

Livrable - Éthique IA

Contexte

Il s'agit dans un document mettant en évidence votre appropriation de la méthodologie proposée pour avoir une démarche éthique traçant vos choix justifiés par rapport aussi aux échanges qu'il y a eu dans votre équipe pendant toutes les phases du projet de la préparation du jeu de données à la proposition de pistes d'amélioration à votre client. En plus de la démarche, vous devrez bien mettre en évidence les décisions que vous avez prises dans l'équipe ainsi que les points de vigilance ou de contrôle à mettre en place par rapport aux problématiques éthiques rencontrées.

Sommaire

Les 7 étapes de l'éthique dans l'IA

- Respect de l'autonomie humaine
- Robustesse technique et sécurité
- Confidentialité et gouvernance des données
- Transparence
- Diversité, non-discrimination et équité
- Bien-être environnemental et sociétal
- Responsabilité

La réflexion lors de la conception de la solution

Application de l'éthique lors de la réalisation de la solution

Regard éthique lors du final de la solution

Les 7 étapes de l'éthique dans l'IA

Respect de l'autonomie humaine

Le but de cette section est de savoir si notre solution va être amené à aboutir à donner des résultats menant à la décision de vie humaine.

Il n'est pas envisageable qu'une IA décide directement de la vie d'une personne. Il faut aussi que le sujet d'études par l'IA (en l'occurrence dans notre cas le personnel de l'entreprise HumanForYou.

Un autre point important, il est important que notre IA n'intervienne pas dans l'autonomie des personnes impliquées, c'est-à-dire qu'elle ne suscite aucune interaction entre les personnes à la suite de l'étude.

Dans notre étude, l'effet escompté est d'obtenir les indicateurs de départs de certains employés, ainsi qu'un modèle de prévision afin de prévoir d'éventuel départs. La société HumanForYou pourra alors ajuster certains critères afin de donner envie aux personnes de rester dans l'entreprise. C'est pour cela que notre IA n'affectera pas l'autonomie humaine. Les employées seront toujours libres de partir ou non de l'entreprise.

Robustesse technique et sécurité

Le principe de la Robustesse technique dans le contexte de l'intelligence artificielle consiste à développer avec une approche préventive et de manière fiable pour limiter au maximum les risques potentiels.

Pour cela, on va faire en sorte que l'IA respecte certaines normes, on va aussi le sécuriser afin d'éviter des attaques sur celui-ci pouvant causer des complications. On essaie aussi d'envisager les vulnérabilités potentielles de notre IA et les options pour les supprimer ou au minimum évaluer les risques et les protocoles à utiliser en cas de problème.

On va aussi se concentrer sur la maintenance et la sécurisation du système, il faut définir un intervalle auquel faire un check-up complet et mettre à jour certaines dispositions si nécessaire. Pour respecter cela, avoir une documentation sur la maintenance, les situations de crise ou encore les plans de secours permet de rendre plus efficace ces différents process.

Pour ce qui est de la robustesse de notre IA, il faut faire attention à comment l'IA va traiter certaines données qui sont considérées comme personnelles. Etant donné que notre IA fonctionnera en local il faudra s'assurer que le système informatique où réside l'IA est sécurisé. Et pour ce qui est des check-up de sécurités et de mise à jour. Il serait bien de revérifier le modèle chaque année pour s'assurer que le modèle préconiser reste le meilleur.

Confidentialité et gouvernance des données

Cette section nous fait prendre conscience de l'utilisation des données, faire attention si les données sont sous la réglementation RGPD.

Il faut faire attention à bien garder les données si jamais elles sont soumises à des réglementations. Définir une personne responsable de la protection de ces données qui aura suivi les différentes phases du développement de l'IA.

Avoir des mécanismes de contrôle du traitement de données & essayer de minimiser l'utilisation des données à caractère personnel.

La provenance des données étant l'Inde, ce pays n'est pas régi par la législation du RGPD mais par le DPB (Data Protection Bill) indien. Le RGPD et le DPB indien partagent des similitudes dans leur objectif de protéger les données personnelles, ils présentent des différences dans leur portée, leur réglementation, leur consentement, leurs droits des personnes concernées et leurs pénalités. Dans notre étude il va donc être important de demander le consentement du personnels avant de pouvoir utiliser ses informations.

Transparence

La transparence est composée de trois éléments

- La traçabilité, est-il possible de retracer le modèle, les règles et les données.
- L'explicabilité, est-il simple d'expliquer les processus techniques du système d'IA.
- La communication, elle permet d'auto-évaluer si les capacités et les limites du système d'IA ont été communiquées aux utilisateurs d'une manière appropriée.

Dans ce contexte, la traçabilité est essentielle pour garantir la transparence de la collecte et de l'utilisation des données. Les données collectées doivent être clairement documentées, et les employés doivent être informés de la collecte de leurs données et de l'utilisation qui en sera faite. Il est important de noter que la collecte de certaines données peut être réglementée par la loi, et que l'entreprise doit être en conformité avec les exigences en matière de confidentialité et de protection des données.

L'explicabilité est également un élément clé de l'éthique de la collecte et de l'utilisation des données. Les employés doivent être en mesure de comprendre pourquoi certaines données sont collectées et comment

elles sont utilisées pour prendre des décisions. Les algorithmes utilisés pour analyser les données doivent être transparents et compréhensibles pour les employés, afin qu'ils puissent avoir confiance dans les résultats de l'analyse.

Enfin, la communication est essentielle pour maintenir la confiance des employés dans le processus d'analyse des données. Les résultats de l'analyse doivent être communiqués de manière claire et accessible, en évitant tout jargon technique qui pourrait rendre la communication opaque. Les employés doivent également être informés des mesures prises pour garantir la confidentialité et la protection de leurs données.

Diversité, non-discrimination et équité

Pour éviter toute discrimination dans le contexte de l'IA on va essayer de créer un ensemble de procédures qui vont concerner l'utilisation des données et la conception algorithmique.

On doit garantir une équité dans le jeu de données et éviter de discriminer sur l'âge, le sexe, les capacités ou encore les caractéristiques.

Les systèmes d'IA ne doivent pas être utilisés pour discriminer les personnes sur la base de leur origine, de leur genre, de leur race, de leur orientation sexuelle, de leur religion ou de leur situation socio-économique.

Globalement pour respecter la diversité durant le développement et l'utilisation d'une IA sur des données utilisateur il faut régulièrement se demander si nos décisions ne vont pas avantager certains groupes et à l'inverse en désavantager d'autres, pour faire cela on peut poser nos questions éthiques à des personnes ne prenant pas directement part au projet.

En outre, nous sommes conscients de la nécessité de respecter les principes de diversité, de non-discrimination et d'équité dans l'analyse de ces données. Nous nous engageons donc à éviter tout biais dans nos analyses et à considérer tous les critères de façon égale. Nous sommes également attentifs à l'identification des variables potentiellement sensibles, telles que le sexe, l'âge, etc., et à veiller à ce que leur inclusion dans l'analyse ne conduise pas à des conclusions discriminatoires ou biaisées.

Bien-être environnemental et sociétal

Les IA doivent être conçues afin de maximiser leur impact social positif et minimiser les impacts négatifs de manière à garantir que les technologies sont utilisées de manière responsable et bénéfique pour la société dans son ensemble. Les concepteurs d'une IA doivent être conscients de l'impact social qu'elle peut créer et être prêt à rendre des comptes concernant les potentiels impacts négatifs.

Dans notre situation, il faut rendre conscience du fait que l'analyse des critères de départ des employés doit être abordée avec prudence, et en gardant à l'esprit les impacts environnementaux et sociétaux potentiels. En effet, il est possible que certaines variables telles que le salaire, la distance avec l'habitation ou la satisfaction au travail soient liées à des problématiques de transport, de consommation d'énergie ou de qualité de vie. Ainsi, l'analyse doit être conduite avec rigueur et en prenant en compte les enjeux environnementaux et sociétaux afin de ne pas accentuer les inégalités.

Enfin, il est essentiel de tenir compte des enjeux sociaux et environnementaux dans la prise de décision concernant les actions à mettre en place pour améliorer le taux de rétention des employés. Il est important de rechercher des solutions qui ne compromettent pas la qualité de vie des employés, qui ne favorisent pas les inégalités sociales ou environnementales et qui respectent les droits de chacun.

Responsabilité

La responsabilité en matière d'utilisation de ces données est bel et bien un enjeu crucial, c'est pourquoi nous devons mettre en place des mécanismes appropriés pour protéger les intérêts des employés, ainsi que l'image de l'entreprise.

Le principe de responsabilité exige la mise en place de mécanismes pour garantir la responsabilité du développement, du déploiement et de l'utilisation des systèmes d'IA. Il est donc essentiel de documenter les risques potentiels pour chaque ensemble de données, ainsi que les mesures de protection à mettre en place en conséquence. Cela garantira que les employés ne subiront pas de préjudice et que l'entreprise ne sera pas confrontée à des répercussions négatives.

Il est également possible d'utiliser d'autres options pour garantir la responsabilité de l'utilisation des données, telles que la création d'un comité responsable des décisions concernant les systèmes d'IA. Ce comité serait composé de personnes compétentes et expérimentées pour prendre des décisions éthiques concernant les données des employés. Il serait également judicieux de former les employés à la compréhension des problématiques éthiques et de définir clairement les rôles de chacun en cas de crise.

En résumé, la responsabilité de l'utilisation des données est primordiale dans l'analyse des critères influençant le taux de démission des employés. Pour garantir la protection des employés et de l'entreprise, il est impératif de documenter les risques potentiels et les mesures de protection à prendre, ainsi que de mettre en place des mécanismes appropriés tels qu'un comité de décision responsable et une formation des employés. Nous sommes déterminés à respecter les normes éthiques les plus élevées dans notre travail et à garantir la sécurité et la confidentialité des données des employés.

La réflexion lors de la conception de la solution

Dans le cas présent, l'objectif est de déterminer quels sont les critères qui influent sur la démission des employés afin que l'entreprise puisse s'améliorer sur ces points.

Il est important de veiller à ce que la collecte et l'utilisation des données respectent la vie privée des employés. Il est donc nécessaire de s'assurer que toutes les données collectées sont anonymisées et que leur utilisation est strictement limitée à l'objectif de l'étude.

Il est également important de s'assurer que les données ne sont pas utilisées de manière discriminatoire. Par exemple, si certaines données, telles que l'âge ou le sexe, peuvent être pertinentes pour l'étude, leur utilisation doit être justifiée et proportionnée à l'objectif poursuivi.

Enfin, il est important de prendre en compte les conséquences potentielles de l'étude sur les employés. Par exemple, si l'étude montre que les employés qui ont des enfants ont tendance à démissionner plus souvent, cela ne doit pas conduire à une discrimination à l'embauche ou à des conditions de travail défavorables pour les employés ayant des enfants.

Pour être plus clair, lors de la conception de la solution pour déterminer les critères qui influent sur la démission des employés, il est essentiel de respecter la vie privée des employés, de ne pas utiliser les données de manière discriminatoire et de prendre en compte les conséquences potentielles de l'étude sur les employés.

Application de l'éthique lors de la réalisation de la solution

L'application de l'éthique lors de la réalisation de l'analyse des données consiste à veiller à ce que le traitement des données soit juste et équitable, tout en respectant la vie privée des employés concernés. Cela implique de prendre en compte plusieurs aspects tels que la sécurité et la confidentialité des données, la non-discrimination, la transparence, et le consentement éclairé des employés.

La sécurité et la confidentialité des données sont essentielles pour protéger les données personnelles des employés. Cela implique de mettre en place des mesures de protection adéquates pour éviter tout accès

non autorisé aux données. De plus, les résultats de l'analyse doivent être communiqués de manière anonyme, sans révéler l'identité ou les informations personnelles des employés.

La non-discrimination est également un aspect crucial de l'éthique de l'analyse des données. Il convient de garantir que l'analyse ne se base pas sur des critères discriminatoires tels que l'âge, le genre ou encore le statut matrimonial. L'analyse doit être juste et équitable pour tous les employés.

La transparence est également importante pour garantir que les employés comprennent pourquoi leurs données sont collectées et comment elles seront utilisées. Les employés doivent être informés de la finalité de l'étude, des types de données collectées et des mesures de sécurité mises en place pour protéger leurs données.

Le consentement éclairé des employés est également un aspect crucial de l'éthique de l'analyse des données. Les employés doivent être informés de la collecte et de l'utilisation de leurs données, et ils doivent donner leur consentement explicite avant que leurs données ne soient collectées et utilisées.

Dans la continuité de cette idée, la mise en application des 7 étapes de l'éthique dans l'intelligence artificielle est le point le plus important lors de la réalisation de la solution

Regard éthique lors du final de la solution

Parmi les résultats obtenus à travers la corrélation entre les attributs étudiés avec le turn-over, certains sortent du lot, et nous pouvons interpréter ces résultats de la manière suivante :

- Plus la **moyenne d'heures travaillé par jour** est élevée, plus il y a de chance d'avoir un départs.
- Beaucoup de **personnes célibataires** ont quittés l'entreprise. Etant donné leur statut c'est dernier peuvent être plus libre.
- Plus les **employés sont jeunes**, plus il y a de chance d'avoir un départ.
- Plus le **nombre d'année travaillés** est faible, plus il y a de chance d'avoir un départ.
- Plus le **nombre d'année passé avec le même manager** est faible, plus il y a de chance d'avoir un départ.
- Plus la **satisfaction au travail** est faible, plus il y a de chance d'avoir un départ.

Ces résultats sont donc très parlants et pertinents, cependant, il est important de différencier les données exploitables des données non exploitables sur le plan éthique. Pour mettre en opposition deux corrélations non exploitables de la même manière, on peut parler de la satisfaction au travail et des personnes célibataires.

En ce qui concerne la satisfaction au travail, nous pouvons simplement imaginer que si les projets dans l'entreprise permettaient de mettre en avant tous les effectifs en essayant de leur donner plus de responsabilités ou plus de confiance, la satisfaction générale est employés serait plus haute, et par conséquent, ils seraient moins amenés à démissionner pour cette raison.

A contrario, pour ce qui est du fait que les personnes célibataires sont plus amenés à démissionner, on ne peut pas prendre la décision de favoriser l'embauche des personnes en concubinage, auquel cas on serait dans un cas de discrimination des personnes célibataires.

Pour conclure, l'interprétation des données est un point primordial dans la prise de décision suite à l'obtention des résultats.