

Problématique

- Comment concevoir et développer une solution informatique efficace et précise pour le diagnostic médical automatique du cancer du sein, en utilisant les langages de programmation C sharp
- Comment surmonter les défis liés à la variabilité des données médicales, à la précision des diagnostics et à l'interopérabilité avec les systèmes de santé existants, afin d'améliorer les taux de détection, les traitements et les résultats pour les patients atteints de cancer [1]

Objectifs

L'objectif principal est de permettre aux étudiants de se familiariser avec les concepts de la programmation orientée objet, de la structure des données, de la gestion des fichiers et des différents algorithmes à travers le développement d'une application complète en C. Cette application aura pour but de :

- Concevoir un algorithme de classification performant qui minimise les erreurs de diagnostic et maximiser la précision du système.
- Réaliser un système automatisé qui réduit la dépendance à l'égard de l'interprétation humaine et permet un traitement plus rapide des résultats.[2]

Technologies à utiliser

Dans le cadre de notre projet, nous avons utilisé les technologies suivantes :

- Le langage de programmation C avec des notions de programmation orientée objet
- CsvHelper : Elle permet la gestion des fichiers CSV.
- Microsoft Visual Studio Community qui servira d'environnement de développement intégré
- Entity Framework Core : C'est un ORM pour la gestion des données
- Windows Presentation Foundation ou WPF : Pour la Création d'interfaces graphiques.[3]

Méthodologie suivie

- Création d'une instance du modèle d'arbre de décision : initialiser un modèle d'arbre de décision à partir de la bibliothèque ou du module que nous utilisons dans notre environnement de programmation. Cela crée une structure de base pour notre modèle, prête à être entraînée sur nos données.
- Entraînement du modèle sur les caractéristiques sélectionnées : Après l'entraînement, on évalue la performance du modèle sur un ensemble de données inconnu appelé ensemble de test en lui fournissant les caractéristiques de cet ensemble et en observant ses prédictions.
- Prédiction sur l'ensemble de test : on évalue la performance du modèle sur un ensemble de test en fournissant ses caractéristiques.
- Évaluation des performances du modèle : on compare ses prédictions sur l'ensemble de test aux données réelles obtenant la précision.[1]

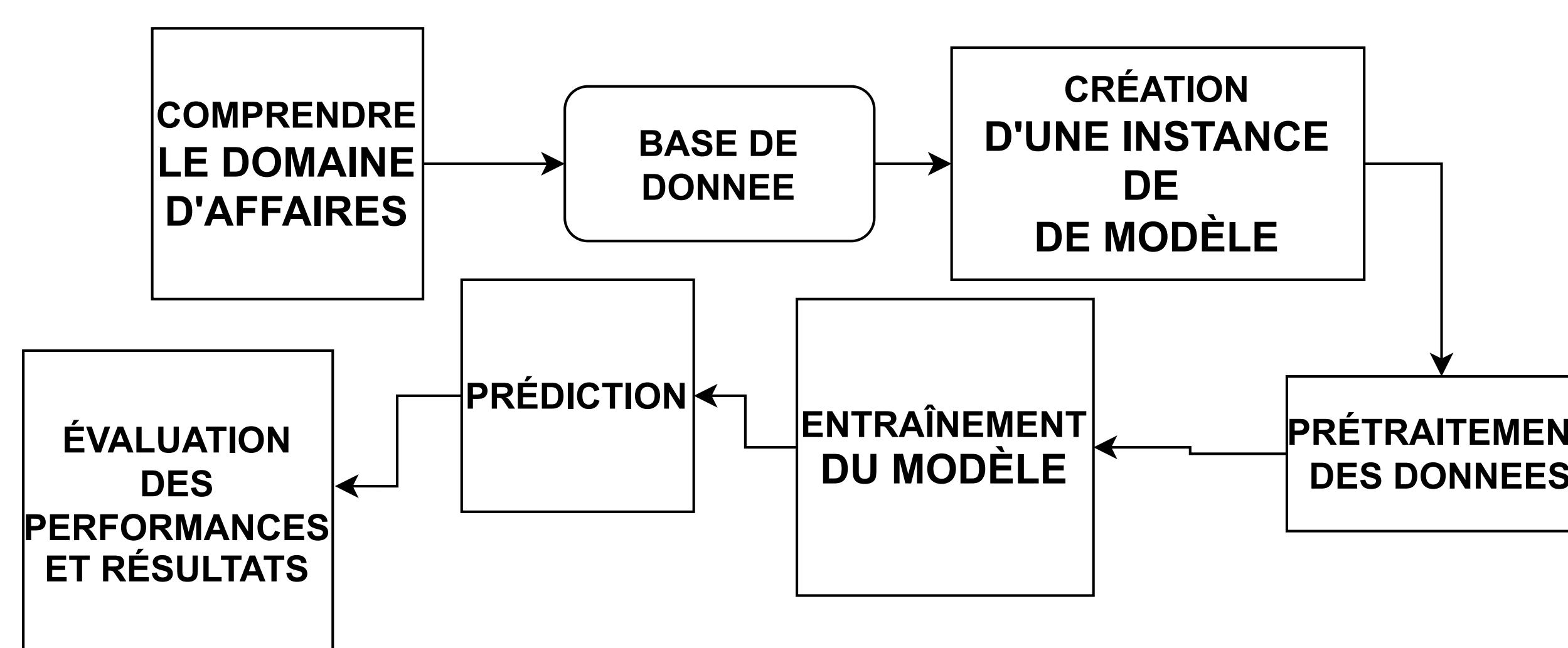


FIGURE 1 – Description du système.

Interface graphique

la figure d'accueil de notre interface graphique est la suivante :



Evaluation

- la base de donnée est un ensemble de caractéristique d'un ensemble de données , potentiellement utilisées dans le cadre d'une analyse prédictif.
- ses caractéristique sont : RaduisWorst,areaWorst ,perimeterWorst,concavePointWorst,concavepointMean...
- Les métriques utilisées pour les mesures de performances sont les suivantes : le taux de reconnaissance (accuracy), la précision, le rappel (recall) et le score F1.

Résultats

representation des resultats obtenue avec notre arbre de decision :

méthodes	accuracy	precision	recall	F1-Score
Resultats	0.94	0.66	1	0.77

Conclusion

l'application permettra d'améliorer la précision des diagnostics, de réduire les erreurs humaines et d'accélérer le processus de diagnostic. Les résultats préliminaires sont encourageants, mais peuvent être améliorés,

- En perspective, On envisage d'utiliser d'autres bases de données afin de valider et confirmer les performances du système proposé

Références

- [1] PALMA Giovanni. Détection automatique des opacités en tomosynthèse numérique du sein. *Thèse de doctorat. Télécom ParisTech*, 2010.
- [2] Mangasarian Olvi Street Nick Wolberg, William and W. Street. *Breast Cancer Wisconsin (Diagnostic)*, chapter DOI : <https://doi.org/10.24432/C5DW2B>. UCI Machine Learning Repository, 1995.
- [3] Yacine Yaddaden. *Creation d'une application console pour le diagnostic médical. TP01*, 2024.