22</b> |
| <b>Annexe .....</b>                             | <b>23</b> |
| o Diagrammes.....                               | 23        |
| o Captures d'Écran.....                         | 28        |
| <b>Table de Matières.....</b>                   | <b>38</b> |



## Résumé

Ce rapport présente le développement d'une plateforme web intelligente destinée à détecter et prévenir les situations de harcèlement scolaire à l'aide de techniques d'intelligence artificielle et de fouille de données. L'application, conçue avec React.js, Node.js et MySQL, offre des interfaces dédiées aux administrateurs, professeurs et parents, favorisant un suivi personnalisé des élèves. Elle permet d'analyser des données simulées pour identifier les élèves à risque grâce à des algorithmes de classification (Random Forest) et de clustering (K-Means). Des visualisations interactives et des recommandations générées par l'IA sont proposées pour accompagner les décisions pédagogiques et psychologiques. Ce rapport décrit l'ensemble du processus de conception, d'implémentation et d'évaluation de la solution, soulignant son potentiel pour améliorer le bien-être scolaire et renforcer la prévention du harcèlement.

## Abstract

This report outlines the development of an intelligent web platform aimed at detecting and preventing school bullying through artificial intelligence and data mining techniques. Built using React.js, Node.js, and MySQL, the application features dedicated interfaces for administrators, teachers, and parents, enabling personalized student monitoring. The system analyzes simulated school data to identify at-risk students using classification (Random Forest) and clustering (K-Means) algorithms. Interactive visualizations and AI-generated recommendations support educational and psychological interventions. This report details the design, development, and evaluation process of the solution, emphasizing its potential to enhance school well-being and support proactive bullying prevention.

---