UNIVERSITÉ DE NANTES

IAE NANTES
ÉCONOMIE & MANAGEMENT

Master Econométrie et Statistiques, parcours Econométrie Appliquée

Projet réalisé dans le cadre du cours de NLP

Détection des textes générés par l'IA : une approche d'analyse textuelle par les modèles de NLP

DUPAU Jasmine

LABRE-BLANC Emma

TRILLAUD Valorys

VILAR VALERA Yava

Encadrées par Monsieur BENNANI Hamza

Année universitaire : 2024-2025

Sommaire

Introduction

Avec l'essor rapide des modèles génératifs comme ChatGPT ou Gemini, la production des textes à été profondément transformée. Leur usage désormais largement banalisé, rend la création automatique de contenus textuels à la fois accessible, rapide et crédible.

La distinction entre les textes rédigés par des humains et ceux générés par des intelligences artificielles devient un enjeu majeur dans de nombreux secteurs comme l'éducation, la communication en entreprise ou la publication scientifique. Une étude rétrospective sur les publications dans la revue *Orthopaedics & Traumatology: Surgery & Research* (OTSR) révèle une augmentation significative de l'utilisation de l'IA dans la rédaction des articles après la sortie publique de ChatGPT (Bisi et al., 2023)[1]. Cette évolution soulève ainsi des préoccupations : comment garantir la fiabilité, la transparence, ou l'originalité d'un contenu lorsque sa source devient incertaine ?

Le sujet de la détection des textes générés par IA est encore jeune, mais il s'inscrit dans une littérature émergente qui interroge les capacités réelles de génération linguistique des modèles et leurs limites en termes de sémantique, de cohérence et de variation syntaxique. Des travaux récents ont ainsi montré que le style d'un texte peut constituer un indicateur pertinent pour identifier une rédaction par IA, notamment à travers des approches stylométriques (Roten et al, 2023)[2].

Dans ce contexte, notre étude porte sur l'analyse textuelle comme levier de détection des textes générés par l'IA, en mobilisant des modèles de traitement automatique du langage naturel (NLP). Le traitement automatique du langage naturel (ou NLP pour Natural Language Processing) regroupe un ensemble de techniques permettant aux machines de comprendre, analyser et générer du langage humain. Dans notre cas, ces outils sont utilisés non pas pour créer du texte, mais pour en analyser la structure linguistique, lexicale et syntaxique afin de repérer des signatures typiques de l'IA.

Ce sujet s'inscrit dans un contexte technologique et social mouvant, marqué par une adoption rapide des outils d'IA, mais aussi par une inquiétude croissante sur leur mésusage : tricherie scolaire, désinformation, perte d'authenticité dans la communication.

D'un point de vue managérial, la capacité à détecter automatiquement des textes générés par IA devient stratégique, notamment pour les RH (recrutement, candidatures), les médias (articles,

[1] https://doi.org/10.1016/j.rcot.2023.09.014
[2] https://doi.org/10.56240/irafpa.cm.v1n1/rot

tribunes) ou les universités. Théoriquement, cette recherche interroge les limites entre langage humain et langage artificiel, tout en explorant les indicateurs linguistiques différenciateurs.

Nos motivations sont doubles : mieux comprendre les spécificités linguistiques des textes IA par rapport aux textes rédigés par des humains d'une part, et évaluer dans quelle mesure les outils actuels de NLP permettent d'automatiser cette détection d'autre part.
Notre problématique est donc la suivante : Peut-on, à travers une analyse linguistique fine basée sur des modèles NLP, distinguer de manière fiable les textes générés par une IA de ceux rédigés par des humains ?

Pour y répondre, nous avons travaillé à partir d'un jeu de données public (Kaggle) déjà annoté, que nous avons préalablement nettoyé.
Nous avons ensuite mené une série d'analyses exploratoires visant à comparer les textes sur différents plans : ponctuation, structure syntaxique, diversité lexicale, rareté des mots, usage des n-grammes, etc.
Enfin, nous avons mis en œuvre des modèles de machine learning pour tester différentes approches de classification automatique.

Notre objectif est à la fois analytique – identifier des motifs récurrents – et applicatif – construire un modèle reproductible de détection.

Ce dossier est structuré en cinq grandes parties. Nous commencerons par présenter le jeu de données utilisé ainsi que les étapes de prétraitement nécessaires à son exploitation (**I**). Ensuite, nous exposerons les méthodes et métriques d'analyse linguistique mobilisées pour répondre à notre problématique (**II**). La troisième partie sera consacrée à une analyse exploratoire des données, permettant d'identifier des tendances initiales (**III**). Finalement nous présenterons les résultats des modélisations de machine learning visant à automatiser la détection des textes générés par IA (**IV**), avant de discuter les résultats obtenus (**V**).

I. Présentation et nettoyage de la base

Dans cette section, nous présenterons les données utilisées dans le cadre de cette étude (**A**) ainsi que les étapes de nettoyage et de prétraitement appliquées afin de les rendre prêtes à la modélisation (**B**).

A. Présentation de la base de données

Notre base de données initiale, provenant du site Kaggle, est composée de 487 235 lignes et 2 variables. Comme indiqué dans le **tableau 1**, la première colonne, intitulée text, contient des extraits d'essais issus de sources variées, tandis que la seconde colonne, generated, est une variable binaire indiquant si le texte a été rédigé par une Intelligence Artificielle (1) ou par un être humain (0). Étant donné l'objectif de notre étude, qui consiste à classifier des textes en fonction de leur origine (IA ou humain), il était essentiel de s'assurer que chaque texte du jeu de données comporte au moins deux mots. Ainsi, nous avons éliminé les entrées ne répondant pas à ce critère. De plus, nous avons pris soin de vérifier qu'il n'y avait pas de valeurs manquantes ni de doublons dans les données, garantissant ainsi leur qualité pour l'analyse. En raison de la volumétrie importante de la base de données, nous avons décidé de réduire sa taille à 10 000 lignes afin d'optimiser le temps d'exécution du code. Pour garantir un échantillon équilibré, nous avons extrait 5 000 observations aléatoires de chaque classe, tout en nous assurant qu'aucun doublon n'était présent dans cette version réduite de la base.

***Tableau 1**: Description de la base de données*

Variable	Valeurs
text	extrait d'essai écrit en anglais
generated	1 = généré par l'IA; 0 = généré par l'humain
Source de données = Kaggle *Nombre d'observations* = 10 000	

B. Pré-traitement des données

Afin d'exploiter efficacement nos données textuelles dans le cadre de modélisations thématiques (comme l'analyse LDA), de prédictions de sentiment et d'autres modèles de machine learning, nous avons créé une nouvelle colonne, texte_propre, qui résulte du prétraitement du texte brut. Cette étape est cruciale en NLP (Natural Language Processing), car elle permet de transformer le texte en une forme exploitable par les algorithmes de machine learning. Le prétraitement commence par la tokenisation, qui consiste à découper le texte en unités de base, appelées tokens, généralement des mots ou des signes de ponctuation. Cette opération facilite l'analyse en

traitant le texte sous forme de séquences distinctes. Nous procédons ensuite à la suppression des stopwords, c'est-à-dire des mots très fréquents mais peu informatifs tels que "the", "is", "and", qui génèrent du bruit dans les données. Leur élimination permet de se concentrer sur les termes plus significatifs, contribuant ainsi à l'amélioration de la précision des modèles. Enfin, nous appliquons la lemmatisation, une opération qui réduit les mots à leur forme de base, ou lemme (par exemple, "running", "ran" et "runs" deviennent "run"). Cela uniformise le vocabulaire et diminue les variations lexicales, améliorant ainsi la qualité des analyses.

II. Méthodologie

Dans cette section, nous détaillons à présent la méthodologie mise en œuvre pour la classification automatique des textes. Nous commencerons par présenter le processus de vectorisation (**A**), qui permet de convertir les données textuelles en représentations numériques, avant de décrire les différents modèles de classification que nous avons testés et les métriques de comparaison utilisées (**B**).

A. Vectorisation

L'objectif de notre travail est de déterminer si un texte a été rédigé par un humain ou généré par une intelligence artificielle. Pour ce faire, nous avons recours à des modèles de classification, qui sont couramment utilisés dans ce type de problématique. Toutefois, ces algorithmes ne peuvent pas traiter directement des données textuelles brutes. Il est donc nécessaire de transformer les textes en représentations numériques à l'aide d'une méthode de vectorisation. Nous avons retenu la technique TF-IDF (Term Frequency – Inverse Document Frequency), qui permet d'évaluer l'importance relative des mots dans un document par rapport à l'ensemble du corpus. Concrètement, cette méthode attribue un poids à chaque mot en fonction de sa fréquence dans le texte (TF) et de sa rareté dans l'ensemble des documents (IDF). Ainsi, un terme fréquent dans un document mais rare dans le corpus recevra un poids élevé, tandis qu'un mot courant dans de nombreux textes sera faiblement pondéré. Cette représentation permet aux modèles d'exploiter efficacement les spécificités lexicales de chaque texte pour améliorer leur capacité de classification.

B. Modèles de classification

a) Présentation des modèles

Dans ce projet, nous avons testé cinq modèles de classification parmi les plus couramment utilisés en apprentissage automatique : la régression logistique, le SVC linéaire, la forêt aléatoire,

le modèle de Bayes naïf multinomial, et le SVC avec noyau non linéaire. Voici une brève description de chacun :

- **Régression logistique**: il s'agit d'un modèle linéaire simple mais performant, souvent utilisé en NLP pour des tâches de classification binaire. Il estime la probabilité d'appartenance à une classe à l'aide de la fonction sigmoïde. Son efficacité repose sur l'hypothèse que les classes sont séparables dans l'espace vectoriel TF-IDF.

- **SVC linéaire**: ce modèle est une version du SVM adaptée aux grandes dimensions, comme celles générées par la vectorisation de textes. Il vise à trouver l'hyperplan qui maximise la séparation entre les classes. Grâce à sa capacité à gérer des espaces de grande dimension, il est bien adapté aux données textuelles.

- **Forêt aléatoire**: modèle d'ensemble basé sur l'agrégation de plusieurs arbres de décision. Chaque arbre est entraîné sur un sous-échantillon du corpus vectorisé. Cette approche permet de capturer des relations non linéaires et d'augmenter la robustesse des prédictions, bien qu'elle soit parfois moins performante que les modèles linéaires sur des données comme les TF-IDF.

- **Modèle multinomial naïf de Bayes**: particulièrement adapté à la classification de textes, ce modèle repose sur le théorème de Bayes et sur l'hypothèse d'indépendance conditionnelle des mots. Il est rapide, efficace et souvent très compétitif dans les tâches de NLP malgré sa simplicité.

- **Modèle SVC (Support Vector Classifier) standard**: partage les bases du SVC linéaire, mais diffère par sa capacité à appliquer des noyaux non linéaires pour projeter les textes vectorisés dans un espace de dimension supérieure, ce qui permet de modéliser des relations complexes entre les documents. Toutefois, ce type de SVC peut se heurter à des difficultés d'optimisation sur des données de grande dimension.

b) Métriques utilisées

Afin de mesurer la performance des modèles de classification, nous avons d'abord divisé notre jeu de données en deux sous-ensembles : un jeu d'entraînement représentant 70 % des observations et un jeu de test couvrant les 30 % restants. Cette répartition vise à entraîner les modèles sur un large échantillon, tout en évaluant leur capacité à généraliser sur des données inédites. L'objectif est d'obtenir un modèle fiable, capable de distinguer efficacement si un texte a

été rédigé par un humain ou généré par une intelligence artificielle, et ce dans des contextes variés.

Pour évaluer les performances, nous nous appuyons principalement sur deux métriques : la précision moyenne (accuracy) et le F1-score moyen :

- **La précision** mesure la proportion de textes correctement classés (qu'ils soient humains ou générés par l'IA). Elle est pertinente ici car notre variable cible "généré_par_IA" est équilibrée, avec autant de textes dans chaque classe.

- **Le F1-score**, qui combine la précision (évite les faux positifs) et le rappel (évite les faux négatifs), est particulièrement utile pour évaluer la qualité de la classification dans chaque classe. Il permet également de détecter un éventuel surajustement en comparant les scores obtenus sur les jeux d'entraînement et de test.

Ainsi, un bon modèle présente une accuracy et un F1-score élevés et proches sur les deux jeux de données, ce qui traduit à la fois une bonne capacité de prédiction et une généralisation satisfaisante.

III. Analyse exploratoire

Avant d'aborder la modélisation, il nous a paru essentiel de réaliser une analyse exploratoire du contenu textuel afin d'identifier les différences éventuelles entre les textes générés par l'intelligence artificielle et ceux rédigés par des humains. Cette exploration repose sur deux axes complémentaires : une analyse syntaxique, principalement centrée sur le texte brut, et une analyse thématique, basée sur les données prétraitées.

Dans un premier temps, nous nous concentrons sur les caractéristiques linguistiques superficielles, telles que la ponctuation, la longueur des textes..., afin d'identifier des schémas d'écriture distinctifs. Ces éléments peuvent fournir des indices sur les styles d'écriture spécifiques à chaque type de rédacteur (**A**). Dans un second temps, nous analysons les données après prétraitement, comme détaillé en section **I.B**, pour approfondir l'examen à travers une analyse thématique et de sentiment, en utilisant des techniques telles que la visualisation des mots les plus fréquents ou la modélisation de sujets (**B**).

A. Analyse syntaxique superficielle

Nous pouvons supposer que l'intelligence artificielle ne possède pas exactement le style d'écriture d'un humain. En effet, les textes générés par l'IA sont produits en réponse à des prompts fournis par les utilisateurs via des interfaces comme ChatGPT, Mistral ou Copilot, et non issus d'une intention discursive spontanée. Cela pourrait expliquer pourquoi les productions de l'IA apparaissent souvent comme mieux structurées ou plus formelles. Toutefois, les récents progrès des modèles de langage rendent de plus en plus difficile la distinction entre textes humains et générés, tant leur fluidité s'améliore.

L'objectif de cette analyse linguistique est donc de confronter ces intuitions à des observations empiriques. Pour cela, nous nous appuyons sur différents indicateurs liés à la ponctuation, à la lisibilité, à la structure grammaticale et à la cohérence des textes (**Tableau 2**). Les résultats principaux sont présentés ci-dessous, tandis que les statistiques détaillées et les définitions des concepts mobilisés figurent **en annexes 1 et 2**.

***Tableau 2** : Résultats de l'analyse linguistique sur les textes bruts*

Concept	Principaux résultats
Ponctuation et taille moyenne des phrases (sur textes bruts)	Les textes humains sont plus longs, avec une ponctuation plus variée et une structure syntaxique plus complexe. Ils comportent notamment davantage de points, de points-virgules et de points d'interrogation, traduisant une tendance naturelle des humains à construire des phrases plus riches, à poser des questions rhétoriques ou à structurer leur discours avec plus de subtilité. En revanche, les textes générés par l'IA se distinguent par une utilisation moyenne plus élevée de virgules, de deux-points et de tirets (dash). Cette prédominance peut s'expliquer par la manière dont les modèles de langage construisent leurs sorties, ils ont tendance à privilégier un enchaînement fluide et linéaire des idées. Ce type de ponctuation, plus "fonctionnelle" que stylistique, permet de produire des phrases claires, cohérentes et logiquement structurées, ce qui correspond bien aux objectifs d'un modèle entraîné pour générer du texte lisible, sans nécessairement reproduire la diversité stylistique humaine. Cela reflète une forme de standardisation linguistique propre aux textes produits par IA, qui cherchent avant tout à éviter l'ambiguïté là où le style humain repose davantage sur la nuance, la spontanéité et l'interaction.
Segmentation en paragraphes (sur textes bruts)	En moyenne, les textes générés par l'IA ont plus de paragraphes par rapport à ceux écrits par des humains. Cela suggère que l'IA a tendance à segmenter davantage le contenu, possiblement pour améliorer la lisibilité. Les humains, quant à eux, utilisent des structures légèrement plus compactes, ce qui peut être synonyme d'une certaine spontanéité dans l'écriture.

Grammaire (sur textes bruts)	Les deux types de textes présentent un nombre similaire de fautes selon l'outil. Il y a en moyenne légèrement moins de fautes pour l'IA et une légère tendance à plus de variabilité chez les humains avec un écart-type et un maximum plus élevés.
Profondeur (sur textes bruts)	Les phrases IA et humaines ont une complexité moyenne similaire.
Lisibilité (sur textes bruts)	En moyenne, les textes humains sont plus faciles à lire que ceux générés par l'IA, avec un score de lisibilité plus élevé, reflétant une syntaxe plus simple et un vocabulaire plus accessible. Ils présentent toutefois parfois des scores très faibles, traduisant une écriture plus spontanée ou moins structurée. Les textes produits par l'IA, quant à eux, sont plus réguliers et formellement rédigés, témoignant d'une production plus standardisée et cohérente.
Part Of Speech (sur textes bruts)	Les deux types de textes utilisent une quantité similaire de noms, ce qui montre que l'IA reproduit bien la structure nominale du langage humain. Les textes humains comportent davantage de verbes et d'adverbes, traduisant un style plus dynamique et nuancé, là où l'IA adopte un ton plus factuel. En revanche, l'IA utilise légèrement plus d'adjectifs, ce qui peut refléter une tendance à surqualifier ou embellir les descriptions, typique des textes générés automatiquement.
Cohérence (sur textes bruts)	Selon ces résultats, les textes humains ont une cohérence lexicale locale légèrement supérieure en moyenne. Cela peut refléter une continuité plus naturelle dans le choix des mots consécutifs. Toutefois, l'écart n'est pas énorme, ce qui souligne que les modèles d'IA parviennent aussi à générer des enchaînements de mots localement cohérents.
Rareté (sur textes pré-traités)	Les textes générés par l'IA contiennent en moyenne un peu plus de mots rares que les textes humains (17.55 contre 17.00). La différence est modeste, mais elle suggère que les textes IA utilisent peut-être un vocabulaire légèrement plus "technique" ou "ciblé".
Diversité (sur textes pré-traités)	Les textes IA ont une diversité lexicale légèrement plus élevée (0.598 contre 0.568). Cela peut vouloir dire que l'IA utilise un vocabulaire un peu plus varié que les humains.

En somme, l'ensemble de ces résultats met en évidence des différences stylistiques notables entre les textes rédigés par des humains et ceux générés par l'IA, et ce sur plusieurs dimensions clés : complexité syntaxique, structure lexicale, organisation du discours, etc. L'analyse de régression révèle que toutes ces différences sont statistiquement significatives[3]. De plus, malgré

[3] Nous avons modélisé le nombre de mots, ponctuations (points, virgules, tirets, points d'exclamation et d'interrogation), adverbes, verbes, adjectifs, d'erreurs d'orthographe, la longueur des phrases et du texte en fonction de la variable binaire "generated". Les résultats ont démontré un impact significatif de cette variable dans tous les cas, indiquant des différences significatives entre les groupes.

l'impossibilité d'inclure tout cet ensemble de variables explicatives à cause de la multicolinéarité, nous avons trouvé à l'aide d'une régression logistique que le nombre de comas, de colonnes, de tirets, de points d'exclamation et de paragraphes sont statistiquement supérieurs chez l'IA, tandis que le nombre de semi-colonnes, de questions et d'adverbes sont supérieurs chez l'humain. Seule la longueur de phrases s'est avérée ne pas impacter significativement le fait que le texte soit rédigé par l'humain ou l'IA. Les résultats détaillés sont disponibles en **annexe 3.**

Certaines observations, comme l'usage plus fréquent d'adjectifs ou de termes techniques par l'IA, trouvent un écho dans des analyses externes, notamment celle du média *Startups Nation* (2023)[4], qui souligne la tendance de ChatGPT à mobiliser un vocabulaire sophistiqué, souvent emprunté à des registres académiques ou professionnels. Cette préférence s'explique par la nature des corpus d'entraînement (articles scientifiques, documents spécialisés), qui influencent fortement le style lexical des modèles. Bien que cela puisse renforcer la précision des réponses sur des sujets complexes, cela tend parfois à alourdir le discours ou à le rendre moins accessible pour un lectorat non expert.

Plus globalement, ces résultats confirment ce que la littérature scientifique récente avance : l'IA ne se contente pas d'imiter la langue humaine, elle la filtre, la structure et la "normalise" à travers des biais statistiques. Comme le soulignent Goar et al. (2023)[5], les modèles génératifs comme ChatGPT reproduisent en priorité les formulations les plus fréquentes de leurs données d'entraînement, ce qui peut nuire à la créativité et à l'originalité du texte.

Ainsi, si la frontière entre style humain et style généré devient de plus en plus subtile en raison des progrès technologiques, certaines caractéristiques distinctives subsistent.

B. Analyse thématique et de sentiment sur les données prétraitées

Dans cette section, nous approfondissons l'analyse des textes après leur prétraitement, en adoptant une approche thématique et sentimentale. Cela permet de dégager des tendances globales et de mieux comprendre les nuances qui caractérisent les textes générés par l'intelligence artificielle par rapport à ceux rédigés par des humains. L'analyse se divise en trois axes complémentaires : analyse de l'occurrence des mots, modélisation de sujets (LDA), analyse de sentiment.

[4] Les mots les plus utilisés par ChatGPT : Analyse du langage du chatbot - Startups Nation – Média des entrepreneurs et Start-ups en France

[5] Goar, V., Yadav, N. S., & Yadav, P. S. (2023). Conversational AI for natural language processing: An review of ChatGPT. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 109-17.

a) Occurrence des mots

Nous commençons par l'analyse de la fréquence des mots. Cette dernière est présentée sous formes de nuages de mots et d'histogrammes qui ont permis de ressortir les mots typiques des deux catégories (humain et et IA).

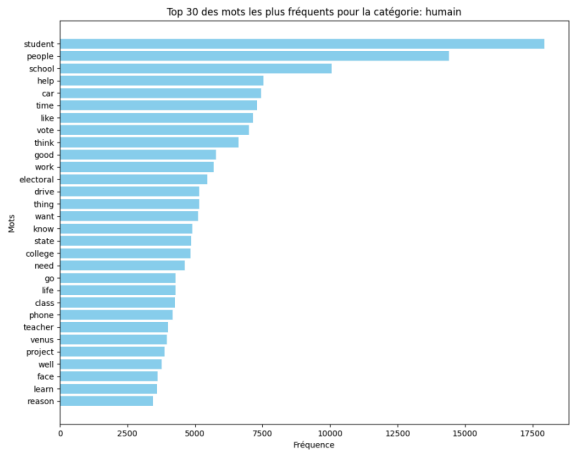
***Figure 1** : Nuages de mots pour les deux classes*

L'observation des nuages de mots (**Figure 1**) met en évidence une première différence notable entre les deux catégories : les textes rédigés par des humains contiennent plusieurs mots très fréquents, souvent répétés (*"people"*, *"think"*, etc.), tandis que les textes générés par l'IA présentent un lexique plus diversifié, avec une fréquence plus uniformément répartie. Une exception notable concerne *"electoral college"*, qui apparaît fréquemment dans les deux corpus. Ces éléments suggèrent une tendance des humains à se répéter davantage, là où l'IA mobilise un vocabulaire plus étendu.

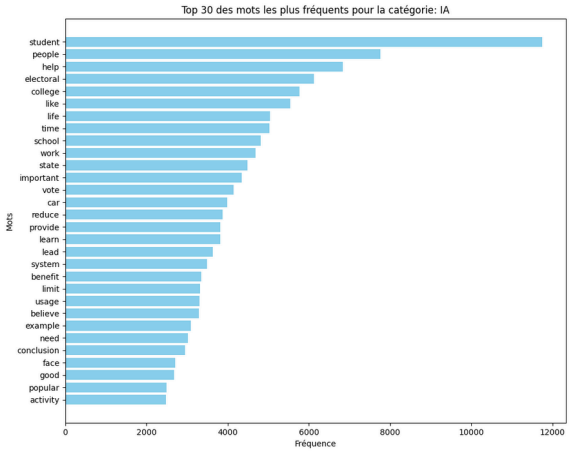
Les histogrammes de fréquence (**Figure 2**) affinent cette analyse : certains mots atteignent jusqu'à 17 500 occurrences dans les textes humains (*"student"*), contre un maximum de 12 000 dans ceux de l'IA. On distingue également des termes caractéristiques de chaque groupe. Les textes humains utilisent plus souvent *"think"*, *"go"*, *"want"*, *"know"* — des verbes d'action ou de cognition, souvent associés à l'expression d'opinions, de désirs ou d'expériences personnelles. Cela reflète un style plus subjectif, introspectif et spontané, typique de l'écriture humaine. À l'inverse, les textes IA privilégient des mots comme *"provide"*, *"example"*, *"important"*, *"reduce"*, révélateurs d'un discours explicatif, structuré et orienté vers l'exposé clair d'idées.

Ces constats suggèrent que certains mots très fréquents et partagés (*"student"*, *"people"*, *"school"*) auront peu d'intérêt pour la modélisation, tandis que les termes plus spécifiques à l'un ou l'autre groupe pourraient se révéler discriminants dans une tâche de classification.

***Figure 2** : Histogrammes de la fréquence des mots selon la catégorie*



Classe : Humain



Classe : IA

b) La LDA

Nous passons à présent à l'identification des thématiques présentes dans les textes. L'objectif est de faire émerger les grands sujets abordés dans chaque catégorie (textes humains vs générés). Pour cela, nous avons eu recours à la méthode Latent Dirichlet Allocation (LDA), un algorithme de modélisation de sujets permettant de détecter automatiquement des groupes de mots co-occurrents reflétant des thématiques latentes définies au nombre de 3 dans notre étude. Pour obtenir des résultats interprétables, nous avons au préalable imposé des règles sur la conversion du texte en numérique avec la fonction *CountVectorizer* telles que : L'exclusion des mots qui apparaissent dans plus de 90% des paragraphes, la suppression des mots contenus dans moins de 2 paragraphes et la non prise en compte des stopwords du dictionnaire anglais. Les résultats sont présentés ci-dessous (**Figure 3**) :

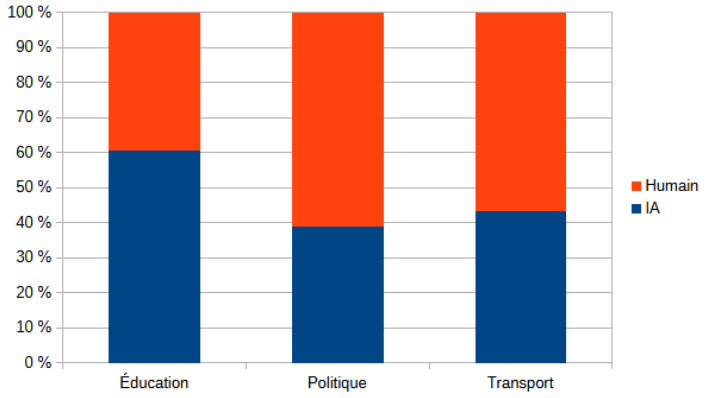
***Figure 3** : Résultats de la LDA*

TOPIC 0	TOPIC 1	TOPIC 2
'technology', 'know', 'good', 'think', 'teacher', 'project', 'work', 'venus', 'learn', 'face', 'life', 'like', 'help', 'people', 'student'	'work', 'popular', 'election', 'candidate', 'activity', 'president', 'people', 'class', 'time', 'state', 'college', 'vote', 'electoral', 'school', 'student'	'road', 'phone', 'world', 'driver', 'pollution', 'transportation', 'city', 'driverless', 'help', 'limit', 'reduce', 'usage', 'drive', 'people', 'car'

A l'observation du top 15 des mots dans chaque topic sur la **Figure 3**, nous avons pu déterminer trois thématiques. La première concerne l'apprentissage, l'éducation et la technologie. En effet, les mots liés à l'école, l'enseignement et la technologie ont une forte présence, accompagnés d'autres termes plus généraux comme “life”, “help” ou encore “people”. Cela suggère que ce sujet est assez large sur l’apprentissage personnel ou scolaire dans un cadre numérique ou scientifique. Le second sujet évoque quant à lui, la politique et le système électoral avec des termes spécifiques à ces thèmes comme “election”, “candidate”, “vote”, “president” et “électoral”. Enfin, le dernier top 15 renvoie à la dimension aux transports dont les problématiques modernes comme la pollution, la mobilité urbaine, les voitures autonomes et la durabilité sont évoquées par “pollution”, “driverless”, “usage” et “reduce”.

Nous avons également examiné la répartition des textes IA et humains selon les thématiques identifiées (**Figure 4**). Il en ressort que la thématique de l’éducation est davantage abordée par les textes générés par l’IA (environ 61 %), contre seulement 39 % pour les textes humains. À l’inverse, la thématique politique est plus fréquemment traitée dans les productions humaines (61 % contre 39 % pour l’IA). Enfin, les textes portant sur les transports présentent une répartition relativement équilibrée entre les deux catégories.

***Figure 4** : Distribution des classes selon les thématiques*



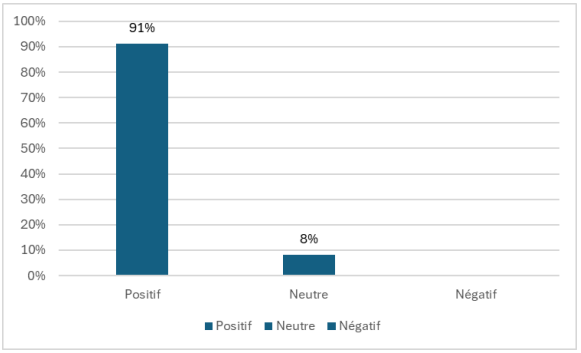
c) Analyse des sentiments

Pour clôturer l'analyse exploratoire, nous avons réalisé l'analyse des sentiments avec la méthode VADER (ou *Valence Aware Dictionary for Sentiment Reasoning*) qui consiste à donner un score de polarité à chacun des textes : score supérieur à 0,05 = texte “positif”, score nul = texte “neutre” et score inférieur à -0,05 = texte “négatif”. L’analyse a été décomposée en deux parties : la première offre une vision globale avec des exemples de paragraphes et la distribution des textes par rapport aux sentiments (**Figure 5**) tandis que la deuxième propose une analyse plus affinée en

déterminant d'une part les sentiments dégagés par les trois thématiques vues précédemment et, d'autre part, les sentiments majoritaires pour chaque type de texte (IA et humain) (**Figure 6**).

Figure 5 : *Exemples de textes et distribution des sentiments sur l'ensemble du jeu de données*

Sentiment	Exemple
Positif	Curfews would also keep teenagers from Getting into trouble.
Neutre	As of today there are To completely driverless cars out. There are only cars that can assist with driving. It will be awhile before driverless cars are out on the roads.
Négatif	His dreams were crashed, shattered, by the new limit on GPA for sports.

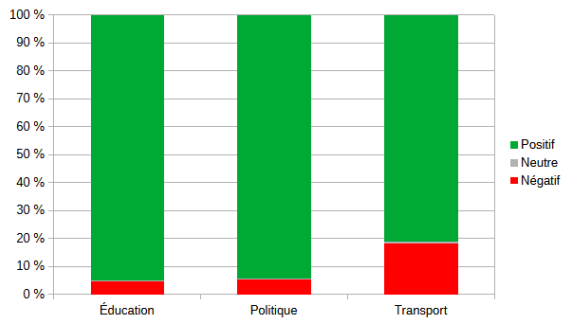


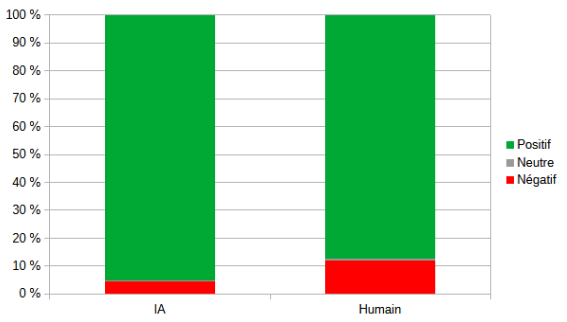
Exemples de paragraphes pour chaque sentiment	Distribution des sentiments

Tout d'abord, nous pouvons observer avec la distribution de la **Figure 5** que le jeu de données est principalement constitué de textes qui dégagent un sentiment positif (91%), suivis de très loin par ceux provoquant un ressenti négatif (8%) et le sentiment neutre est très peu représenté voire inexistant en termes de proportions (0%). Cela pourrait être dû aux bornes choisies pour classifier les scores de polarité, qui ne laissent que peu d'espace pour les cas ambigus. Il est également possible que la nature des sujets abordés, ou encore le style rédactionnel dominant, contribue à cette distribution fortement déséquilibrée.

Figure 6 *: Distribution des sentiments selon les thématiques et les classes*

Thématiques	Classes





*Les statistiques indiquées se trouvent dans plusieurs tableaux présentés dans l'**annexe 4**.*

Concernant les thématiques, nous observons sur la **figure 6** que celles liées à l'éducation (topic 0) et à la politique (topic 1) présentent une très forte prédominance de sentiments positifs, avec respectivement 95% et 94% de textes à tonalité positive. Cela suggère que ces thèmes sont majoritairement abordés de manière valorisante ou neutre, en particulier pour l'éducation, souvent traitée sous l'angle de l'apprentissage, de la réussite ou de l'innovation. La politique, bien qu'étant un sujet potentiellement conflictuel, est ici décrite de manière étonnamment modérée, probablement sous une forme descriptive ou institutionnelle. Cependant, la thématique traitant du transport et de l'environnement (topic 2) se distingue nettement par une tonalité plus contrastée : si les textes restent majoritairement positifs (81%), la proportion de sentiments négatifs y est plus élevée (18% contre 5% pour les autres), ce qui en fait le topic le plus polarisé émotionnellement. Cette différence peut s'expliquer par la nature du sujet, qui englobe des problématiques critiques comme la pollution, les limitations de circulation ou les inquiétudes liées au changement climatique, souvent source de préoccupations.

Enfin, nous avons clôturé l'analyse des sentiments en les confrontant aux catégories textes générés par l'humain et textes générés par l'IA (**Figure 6**).
À la lecture du tableau, nous pouvons observer que les textes générés par l'IA présentent une proportion plus faible de sentiments négatifs (5 % contre 12 % pour les textes humains) et une proportion plus élevée de sentiments positifs (95 % contre 87 %). Ainsi, les productions de l'IA apparaissent globalement plus positives et moins polarisées que celles des humains. Cette différence peut s'expliquer par la nature même des modèles génératifs comme ChatGPT, qui sont entraînés sur de vastes corpus visant souvent à produire un langage poli, consensuel et non offensant (notamment pour éviter les propos sensibles ou conflictuels). Cela les conduit à générer des textes au ton plus modéré, voire positivement biaisé. À l'inverse, les textes humains, souvent ancrés dans des expériences ou opinions personnelles, laissent davantage place à l'expression de frustrations, critiques ou émotions négatives, ce qui explique leur tonalité plus contrastée.

IV. Présentation des résultats

Dans cette section, nous présentons les résultats issus de nos différentes modélisations. Nous débuterons par une comparaison des performances des modèles développés, dans le but d'identifier celui qui s'est révélé le plus efficace pour distinguer les textes rédigés par des humains de ceux générés par une intelligence artificielle (**A**). Nous poursuivrons ensuite par une analyse approfondie des facteurs ayant permis au meilleur modèle d'atteindre ses performances, en mettant en lumière les éléments textuels les plus discriminants (**B**).

A. Comparaison des performances des modèles

Dans cette section, nous comparons les performances des différents modèles utilisés pour la classification des textes, à savoir : la régression logistique, le Linear SVC, la forêt aléatoire, le multinomial Naïve Bayes (NB) et le SVC. Les résultats obtenus montrent une performance globalement très satisfaisante en termes de qualité prédictive. Le **tableau 6** présente les métriques de précision des modèles testés sur les jeux d'entraînement (train) et de test. Cette métrique indique le pourcentage de textes correctement classés en fonction de leur origine, qu'ils soient générés par l'intelligence artificielle (IA) ou par un humain.

En termes de précision, tous les modèles affichent des performances élevées, largement supérieures à 90 % à l'exception du modèle SVC, qui se distingue par des résultats nettement moins performants. Le modèle SVC présente en effet une précision de seulement 50 % sur le jeu d'entraînement et 49 % sur le jeu de test, ce qui témoigne d'une efficacité particulièrement limitée et d'une incapacité à bien classifier les textes.

***Tableau 6** : Précision des modèles de classification*

	R. Logistique	**Linear SVC**	**Forêt aléatoire**	**Multinomial NB**	**SVC**
Train	0,977	0,997	1,000	0,946	0,505
Test	0,968	0,979	0,966	0,938	0,488

L'analyse comparative des résultats entre les jeux d'entraînement et de test ne met en évidence aucun signe majeur de surajustement. Bien que les performances sur le jeu de test soient légèrement inférieures à celles obtenues sur le jeu d'entraînement, les écarts demeurent faibles et acceptables. Il est également important de noter que la métrique f1-score a été évaluée et qu'elle est quasiment identique à celle de la précision, ce qui est attendu dans le cas d'un jeu de données équilibré. Par conséquent, nous avons choisi de ne pas l'inclure dans le tableau

Le rapport de classification montre que la qualité des prédictions est équilibrée entre les deux classes pour la majorité des modèles, à l'exception du modèle SVC. Ce dernier échoue complètement à prédire correctement la classe « humain », avec un f1-score et une précision de 0 pour cette catégorie. En effet, le modèle SVC semble prédire systématiquement la classe "IA", indépendamment de la véritable classe des textes, ce qui entraîne un taux de faux positifs extrêmement élevé. Tous les textes humains sont classés à tort comme générés par l'IA, tandis que les textes générés par l'IA sont correctement identifiés. Ce phénomène pourrait être dû à la manière dont SVC avec noyau linéaire gère la complexité des données textuelles. Avec les vecteurs TF-IDF, les textes sont représentés par un grand nombre de caractéristiques. Le modèle

SVC, même avec un noyau linéaire, semble avoir des difficultés à bien gérer cette complexité par rapport à d'autres modèles, qui montrent des résultats bien plus satisfaisants.

Enfin, si l'on devait sélectionner un modèle optimal pour cette tâche de classification, ce serait le Linear SVC, qui présente la précision la plus élevée parmi tous les modèles testés. **Le tableau 7** montre la matrice de confusion générée par Linear SVC sur l'échantillon de test. Nous observons que seulement 29 textes sur 1537 ont été faussement prédits comme étant générés par l'IA alors qu'ils sont en réalité d'origine humaine, tandis que seulement 33 textes sur 1463 ont été classés à tort comme étant d'origine humaine alors qu'ils ont été générés par l'IA. Ce modèle combine efficacement la puissance des SVM avec la stabilité du noyau linéaire, qui est particulièrement bien adapté aux représentations TF-IDF des textes.

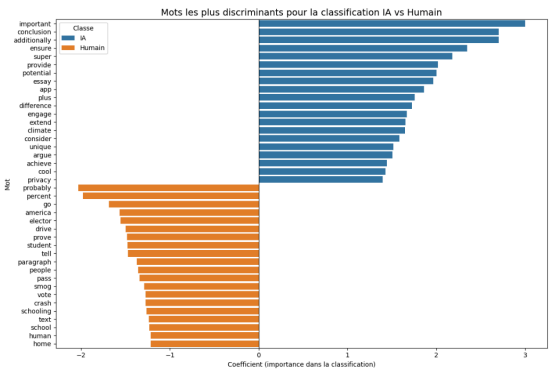
***Tableau 7** : Matrice de confusion du modèle linear SVC*

Réel /Prédit	Etre humain	Intelligence Artificielle
Etre humain	1508	29
Intelligence Artificielle	33	1430

B. Analyse de l'importance des mots dans la classification

Pour clôturer notre étude, nous avons souhaité identifier les mots ayant le plus contribué à la classification des textes par le modèle Linear SVC. Cette analyse permet de mieux comprendre les caractéristiques linguistiques qui distinguent les textes générés par une intelligence artificielle de ceux écrits par des humains.

***Figure 7** : Importance des mots selon le modèle Linear SVC*



Comme nous pouvons le constater à l'aide du graphique ci-dessus (**Figure 7**), les mots les plus fortement associés aux textes générés par l'IA incluent par exemple "important", "conclusion", "additionally", "ensure", "provide", "potential", ou encore "argue". Ces termes relèvent d'un registre formel et structuré, typique du langage académique. Ils traduisent une volonté de construction logique et claire, souvent recherchée dans les réponses produites par les IA. D'autres termes comme "essay", "privacy", "climate" ou "unique" évoquent des thématiques fréquentes dans les contextes scolaires ou scientifiques. Enfin, la présence de mots comme "super", "cool", "app", "plus" souligne une tentative d'insuffler un enthousiasme artificiel ou une touche technologique, souvent observée dans des contenus générés automatiquement. Ces éléments corroborent les observations issues d'analyses stylistiques de ChatGPT, qui indiquent un usage très marqué de connecteurs logiques et de tournures comme « il est important de noter » ou « en outre », conférant au discours une structure rigide, parfois perçue comme peu naturelle ou prévisible.

En contraste, les mots les plus associés aux textes humains, tels que "probably", "go", "tell", "home", "vote", "school", "drive" ou "people", traduisent une écriture plus spontanée, plus incarnée, souvent marquée par l'incertitude, l'opinion personnelle ou l'expérience vécue. D'autres termes comme "percent", "smog", "crash", "elector" ou "student" renvoient à des références concrètes et contextualisées, souvent ancrées dans une réalité socio-politique ou personnelle. Ce vocabulaire, plus spécifique et nuancé, contraste avec le langage plus générique et "lissé" de l'IA.

Ces résultats sont cohérents avec des observations linguistiques précédentes : les humains utilisent un langage plus varié, moins formaté, parfois moins rigoureux mais plus riche en signaux émotionnels, subjectifs ou contextuels. À l'inverse, les IA tendent à produire des textes bien structurés, standardisés, qui cherchent la clarté au détriment parfois de la spontanéité. Ce décalage dans le style lexical a permis au modèle de mieux distinguer les deux types d'auteurs, en s'appuyant sur les occurrences de mots emblématiques de chaque registre.

V. Conclusion et discussion

Pour conclure, cette étude avait pour objectif d'identifier et de caractériser les différences stylistiques et thématiques entre des textes rédigés par des humains et d'autres générés par une intelligence artificielle. Pour ce faire, nous avons utilisé un jeu de données disponible sur Kaggle, composé d'un corpus équilibré de textes humains et générés automatiquement.

Nous avons d'abord prétraité les données, puis conduit une analyse exploratoire en deux volets : sur les textes bruts pour certaines métriques (ponctuation, paragraphes, lisibilité, cohérence…), et sur les textes nettoyés pour des analyses plus poussées (analyse thématique, fréquence lexicale, sentiment). Il en ressort des différences notables entre les deux types de productions, notamment sur la ponctuation, la nombre des paragraphes, ou encore la variété grammaticale (verbes, adjectifs, adverbes). D'un point de vue thématique, l'IA tend à aborder plus fréquemment les sujets liés à l'éducation, tandis que les textes humains évoquent davantage des thématiques politiques. Par ailleurs, l'analyse de sentiment révèle que les textes humains sont plus enclins à exprimer des émotions négatives, contrairement aux textes générés qui adoptent un ton plus neutre ou positif, probablement en raison de la tendance des modèles à éviter les propos polarisants.

L'analyse de l'occurrence des mots a mis en évidence des termes caractéristiques de chaque catégorie : un vocabulaire plus introspectif et subjectif chez les humains ("think", "go", "want"), face à un lexique plus formel et structuré chez l'IA ("provide", "conclusion", "important"). Ces différences suggèrent que certains mots peuvent constituer des indicateurs discriminants utiles pour la classification automatique.

Dans un second temps, nous avons entraîné plusieurs modèles de classification à l'aide de représentations TF-IDF pour prédire l'origine (humaine ou générée) des textes. Les performances obtenues sont globalement satisfaisantes, et le meilleur modèle s'est avéré être le **LinearSVC**, probablement en raison de sa capacité à bien exploiter les dimensions linéaires du TF-IDF, tout en restant robuste face à la forte dispersion du vocabulaire. L'analyse des poids attribués aux mots par ce modèle a confirmé nos observations précédentes : certains termes utilisés préférentiellement par l'IA ou les humains se révèlent très discriminants.

Cette étude nous a ainsi permis de confirmer l'existence de différences tangibles entre écriture humaine et générée, tant sur le plan lexical que stylistique et thématique. Nos résultats sont en cohérence avec ceux d'études récentes (comme Goar et al., 2023), qui soulignent la tendance des modèles de langage à standardiser les productions textuelles, souvent au détriment de la spontanéité, de la subjectivité ou de la créativité.

Discussion

Cependant, notre travail présente certaines limites. D'abord, la base de données utilisée, bien que pertinente, ne reflète pas l'ensemble des contextes ou registres d'écriture (professionnels, artistiques, conversationnels, etc.). Par ailleurs, la qualité des textes humains peut varier fortement selon les auteurs, introduisant un bruit difficile à contrôler. Enfin, notre approche repose principalement sur des représentations statistiques (comme TF-IDF), qui ne capturent pas toute la richesse sémantique des textes.

Pour aller plus loin, plusieurs pistes d'approfondissement sont envisageables. D'abord, il serait pertinent d'intégrer directement au modèle certaines variables discriminantes issues de l'analyse exploratoire, comme le nombre de mots rares ou la complexité de la syntaxe, afin d'évaluer leur pouvoir explicatif complémentaire aux représentations classiques de type TF-IDF. Par ailleurs, on pourrait tester des modèles plus avancés, comme les réseaux de neurones basés sur des représentations contextuelles (BERT, RoBERTa), pour capturer des nuances plus fines. Enfin, l'étude gagnerait à être élargie à des corpus plus variés (dialogues, tweets, récits littéraires, etc.) et à une analyse stylistique plus fine, incluant par exemple la cohérence discursive, les niveaux de lecture, ou la détection de marqueurs narratifs.

VI. Bibliographie/sitographie

Articles

Bisi, T., Risser, A., Clavert, P., Migaud, H., & Dartus, J. (2023). "What is the rate of text generated by artificial intelligence over a year of publication in Orthopedics & Traumatology: Surgery & Research? Analysis of 425 articles before versus after the launch of ChatGPT in November 2022". Orthopaedics & Traumatology: Surgery & Research : https://doi.org/10.1016/j.rcot.2023.09.014

Goar, V., Yadav, N. S., & Yadav, P. S. (2023). Conversational AI for natural language processing: An review of ChatGPT. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 109-17.

Roten, C.-A., Nicollerat, S., Pousaz, L., & Genilloud, G. (2023). "Détecter par stylométrie la fraude académique utilisant ChatGPT". Les Cahiers de l'IRAFPA : https://doi.org/10.56240/irafpa.cm.v1n1/rot

Site

Les mots les plus utilisés par ChatGPT : Analyse du langage du chatbot - Startups Nation – Média des entrepreneurs et Start-ups en France

VII. Annexes

Annexe 1 : Méthodologie des indicateurs linguistiques extraits des textes bruts

Cette annexe présente les différentes mesures stylistiques utilisées pour analyser les textes dans leur forme brute, c'est-à-dire avant tout prétraitement. Ces indicateurs visent à mettre en lumière les différences structurelles et linguistiques entre les textes générés par une intelligence artificielle et ceux produits par des humains.

- **Ponctuation et taille moyenne des phrases**

Ponctuation : Comptage brut du nombre de signes de ponctuation présents dans chaque texte, ventilé par type : points, virgules, deux-points, points-virgules, tirets, points d'interrogation et d'exclamation.
Taille moyenne des phrases : Calculée comme le nombre moyen de mots par phrase, en utilisant une segmentation simple sur le point (.) comme séparateur.

- **Segmentation en paragraphe**

Nombre de paragraphes par texte basé sur le nombre de sauts de ligne (\n) pour estimer le découpage structurel du texte.

- **Grammaire**

Évaluation de la qualité grammaticale des textes. Utilisation de la bibliothèque SpellChecker pour détecter les mots non reconnus dans le dictionnaire. Pour chaque texte, comptage des mots considérés comme mal orthographiés.

- **Profondeur**

Calcul de la profondeur maximale des dépendances syntaxiques via la bibliothèque spaCy. Pour chaque mot, on mesure le nombre de niveaux nécessaires pour remonter jusqu'à la racine de la phrase. Un score élevé reflète une structure grammaticale plus complexe.

- **Indice de lisibilité**

Évaluation de la facilité de lecture d'un texte. Calcul avec la bibliothèque textstat de l'indice Flesch, basé sur : la longueur moyenne des phrases et la longueur moyenne des mots (en syllabes). Plus le score est élevé, plus le texte est simple à comprendre (score typique pour l'anglais : >60 = facile ; <30 = difficile).

- **Part of speech**

Comptage des catégories grammaticales : noms, verbes, adjectifs, adverbes. Traitement avec spaCy pour étiqueter chaque mot selon sa catégorie grammaticale (nom, verbe, adjectif, adverbe...).

- **Cohérence**

Calcul d'un score de cohérence basé sur la similarité sémantique entre chaque paire de mots consécutifs, à partir de leurs vecteurs (spaCy – modèle en_core_web_md).
Le score correspond à la moyenne des similarités cosinus ; plus il est élevé, plus le texte présente des enchaînements logiques et fluides entre les mots.

- **Rareté des mots**

Cette métrique mesure la fréquence d'apparition des mots dans un texte par rapport à l'ensemble du corpus. Elle utilise le **TF-IDF**, une technique qui permet d'extraire les mots rares d'un texte, c'est-à-dire ceux qui apparaissent fréquemment dans un document mais peu dans l'ensemble du corpus. Plus un mot a un score TF-IDF élevé, plus il est considéré comme rare. Pour cette analyse, un seuil de 0,01 est fixé pour considérer un mot comme rare, ce qui signifie qu'un mot doit apparaître dans moins de 1% des documents pour être comptabilisé comme rare.

- **Diversité des mots** :

La diversité mesure la richesse lexicale d'un texte, en calculant la proportion de mots uniques par rapport au nombre total de mots dans un texte. Un score plus élevé indique une plus grande variété de vocabulaire, tandis qu'un score faible suggère une répétition plus fréquente de mots. Cette mesure permet de quantifier l'étendue du vocabulaire utilisé dans les textes et d'étudier la diversité lexicale des écrits humains et générés par l'IA.

Annexe 2 : Statistiques comparatives entre les textes IA et humain réalisées sur le texte brut

Statistiques de ponctuation

generated	length	punct	dot_count	comma_count	semicolon_count	colon_count	dash_count	question_count	exclam_count	avg_sentence_len
0	417.9882	48.5464	20.9316	15.7516	0.2356	0.1876	0.4458	0.7936	0.3000	22.278269
1	342.5336	46.3464	17.6670	20.5114	0.0554	0.5108	0.9158	0.3266	0.4216	19.924035

Statistiques sur les paragraphes

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	9.6778	6.446821	1.0	7.0	9.0	11.0	93.0
1	5000.0	11.3242	5.880273	1.0	9.0	11.0	13.0	63.0

Statistiques sur la grammaire

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	37.5840	18.903090	1.0	24.0	34.0	48.0	156.0
1	5000.0	36.7404	15.364134	0.0	27.0	34.0	45.0	133.0

Indice de lisibilité

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	63.017572	16.204908	-357.170109	56.474462	64.290913	71.872778	93.514328
1	5000.0	47.018638	16.681473	-23.360000	35.781300	46.060672	57.570926	96.623335

Analyse de la profondeur

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	9.6570	2.180803	5.0	8.0	9.0	11.0	22.0
1	5000.0	9.6068	2.094152	3.0	8.0	9.0	11.0	24.0

Analyse du nombre de noms

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	86.7456	43.466515	14.0	55.0	78.0	108.0	294.0
1	5000.0	84.3490	37.891914	1.0	56.0	80.0	106.0	372.0

Analyse du nombre de verbes

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	60.6766	28.142424	8.0	40.0	55.0	75.0	210.0
1	5000.0	47.5382	16.790668	0.0	36.0	47.0	57.0	142.0

Analyse du nombre d'adjectifs

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	29.3982	16.312770	1.0	17.0	26.0	38.0	119.0
1	5000.0	32.7326	15.361038	0.0	23.0	30.0	40.0	182.0

Analyse du nombre d'adjectifs

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	20.9078	11.700610	0.0	13.0	19.0	27.0	86.0
1	5000.0	15.5080	7.190402	0.0	10.0	15.0	20.0	50.0

Analyse de la cohérence

generated	count	mean	std	min	25%	50%	75%	max
0	5000.0	0.378807	0.044378	0.204238	0.349176	0.379129	0.408867	0.544791
1	5000.0	0.362109	0.049490	0.134622	0.327237	0.356874	0.390725	0.572465

Annexe 3 : Résultats de l'analyse de régression logistique

Variable dépendante: Écrit par l'IA (1) ou l'humain (0)

Variables explicatives	Coefficient	Ecart-type
Nombre de commas	0.1020***	0.003
Nombre de semi-colonnes	-0.9285***	0.068
Nombre de colonnes	0.0883***	0.031
Nombre de tirets	0.0877***	0.017
Nombre de questions	-0.3421***	0.026
Nombre d'exclamations	0.2253***	0.022
Longueur moyenne de phrase	0.0020	0.002
Nombre de paragraphes	0.0705***	0.005
Indice de lisibilité	-0.0161***	0.001
Nombre d'adverbes	-0.0956***	0.004

Note: *** significatif au seuil de 1%

Annexe 4 : Répartition des classes selon les thématiques

Topic	Proportion (arrondi)
0	**IA :** 61%**,** Humain : 39%
1	**IA :** 39%**,** Humain : 61%
2	**IA :** 43%**,** Humain : 57%

Annexe 5 : Statistiques comparatives entre les textes IA et humain réalisées sur les sentiments

Répartition des sentiments

Polarité des textes	Répartition
Positif	9 133
Neutre	38
Négatif	829

Répartition des sentiments selon les classes

Classe	Proportion (arrondi)
IA	**Positif : 95%, Négatif : 5%**, Neutre : 0%
Humain	**Positif : 87%, Négatif : 12%**, Neutre : 1%

Répartition des sentiments selon les thématiques

Topic	Proportion (arrondi)
0	**Positif : 95%,** Négatif : 5%, Neutre : 0%
1	**Positif : 94%,** Négatif : 5%, Neutre : 0%
2	Positif : 81%, **Négatif : 18%**, Neutre : 1%

Table des matières