

# Livrable I.A. Éthique

BENARD Kevin

CLERC Jules

MEHRENBARGER Maxime

MONTET Thomas

17 mars 2022

MENHAJ Lamyae

CESI FISE INFO - NICE

Projet Intelligence Artificielle

# Table des matières

1) Etude de l'éthique .....	3
1.1 Introduction.....	3
1.2 Traitement de données.....	3
1.3 Choix de l'algorithme d'IA.....	3
1.4 Analyse de l'impact de l'intelligence artificielle.....	5
1.5 Questions & Réponses .....	5
2) Bibliographie .....	8
3) Glossaire .....	9

# 1) Éthique de l'intelligence artificielle

## 1.1 Introduction

Une entreprise nous a contactés, pour déterminer les facteurs ayant le plus d'influence sur le taux élevé de [turn-over<sup>1</sup>](#) dans l'entreprise, et proposer des modèles afin d'avoir des pistes d'amélioration pour donner à leurs employés l'envie de rester.

L'entreprise nous a fourni des données collectées sur l'année 2015 concernant les employés, en format CSV. Ces données, envoyées par les ressources humaines contiennent, les informations sur la dernière évaluation de chaque employé fait par son manager, sur l'enquête de qualité de vie au travail et sur les horaires de travail d'arrivée et de départ de chaque employé.

La première phase du projet a été de générer et de traiter les jeux de données. On a commencé par importer toutes les données récupérées dans des trames de données ([dataframes<sup>2</sup>](#) en anglais) pour permettre de mieux modifier, manipuler les données.

## 1.2 Traitement de données

Nous avons donc commencé par trier les données en retirant celle qui ne servait pas pour le traitement de donnée (Over18, EmployeeCount, StandardHours).

Tout d'abord 'Over18' indique si l'employé à plus de dix-huit ans, tous les salariés ayant plus de 18 ans, cette donnée n'est pas utile ; Ensuite 'EmployeeCount' informe du nombre d'employés par ligne de donnée, elle nous permet d'être sûrs que tous les employés étaient comptés dans les effectifs de 2015, mais elle n'est pas nécessaire à notre analyse ; 'StandardHours' donne le nombre d'heures standard qu'un employé doit travailler, il est identique à tous les employés, et peut ainsi être retiré de la trame de données.

Une fois les données sur les employés triées, nous avons ensuite géré les données des autres fichiers : Les horaires de travail d'arrivée et de départ, les deux fichiers sont fusionnés pour calculer la moyenne d'heures travaillées, et déterminer le nombre de jours d'absence pour chaque employé.

## 1.3 Choix de l'algorithme d'IA

En ce qui concerne, la sécurité des données via des attaques numériques, avec le choix des informations dans les jeux de données, cela nous permet d'avoir des informations moins précises au sujet de l'employé et de rester transparent sur l'utilisation des données. Tout dépend des

informations saisies dans les données, que l'humain a décidé de traiter qui pourrait menacer l'identité de l'employé.

De même, pour éviter les préjugés injustifiés, nous avons fait le choix de ne pas garder les données sur le sexe des employés dans notre [analyse de données](#)<sup>4</sup> dans l'objectif d'éviter l'exclusion, l'équité et les discriminations non-intentionnelles afin de positionner tous les employés sur un pied d'égalité.

Après avoir traité chaque donnée, nous avons simplement regroupé toutes les données sur un seul jeu de donnée pour pouvoir passer à l'implémentation des modèles. La deuxième phase du projet a été de réaliser des algorithmes d'[IA](#)<sup>3</sup> afin d'analyser les données de l'entreprise et de déterminer quelles sont les facteurs qu'influe sur le taux de turn-over dans l'entreprise.

Nous savons que nous étions dans le cas d'algorithme de [classification](#)<sup>5</sup>, grâce au graphe ci-dessous de [scikit-learn](#)<sup>6</sup>, nous pouvions voir vers quel type d'algorithme s'orienter en fonction de nos jeux de données.

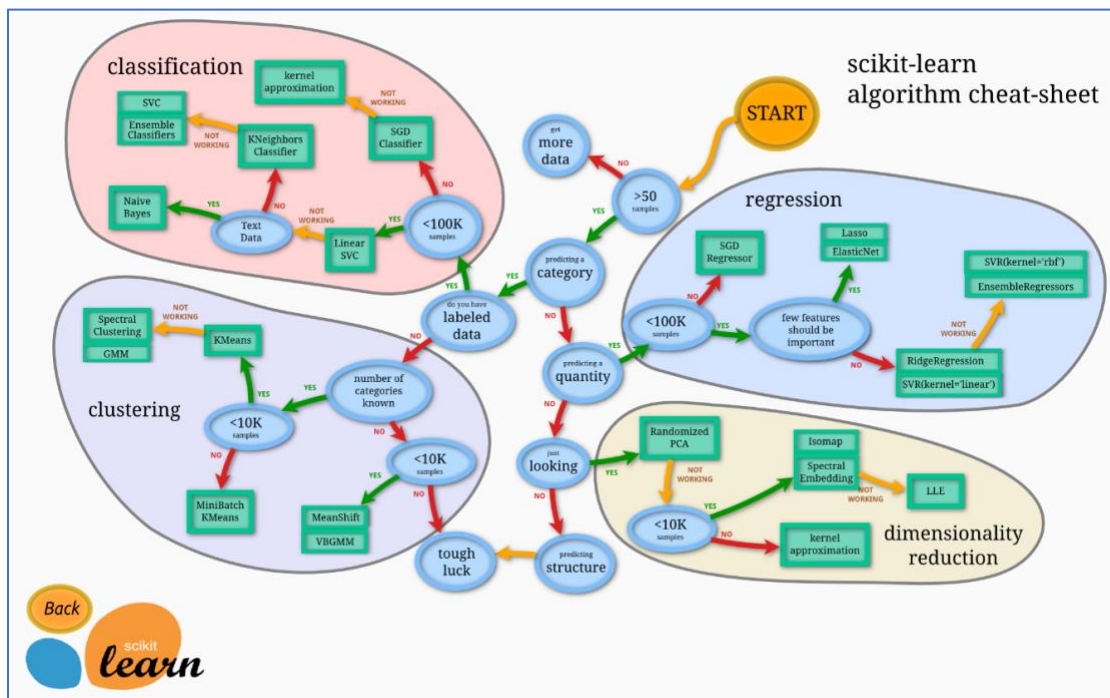


Figure 1: Carte du choix du modèle de scikit learn

## 1.4 Analyse de l'impact de l'intelligence artificielle

Pour la phase de test, nous avons donc choisi de faire une comparaison de trois modèles dans le processus de classification :

- La Régression logistique
- La C-Support Vector Classification
- Le Random Forest Classifier

En effectuant les tests avec chaque modèle, nous pouvions déterminer le nombre de personnes qui allait partir de l'entreprise en fonction des données que nous avons gardé dans nos jeux de données. Et de comparer lequel des trois algorithmes à des meilleures performances pour réaliser cette étude.

## 1.5 Questions / Réponses

### Agence humaine et autonomie

*Le système d'IA, est-il conçu pour interagir, guider ou prendre des décisions par des utilisateurs humains qui affectent les humains ou la société ?*

L'IA va simplement effectuer des prédictions sur des jeux de données afin d'analyser en détail chaque donnée qui peut avoir un impact sur la démission de l'employé, et grâce aux résultats obtenus, proposer des solutions pour y remédier. Donc l'IA n'interagira pas directement avec les humains, mais juste informé des éléments, facteurs du nombre d'attrition.

### Confidentialité et gouvernance des données

*Votre système d'IA est-il formé ou a-t-il été développé en utilisant ou en traitant des données personnelles (y compris des catégories spéciales de données personnelles) ?*

En ce qui concerne, la sécurité des données via des attaques numériques, avec le choix des informations dans les jeux de données, cela nous permet d'avoir des informations moins précises au sujet de l'employé et de rester transparent sur l'utilisation des données. Tout dépend des informations saisies dans les données, que l'humain a décidé de traiter qui pourrait menacer l'identité de l'employé. Dans notre cas, l'âge ou le sexe de l'employé qui n'ai pas nécessairement utile d'être traité et permet d'éviter de dévoiler l'identité de l'employé.

## **Diversité, non-discrimination et équité**

*Avez-vous établi une stratégie ou un ensemble de procédures pour éviter de créer ou de renforcer des biais injustes dans le système d'IA, tant en ce qui concerne l'utilisation des données d'entrée que pour la conception de l'algorithme ?*

De même, pour éviter les préjugés injustifiés, nous avons fait le choix de ne pas garder les données sur le sexe des employés dans notre analyse de données dans l'objectif d'éviter l'exclusion, l'équité et les discriminations non-intentionnelles afin de positionner tous les employés sur un pied d'égalité.

## **Bien-être sociétal et environnemental**

*Y a-t-il des impacts négatifs potentiels du système d'IA sur l'environnement ?*

Comme notre jeu de donnée possède une quantité correcte de donnée, cela n'entraîne pas d'empreinte carbone lié à la consommation d'énergie de l'IA. L'IA ne réalise pas de chose complexe comme générer ou reconnaître des mots et des phrases très proches du langage humain (NLP) qui lui peut émettre en moyenne autant de dioxyde de carbone que cinq voitures américaines durant toute leur durée de vie.

Pour donner suite à cela, avec des [matrices de corrélation](#)<sup>7</sup>, nous avons proposé des solutions pour réduire l'attrition des employés dans l'entreprise. Par exemple, 50,4 % des personnes qui partent de l'entreprise sont célibataires contre 28.5 % pour les personnes restantes dans l'entreprise. Le profil-type d'une personne souhaitant démissionner, est un jeune et nouvel employé célibataire, voyageant fréquemment et ne pouvant donc pas facilement s'intégrer à l'entreprise. Comme solution pour l'entreprise, elle peut organiser des événements entre employés de la société ou bien des speed dating afin de créer des événements conviviaux pour créer des liens entre les employés de l'entreprise.

## 2) **Bibliographie**

Scikit-learn:

<https://scikit-learn.org/stable/>

Kaggle:

<https://www.kaggle.com>

Article impact IA sur l'environnement :

<https://www.journaldunet.com/solutions/dsi/1496761-impact-de-l-ia-sur-l-environnement-demeler-le-vrai-du-faux/>

Ressource Moodle, exemple d'éthique :

<https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>



### 3) Glossaire

**Turn-over<sup>1</sup>** : Le turnover caractérise la rotation, le renouvellement des employés dans une même entreprise, par démissions, licenciements ou ruptures à l'amiable. Il se calcule en général par un taux exprimé en pourcentage.

**Dataframe<sup>2</sup>** : Un DataFrame est un ensemble de Series Pandas indexées par une valeur. Dans cet article, nous allons présenter la structure des DataFrames puis, se pencher sur ses différents attributs et méthodes de base en expliquant leur utilité et leur fonctionnement.

**Intelligence artificielle<sup>3</sup>** : L'IA désigne la possibilité pour une machine de reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité.

**Analyse des données<sup>4</sup>** : L'analyse des données est une famille de méthodes statistiques dont les principales caractéristiques sont d'être multidimensionnelles et descriptives. Dans l'acception française, la terminologie « analyse des données » désigne donc un sous-ensemble de ce qui est appelé plus généralement la statistique multivariée.

**Classification<sup>5</sup>** : Les algorithmes de classification sont des algorithmes de tri qui permettent de choisir lequel est le plus adapté au problème traité tout en tenant compte des contraintes imposées.

**Scikit-learn<sup>6</sup>** : Scikit-learn est la principale bibliothèque d'outils dédiés au machine learning et à la data-science dans l'univers Python.

**Matrice de corrélation<sup>7</sup>** : Une matrice de corrélation est un tableau indiquant les coefficients de connexion entre les facteurs. Chaque cellule du tableau indique le lien entre les deux facteurs. Une grille de connexion est utilisée pour présenter des informations, comme contribution à une enquête plus approfondie, et comme indication pour des examens de pointe.