|  |
| --- |
| Lorenzo BARBEY  Axel FEVEZ |

|  |
| --- |
| ISIMA |
| Intelligence Artificielle |
| PROJET IA |

[I. Consigne : 1](#_Toc153458521)

[II. Thème 1](#_Toc153458522)

[Objectif 1](#_Toc153458523)

[III. Modèle 2](#_Toc153458524)

[Machine Learning (ML) 2](#_Toc153458525)

[Représentation des Connaissances (RC) 2](#_Toc153458526)

[Pourquoi pas Contrainte de Satisfaction de Problèmes (CSP) ? 2](#_Toc153458527)

[IV. Approche 3](#_Toc153458528)

[Machine Learning (ML) 3](#_Toc153458529)

[Représentation des Connaissances (RC) 4](#_Toc153458530)

[V. Ressources 5](#_Toc153458531)

# Consigne :

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

# Thème

**Détection de texte généré par IA**

## Objectif

A partir d’un texte en entrée, retourner un booléen *True* si le texte a été généré par une IA et *False* par un humain.

# Modèle

## Machine Learning (ML)

L'apprentissage automatique est crucial pour entraîner votre modèle à reconnaître les caractéristiques spécifiques des textes générés par l'IA. Vous pourriez utiliser des approches supervisées avec des ensembles de données annotés pour l'entraînement.

## Représentation des Connaissances (RC)

La représentation des connaissances peut être utile pour structurer les informations sur les modèles de langage utilisés par les IA pour générer des textes. Cela peut aider à élaborer des mécanismes de détection basés sur la compréhension des schémas et des caractéristiques spécifiques aux textes générés par IA.

## Pourquoi pas Contrainte de Satisfaction de Problèmes (CSP) ?

Bien que le Contrainte de Satisfaction de Problèmes (CSP) soit une technique puissante pour résoudre des problèmes structurés par des contraintes, son application directe à la détection de textes générés par IA peut être moins adaptée. La détection de ces textes est souvent un problème probabiliste complexe, où les modèles d'apprentissage automatique excèdent en apprenant des modèles à partir de données. Les caractéristiques des textes générés par IA peuvent être difficiles à exprimer en termes de contraintes formelles, et les approches basées sur le CSP peuvent être moins flexibles pour capturer la variété des structures linguistiques. De plus, les besoins d'entraînement continu et d'adaptation rapide aux évolutions des modèles linguistiques rendent le CSP moins adapté à ces exigences. En fin de compte, le choix entre CSP et d'autres approches dépend des spécificités du problème et des objectifs de détection de textes générés par IA.

# Approche

Vous pourriez envisager une approche combinée en utilisant ces branches de l'IA. Par exemple, entraîner un modèle d'apprentissage automatique pour détecter des anomalies ou des caractéristiques spécifiques aux textes générés par IA.

L'efficacité de votre modèle dépendra largement de la qualité de vos données d'entraînement, de la représentativité de ces données par rapport aux textes que vous souhaitez détecter, et de la sophistication de votre approche d'analyse des caractéristiques.

## Machine Learning (ML)

Le Machine Learning (ML) joue un rôle clé dans le développement d'une IA visant à détecter les textes générés par une autre IA. Voici comment le ML peut être intégré dans ce processus :

### Collecte des données

Pour entraîner un modèle de détection, vous aurez besoin d'un ensemble de données qui comprend à la fois des exemples de textes générés par IA et des textes non générés par IA. Ces données doivent être représentatives de ce que le modèle rencontrera dans le monde réel.

### Prétraitement des données

Le prétraitement des données est crucial pour rendre les données adaptées à l'entraînement du modèle. Cela peut inclure la tokenisation, le nettoyage des données, la suppression des bruits indésirables, etc.

### Choix de l'algorithme

Sélectionnez un algorithme d'apprentissage automatique approprié en fonction de la nature de votre problème. Les algorithmes de classification, tels que les machines à vecteurs de support (SVM), les arbres de décision, ou les réseaux neuronaux, peuvent être utilisés pour la détection de texte.

### Entraînement du modèle

Utilisez votre ensemble de données annoté pour entraîner le modèle. Le modèle apprendra à partir des exemples fournis, en ajustant ses paramètres pour mieux discriminer entre les textes générés par IA et les autres.

### Validation et ajustement

Évaluez la performance de votre modèle sur un ensemble de données de validation distinct. Si nécessaire, ajustez les hyperparamètres ou choisissez un modèle différent pour améliorer les performances.

### Tests en conditions réelles

Testez le modèle sur des données réelles pour évaluer sa performance dans des situations du monde réel. Assurez-vous que le modèle généralise bien et peut traiter des textes générés par IA provenant de différentes sources et contextes.

## Représentation des Connaissances (RC)

La représentation des connaissances (RC) peut jouer un rôle important dans le développement d'une IA visant à détecter les textes générés par une autre IA. Voici quelques façons dont la RC peut être intégrée dans ce processus :

### Modélisation des Caractéristiques des Textes Générés par IA

En utilisant la RC, vous pouvez modéliser les caractéristiques spécifiques des textes générés par IA. Cela peut inclure la structure syntaxique, la sémantique, les schémas récurrents ou tout autre aspect particulier aux textes générés par des modèles linguistiques automatisés.

### Construction de Graphes de Connaissances

Vous pouvez construire des graphes de connaissances qui représentent les relations sémantiques entre les entités dans les textes. Cela permet à l'IA de comprendre les liens conceptuels entre les mots, phrases ou idées, facilitant ainsi la détection de schémas spécifiques associés aux textes générés par IA.

### Développement de Modèles Hybrides

Combiner des approches basées sur la RC avec des techniques d'apprentissage automatique peut créer des modèles hybrides puissants. Par exemple, un modèle de détection basé sur la RC peut être combiné avec des algorithmes de ML pour renforcer la capacité de généralisation et d'adaptation aux variations.

### Utilisation de Logiques de Description

Les logiques de description peuvent être utilisées pour représenter formellement les connaissances sur les structures et les propriétés des textes générés par IA. Ces représentations formelles facilitent la manipulation et la comparaison des connaissances, renforçant ainsi la capacité de détection.

### Explication des Décisions du Modèle

La RC peut également être utilisée pour rendre les décisions du modèle plus compréhensibles. Cela peut être crucial dans des domaines tels que la détection de textes générés par IA, où la transparence du processus décisionnel est importante.

# Ressources

* Consigne
* Chat GPT