# **Éthique de l’intelligence artificielle**

**Sommaire :**

***I - Introduction***

***II - Risque par rapport au modèle économique***

***III - Risque juridique***

***IV- Risques stigmatisation et discrimination***

***V - Récapitulatif des mesures à prendre faces aux risques éthiques***

***VI - Conclusion : L’éthique, Un élément essentiel de légitimité***

***I - Introduction***

Les préoccupations sur l’éthique des algorithmes doivent être prises très au sérieux. Même en dehors de toute considération éthique, si l’IA veut intégrer les meilleures normes de “best practice” et de performance, l’IA doit intégrer l’éthique dès la conception des algorithmes (« *ethics by design »*). La performance éthique de l’algorithme (absence de discrimination, respect des individus…) doit figurer parmi les critères de performance, au même titre que l’exactitude des prédictions.

Cependant, l’éthique ne va pas de soi et comporte des difficultés. Un travail d'éclaircissement est nécessaire, car l’éthique est un sujet ouvert qui ne comporte pas de solution. Parfois même, une éthique rentre en contradiction avec une autre. C’est le cas par exemple du conflit entre éthique individuelle et de groupe. D’une part, il est impossible de garantir à la fois une équité individuelle – par exemple, des individus égaux doivent être traités également indépendamment de leur origine ethnique ou de leur genre – et d’autre part une équité de groupe – des individus par essence différents, en matière de genre par exemple, doivent être traités de manière inégale pour rétablir une équité sociale. L’algorithme devra donc être guidé par des choix politiques.

De plus, l’éthique est temporelle. Les valeurs morales que nous acceptons aujourd’hui évoluent rapidement ; ce que nous étions prêts à accepter hier n’est plus valable aujourd’hui. Les algorithmes d’IA devront donc intégrer cette contrainte temporelle à chaque étape, de la conception à leur déploiement sur les marchés.

Dans cette présentation, nous vous présentons une liste non exhaustive des risques encourues en termes d’éthique. Les sous parties seront systématiquement une explication en détail du risque suivi d’une explications quant à l’alignement de notre application avec les normes éthiques, voire judiciaires.

***II - Risque par rapport au modèle économique prédateur***

***a) Risque d’entrer dans le capitalisme de surveillance***

Le premier risque éthique majeur associé à notre application est le potentiel de contribuer à la dynamique du capitalisme de surveillance.

Le capitalisme de surveillance, également connu sous le nom d'économie de la surveillance, fait référence à un modèle économique dans lequel les entreprises collectent et exploitent massivement les données personnelles des utilisateurs pour générer des profits. Ce concept a été popularisé par Shoshana Zuboff, une professeure émérite de la Harvard Business School, dans son livre The Age of Surveillance Capitalism.

Dans le capitalisme de surveillance, les entreprises utilisent des technologies telles que la collecte de données, l'analyse de données, l'intelligence artificielle et le machine learning pour surveiller les activités en ligne des individus, recueillir des données personnelles et comportementales, et les utiliser à des fins commerciales. Ces données sont souvent utilisées pour cibler de manière personnalisée la publicité, influencer les choix des utilisateurs, et créer des modèles de comportement et des profils détaillés.

Le capitalisme de surveillance soulève des préoccupations majeures en termes de vie privée, de consentement, de manipulation et de pouvoir des entreprises. Les utilisateurs peuvent avoir peu de contrôle sur la collecte, l'utilisation et le partage de leurs données personnelles. De plus, les modèles de comportement prédictifs peuvent influencer les choix des utilisateurs de manière subtile, voire manipulatrice.

En adoptant une approche centrée sur la collecte massive de données personnelles des utilisateurs et leur utilisation pour générer des profits par le biais de publicités ciblées, l’application pourrait s'inscrire dans le cadre du capitalisme de surveillance. Cela signifie que les principaux objectifs commerciaux sont axés sur la collecte et l'exploitation de données personnelles, plutôt que sur la création de valeur ajoutée pour les utilisateurs ou la société dans son ensemble.

***b) Réponse***

Par conséquent, il est essentiel de réfléchir attentivement aux implications éthiques d'une telle approche et de considérer des alternatives plus respectueuses de la vie privée et de l'autonomie des utilisateurs. Nous allons voir plus bas que nous avons un business modèle qui respecte davantage la confidentialité des données, qui est transparent dans la collecte et l'utilisation des données, et qui donne aux utilisateurs un réel contrôle sur leurs informations personnelles.

***III - Risque juridique***

***a ) Juridique : Risque par rapport au respect de la confidentialité et de la protection des données***

Notre application doit respecter les lois et réglementations en vigueur concernant la protection des données, telles que le RGPD. Cela signifie obtenir un consentement éclairé des utilisateurs, sécuriser les données de manière appropriée et garantir qu'elles ne sont pas utilisées à des fins non autorisées.

Crypter les données, offrir le droit de supervision et de suppression des données sont des mesures importantes pour se conformer au règlement général sur la protection des données (RGPD), mais elles ne sont pas les seules exigences.

Voici quelques éléments supplémentaires à prendre en compte pour assurer la conformité au RGPD :

1. Base légale : Vous devez avoir une base légale appropriée pour collecter, traiter et stocker les données personnelles. Cela peut inclure le consentement explicite de l'individu, l'exécution d'un contrat, le respect d'une obligation légale, la protection des intérêts vitaux de la personne concernée ou l'intérêt légitime de l'entreprise.
2. Transparence : Vous devez informer les individus de manière claire et concise sur la collecte, le traitement et l'utilisation de leurs données personnelles. Cela peut être réalisé à l'aide d'une politique de confidentialité ou d'une déclaration de protection des données.
3. Minimisation des données : Vous devez limiter la collecte et le traitement des données personnelles aux seules informations nécessaires pour atteindre l'objectif spécifié. Il est important de ne pas conserver les données plus longtemps que nécessaire.
4. Sécurité des données : Outre le cryptage, vous devez mettre en place des mesures de sécurité appropriées pour protéger les données personnelles contre tout accès non autorisé, la divulgation, l'altération ou la destruction accidentelle ou illicite.
5. Transferts internationaux : Si vous transférez des données personnelles en dehors de l'Union européenne, vous devez vous assurer que vous avez des mécanismes légaux appropriés en place pour garantir un niveau adéquat de protection des données.
6. Droits des individus : Outre le droit de supervision et de suppression, les individus ont d'autres droits en vertu du RGPD, tels que le droit d'accès, de rectification, de limitation du traitement et de portabilité des données. Vous devez être prêt à répondre à ces demandes de manière appropriée.

D’autant plus que le RGPD n’est pas la seule instance de droit. Le règlement général sur la protection des données (RGPD) a donné des ailes à la Commission européenne. Après le RGPD, il y aura le Data Act ; la directive Open Data existe depuis 2019, le Data Governance Act, appliqué depuis juin (avec un délai de mise en conformité de 15 mois), et d’autres se préparent.

Les sanctions peuvent être lourdes dans le cas contraire, comme l’atteste cet article :

<https://www.lemonde.fr/pixels/article/2023/04/05/chatgpt-premieres-plaintes-aupres-de-la-cnil-contre-le-logiciel-d-intelligence-artificielle_6168425_4408996.html>

<https://www.lemonde.fr/idees/article/2023/06/06/les-systemes-d-intelligence-artificielle-vont-amplifier-les-biais-de-genre-dans-tous-les-domaines_6176418_3232.html#xtor=AL-32280270-%5Bdefault%5D-%5Bios%5D>

En effet, nous risquons des procédures à notre encontre suite à des plaintes portant sur l’utilisation des données personnelles déposées auprès de la CNIL.

Cependant, Le RGPD offre paradoxalement une immense opportunité de collecte pour les publicitaires et autres sociétés de services marketing.

***b ) Réponse***

Pour respecter la confidentialité et la protection des données des utilisateurs, voici quelques mesures que votre application pourrait prendre :

1. Consentement éclairé : Obtenir le consentement éclairé des utilisateurs avant de collecter et d'utiliser leurs données personnelles.
2. Politique de confidentialité : Fournir une politique de confidentialité claire et facilement accessible, qui explique en détail quelles données sont collectées, comment elles sont utilisées, et quelles mesures de sécurité sont mises en place pour les protéger.
3. Minimisation des données : Collecter uniquement les données nécessaires à des fins spécifiques et légitimes.
4. Sécurité des données : Mettre en place des mesures de sécurité robustes pour protéger les données des utilisateurs contre les accès non autorisés, les pertes ou les fuites.
5. Partage de données limité : Envisager de partager les données des utilisateurs avec des tiers, assurez-vous d'obtenir leur consentement explicite pour cela.
6. Durée de conservation des données : Déterminer une période de conservation appropriée pour les données collectées.
7. Transparence sur les pratiques de collecte et d'utilisation des données : Informez les utilisateurs de manière transparente sur les pratiques de collecte et d'utilisation des données de votre application.
8. Accès et contrôle des utilisateurs : Donner aux utilisateurs la possibilité d'accéder à leurs propres données, de les corriger si elles sont inexactes et de les supprimer si elles le souhaitent.

***IV- Risques stigmatisation et discrimination***

***a) Risque de reproduction des stéréotypes***

<https://www.lemonde.fr/pixels/article/2018/03/28/intelligence-artificielle-ce-qu-il-faut-retenir-du-rapport-de-cedric-villani_5277697_4408996.html>

Si l’on n’y prend garde, l’intelligence artificielle reproduira nos stéréotypes de genre

Il en est ainsi des inégalités entre les femmes et les hommes, qui sont inscrites dans les données dont se nourrissent les algorithmes et dont l’IA est incapable de reconnaître le caractère discriminant. Si l’on n’y prend garde, les algorithmes produisent nos stéréotypes de genre et perpétuent la place inégale des femmes par rapport aux hommes dans nos sociétés, ce constat étant valable pour d’autres formes d’inégalités, et encore plus pour les inégalités croisées.

*« L’IA ne peut pas être une nouvelle machine à exclure »*, écrit M. Villani. Entre les biais de certains programmes – qui reproduisent le racisme et le sexisme humain par exemple, une personne au cheveux long est une femme. Le manque de mixité dans les équipes de recherche et les risques sociaux que peuvent représenter ces programmes, l’IA peut inquiéter.

***b ) Qu'est-ce que les biais de l'IA ? Exemple***

Dans le cadre du projet Gender Shades, les chercheurs Inioluwa Deborah Raji de l’Université de Toronto et Joy Buolamwini du MIT ont **testé les technologies de reconnaissance faciale** de Microsoft, IBM, et de l’entreprise chinoise Face++.

À travers ce passage en revue, ils se sont aperçus que chacune de ces technologies se révélait **plus performante sur des visages à peau blanche** que sur des visages à la peau plus sombre. Il s’agit-là d’un simple exemple parmi tant d’autres, tristement révélateur de ce mal qui corrompt l’IA.

Les deux chercheurs ont aussi examiné des systèmes de classification par genre. Là encore, les trois systèmes se sont avérés bien **moins performants sur les visages de femmes** que sur les visages d’hommes. Une femme à la peau foncée sera mal catégorisée dans 34,7% des cas, alors que le taux d’erreur est de 0,8% pour un homme à la peau claire.

Il est donc indéniable que **les femmes et les personnes à la peau sombre sont discriminées** par l’intelligence artificielle. Or, cette discrimination peut avoir des conséquences potentiellement dramatiques.

De plus, les différences culturelles et d’expressions en fonction de la couleur de peaux pourraient être qualifiées d’*« atypiques »*, créant un risque de stigmatisation et de discrimination pour les personnes qui ne correspondent pas au shéma classique. Par exemple, un homme aux cheveux longs pourraient être pris pour une femme.

***c ) Quelles sont les causes des biais de l'IA?***

L’intelligence artificielle (IA) fait désormais partie intégrante de notre quotidien (industrie, emploi, santé, consommation, etc.). La dynamique se situe dorénavant autour de la nature des activités et des tâches qu’il est possible de lui confier. Cette technologie, par sa capacité à exécuter de manière plus qualitative et plus rapide des actions jusqu’ici réalisées par des humains, soulève de nombreux questionnements.

En cause, le *« machine learning »* (« apprentissage automatique »), qui engendre des décisions inexplicables, non contrôlées et prises en compte sans recours, ce qui déconcerte. Plus particulièrement, l’IA connaît une limite grandissante en termes de discrimination, que l’on qualifie de « biais ». Le biais discriminatoire est le fait pour l’intelligence artificielle de fausser le résultat obtenu et de discriminer les personnes, de manière volontaire ou non, en fonction de leur genre, de leur origine ethnique, ou encore de leur orientation sexuelle.

L’origine des biais est multiple : ils peuvent tout d’abord être causés par le programmateur lui-même qui, en concevant l’algorithme, y insère ses propres biais, caractérisés par ses croyances, goûts, idéologies, etc. Ils peuvent aussi être issus des données contenues dans l’algorithme, c’est-à-dire des données à partir desquelles l’intelligence artificielle apprend\*.\*

Aujourd’hui, le phénomène va encore plus loin avec l’utilisation de systèmes de réseaux de neurones profonds. En effet, les informaticiens ne sont plus capables d’expliquer quels ont été les critères de décision et par conséquent, d’expliquer le résultat en lui-même. Des liens sont faits par l’IA, sans que l’on puisse savoir si une logique précise a été suivie.

Les biais induits par la machine provoquent des atteintes inévitables aux droits fondamentaux : les principes d’égalité de traitement et de non-discrimination des citoyens font partie des valeurs fondamentales du modèle social français consacrées par la Constitution. La discrimination est une atteinte à l’égalité des droits proclamée à l’article 1er de la Déclaration universelle des droits de l’homme.

La cause des biais discriminatoires est à chercher directement dans **les ensembles de données sur lesquels les IA sont entraînées**. En inspectant deux des principaux « [**datasets**](https://datascientest.com/public-datasets-machine-learning) » massivement utilisés pour nourrir les systèmes de reconnaissance faciale, on constate très rapidement qu’ils sont majoritairement composés de portraits de personnes à la peau claire : 79,6% pour l’ensemble de données IJB-A, et 86,2% pour Adience.

On distingue **deux catégories de biais de l’IA** : les biais algorithmiques, et les biais sociétaux. Dans le premier cas, l’intelligence artificielle a été entraînée sur des données biaisées. C’est précisément le problème des systèmes de reconnaissance faciale nourris aux données d’IJB-A ou Adience.

Dans le deuxième cas, celui des biais sociétaux, l’intelligence artificielle est biaisée par des préjugés et **des stéréotypes profondément ancrés dans l’inconscient collectif**. Ce facteur est encore plus redoutable, car il est très difficile de s’en défaire ou même d’en prendre conscience…

Dans la grande majorité des situations, les biais de l’intelligence artificielle sont directement hérités des Data Scientists qui la créent. À l’instar d’un enfant dont la vision sera inconsciemment influencée par l’éducation prodiguée par ses parents, l’IA peut reproduire les mêmes discriminations que ses créateurs.

Ainsi, les biais sociétaux reflètent **une intolérance sociale ou une discrimination institutionnelle** profondément enracinée dans les mœurs. Les algorithmes et les données pourront sembler neutres, mais l’IA se révélera discriminatoire à l’usage.

Il est **très difficile d’identifier et de tracer les biais sociétaux**. En 2014, une étude avait révélé que Google Ads proposait davantage de pubs pour des métiers hautement rémunérés aux hommes qu’aux femmes. Ce biais reflète un sexisme encore trop présent dans nos sociétés.

Google se contente d’autoriser les publicitaires à cibler uniquement des hommes pour leurs annonces, mais ce sont les entreprises qui sont coupables de ce ciblage publicitaire discriminatoire. Il est donc **difficile de déterminer qui est responsable**.

Parmi les autres exemples d’outils d’IA biaisés, on peut citer le **système de recrutement automatisé** développé par Amazon en 2014. Cette intelligence artificielle était censée sélectionner les meilleurs CV parmi les candidats à un poste à pourvoir.

Malheureusement, le système s’est révélé discriminatoire à l’égard des femmes. Pour cause, il avait été entraîné sur **les données historiques des employés d’Amazon** : en grande majorité des hommes blancs.

Notre application doit être conçue de manière à minimiser les biais et les discriminations. Cela implique de s'assurer que les recommandations et les publicités personnalisées sont basées sur des critères objectifs et pertinents, et qu'elles ne favorisent pas indûment certains groupes ou ne perpétuent pas les préjugés existants. En règle générale, cette discrimination reflète celles que l’on déplore dans notre société à l’égard des couleurs de peaux, du genre, de l’âge ou de la nationalité.

Tant que ce problème ne sera pas résolu, **l’intelligence artificielle ne pourra révéler tout son potentiel** et risque même de s’avérer néfaste et dangereuse pour la société. À travers ce dossier, découvrez les causes et les conséquences des biais de l’IA et les solutions pour les faire disparaître…

***d ) Réponse : Une intelligence artificielle éthique***

Le développement d’une intelligence artificielle éthique est un enjeu primordial pour les entreprises. Or, pour inspirer confiance à ses salariés et à ses clients, une entreprise doit s’assurer que le recours à l’intelligence artificielle s’accompagne de l’élimination de tout biais discriminatoire.

Pour faire face aux risques de stigmatisation et de discrimination, notre application peut prendre plusieurs mesures :

1. Collecte de données objectives : Assurer que notre application collecte et utilise des données objectives et pertinentes pour ses recommandations et publicités personnalisées.
2. Évaluation et correction des biais : Mettre en place des mécanismes pour évaluer et atténuer les biais dans les recommandations et les publicités générées par votre application.
3. Diversité et inclusion dans les données d'entraînement : Veiller à ce que les données d'entraînement utilisées pour alimenter le modèle d'IA soient diversifiées et représentatives de la population.
4. Transparence et applicabilité : Rendre les algorithmes et les processus de recommandation transparents et explicables.
5. Mécanismes de signalement et de révision : Mettre en place des mécanismes permettant aux utilisateurs de signaler tout contenu inapproprié, offensant ou discriminatoire.
6. Surveillance continue : Surveiller en permanence les résultats et les effets de votre application pour identifier tout signe de stigmatisation ou de discrimination.

***V - Explication de notre traitement de la data***

I - Alignée avec le RGPD

Nous allons expliquer ici pourquoi notre application n'est pas soumise au RGPD.

Untitled

1. Nous on transmet pas de data, on leur affiche seulement le résultat de notre analyse

Nous avons deux types de clients. D'un côté les JCDecaux pour mettre à disposition automatiquement leurs écrans, et de l'autre côté des Havas ou Publicis qui ont des clients qui veulent diffuser des campagnes de communication. Ils renseignent sur la WebApp leurs objectifs, le profil cible etc. Une fois ces renseignements enregistrés, nous traitons la data à travers nos différentes IA. Une fois l’analyse faite, nous leur affichons les résultats de nos analyses sur le site. Ils n’ont qu’un accès en readonly, ils ne peuvent que constater le produit de notre analyse. Ni la data ni le processus d’analyse ne leur sont accessibles.

Ainsi, nous ne transmettons pas de données que nous avons récolté, on leur affiche seulement le résultat de notre analyse, comme dans DataDog. Nous n’avons pas un modèle de capitalisme de surveillance.

En effet, ces analyses de prédiction ne sont pas négociées sur un nouveau marché de comportements. Il n’y a pas d’opérations commerciales puisqu’on ne révèle ni ne vendons la data à des tiers tels que Havas ou JCDecaux.

2 ) Cryptage des données + suppression des méta datas

Le display kiosk est composé de l’écran + caméra + screen. C’est ici que l’on collecte les données concernant les usagers sur leurs localisation, la date, leurs préférences en fonction de leur niveau d’attention. Ces données vont ensuite être utilisées afin de créer des profils détaillés et des prédictions comportementales.

Dans un premier temps, nous allons crypter la données à travers une librairie python. La clé sera stockée dans un environnement auquel personne n’a accès. Donc personne ne pourra décrypter les données. Seul un user créer automatiquement sur docker et activable à travers GitHub Action peut démarrer le programme. Ce sera le seul à avoir accès à la clé de cryptage des données.

Dans un deuxième temps, les métadatas ( lié au lieu, la date etc ) seront supprimés.

Ainsi, les données sont soit anonymisées soit détruites, si bien que plus personne n’a accès aux données qui ont permis l’analyse. Ainsi, nous ne pouvons vendre aux annonceurs ou aux tiers la data, puisqu’elle a été supprimée. De plus, la vie privée n’est pas affectée.

En conclusion, les objectifs commerciaux ne sont pas axés sur la collecte et l’exploitation de données personnelles, puisqu’on supprime les données anonymisées une fois analysées ( sauf cas où ne parvenons pas à analyser la data, mais dans ce cas-là ça rentre dans une boucle jusqu’à l’avoir analysé ).

Cependant, on doit bien admettre qu’on essaye d’influencer les choix des utilisateurs de manière subtile, voire manipulatrice.

3 ) Modèle consenti et ciblé

D’ailleurs, ce qui pose problème, ce sont les caméras de surveillance algorithmique non ciblées présentes dans des zones accessibles au public. Or dans notre cas, les caméras de vidéosurveillances seront à la fois ciblées et présentes dans des lieux privés où l’on a recueilli le consentement des usagers. Ainsi, les peurs concernant les exercices des libertés civiques fondamentales, et notamment le droit à la liberté de réunion, d’association et d’expression ne sont ici pas justifiées.

4) On exclue de l’analyse des caractéristiques lié à la couleur de peau

Il n'est pas explicitement illégal d'analyser la couleur de peau dans les statistiques en France. Cependant, la collecte de données basée sur la race ou l'origine ethnique est réglementée en France en vertu de la loi Informatique et Libertés de 1978, ainsi que du règlement général sur la protection des données (RGPD) de l'Union européenne.

En France, la collecte de données sensibles, y compris celles liées à la race ou à l'origine ethnique, est généralement interdite, sauf dans certains cas spécifiques prévus par la loi. Par conséquent, les organismes publics et privés sont tenus de respecter des principes stricts de protection des données lorsqu'ils collectent, traitent et analysent des informations sur la race ou l'origine ethnique des individus.

Cependant, il est important de noter que l'utilisation de catégories comme la race dans les statistiques peut être délicate et controversée. En France, la Constitution et la législation ont pour principe de ne pas reconnaître les différences raciales ou ethniques, afin de promouvoir l'égalité et l'unité de tous les citoyens. Par conséquent, l'utilisation de critères raciaux dans les statistiques peut être considérée comme allant à l'encontre de ce principe.

Dans le cadre de notre projet, nous ne prenons pas en compte la couleur de peau, la race ou l’origine ethnique de la personne. Ainsi, nous sommes dans les clous au niveau du RGPD.

5 ) Liste des exigences RGPD avec lesquelles on est au norme

a ) Aux normes

Au vu de ce qu’on a vu précédemment, voici la liste des exigences RGPD avec lesquelles nous sommes au normes :

* Cryptage des données
* Droit à la supervision + Droits des individus
* Suppression des données
* Base légale
* Transparence
* Minimisation des données
* Sécurité des données

b) Spécification sécurité des données

Voici le guide de la sécurité des données personnelles.

* sensibiliser les utilisateurs → on rappelle aux utilisateurs leurs droits
* authentifier les utilisateurs → il y a bien une authentification
* gérer les habilitations → On limite bien les accès aux seules données dont un utilisateur a besoin.
* tracer les opérations et gérer les incidents → Pour tracer les opérations et prévoir des procédures pour gérer les incidents afin de pouvoir réagir en cas de violation de données, nous allons créer du code dans le Back End.
* sécuriser les postes de travail → antivirus
* sécuriser l’informatique mobile →
* protéger le réseau informatique interne → nous allons autoriser uniquement les fonctions réseau nécessaires aux traitements mis en oeuvre.
* sécuriser les serveurs → Renforcer les mesures de sécurité appliquées aux serveurs.
* sécuriser les sites web → S’assurer que les bonnes pratiques minimales sont appliquées aux sites web.
* sauvegarder et prévoir la continuité d’activité → Nous allons effectuer des sauvegardes régulières pour limiter l’impact d’une disparition ou d'une altération non désirée de données.
* archiver de manière sécurisée → Nous allons archiver les données qui ne sont plus utilisées au quotidien mais qui n’ont pas encore atteint leur durée limite de conservation, par exemple parce qu’elles sont conservées afin d’être utilisées en cas de contentieux.
* encadrer les développements informatiques → intégrer sécurité et protection des données personnelles au plus tôt dans les projets.

II - Mesures pour éviter les biais

### **IA Gender Détection**

Pour la détection du genre, nous utiliserons un modèle de type ResNet, qui est réputé pour sa précision dans le classement des images, mais sera spécifiquement entraînée pour classer les images de visages en deux catégories : homme et femme. Le transfert d'apprentissage sera à nouveau utilisé pour adapter un modèle pré-entraîné à nos besoins. Les données d'entraînement proviendront de bases de données publiques de visages, avec des étiquettes de genre, ainsi que de nos propres données recueillies sur site.

Le problème c’est qu’actuellement, ces modèles comportent certains biais. En effet, puisque l’IA est basée sur la statistique, elle va ressortir le résultat le plus probable en fonction des caractéristiques de la personne analysée. Ce phénomène peut créer des stéréotypes. Si la personne est un homme aux cheveux long, alors l’IA estimera que c’est une femme.

b ) Notre action

Nous allons intervenir dans le code afin de rectifier ce biais.

1 ) Nous allons ajouter des layers au modèle ResNet afin de corriger ce biais.

2 ) Ajouter des conditions dans le backend. Par exemple, dès qu’une personne a des cheveux longs, nous allons vérifier si elle possède une barbe ou pas. Si elle a des cheveux mais pas de barbe, nous partons du principe que c’est une femme. Si la personne possède des cheveux longs et une barbe, alors nous partons du principe que c’est un homme.

***V - Récapitulatif des mesures à prendre faces aux risques éthiques***

1. Respect de la confidentialité et de la protection des données : Notre application doit respecter les lois et réglementations en vigueur concernant la protection des données, telles que le RGPD. Cela signifie obtenir un consentement éclairé des utilisateurs, sécuriser les données de manière appropriée et garantir qu'elles ne sont pas utilisées à des fins non autorisées. Ainsi, on anonymisera les données, on recherchera le consentement des ayants droit, on fera attention à la propriété intellectuelle, bref, on contrôlera leur réutilisation, dit le Data Governance Act.
2. Transparence et applicabilité : Notre application doit être transparente quant à la collecte et à l'utilisation des données des utilisateurs. Les utilisateurs doivent être informés de manière claire et compréhensible sur la manière dont leurs données sont utilisées, y compris leur vente à des tiers. De plus, il est important de pouvoir expliquer les décisions prises par le système d'IA de manière compréhensible aux utilisateurs.
3. Évitement du biais et de la discrimination : Notre application doit être conçue de manière à minimiser les biais et les discriminations. Cela implique de s'assurer que les recommandations et les publicités personnalisées sont basées sur des critères objectifs et pertinents, et qu'elles ne favorisent pas indûment certains groupes ou ne perpétuent pas les préjugés existants.
4. Consentement volontaire et possibilité de retrait : Les utilisateurs doivent avoir la possibilité de donner leur consentement volontaire pour l'utilisation de leurs données, et ils doivent également pouvoir retirer leur consentement à tout moment. Il est important de respecter le choix des utilisateurs et de leur donner le contrôle sur leurs données.
5. Impact social et responsabilité : Votre application doit être évaluée pour comprendre son impact social global. Il est important de se demander si l'utilisation des données des utilisateurs à des fins de publicité personnalisée est bénéfique pour la société dans son ensemble et si cela peut avoir des conséquences néfastes sur les individus ou les groupes.

***VI - Conclusion : L’éthique, Un élément essentiel de légitimité***

Le développement d’une IA éthique est donc devenu primordial pour les entreprises à tous les niveaux, de la chaîne de production jusqu’à la relation client. Tous les métiers sont confrontés au même défi, concilier transparence juridique et technologique, protection de la vie privée des utilisateurs et sécurité. Ce n’est qu’en donnant ce gage de confiance que l’IA deviendra un réel avantage pour les entreprises.

Toutes doivent définir un cadre de responsabilité aussi bien numérique que moral. En effet, une entreprise voulant faire de l’IA un atout de compétitivité doit inspirer la confiance, tant de ses clients que de ses collaborateurs.

L’éthique est un élément essentiel à la légitimité d’une entreprise. Pour ce faire, chacune d’elles devra démontrer que les bénéfices engendrés par son avancement technologique respectent les droits des personnes qu’elle impacte et prennent en considération les conséquences sur autrui.

Les entreprises devront mettre en avant leur capacité à gérer leur « capital éthique » et à préserver une réputation positive face à leurs concurrentes. Une régulation de l’utilisation de l’intelligence artificielle est indispensable et doit être prônée par toute entreprise. Par exemple, une charte éthique d’usage interne des données, une confidentialité des données conforme au RGPD et à la réglementation sur la régulation des algorithmes contribuent à assurer un avantage compétitif pour toute entreprise faisant appel à l’intelligence artificielle.