



Travaux pratiques de Machine Learning: Prédire la classification des personnes malades et non malade en choisissant un modèle



Introduction

L'apprentissage automatique (Machine Learning) est un sous-domaine de l'intelligence artificielle qui permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés. Il utilise des algorithmes pour trouver des modèles dans les données et faire des prédictions ou prendre des décisions sans intervention humaine.

Problématique

La prédiction de la classification est un problème courant en apprentissage automatique. L'objectif est de prédire la classe d'une observation en utilisant un modèle entraîné sur des données étiquetées. Il existe plusieurs défis et problèmes liés à la prédiction de la classification, tels que:

- **Le déséquilibre des classes:** lorsque les classes ne sont pas équilibrées dans les données d'entraînement, cela peut entraîner un modèle biaisé qui prédit mieux la classe majoritaire.
- **Le surapprentissage:** lorsque le modèle est trop complexe et mémorise les données d'entraînement au lieu de généraliser à partir d'elles, il peut avoir de bonnes performances sur les données d'entraînement mais de mauvaises performances sur les données de test.
- **La sélection des caractéristiques:** le choix des caractéristiques à utiliser pour entraîner le modèle peut avoir un impact important sur ses performances. L'utilisation de caractéristiques non pertinentes ou redondantes peut réduire l'efficacité du modèle.
- **La validation croisée:** pour évaluer les performances du modèle de manière fiable, il est important d'utiliser des techniques telles que la validation croisée pour estimer l'erreur de généralisation du modèle.

Objectif

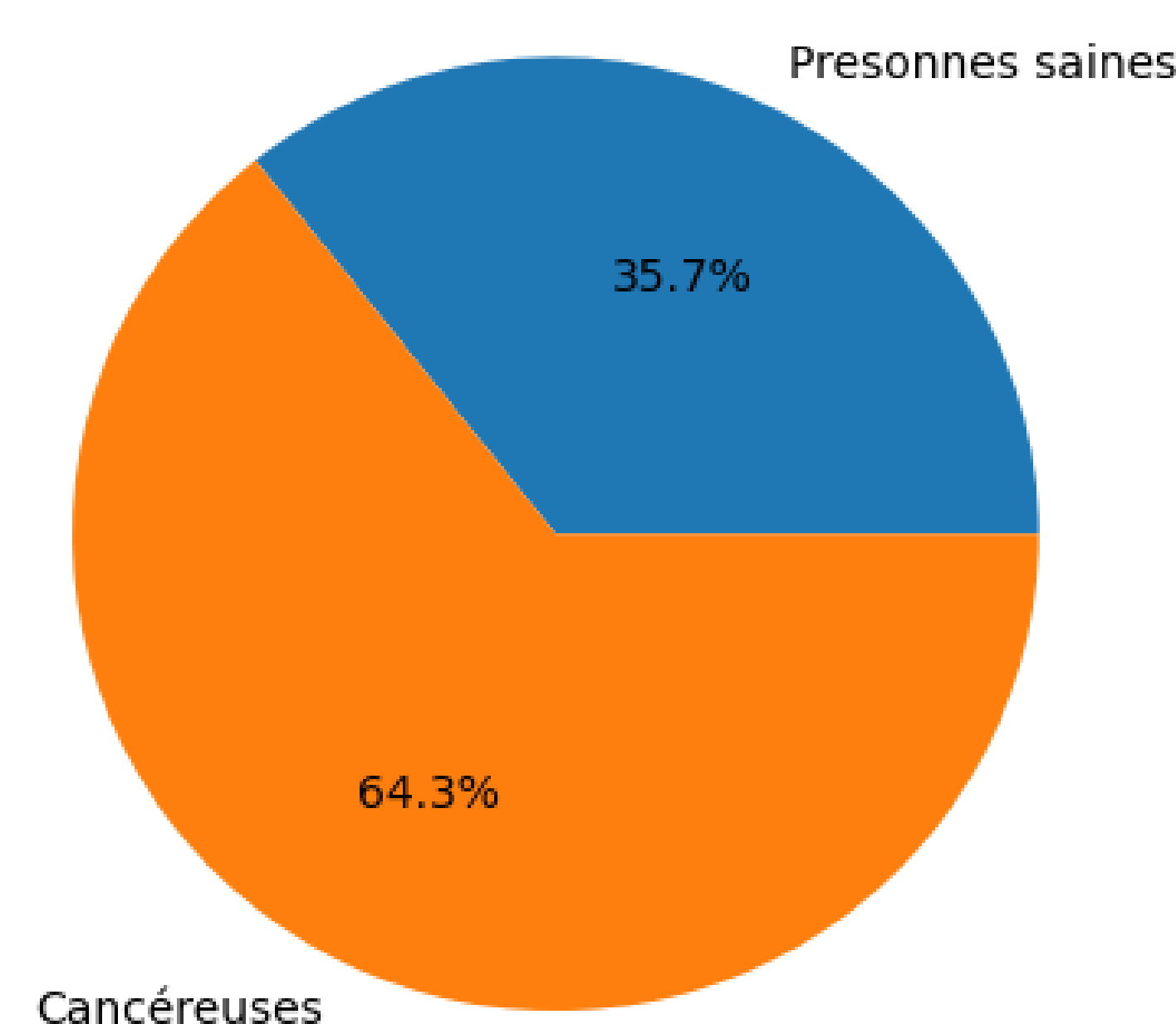
Les objectifs de la prédiction de la classification des personnes atteintes du cancer et celles qui ne sont pas atteintes peuvent varier en fonction du contexte, on peut citer quelques un comme:

1. **Améliorer le diagnostic des patients:** L'un des principaux objectifs de la prédiction de la classification des personnes atteintes du cancer et celles qui ne sont pas atteintes est d'aider les médecins à diagnostiquer les patients plus rapidement et avec plus de précision. En utilisant un modèle d'apprentissage automatique entraîné sur des données historiques, les médecins peuvent obtenir une prédiction rapide et fiable de l'état de santé d'un patient en se basant sur ses caractéristiques médicales.
2. **Identifier les facteurs de risque:** La prédiction de la classification des personnes atteintes du cancer et celles qui ne sont pas atteintes peut également être utilisée pour identifier les facteurs de risque associés à la maladie. En analysant les caractéristiques des patients qui sont prédits comme étant atteints du cancer, les chercheurs peuvent identifier les facteurs qui augmentent le risque de développer la maladie. Cela peut aider à développer des stratégies de prévention pour réduire l'incidence du cancer.
3. **Améliorer la gestion des soins de santé:** La prédiction de la classification des personnes atteintes du cancer et celles qui ne sont pas atteintes peut également être utilisée pour améliorer la gestion des soins de santé. En identifiant les patients à haut risque de développer un cancer, les hôpitaux et les cliniques peuvent mettre en place des programmes de dépistage et de prévention pour réduire l'incidence de la maladie. Cela peut également aider à réduire les coûts des soins de santé en évitant des traitements coûteux pour les patients atteints du cancer.

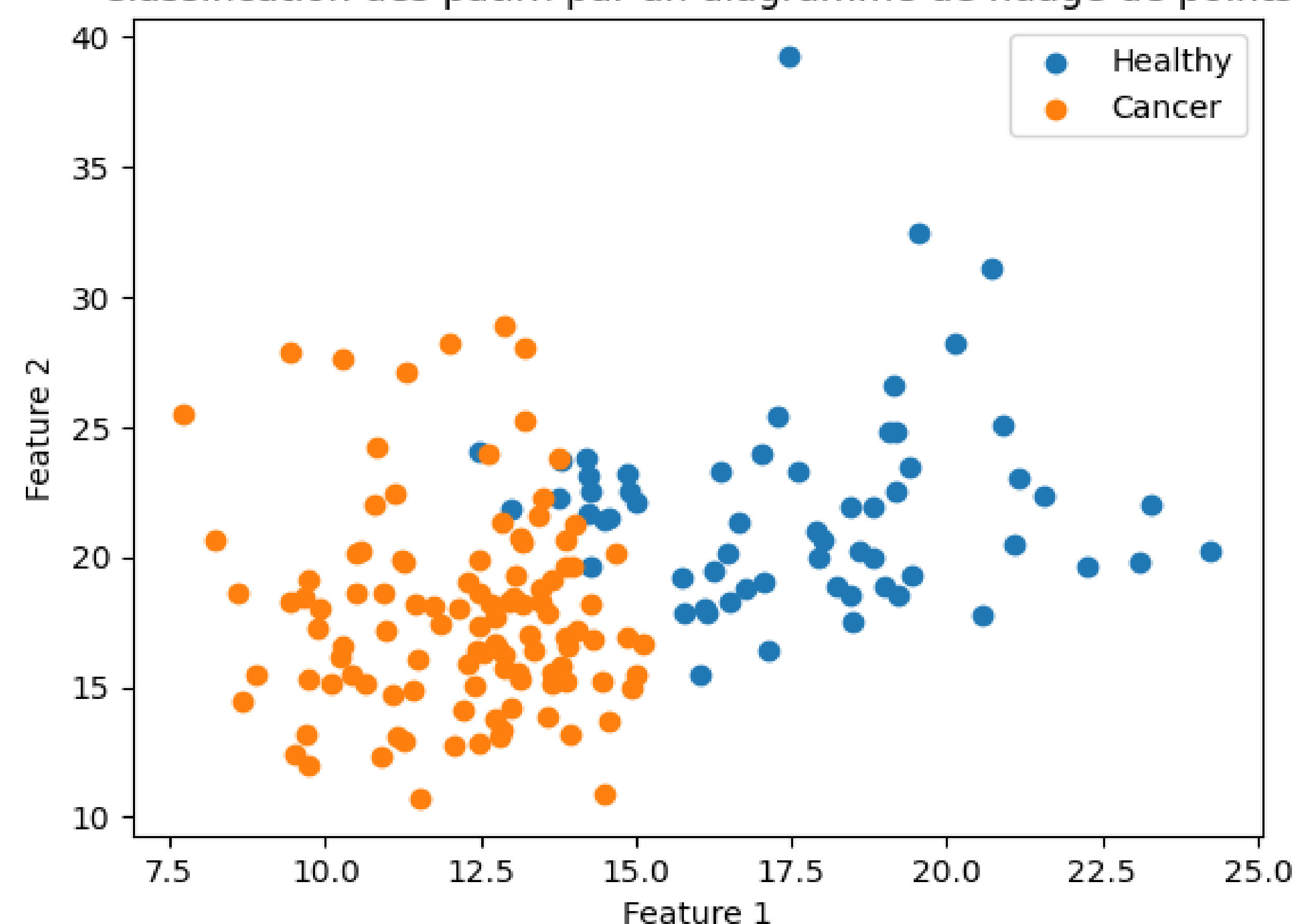
Cet en ce sens que nous nous basons sur l'apprentissage automatique pour prédire la classification des personnes cancéreuses et non cancéreuses. Cela permettrait au différents parties de prendre une décision vis-à-vis de résultat obtenus.

Plus détaillé dans la partie conception

Classification des pertients sous forme d'un diagramme circulaire



Classification des patirnn par un diagramme de nuage de points



Conclusion

La prédiction de classification des personnes cancéreuses et non cancéreuses est un domaine important de la recherche en apprentissage automatique qui peut avoir un impact significatif sur la santé des patients. En utilisant des données médicales pour entraîner des modèles fiables, il est possible d'aider les médecins à diagnostiquer le cancer plus tôt et à commencer le traitement plus rapidement. Cependant, il existe de nombreux défis et problèmes liés à la prédiction de la classification des personnes cancéreuses et non cancéreuses, et il est important d'utiliser des méthodes rigoureuses pour évaluer les performances des modèles avant de les utiliser en pratique.

Réalisé par:

1. YASSINE ZAKARIA IBETH
2. FOURISSOU ABDELKERIM LONA WINAMOU
3. MIMBETINAN DJIMTEBAYE SUCCES
4. MBAIHONDOUM FRANCIS
5. BAYTA BOURKOU ELIE