

République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche Scientifique
Université de Mohamed El Bachir El Ibrahimi de Bordj Bou Arréridj
Faculté des Mathématiques et d'Informatique
Département d'informatique



MEMOIRE

Présenté en vue de l'obtention du diplôme

Master en informatique

Spécialité : Ajoutez votre spécialité ici

THEME

Classification Des IRM Cérébrales Pathologique Avec SVM Optimisé

Présenté par :

AZZA MOHAMMED ELAMIN

MEKKAS LYES

Soutenu publiquement le : jr/mm/aaaa

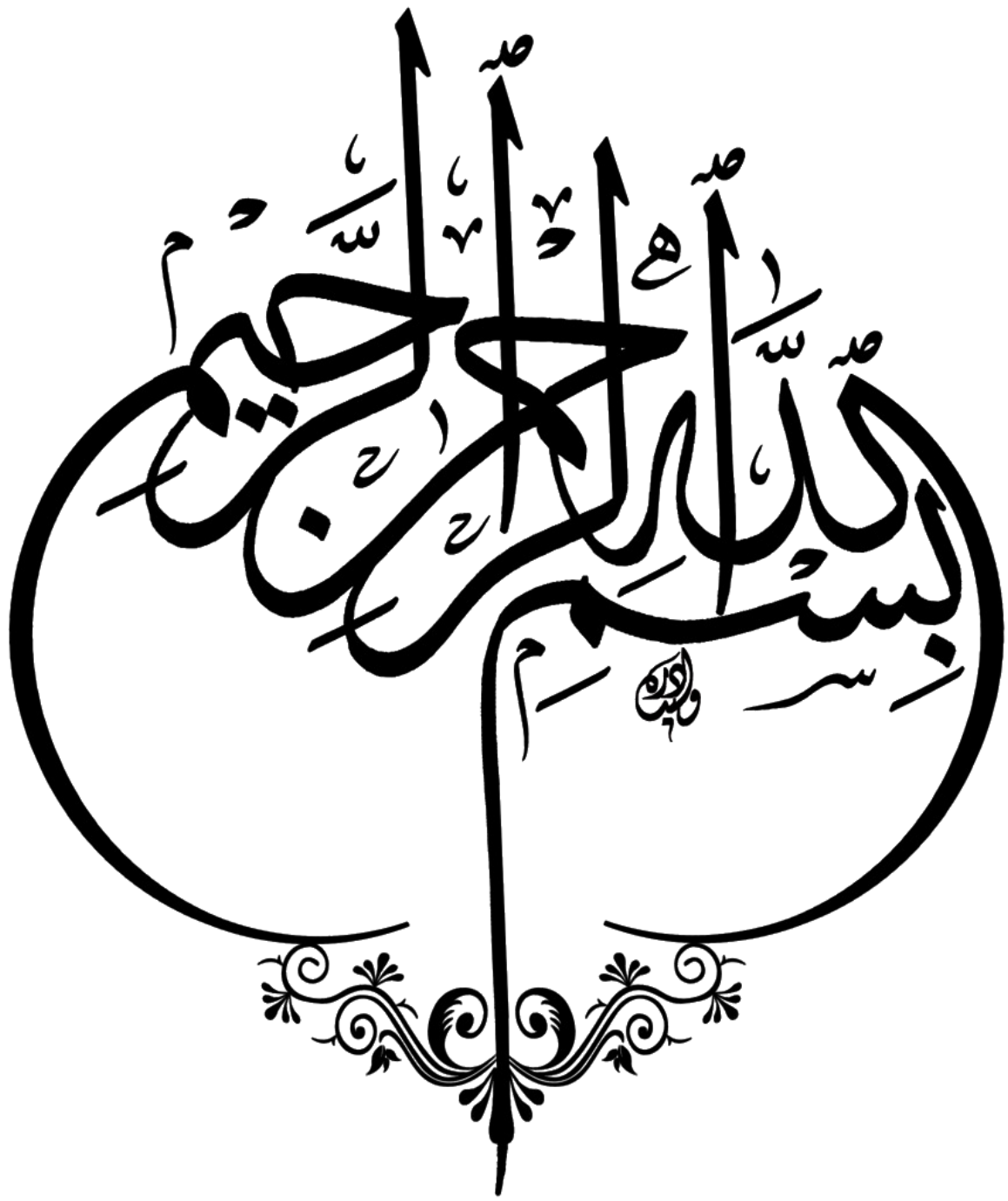
Devant le jury composé de :

Président : Hadj Barkat

Examineur : Benabid Sonia

Encadreur : Chellakh Hafida

2022/2023



Dédicace

Je souhaite dédier cette mémoire à vous, ma famille chérie, qui avez été mes piliers tout au long de ce parcours académique.

À mes chers parents, mon frère et ma sœur,

C'est avec une immense gratitude que je dédie cette mémoire à vous, ma famille bien-aimée. Votre soutien indéfectible, votre amour inconditionnel et votre présence constante ont été les fondements de mon parcours académique.

À mes merveilleux parents

Vous êtes mes héros. Votre dévouement, vos sacrifices et votre persévérance ont été les moteurs qui m'ont poussé à aller de l'avant. Votre confiance en moi m'a donné la force de surmonter les obstacles et d'atteindre mes objectifs. Votre amour incommensurable a illuminé chaque pas de ma vie.

Mon cher frère et ma chère sœur

Vous êtes mes compagnons de vie et mes meilleurs amis. Votre présence joyeuse, votre encouragement constant et votre soutien inconditionnel m'ont donné la force de me dépasser. Nos liens indéfectibles sont une source infinie de bonheur et de réconfort.

Cette dédicace est un témoignage de ma reconnaissance éternelle envers vous tous. Votre foi en moi a été le moteur de ma réussite. Votre amour et votre soutien inébranlables ont été les piliers sur lesquels j'ai construit ma confiance et ma détermination.

Puissent ces mots exprimer toute l'affection, l'admiration et la reconnaissance que je ressens à votre égard. Vous êtes ma famille, ma source d'inspiration et ma raison d'être. C'est avec une profonde gratitude que je célèbre cette réalisation avec vous.

Avec tout mon amour et ma reconnaissance,

AZZA MOHAMMED ELAMIN

Dédicace

Je dédie humblement ce travail et exprime ma profonde gratitude :

À ma mère bien-aimée

Malgré ton absence, ton amour et ton soutien continuent de m'entourer. Ta présence dans ma vie a façonné ma personne et m'a guidé dans mes choix. Je ne pourrai jamais assez te remercier pour tout ce que tu as fait. Que Dieu te garde dans Sa miséricorde.

À mon père bien-aimé

Tu as toujours été là pour me soutenir et m'encourager. Ce travail est un humble marque de ma gratitude et de mon affection envers toi. Que Dieu te garde.

À mes frères, Farouk, Walid, et mes sœurs bien-aimés, Que Dieu vous préserve.

À toute mes meilleurs amis

Je suis honoré(e) d'avoir votre amitié. Vous avez été des piliers dans ma vie.

À toute ma famille

Qu'elle soit proche ou éloignée, Que Dieu vous accorde santé, bonheur, courage et surtout réussite.

À tous ceux que j'aime et qui me sont chers, Votre présence dans ma vie a été une bénédiction. Que Dieu vous protège et vous guide.

Avec tout mon amour et ma reconnaissance.

Mekkas Lyes

Remerciement

Nous tenons tout d'abord à exprimer notre profonde gratitude envers **Allah, le Tout-Puissant et Miséricordieux**, pour nous avoir accordé la force, le courage et la patience nécessaires pour mener à bien ce modeste travail.

Nous souhaitons également exprimer notre sincère reconnaissance envers notre Encadrante, **Mme CHELLAKH Hafida**, pour ses précieux conseils, son assistance inestimable tout au long de cette période de travail et pour son soutien constant et ses orientations.

Nos remerciements vont également à l'ensemble des membres honorables du jury qui ont accepté d'évaluer ce travail avec bienveillance.

Nous tenons à remercier chaleureusement nos parents et nos proches pour leur amour, leur soutien constant et leur encouragement indéfectible tout au long de notre parcours académique.

Enfin, nous souhaitons exprimer notre gratitude à nos amis pour leur présence et leur soutien précieux.

Résumé :

L'imagerie par résonance magnétique (IRM) est largement utilisée pour l'évaluation des tumeurs cérébrales, fournissant des images détaillées de la structure et de l'activité du cerveau. Les techniques d'intelligence artificielle (IA) et d'apprentissage automatique, telles que les machines à vecteurs de support (SVM) et les caractéristiques Histogramme des Gradients Orientés (HOG), ont montré leur efficacité dans l'amélioration de la détection des tumeurs. Cette recherche explore des méthodes avancées, notamment l'optimisation des SVM à l'aide de l'optimisation par essaim de particules (PSO) et l'extraction optimisée des caractéristiques HOG. L'objectif est de développer un outil automatisé pour la détection précise et efficace des tumeurs cérébrales. L'intégration de l'optimisation des SVM basée sur le PSO et de l'extraction des caractéristiques HOG raffinée vise à améliorer la précision diagnostique et les résultats des patients dans la prise en charge des tumeurs cérébrales.

Mots-clefs : IRM; tumeurs cérébrales; extraction de caractéristiques; optimisation; classification

Abstract:

Magnetic Resonance Imaging (MRI) is widely used for brain tumor evaluation, providing detailed images of brain structure and activity. AI and machine learning techniques, such as Support Vector Machines (SVM) and Histogram of Oriented Gradients (HOG) features, have shown promise in improving tumor detection. This research explores advanced methods, including SVM optimization using Particle Swarm Optimization (PSO) and optimized HOG feature extraction. The goal is to develop an automated tool for accurate and efficient brain tumor detection. The integration of PSO-based SVM optimization and refined HOG feature extraction aims to enhance diagnostic accuracy and improve patient outcomes in brain tumor detection.

Keywords : MRI ; brain tumors; feature extraction; optimization; classification.

ملخص :

يستخدم التصوير بالرنين المغناطيسي (MRI) على نطاق واسع لتقييم ورم الدماغ، مما يوفر صوراً مفصلة لبنية الدماغ ونشاطه. أظهرت تقنيات الذكاء الاصطناعي والتعلم الآلي، مثل ميزات آلات المتجهات الداعمة (SVM) وميزات الرسم البياني للتدرجات الموجهة (HOG) أنها تقنيات واعدة في تحسين اكتشاف وتحديد الأورام الدماغية. يستكشف هذا البحث الأساليب المتقدمة، بما في ذلك تحسين SVM باستخدام تحسين حشد الجسيمات (PSO) واستخراج ميزة HOG المحسّن. الهدف هو تطوير أداة آلية للكشف الدقيق والفعال عن ورم الدماغ. يهدف دمج تحسين SVM القائم على PSO واستخراج ميزة HOG المكررة إلى تعزيز دقة التشخيص وتحسين نتائج المرضى في تشخيص وتحديد الأورام الدماغية.

الكلمات الرئيسية: MRI؛ أورام الدماغ؛ ميزة استخراج؛ تحسين؛ تصنيف.

Table des matières

Liste des abréviations	i
Liste des figures.....	ii
Liste des tableaux.....	v
Introduction générale :.....	1
1. Chapitre 1 : Notion D'étude.....	4
1.1 Introduction :	4
1.2 Le cerveau :	4
1.2.1 Anatomie cérébrale :.....	5
1.2.2 Le liquide céphalo rachidien (LCR) :.....	6
1.2.3 La matière grise :.....	6
1.2.4 La matière blanche :	6
1.3 Tumeurs :.....	6
1.3.1 Les types de Tumeurs :.....	7
1.3.2 Les tumeurs cérébrales :.....	7
1.3.3 Les Tumeurs cérébrales bénignes :	8
1.3.4 Tumeurs cérébrales malignes :	9
1.4 Imagerie par résonance magnétique (IRM)	10
1.4.2 La détection d'une tumeur cérébrale par l'IRM :	12
1.5 Intelligence Artificielle (AI) :.....	12
1.5.1 Les domaines de l'intelligence artificielle :	13
1.5.2 L'intelligence artificielle dans le secteur de santé :.....	14

1.6 Machine Learning (apprentissage automatique) :	15
1.6.1 Le Principe général du fonctionnement de machine Learning :	15
1.6.2 Les avantages de Machine Learning :	16
1.6.3 L'utilité de la machine Learning :	17
1.7 Le Deep Learning (apprentissage profond) :	17
1.7.1 Le fonctionnement de Deep Learning :	17
1.7.2 Caractéristiques de Le Deep Learning :	18
1.7.3 Les points forts et faibles de Le Deep Learning :	19
1.8 La différence entre Le Machine Learning et Le Deep Learning :	20
1.9 Conclusion :	22
2. Chapitre 2 : La classification	24
2.1 Introduction :	24
2.2 La classification :	24
2.2.1 Apprentissage supervisé :	25
2.2.1.1 La Classification et de la Régression :	26
2.2.2 Apprentissage non supervisé :	27
2.2.3 Apprentissage semi-supervisé :	28
2.2.4 Apprentissage par renforcement :	28
2.3 La classification d'images :	29
2.3.1 Classification des images et l'apprentissage machine :	29
2.3.2 Les étapes principales du processus de classification des images :	30
2.4 HOG (Histogram of Oriented Gradients) :	31
2.4.2 Les principales caractéristiques de l'algorithme HOG :	33
2.4.3 Les avantages de l'algorithme HOG :	33
2.4.4 Les paramètres de HOG :	34
2.4.4.1 Orientation (nombre de bins d'orientation) :	34
2.4.4.2 Cellules par bloc (cells per block) :	35
2.4.4.3 Pixels par cellule (pixels per cell) :	35

2.5 L'optimisation de HOG :	36
2.6 L'analyse en composantes principales du noyau (Kernel-PCA) :	37
2.7 K-Nearest Neighbors (KNN ou méthode des k plus proches voisins) :	38
2.7.1 Le Principe général de KNN :	39
2.7.2 Fonctionnement de KNN :	39
2.7.3 Les avantages et les inconvénients de l'algorithme KNN :	40
2.8 Machine à vecteurs de support (SVM) :	41
2.8.1 Fonctionnement de SVM :	42
2.8.2 Notions de base (l'hyperplan, la marge, Les vecteurs de support, les classes) :	43
2.8.2.1 L'hyperplan :	43
2.8.2.2 La marge :	44
2.8.2.3 Les vecteurs de support :	44
2.8.2.4 Les classes :	45
2.8.3 Le problème de discrimination linéaire :	45
2.8.3.1 La marge d'un classifieur :	46
2.8.3.2 Cas séparable :	46
2.8.3.3 Cas non-séparable :	49
2.8.4 Le cas non linéaire :	50
2.8.4.1 L'astuce du noyau :	50
2.8.4.2 Fonctions du noyau :	51
2.8.6 Avantages des machines à vecteurs de support :	53
2.8.7 Inconvénients des machines à vecteurs de support :	54
2.9 L'optimisation par essaims de particules (PSO) :	54
2.9.1 Les paramètres de PSO :	55
2.9.2 Les avantages de L'optimisation par essaims de particules (PSO) :	56
2.9.3 SVM et PSO :	57
2.10 L'évaluation de performance :	58
2.10.1 Matrice de confusion :	59
2.10.2 La surface sous la courbe ROC (AUC, Area Under the Curve) :	61

2.11 Conclusion :	63
3. Chapitre 3 : Résultats et Analyse	65
3.1 Introduction :	65
3.2 Matériel :	65
3.3 Environnement de travaille :	65
3.3.1 Python.....	65
3.3.2 Jupyter Notebook :	66
3.3.3 Les bibliothèques	67
3.4 Les bases des données utilisées :	67
3.4.1 Base des données 1 :	67
3.4.2 Base des données 2 :	68
3.5 Les Modèles et résultats :	68
3.5.1 Modèle 1 Extracteur HOG avec les Classificateurs SVM et KNN :	68
3.5.2 Modèle 2 Extracteur HOG et Kernel-PCA avec les Classificateurs SVM et KNN:.....	70
3.5.4 Modèle 4 Extracteur HOG et Kernel-PCA avec Classificateurs SVM (optimisé par PSO) et KNN :	74
3.5.5 Modèle 5 Extracteur HOG optimisé avec les Classificateurs SVM et KNN :	75
3.5.6 Modèle 6 Extracteur HOG optimisé avec Classificateurs (SVM optimisé par PSO) et KNN :	76
3.6 Résultat :	79
3.6.1 Modèle 1 : Extracteur HOG avec les Classificateurs SVM et KNN :.....	79
3.6.2 Modèle 2 : Extracteur HOG et Kernel-PCA avec les Classificateurs SVM et KNN :	81
3.6.3 Modèle 3 : Extracteur HOG avec Classificateurs SVM (optimisé par PSO) KNN :	83
3.6.4 Modèle 4 : Extracteur HOG et Kernel-PCA avec Classificateurs SVM optimisé par PSO et KNN :	86
3.6.5 Modèle 5 : Extracteur HOG optimisé avec les Classificateurs SVM et KNN :..	89

3.6.6 Modèle 6 : Extracteur HOG optimisé avec Classificateurs SVM optimisé par PSO et KNN :	91
3.7 Comparaison des résultats entre les six modèles :	94
3.7.1 Comparaison des résultats entre les modèles de SVM et les modèles de SVM optimisé par PSO :	95
3.7.2 Comparaison des résultats entre les modèles de HOG et les modèles de HOG optimisé :	96
3.8 Conclusion :	98
Conclusion générale :	100
Les Références :	101

Liste des abréviations

AI Artificial Intelligence

HOG Histogram of Oriented Gradients

IRM Magnetic Resonance Imaging

KNN K-Nearest Neighbors

ML Machine Learning

PCA Principal Component Analysis

PSO Particle Swarm Optimization

RBF Radial Basis Function

ROC Receiver Operating Characteristic

AUC Area Under Curve

SVC Support Vector Classification

NU-SVC Non-Uniform Support Vector Classification

ACC Accuracy (exactitude)

Poly polynomiale

SVM Support Vector Machines

TN True Negative

TP True Positive

FN False Negative

FP False Positive

Liste des figures

Figure 1: Anatomie horizontale d'un cerveau humain [06]	5
Figure 2: Anatomie verticale d'un cerveau humain [06]	5
Figure 3: IRM d'une tumeur cérébrale [08]	8
Figure 4: Scanner IRM [14]	11
Figure 5: l'intelligence artificielle et ses sous-domaines	14
Figure 6: la différence de Machine Learning et Deep Learning [27]	22
Figure 7: le processus de classification	25
Figure 8: la différence entre la régression et la classification [28]	26
Figure 9: les étapes de HOG	33
Figure 10: Le paramètre de l'orientation du HOG [31]	35
Figure 11 : la distribution des pixels dans une cellule d'un bloc [32]	36
Figure 12 : le processus de classification d'images	38
Figure 13 : le fonctionnement de KNN [36]	40
Figure 14 : fonctionnement de SVM	42
Figure 15 : kernel function	43
Figure 16 : la marge est la distance	44
Figure 17 : les vecteurs de support	45
Figure 18 : les types de problèmes de discrimination binaire	46
Figure 19 : Marge maximale pour séparer deux classes	47
Figure 20 : noyau linéaire [43]	51

Figure 21 : noyau polynomial degré 3 [43]	51
Figure 22 : radial basis function [43]	52
Figure 23 : Noyau Tangent Hyperbolique [43].....	53
Figure 24 : La structure de fonctionnement de PSO avec le SVM.....	58
Figure 25 : matrice de confusion	60
Figure 26 : exemple de courbe ROC (AUC) [55]	63
Figure 27 : Logo python [56].....	66
Figure 28: Logo jupyter [57].....	67
Figure 29: l'architecture de model-1.....	70
Figure 30: architecture de model-2	72
Figure 31 :l'architecture de model-3	74
Figure 32: architecture de model-4	75
Figure 33: Architecture de model-5	76
Figure 34: architecture de model-6	78
Figure 35 : matrice de confusion et le corbe Roc de HOG avec C-SVC Linear pour dataset-1.....	79
Figure 36 : matrice de confusion et le courbe Roc de HOG avec C-SVC linear pour dataset-2.....	80
Figure 37 : matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC poly pour dataset-1	81
Figure 38: matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC RBF pour dataset-2.....	82

Figure 39 : matrice de confusion et le courbe Roc de HOG avec NU-SVC poly pour dataset-1	85
Figure 40 : matrice de confusion et le courbe Roc de HOG avec C-SVC RBF pour dataset-2.....	85
Figure 41 : matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC Linear pour dataset-1	87
Figure 42: matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC Linear pour dataset-2.....	88
Figure 43: matrice de confusion et le courbe Roc de HOG optimisé avec C-SVC Poly pour dataset-1.....	89
Figure 44 : matrice de confusion et le courbe Roc de HOG optimisé avec C-SVC Linear pour dataset-2.....	90
Figure 45 : matrice de confusion et le courbe Roc de HOG optimisé avec NU-SVC Poly pour dataset-1.....	92
Figure 46: matrice de confusion et le courbe Roc de HOG optimisé avec C-SVC Poly pour dataset-1.....	93
Figure 47 : Un diagramme à barres indique la meilleure accuracy entre les modèles...	94

Liste des tableaux

Tableau 1: les points forts et les points faibles de DL.....	19
Tableau 2: Tableaux des paramètres de PSO.....	79
Tableau 3 :Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracture HOG pour dataset-1.....	79
Tableau 4: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracture HOG pour dataset-2.....	80
Tableau 5: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG et Kernel-PCA pour dataset-1	81
Tableau 6: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG et Kernel-PCA pour dataset-2	82
Tableau 7: les Paramètres choisis par PSO pour le classificateur SVM	83
Tableau 8: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG pour dataset-1	84
Tableau 9: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG pour dataset-2	85
Tableau 10: les paramètres choisis par PSO	86
Tableau 11: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG et Kernel-PSA pour dataset-1.....	87
Tableau 12: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG et Kernel-PSA pour dataset-2.....	88
Tableau 13: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG optimisé pour dataset-1	89

Tableau 14: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG optimisé pour dataset-2	90
Tableau 15: les paramètres choisis par PSO	91
Tableau 16: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG optimisé pour dataset-1.....	92
Tableau 17: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG optimisé pour dataset-2.....	93
Tableau 18 :Tableaux de comparaison entre les modèles avec SVM et les modèle SVM-PSO.....	96
Tableau 19 : Tableaux de comparaison entre les modèles avec HOG et les modèle avec HOG optimisé.....	97

INTRODUCTION GÉNÉRALE

Introduction générale :

Contexte :

L'imagerie par résonance magnétique (IRM) est une technique d'imagerie médicale largement utilisée pour l'évaluation des tumeurs cérébrales. Offrant des images détaillées pour une détection précoce et une évaluation précise. Son utilisation combinant des champs magnétiques et des ondes radio est cruciale pour la prise en charge des tumeurs cérébrales, permettant aux professionnels de la santé de prendre des décisions éclairées pour des soins optimaux. [01] Les tumeurs cérébrales posent des défis majeurs en termes de diagnostic et de traitement. Une détection précoce et précise est cruciale pour améliorer les résultats du traitement et la qualité de vie des patients. Les techniques avancées de classification d'images permettent d'analyser et d'interpréter les caractéristiques visuelles extraites des images, facilitant ainsi une détection précise des tumeurs cérébrales. [02] Les SVM (machines à vecteurs de support) sont couramment utilisées pour la classification d'images médicales en raison de leur capacité à séparer efficacement des ensembles de données complexes. L'HOG (histogramme des gradients orientés) est une méthode bien établie pour extraire des caractéristiques visuelles pertinentes à partir d'images, ce qui facilite l'identification et la classification des tumeurs cérébrales en capturant des informations sur leur forme et leur texture. [03]

La combinaison des SVM et des caractéristiques HOG permet de développer un système de détection précis et efficace des tumeurs cérébrales. Cette approche exploite les avantages des SVM en termes de classification robuste et des caractéristiques HOG en termes de représentation visuelle des motifs. Cette recherche vise à fournir un outil automatisé de détection des tumeurs cérébrales, améliorant ainsi la précision, la rapidité et l'efficacité du diagnostic.

Problématique :

La recherche se concentre sur l'optimisation de l'utilisation des SVM et des caractéristiques HOG dans l'analyse des images IRM cérébrales pour la détection et la classification précise des tumeurs cérébrales. L'accent est mis sur l'amélioration de la précision diagnostique en ajustant les paramètres SVM et en optimisant l'extraction des caractéristiques HOG. De plus, il est essentiel d'exploiter pleinement les informations détaillées fournies par les images IRM pour une détection précoce et une classification précise des différents types de tumeurs. L'objectif global est de développer des solutions efficaces pour obtenir une détection et une classification plus précises en optimisant l'utilisation des SVM et des caractéristiques HOG. Cela permettra d'améliorer la précision diagnostique, de réduire le temps de traitement et de tirer pleinement parti des informations fournies par les images IRM du cerveau.

L'objectif de notre recherche :

Notre objectif est l'exploiter les avancées de l'intelligence artificielle et de l'apprentissage automatique pour améliorer la détection des tumeurs cérébrales à partir d'images médicales, en utilisant les techniques d'optimisation au niveau de la classification et de l'extraction de la caractéristique.

Nous cherchons à développer un modèle automatisé et précis qui peut aider les professionnels de la section de santé à diagnostiquer les tumeurs cérébrales de manière plus rapide et plus fiable. Les résultats de cette recherche auront un impact significatif sur le domaine médical en facilitant la détection précoce des tumeurs cérébrales, ce qui conduira à des traitements plus efficaces et à de meilleurs résultats cliniques pour les patients.

Organisation de notre thèse :

Chapitre 1 : Ce chapitre présente les concepts généraux et les définitions des tumeurs cérébrales, de l'IRM, de l'intelligence artificielle, du deep learning et du machine learning, ainsi que leur relation avec l'imagerie médicale et les maladies du cerveau.

Chapitre 2 : Ce chapitre examine le traitement et la classification des images, en mettant l'accent sur l'optimisation des SVM et des caractéristiques HOG. Il explore les méthodes d'évaluation de performance pour mesurer l'efficacité de la classification.

Chapitre 3 : Ce chapitre présente les résultats obtenus en comparant différents modèles proposés. Il décrit également l'environnement du travail utilisé pour mener les expériences et effectuer les comparaisons des résultats.

CHAPITRE 1

NOTIONS D'ÉTUDE

1. Chapitre 1 : Notion D'étude

1.1 Introduction :

Les tumeurs cérébrales peuvent avoir des conséquences graves sur la santé des patients, nécessitant une détection précoce pour une intervention rapide et efficace. L'imagerie par résonance magnétique (IRM) est une méthode courante pour diagnostiquer ces tumeurs, cependant, l'analyse des images IRM peut être difficile pour les médecins en raison de la similarité des tumeurs avec les tissus normaux.

Dans ce chapitre on va commencer par une présentation de l'IRM et de son fonctionnement, suivi d'une définition de la pathologie et des tumeurs cérébrales. Nous verrons également l'importance de l'intelligence artificielle (IA) dans le domaine médical, en particulier dans l'imagerie médicale, pour le diagnostic et la prédiction des tumeurs. Enfin, nous aborderons l'impact de l'IA sur le domaine médical et ses techniques d'apprentissage automatique et profond.

1.2 Le cerveau :

Le cerveau est un organe vital du système nerveux central situé dans la boîte crânienne. Il est responsable de la régulation de la plupart des fonctions corporelles et cognitives, y compris la perception sensorielle, la pensée, l'apprentissage, la mémoire, la motricité, l'émotion et la régulation du système endocrinien. Le cerveau est constitué de plusieurs parties, chacune ayant des fonctions spécifiques. Les principales structures du cerveau sont le cerveau antérieur (ou cerveau limbique), le cerveau moyen et le cerveau postérieur (ou cervelet). [04] Chacune de ces parties est divisée en de nombreuses régions qui ont des rôles différents. Le cerveau est un organe extrêmement complexe et sa compréhension complète est encore largement incomplète. Cependant, la recherche scientifique continue à découvrir de nouvelles informations sur le fonctionnement du cerveau et son rôle dans la cognition et le comportement humain. [05]

1.2.1 Anatomie cérébrale :

Le cerveau est une masse ovoïde elle représente la partie la plus volumineuse du système nerveux, il est composé en premier du télencéphale qui a son tour est subdivisé en deux parties hémisphérique droit et gauche, et en deuxième composé du diencéphale qui réunit les deux hémisphères cérébraux par le corps calleux et par le trigone cérébral [06].

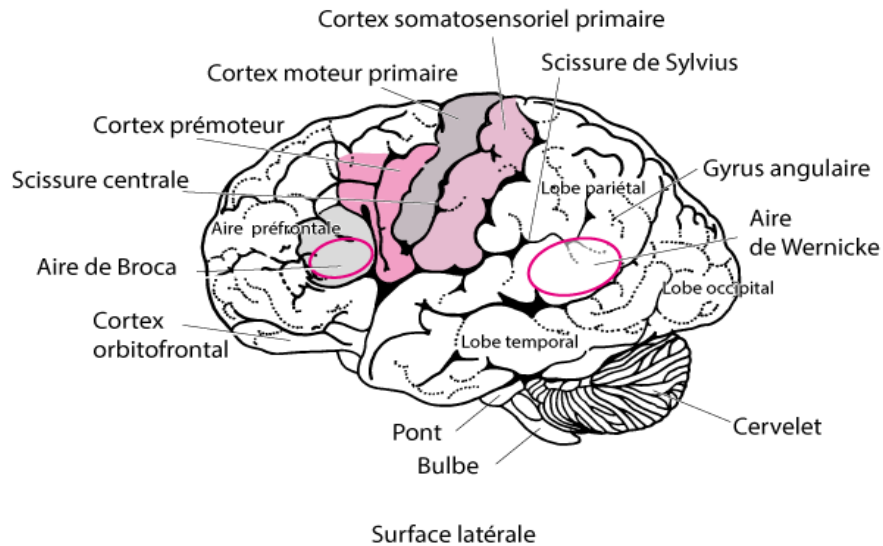


Figure 1: Anatomie horizontale d'un cerveau humain [06]

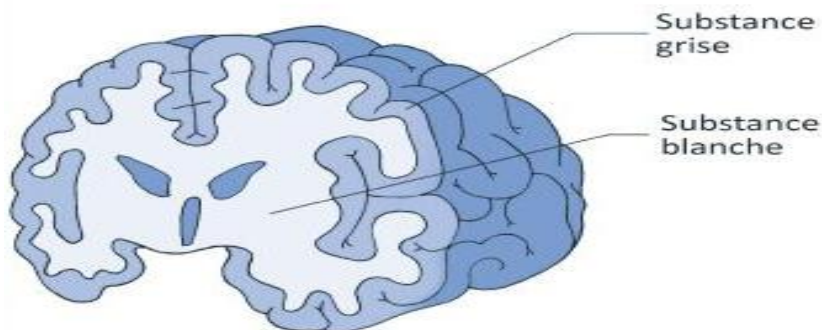


Figure 2: Anatomie verticale d'un cerveau humain [06]

1.2.2 Le liquide céphalo rachidien (LCR) :

Liquide qui remplit les ventricules et qui beigne le cerveau et joue un rôle essentiel dans les échanges entre le sang et les tissus nerveux [06].

1.2.3 La matière grise :

Qui se trouve également dans la moelle épinière, le tronc cérébral, le cervelet, elle est composée du cortex et de noyaux gris centraux [06] .

- **Le cortex :**

qui est constitué de 70% de l'ensemble des neurones du cerveau il est le centre de la conscience, des plis sont créés et sont appelé sillons [06].

- **Les noyaux centraux :**

Les noyaux centraux quant à eux représente une plus faible densité en neurone que le cortex on distingue ainsi les noyaux du télencéphale dit noyaux caudés et lenticulaires et les noyaux diencephale dit thalamus qui est le plus volumineux des noyaux centraux et l'hypothalamus [06]

1.2.4 La matière blanche :

Constitué principalement de fibre à myéline celle-ci se situe dans la partie intense du cerveau, la fibre de myéline recouvre les axons des neurones pour en accélérer la conduction, ils s'assemblent en faisceaux pour établir des connexions avec d'autres groupes de neurones [06].

1.3 Tumeurs :

Une tumeur est une masse anormale de tissu qui se forme lorsque les cellules dans un certain endroit du corps commencent à se multiplier sans contrôle. Les tumeurs peuvent être bénignes ou malignes, ce qui dépend de leur capacité à se propager et à causer des dommages aux tissus environnants. Les tumeurs bénignes sont généralement inoffensives et peuvent souvent être enlevées sans causant de dommages supplémentaires, tandis que les tumeurs malignes peuvent se propager à d'autres parties du corps et causer des dommages importants.[07]

1.3.1 Les types de Tumeurs :

- **Les Tumeurs cancéreuses :**

Les tumeurs cancéreuses sont des masses anormales de cellules qui se développent de manière incontrôlée et envahissent les tissus environnants. Elles peuvent être bénignes ou malignes. Les tumeurs malignes peuvent invader les tissus voisins, envoyer des cellules à d'autres parties du corps via le sang ou la lymphe et compromettre les fonctions vitales. Le cancer est souvent traité avec une combinaison de chirurgie, radiothérapie et chimiothérapie.

- **Les Tumeurs non cancéreuse**

Les tumeurs non cancéreuses sont des masses anormales de tissus qui ne sont pas malignes. Elles ne se propagent pas à d'autres parties du corps et ne mettent généralement pas la vie en danger, mais elles peuvent néanmoins causer des symptômes gênants ou provoquer des complications. Certaines tumeurs non cancéreuses peuvent également devenir malignes si elles ne sont pas traitées à temps.

- **Les Tumeurs intermédiaires**

Les tumeurs intermédiaires sont des tumeurs malignes qui se situent entre les tumeurs bénignes et les tumeurs malignes les plus graves. Elles peuvent se développer lentement et sont généralement considérées comme moins agressives que les tumeurs malignes les plus graves, mais elles peuvent toujours causer des problèmes de santé importants et peuvent évoluer vers des tumeurs plus graves si elles ne sont pas traitées.

1.3.2 Les tumeurs cérébrales :

Une tumeur cérébrale est une croissance anormale de cellules dans le cerveau. Les tumeurs peuvent être bénignes ou malignes. Les tumeurs bénignes sont généralement non cancéreuses et peuvent être facilement enlevées par une intervention chirurgicale. Les tumeurs malignes sont plus agressives et peuvent se propager à d'autres parties du corps. Les symptômes d'une tumeur au cerveau peuvent inclure des maux de tête, des nausées, des vomissements, des convulsions, des pertes de mémoire et des changements dans le comportement. Les traitements pour les tumeurs au cerveau peuvent inclure la chirurgie, la radiothérapie et la chimiothérapie [08].

Les tumeurs cérébrales peuvent causer des problèmes des façons suivantes :

- En envahissant directement et en détruisant le tissu cérébral
- En exerçant une pression directe sur des tissus voisins
- En augmentant la pression dans le crâne (pression intracrânienne), parce que la tumeur prend de la place et que le crâne ne peut pas s'étendre pour faire face à cela
- En causant une accumulation de liquide dans le cerveau
- En bloquant la circulation normale du liquide céphalorachidien dans les espaces du cerveau, provoquant ainsi l'expansion de ces espaces
- En causant des saignements

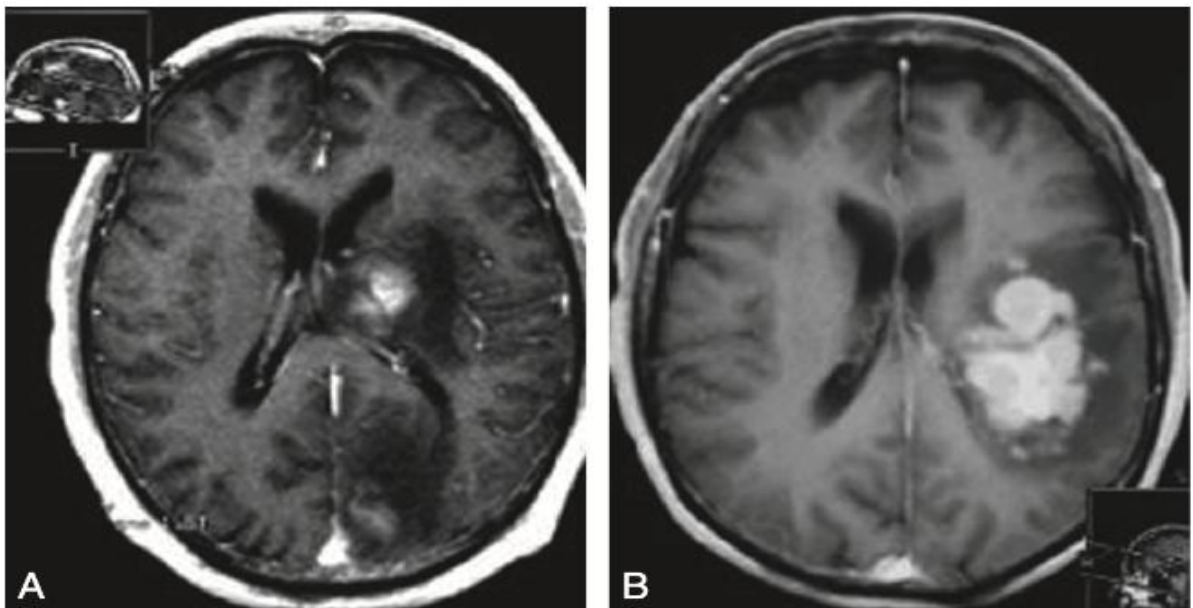


Figure 3: IRM d'une tumeur cérébrale [08]

1.3.3 Les Tumeurs cérébrales bénignes :

Les tumeurs cérébrales bénignes sont des tumeurs qui ne se propagent pas à d'autres parties du corps et qui peuvent être traitées avec succès par chirurgie. Elles comprennent des tumeurs comme les hémangioblastomes, les méningiomes, les adénomes hypophysaires, les pinéalomes, etc. [09]

- Les Méningiomes : Sont des tumeurs habituellement bénignes des méninges qui peuvent comprimer le tissu cérébral adjacent. Ils sont parmi les tumeurs intracrâniennes les plus fréquentes. Les méningiomes sont les seules tumeurs cérébrales qui sont plus fréquentes chez la femme. Ces tumeurs ont tendance à survenir entre 40 et 60 ans. [10]
- les Adénomes hypophysaires : Un adénome hypophysaire est une tumeur bénigne qui s'est développée dans l'hypophyse. L'hypophyse est une petite glande située sous le cerveau qui contrôle la production d'hormones comme le cortisol, les hormones thyroïdiennes, l'hormone de croissance, la prolactine, les hormones sexuelles, l'ocytocine et l'hormone antidiurétique. [11]
- les Pinéalomes : On appelle les tumeurs de la région pinéale proviennent des cellules pinéales de la glande pinéale elle-même. Si les cellules se divisent rapidement, elles sont considérées comme des pinéaloblastomes, alors que les tumeurs des cellules à croissance lente sont appelées pinéalomes.

1.3.4 Tumeurs cérébrales malignes :

Une tumeur maligne du cerveau, d'autre part, se développe souvent rapidement en taille et se développe de manière destructrice dans les tissus environnants, montrant une croissance dite infiltrante. Les tumeurs cérébrales malignes les plus courantes sont:

- Gliomes : Le gliome est un nom général pour les tumeurs issues des cellules gliales, le tissu conjonctif ou de soutien du cerveau. Il existe de nombreux types de gliomes, y compris les astrocytomes, les oligodendrogliomes et les épendymomes. Le type de gliome le plus fréquent est l'astrocytome. Les gliomes représentent environ 60% de toutes les tumeurs cérébrales primaires et sont souvent malignes. [12]
- métastases : On les appelle parfois tumeurs cérébrales métastatiques. Le cancer qui prend naissance dans le cerveau porte le nom de cancer primitif du cerveau. Les métastases cérébrales sont bien plus courantes que le cancer primitif du cerveau. Les métastases cérébrales sont des tumeurs secondaires qui se propagent à partir d'une autre tumeur primaire située ailleurs dans le corps [13]

1.4 Imagerie par résonance magnétique (IRM)

IRM signifie imagerie par résonance magnétique. Elle utilise un champ magnétique (aimant) et des ondes radio. Aucune radiation ionisante n'est émise, son principe consiste à réaliser des images du corps humain grâce aux nombreux atomes d'hydrogène qu'il contient. Placés dans un puissant champ magnétique, tous les atomes d'hydrogène s'orientent dans la même direction : ils sont alors excités par des ondes radio durant une très courte période (ils sont mis en résonance). A l'arrêt de cette stimulation, les atomes restituent l'énergie accumulée en produisant un signal qui est enregistré et traité sous forme d'image par un système informatique [14].

L'IRM étudie avec une grande précision de nombreux organes tels que le cerveau, la colonne vertébrale, les articulations et les tissus mous. Elle est d'une grande utilité lorsqu'une analyse très fine est nécessaire et que certaines lésions ne sont pas visibles sur les radiographies standards, l'échographie ou le scanner. Elle permet de faire des images en coupes dans différents plans et de reconstruire en trois dimensions la structure analysée.

Elle recherche :

- Au niveau du cerveau : des lésions infectieuses ou inflammatoires, des anomalies des vaisseaux, ainsi que des tumeurs ;
- Au niveau de la colonne vertébrale : des hernies discales



Figure 4: Scanner IRM [14]

- Au niveau des articulations : des lésions ligamentaires ou méniscales.

1.4.1 L'identification d'une tumeur cérébrale par l'IRM :

L'IRM (Imagerie par Résonance Magnétique) peut aider à identifier une tumeur cérébrale en produisant des images détaillées des structures internes du cerveau. Voici comment l'IRM peut identifier une tumeur cérébrale :

Les images de l'IRM peuvent montrer des zones anormales dans le cerveau, telles que des masses ou des lésions qui peuvent indiquer une tumeur cérébrale.[15]

Les tumeurs cérébrales peuvent avoir des caractéristiques distinctives sur les images de l'IRM, telles que leur forme, leur taille, leur texture et leur prise de contraste avec l'injection d'un produit de contraste.[16]

Les images de l'IRM peuvent aider à déterminer la localisation de la tumeur dans le cerveau, ce qui peut être important pour la planification du traitement.

L'IRM peut également être utilisée pour surveiller la croissance ou la régression d'une tumeur cérébrale au fil du temps. [17]

En somme, l'IRM est une technique d'imagerie très utile pour identifier les tumeurs cérébrales, car elle fournit des images très détaillées de l'intérieur du cerveau, ce qui permet aux médecins de détecter les zones anormales et de déterminer la nature et l'étendue de la tumeur.

1.4.2 La détection d'une tumeur cérébrale par l'IRM :

L'IRM (imagerie par résonance magnétique) peut détecter les tumeurs cérébrales en créant des images détaillées des tissus cérébraux. Les tumeurs cérébrales peuvent être détectées en raison de leur différence de signal par rapport aux tissus normaux du cerveau.[18] Voici comment l'IRM détecte une tumeur cérébrale :

1. L'IRM utilise un champ magnétique puissant pour aligner les noyaux d'hydrogène dans les tissus cérébraux.
2. Des ondes radio sont émises pour perturber l'alignement des noyaux d'hydrogène.
3. Lorsque les noyaux d'hydrogène reviennent à leur état d'alignement initial, ils émettent des signaux radio qui sont détectés par l'IRM et convertis en images.
4. Les tumeurs cérébrales peuvent apparaître comme des zones de signal anormal sur les images IRM en raison de leur densité et leur composition différentes des tissus cérébraux normaux.
5. Les radiologues ou les neurologues spécialisés peuvent examiner les images pour identifier les zones anormales qui pourraient indiquer la présence d'une tumeur cérébrale.

1.5 Intelligence Artificielle (AI) :

L'intelligence artificielle (IA, ou AI en anglais pour Artificial Intelligence) consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle. L'IA se retrouve implémentée dans un nombre grandissant de domaines d'application.

La notion voit le jour dans les années 1950 grâce au mathématicien Alan Turing. Dans son livre *Computing Machinery and Intelligence*, ce dernier soulève la question d'apporter aux machines une forme d'intelligence. Il décrit alors un test aujourd'hui connu sous le nom « Test de Turing » dans lequel un sujet interagit à l'aveugle avec un autre humain, puis avec une machine programmée pour formuler des réponses sensées. Si le sujet n'est pas capable de faire la

différence, alors la machine a réussi le test et, selon l'auteur, peut véritablement être considérée comme « intelligente ».

De Google à Microsoft en passant par Apple, IBM ou Facebook, toutes les grandes entreprises dans le monde de l'informatique planchent aujourd'hui sur les problématiques de l'intelligence artificielle en tentant de l'appliquer à quelques domaines précis. Chacun a ainsi mis en place des réseaux de neurones artificiels constitués de serveurs et permettant de traiter de lourds calculs au sein de gigantesques bases de données.

1.5.1 Les domaines de l'intelligence artificielle :

L'intelligence artificielle est un domaine très vaste et complexe qui englobe de nombreux sous-domaines. Les principaux domaines de l'intelligence artificielle comprennent :

- Apprentissage automatique (Machine Learning) : cela comprend l'utilisation d'algorithmes pour apprendre à partir des données et effectuer des tâches sans être explicitement programmés.
- Vision par ordinateur : cela comprend l'utilisation de techniques informatiques pour analyser des images et extraire des informations utiles.
- Traitement du langage naturel (NLP) : cela comprend l'utilisation d'algorithmes pour analyser le langage humain et en extraire des informations utiles.
- Robotique : cela comprend la conception, la construction et le contrôle de robots autonomes ou semi-autonomes.
- Systèmes experts : cela comprend l'utilisation d'algorithmes pour simuler le raisonnement humain et prendre des décisions basées sur les données disponibles.
- Apprentissage profond (Deep Learning) : cela comprend l'utilisation de réseaux neuronaux profonds pour apprendre à partir des données et effectuer des tâches sans être explicitement programmé.
- La reconnaissance du visage : est une technologie à même d'identifier ou de vérifier un sujet au moyen d'une image faciale, une vidéo ou tout élément audiovisuel de son visage.

- La reconnaissance de l'écriture manuscrite (en anglais, handwritten text recognition ou HTR) est un traitement informatique qui a pour but de traduire un texte écrit en un texte

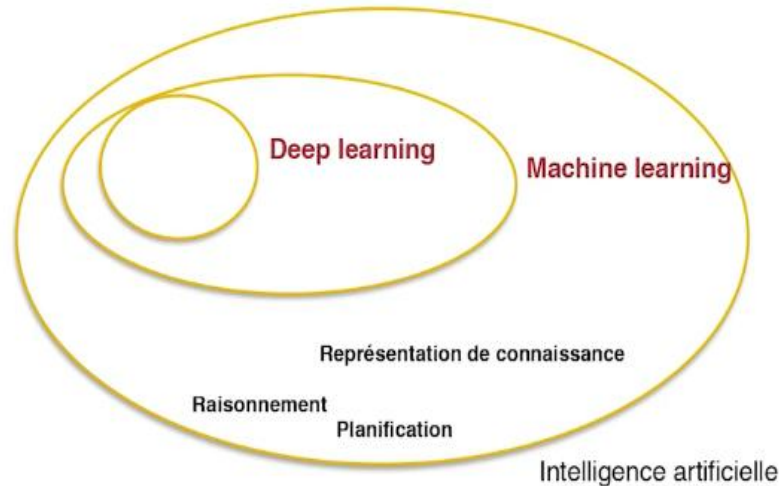


Figure 5: l'intelligence artificielle et ses sous-domaines

codé numériquement.

1.5.2 L'intelligence artificielle dans le secteur de santé :

L'intelligence artificielle (IA) est de plus en plus utilisée dans le secteur de la santé pour améliorer les diagnostics, la recherche médicale et les soins aux patients. Elle peut aider à traiter des quantités massives de données pour découvrir de nouvelles associations et tendances, ainsi qu'à fournir des analyses précises en temps réel pour aider les médecins à prendre des décisions informées. Cependant, l'IA présente également des défis en termes de protection de la vie privée et d'éthique, et il est important de veiller à ce que ses applications soient utilisées de manière responsable,

L'intelligence artificielle consiste à utiliser des modèles d'apprentissage automatique pour rechercher des données médicales et découvrir des informations permettant d'améliorer les résultats en matière de santé. Grâce aux récents progrès de l'informatique et des sciences de l'information, l'intelligence artificielle fait partie intégrante des soins de santé modernes. Les algorithmes d'IA et d'autres applications alimentées par l'intelligence artificielle sont utilisés pour soutenir les professionnels de la santé dans les environnements cliniques et les recherches en cours. [19]

1.6 Machine Learning (apprentissage automatique) :

Le Machine Learning est une sub-discipline de l'intelligence artificielle qui permet à des systèmes informatiques d'apprendre sans être explicitement programmés. Cela signifie que plutôt que de suivre des instructions détaillées pour effectuer une tâche spécifique, un système de machine learning peut développer des modèles basés sur des données d'entraînement et les utiliser pour effectuer des tâches complexes, telles que la reconnaissance de la parole, la reconnaissance d'images et la prédiction de la performance. [20]

Les algorithmes de machine learning sont largement utilisés dans de nombreux domaines, notamment la finance, la santé, la défense et le divertissement. Les algorithmes de machine learning peuvent être classés en trois grandes catégories : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage renforcé. [21]

Le machine learning est un domaine en constante évolution qui a un impact considérable sur de nombreux domaines et qui est en train de changer la façon dont les systèmes informatiques interagissent avec le monde qui les entoure. Les développements futurs dans ce domaine pourraient avoir un impact considérable sur la vie quotidienne et la société dans son ensemble.

1.6.1 Le Principe général du fonctionnement de machine Learning :

Le fonctionnement du Machine Learning est basé sur l'apprentissage automatique à partir de données. Les algorithmes de Machine Learning sont formés à partir de données d'entraînement et utilisent ces données pour développer des modèles qui peuvent effectuer des tâches complexes telles que la reconnaissance d'images, la traduction automatique de la parole et la prédiction de la demande de produits. [22] le Machine Learning se décline sous différents types de modèles, qui emploient chacun des techniques algorithmiques différentes. Selon la nature des données et le résultat souhaité, l'un de ces quatre modèles d'apprentissage peut être :

Supervisé, non supervisé, semi-supervisé ou par renforcement. Dans chacun de ces modèles, une ou plusieurs techniques algorithmiques peuvent être appliquées. Tout dépend des ensembles des données qui seront utilisés et de l'objectif visé au niveau des résultats. Par nature, les algorithmes de Machine Learning sont conçus pour classer des éléments, repérer des patterns, prévoir des résultats et prendre des décisions éclairées. Les algorithmes peuvent être mis en œuvre

individuellement ou en groupe dans le but d'atteindre la plus grande précision possible lorsque les données utilisées sont complexes et imprévisibles. [23]

1.6.2 Les avantages de Machine Learning :

- **Automatisation du processus de prise de décision :**

L'apprentissage automatique peut automatiser les processus de prise de décision complexes en se basant sur les données historiques. Cela peut permettre de gagner du temps et de l'efficacité dans des domaines tels que la finance, la santé, la production et bien d'autres.

- **Capacité à traiter des quantités énormes de données :**

L'apprentissage automatique peut traiter des quantités énormes de données en temps réel et en découvrir des relations complexes qui pourraient être difficiles à détecter pour les humains.

- **Amélioration continue avec l'ajout de données supplémentaires :**

Les modèles d'apprentissage automatique peuvent continuer à s'améliorer au fil du temps en fonction de l'ajout de nouvelles données. Cela permet aux modèles de s'adapter aux nouvelles tendances et aux nouvelles informations.

- **Capacité à détecter des relations complexes dans les données :**

L'apprentissage automatique peut détecter des relations complexes dans les données qui peuvent être difficiles à détecter pour les humains. Cela peut permettre de prendre des décisions informées en se basant sur des relations complexes dans les données.

- **Performances supérieures à celles des algorithmes traditionnels dans de nombreux cas :**

Dans de nombreux cas, les modèles d'apprentissage automatique peuvent dépasser les performances des algorithmes traditionnels en matière de précision et de rapidité. Cela peut être

particulièrement utile dans des domaines tels que la reconnaissance de la parole, la reconnaissance d'images, et la classification de données.

- **Possibilité de travailler sans intervention humaine :**

Dans certains cas, les modèles d'apprentissage automatique peuvent fonctionner en continu sans intervention humaine, ce qui peut être utile dans des applications telles que la surveillance, la détection d'anomalies et la maintenance prédictive.

1.6.3 L'utilité de la machine Learning :

La machine Learning est utilisée dans de très nombreux domaines :

- L'automobile, avec notamment la voiture autonome.
- Les biens consommations, via la recommandation de produits.
- La finance, via des modèles de prédiction des risques financiers notamment.
- Le transport, via les applications de calcul d'itinéraires.
- La santé, via des assistants de diagnostic.

1.7 Le Deep Learning (apprentissage profond) :

Le Deep Learning est une méthode d'IA issue du concept de machine learning. Cette méthode d'apprentissage dit profond se base plus spécifiquement sur la notion de réseau de neurones artificiel.

De structure non linéaire, un réseau de neurones artificiel se présente sous la forme d'un réseau d'ensembles d'unités d'exécution d'information (représentant les neurones) superposées en couches et liés entre eux via des connecteurs (les synapses). Partant de-là, il traite l'information à travers des modèles de propagation d'activations de ces unités, s'actionnant au-delà d'un certain seuil.

1.7.1 Le fonctionnement de Deep Learning :

Le Deep Learning fonctionne en utilisant des réseaux de neurones artificiels pour modéliser les relations complexes dans les données. Les réseaux de neurones sont formés sur des données d'entraînement en utilisant des algorithmes d'optimisation pour ajuster les poids des connexions entre les neurones.

Le processus de formation d'un modèle de Deep Learning commence par l'acquisition de données d'entraînement qui représentent les relations complexes que le modèle doit apprendre. Les données d'entraînement sont alors utilisées pour ajuster les poids des connexions entre les neurones dans le réseau de neurones. Cet ajustement est effectué en utilisant des algorithmes d'optimisation tels que la descente de gradient. Les algorithmes d'optimisation comparent les résultats prédits par le modèle aux valeurs réelles pour les données d'entraînement et ajustent les poids des connexions en conséquence. [24]

Une fois le modèle formé, il peut être utilisé pour effectuer des prédictions sur de nouvelles données en utilisant les poids ajustés. Pour ce faire, les nouvelles données sont entrées dans le réseau de neurones, et les activations des neurones sont calculées en utilisant les poids ajustés. Les activations finales peuvent être utilisées pour effectuer une tâche spécifique, telle que la classification ou la régression.

En résumé, le Deep Learning fonctionne en utilisant des réseaux de neurones pour modéliser les relations complexes dans les données, en les formant sur des données d'entraînement en utilisant des algorithmes d'optimisation pour ajuster les poids des connexions, et en utilisant les poids ajustés pour effectuer des prédictions sur de nouvelles données.[25]

1.7.2 Caractéristiques de Le Deep Learning :

- **Utilisation de réseaux de neurones profonds :**

Le Deep Learning utilise des réseaux de neurones profonds pour modéliser les relations complexes dans les données.

- **Capacité à traiter des quantités énormes de données :**

Le Deep Learning est capable de traiter de grandes quantités de données, ce qui en fait un choix populaire pour les applications impliquant de grandes quantités de données.

- **Modélisation de relations complexes :**

Le Deep Learning est conçu pour modéliser des relations complexes dans les données en utilisant des réseaux de neurones profonds.

- **Performances supérieures pour certaines tâches d'apprentissage automatique :**

Le Deep Learning a montré de bonnes performances pour certaines tâches d'apprentissage automatique, notamment la reconnaissance d'images et la reconnaissance de la parole.

- **Capacité à capturer des caractéristiques complexes :**

Le Deep Learning est capable de capturer des caractéristiques complexes dans les données, ce qui en fait un choix populaire pour les applications impliquant des images, des signaux et d'autres formes de données complexes.

1.7.3 Les points forts et faibles de Le Deep Learning :

Tableau 1: les points forts et les points faibles de DL

Les Points forts	Les Points Faibles
Capacité à traiter de grandes quantités de données	Nécessité de grandes quantités de données d'entraînement
Modélisation de relations complexes	Temps de formation relativement long
Performances supérieures pour certaines tâches d'apprentissage automatique	Complexité du modèle
Capacité à capturer des caractéristiques complexes	Nécessité de ressources informatiques importantes
Utilisation de réseaux de neurones profonds	Difficulté à comprendre les décisions prises par

	le modèle
Flexibilité pour les différentes tâches d'apprentissage automatique	Sensibilité aux données d'entraînement bruitées ou biaisées
Possibilité de travailler avec des données non structurées	Tendance à surapprendre les données d'entraînement
Capacité d'auto-apprentissage	Besoin de méthodes de validation rigoureuses pour éviter l'overfitting
Possibilité de travailler avec des données incomplètes ou manquantes	Besoin de méthodes de normalisation et de prétraitement des données
Maturité croissante de la technologie et de la communauté de développeurs.	Manque de garanties sur la sécurité et la confidentialité des données.

1.8 La différence entre Le Machine Learning et Le Deep Learning :

Le Machine Learning (apprentissage automatique) et le Deep Learning sont deux domaines liés de l'intelligence artificielle qui ont suscité beaucoup d'intérêt et de développement ces dernières années. Bien qu'ils soient souvent utilisés de manière interchangeable, il existe des différences clés entre ces deux approches qui les rendent uniques et les rendent appropriées pour différents types de tâches. [26]

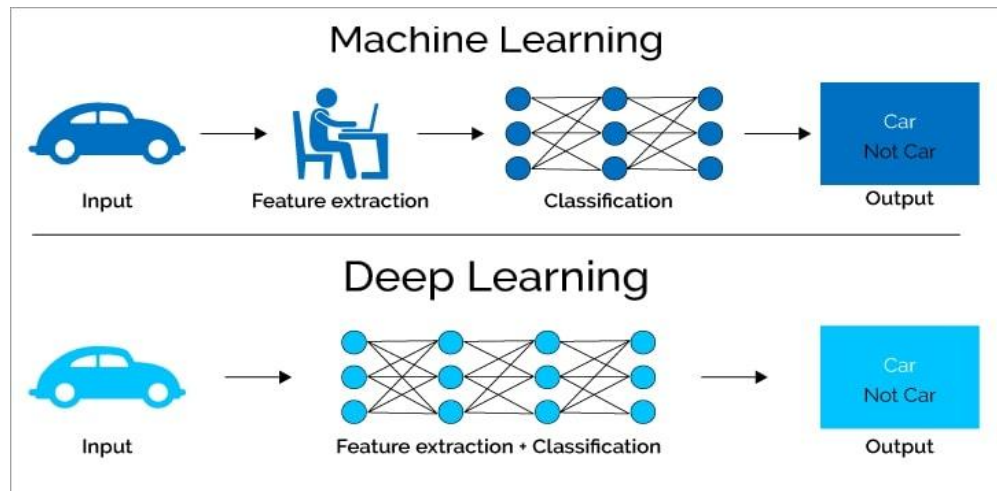
Tout d'abord, le Machine Learning est un sous-domaine de l'intelligence artificielle qui vise à construire des algorithmes qui peuvent apprendre de nouvelles connaissances à partir de données. Le Deep Learning, d'autre part, est une approche plus spécifique du Machine Learning qui utilise des réseaux de neurones profonds pour résoudre des problèmes complexes. Les réseaux de neurones profonds sont des modèles mathématiques qui sont conçus pour imiter le fonctionnement du cerveau en utilisant des couches de neurones interconnectées pour effectuer des tâches complexes.

En second lieu, le Machine Learning est généralement utilisé pour résoudre des tâches de classification et de régression, telles que la reconnaissance d'images, la prédiction des prix des actions et la classification de textes. Le Deep Learning, quant à lui, est particulièrement utile pour les tâches d'apprentissage profond, telles que la reconnaissance de la parole, la reconnaissance d'images et la traduction automatique. Le Deep Learning a également été utilisé avec succès pour les jeux de vidéo, la robotique et les systèmes recommandation. [27]

Enfin, l'un des principaux avantages du Deep Learning est sa capacité à gérer des quantités énormes de données, ce qui en fait un choix populaire pour les applications qui nécessitent une analyse de grandes quantités de données. Cependant, la complexité des modèles de Deep Learning et la nécessité de grandes quantités de données d'entraînement peuvent également être considérées comme des points faibles. [27] le Machine Learning et le Deep Learning sont deux approches importantes de l'intelligence artificielle qui ont chacune leur propre conjoncture d'utilisation. Le Machine Learning est utile pour des tâches simples telles que la classification et la régression, tandis que le Deep Learning est plus adapté aux tâches complexes telles que la reconnaissance de la parole et la reconnaissance d'images. Les développeurs et les chercheurs doivent prendre en compte les différences clés entre ces deux approches pour choisir la plus appropriée pour leur projet.

Un réseau de neurones artificiels (RNA) est un système informatique qui est conçu pour fonctionner de manière similaire à un cerveau humain. Il est basé sur un ensemble de nœuds interconnectés, appelés "neurones artificiels", qui sont conçus pour traiter et transmettre des informations. Les neurones artificiels sont connectés entre eux par des arcs, qui représentent des liens de communication .[27] Leur fonctionnement est basé sur des algorithmes de reconnaissance de modèles et d'apprentissage automatique, qui permettent au RNA de s'adapter à de nouvelles informations en modifiant les poids associés à chaque lien de communication.

Les RNA peuvent être utilisés dans une variété d'applications, notamment la reconnaissance de la parole, la reconnaissance d'images, la classification et la prédiction. Ils sont souvent utilisés dans le cadre du deep learning, une sous-discipline du machine learning qui implique l'utilisation de réseaux de neurones profonds pour résoudre des problèmes complexes.



1.

9 **Figure 6: la différence de Machine Learning et Deep Learning [27]**

Conclusion :

En conclusion, les tumeurs cérébrales sont des affections graves qui peuvent avoir des conséquences dévastatrices sur la santé humaine. L'imagerie par résonance magnétique (IRM) est une méthode largement utilisée pour diagnostiquer ces tumeurs, mais l'analyse des images peut être difficile pour les médecins. L'utilisation de techniques d'intelligence artificielle, telles que la classification SVM, peut aider à identifier les tumeurs cérébrales à partir d'images IRM. Cette approche offre des avantages significatifs pour la détection précoce et la surveillance des tumeurs cérébrales. Dans ce chapitre nous avons également vu l'importance de l'IA dans le domaine médical, en particulier dans l'imagerie médicale, pour le diagnostic et la prédiction des tumeurs. En somme, ce chapitre vise à présenter une vue d'ensemble sur l'utilisation de l'IRM et de l'IA pour la détection et le diagnostic des tumeurs cérébrales.

CHAPITRE 2

LA CLASSIFICATION

2. Chapitre 2 : La classification

2.1 Introduction :

Le présent chapitre explore en détail la classification des images et l'utilisation de l'algorithme de la machine à vecteurs de support (SVM) pour la classification d'images. Nous examinerons les principes de base de SVM, ainsi que les techniques de sélection de noyaux et de paramètres. Nous aborderons également l'optimisation de SVM en utilisant l'algorithme d'optimisation par essaim de particules (PSO), et nous terminerons par une évaluation des performances de SVM à travers différentes mesures de performance.

2.2 La classification :

La classification est une tâche d'apprentissage automatique qui consiste à assigner une étiquette ou une classe à une entrée donnée. Cela peut être considéré comme une forme de résolution de problèmes de reconnaissance de catégories. La classification est souvent utilisée pour des tâches telles que la reconnaissance d'images, la reconnaissance de la parole et la classification de courrier électronique en tant que spam ou non spam.

Le processus de classification se déroule en plusieurs étapes :

❖ Prétraitement des données :

Les données d'entrée sont prétraitées pour être en conformité avec le modèle de classification.

❖ Entraînement du modèle

Un modèle de classification est formé à partir des données d'entraînement en utilisant un algorithme d'apprentissage automatique.

❖ Évaluation du modèle :

Le modèle formé est évalué en utilisant des données de validation pour déterminer son niveau de performance.

❖ Prédiction :

Le modèle formé peut être utilisé pour prédire la classe d'une entrée donnée en dehors de l'ensemble de données d'entraînement.

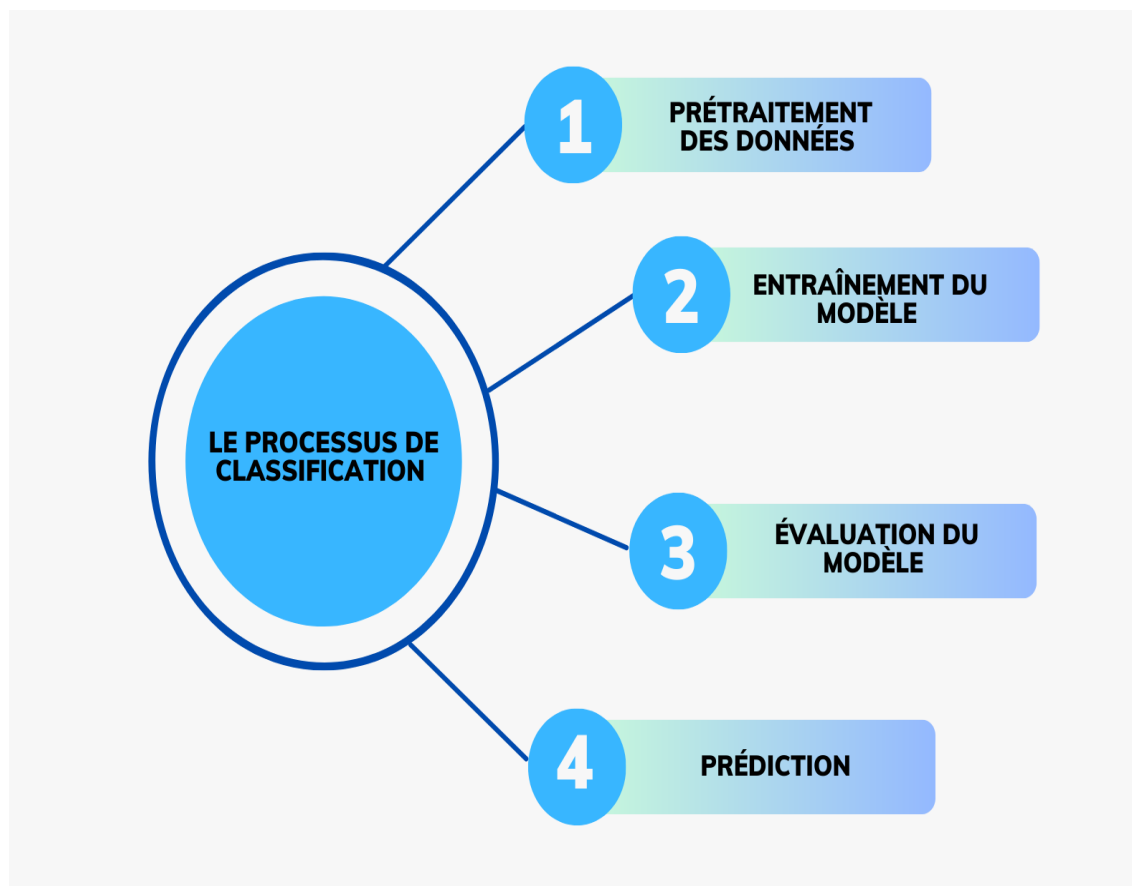


Figure 7: le processus de classification

2.2.1 Apprentissage supervisé :

Dans l'apprentissage supervisé, les data scientists alimentent les algorithmes avec des jeux de données d'entraînement étiquetées, et définissent les variables qu'ils veulent que l'algorithme évalue pour les corrélations. Les entrées et les sorties sont toutes spécifiées dans cette phase de formation. Par exemple, quand un data scientist essaie d'entraîner un algorithme pour déduire si une image contient un chat, en utilisant l'apprentissage supervisé, il crée une étiquette pour chaque image utilisée dans les données de formation indiquant si l'image contient un chat ou non. [28]

En classification supervisée:

Classification

- on connaît déjà le nombre de groupes qui existent dans la population;
- on connaît le groupe auquel appartient chaque observation de la population;
- on veut classer les observations dans les bons groupes à partir de différentes variables.

On peut ensuite utiliser une règle de classification pour prédire les groupes auxquels appartiennent de nouvelles observations. Des exemples classiques d'applications sont:

- identifier si une transaction bancaire est frauduleuse ou pas;
- reconnaître des chiffres écrits à la main;
- identifier le type de cancer dont souffre un patient.

2.2.1.1 La Classification et de la Régression :

L'apprentissage supervisé est généralement effectué dans le contexte de la classification et de la

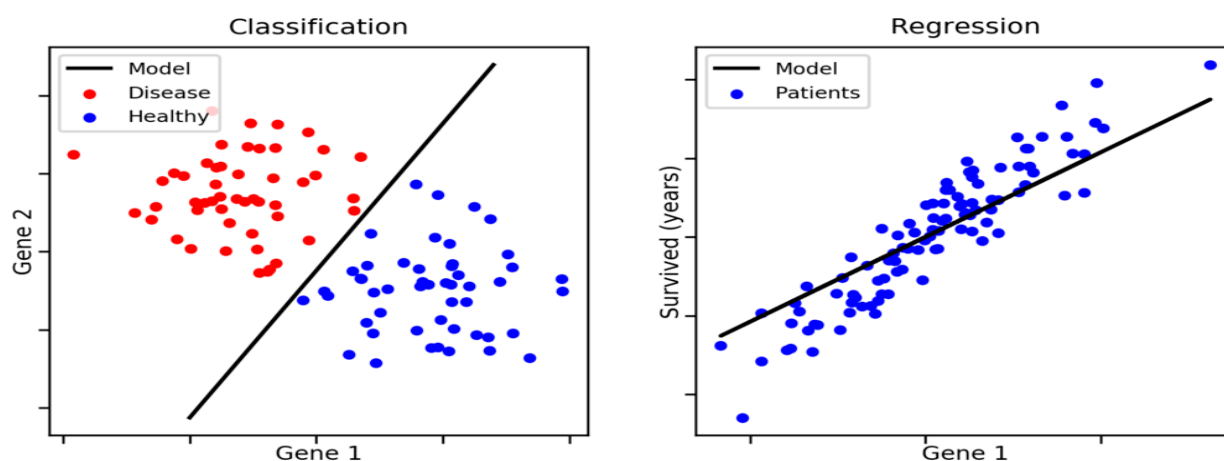


Figure 8: la différence entre la régression et la classification [28]

régression.

- **Classification:**

Un problème de classification survient lorsque la variable de sortie est une catégorie, telle que «rouge», «bleu» ou «maladie» et «pas de maladie». Exemples :

- ❖ En finance et dans le secteur bancaire pour la détection de la fraude par carte de crédit (fraude, pas fraude).
- ❖ Détection de courrier électronique indésirable (spam, pas spam).
- ❖ Dans le domaine du marketing utilisé pour l'analyse du sentiment de texte (heureux, pas heureux).

❖ En médecine, pour prédire si un patient a une maladie particulière ou non.

- **Régression:**

Un problème de régression se pose lorsque la variable de sortie est une valeur réelle, telle que:

❖ Prédire le prix de l'immobilier

❖ Prédire le cours de bourse

Certains types courants de problèmes fondés sur la classification et la régression incluent la prévision et la prévision de séries temporelles, respectivement.

2.2.2 Apprentissage non supervisé :

Avec l'apprentissage non supervisé, l'algorithme est entraîné sur des données non étiquetées. Il parcourt les ensembles de données à la recherche de toute connexion significative. Cette approche est utile lorsque vous ne savez pas ce que vous recherchez. Si vous montriez à cet algorithme plusieurs milliers ou millions d'images, il pourrait en arriver à classer un sous-ensemble d'images (un cluster), par exemple celles où les humains reconnaîtraient des félins.

En revanche, un algorithme entraîné sur des images labélisées de chats dans un jeu de comparaison avec des images de chien sera capable d'identifier des images de chat avec un degré de confiance élevé. Mais si le projet d'apprentissage supervisé nécessite un million d'images étiquetées pour développer le modèle, la prédiction générée par le modèle exige beaucoup d'efforts humains (ou autant d'astuces). [28]

En contrepartie en classification non supervisée :

- on ne connaît souvent pas le nombre de groupes qui existent dans la population.
- on ne connaît pas le groupe auquel appartient chaque observation de la population.

Classification

- on veut classer les observations dans des groupes homogènes à partir de différentes variables.

Les applications typiques sont nombreuses. Par exemple:

- en biologie : l'élaboration de la taxonomie animale.
- en psychologie : la détermination des types de personnalités présents dans un groupe d'individus.
- en text mining : le partitionnement de courriels ou de textes en fonction du sujet traité.

Il existe plusieurs familles de méthodes de classification non supervisée. Les plus communes sont :

- la classification hiérarchique ;
- la classification non hiérarchique, par exemple la méthode des k-moyennes (*k-means*);
- la classification basée sur une densité ;
- la classification basée sur des modèles statistiques/probabilistes, par exemple un mélange de lois normales.

2.2.3 Apprentissage semi-supervisé :

Les data scientists peuvent prendre une sorte de raccourci appelé apprentissage semi-supervisé qui combine les deux approches. L'apprentissage semi-supervisé décrit un flux de travail spécifique dans lequel des algorithmes d'apprentissage non supervisé sont utilisés pour générer automatiquement des étiquettes, qui peuvent être introduites dans les algorithmes d'apprentissage supervisé. Reprenons notre exemple de la reconnaissance d'images. Dans cette approche, les humains étiquettent manuellement certaines images, l'apprentissage non supervisé devine les étiquettes pour d'autres, puis toutes ces étiquettes et images sont introduites dans les algorithmes d'apprentissage supervisé pour créer un modèle d'IA. [28]

L'apprentissage semi-supervisé a un grand avantage financier. Il permet de baisser le coût de l'étiquetage des grands ensembles de données utilisés dans le cadre du machine learning. « Si vous pouvez faire en sorte que les humains étiquettent 0,01 % de vos millions d'échantillons, l'ordinateur peut alors exploiter ces étiquettes pour augmenter de manière significative sa précision prédictive ».

2.2.4 Apprentissage par renforcement :

Généralement utilisé pour apprendre à une machine à exécuter une séquence d'étapes, l'apprentissage par renforcement est différent de l'apprentissage supervisé et non supervisé. Les scientifiques programment un algorithme pour effectuer une tâche, en lui donnant des indices positifs ou négatifs au fur et à mesure qu'il travaille sur la façon d'effectuer la tâche. Le programmeur fixe les règles pour les récompenses, mais laisse à l'algorithme le soin de décider lui-même des étapes à suivre pour maximiser la récompense, et donc accomplir la tâche.

2.3 La classification d'images :

est un domaine très actif de la recherche en intelligence artificielle et en apprentissage automatique. Le but de la classification d'images est de déterminer la catégorie à laquelle appartient une image donnée en utilisant des algorithmes de machine learning. Les images peuvent être classées en de nombreuses catégories différentes, telles que les objets, les animaux, les paysages, les visages, etc. [29] il existe plusieurs méthodes couramment utilisées pour la classification d'images, notamment les arbres de décision, les k-plus proches voisins (KNN) et les réseaux de neurones à rétro propagation. Chacune de ces méthodes a ses propres avantages et inconvénients, et le choix de la méthode dépend du problème à résoudre et des caractéristiques des données.

2.3.1 Classification des images et l'apprentissage machine :

Les méthodes manuelles se sont avérées très difficiles à appliquer pour des tâches en apparence très simples comme la classification des images, la reconnaissance d'objets dans les images ou la reconnaissance vocale. Les données venant du monde réel les échantillons d'un son ou les pixels d'une image sont complexes, variables et entachées de bruit. Pour une machine, une image est un tableau de nombres indiquant la luminosité (ou la couleur) de chaque pixel, et un signal sonore une suite de nombres indiquant la pression de l'air à chaque instant. Comment une machine peut-elle transcrire la suite de nombres d'un signal sonore en série de mots tout en ignorant le bruit ambiant, l'accent du locuteur et les particularités de sa voix ? Comment une machine peut-elle identifier un chien ou une chaise dans le tableau de nombres d'une image quand l'apparence d'un chien ou d'une chaise et des objets qui les entourent peut varier infiniment ? Il est virtuellement impossible d'écrire un programme qui fonctionnera de manière robuste dans toutes les situations. C'est là qu'intervient l'apprentissage machine (que l'on appelle aussi apprentissage automatique). C'est l'apprentissage qui anime les systèmes de toutes les grandes entreprises d'Internet. Elles l'utilisent depuis longtemps pour filtrer les contenus indésirables, ordonner des réponses à une recherche, faire des recommandations, ou sélectionner les informations intéressantes pour chaque utilisateur.

2.3.2 Les étapes principales du processus de classification des images :

Le processus de classification d'images peut être divisé en plusieurs étapes, telles que :

- **Collecte des données :** La première étape consiste à collecter les images qui seront utilisées pour la classification. Cela peut être fait à partir de sources en ligne, d'une base de données interne ou en créant un ensemble de données en utilisant un appareil photo.
- **Prétraitement des images :** Les images collectées peuvent avoir des caractéristiques telles que des dimensions différentes, des niveaux d'intensité variés, du bruit, des distorsions, etc. Par conséquent, il est nécessaire de prétraiter les images pour normaliser les caractéristiques des images et réduire les variations inutiles. Les techniques de prétraitement des images comprennent le redimensionnement, le recadrage, la normalisation, la suppression du bruit, etc.
 - **Redimensionnement :** Les images peuvent être redimensionnées pour un traitement plus rapide et pour garantir que toutes les images ont la même taille.
 - **Normalisation :** Les pixels des images peuvent être normalisés pour qu'ils soient tous compris dans la même plage de valeurs, par exemple entre 0 et 1. Cela facilite le traitement des images par le modèle SVM.
 - **Application de filtres :** Les filtres peuvent être appliqués pour améliorer la qualité de l'image et/ou extraire des caractéristiques spécifiques de l'image. Les filtres peuvent inclure des filtres de flou, des filtres de contour, des filtres de seuillage, etc.
 - **Segmentation :** La segmentation est le processus de séparation d'une image en différentes régions, telles que des objets, des contours ou des textures. La segmentation peut être utilisée pour extraire des caractéristiques spécifiques de l'image ou pour éliminer les informations non pertinentes.
 - **Augmentation de données :** L'augmentation de données est une technique utilisée pour augmenter la taille de l'ensemble de données en générant de nouvelles images à partir des images existantes. Cela peut être fait en effectuant des rotations, des translations, des zooms, des changements de luminosité ou des déformations sur les images existantes.

L'objectif du prétraitement des images est de réduire les variations dans les images et d'extraire les caractéristiques pertinentes pour la classification. Les techniques de prétraitement peuvent varier en fonction du type d'image et des objectifs de classification.

- **Extraction des caractéristiques :** L'étape suivante consiste à extraire des caractéristiques à partir des images prétraitées. Les caractéristiques sont des valeurs numériques qui capturent des informations sur les propriétés visuelles de l'image. Les techniques courantes d'extraction de caractéristiques comprennent les descripteurs locaux tels que HOG (histogrammes de gradient orienté), LBP (patterns binaires locaux), descripteurs basés sur des réseaux de neurones tels que CNN (réseaux de neurones convolutionnels), etc.

2.4 HOG (Histogram of Oriented Gradients) :

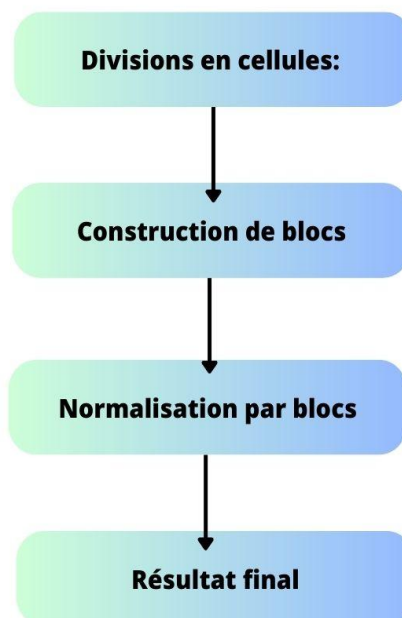
Les HOG ont été proposés par Navneet Dalal et Bill Triggs, chercheurs à l'inriade Grenoble, à la conférence CVPR de juin 2005 [30]. L'histogramme des dégradés orientés est un descripteur de caractéristique utilisé dans la vision par ordinateur et le traitement d'image pour la détection d'objets.[30] l'idée essentielle derrière l'histogramme de gradient orienté c'est que l'apparence locale et la forme d'objet dans une image peut être décrite par la distribution d'intensité des gradients ou de direction des contours. Cette méthode est similaire à celle des histogrammes d'orientation de bord, des descripteurs de transformation d'entités invariables à l'échelle et des contextes de forme, mais diffère en ce qu'elle est calculée sur une grille dense de cellules uniformément espacées. Le descripteur HOG maintient quelques avantages clés par rapport aux autres méthodes. Puisque le descripteur histogramme de gradient orienté opère sur les cellules localisées, la méthode maintient l'invariance à des transformations géométriques et photométriques, ces changements ne feront leur apparition que dans les larges régions d'espaces.[30]

2.4.1 les étapes de L'algorithme HOG:

- **Divisions en cellules :** Tout d'abord, l'image est divisée en petites cellules, généralement de 8x8 pixels. Pour chaque cellule, un histogramme des gradients orientés est calculé. Les gradients orientés sont les vecteurs qui indiquent la direction et l'intensité des changements de luminosité dans l'image.

Classification

- **Construction de blocs :** Les cellules sont ensuite regroupées en blocs plus grands, qui sont utilisés pour représenter l'apparence d'une zone plus grande dans l'image. Les histogrammes de chaque cellule sont concaténés pour former un histogramme global pour chaque bloc.
- **Normalisation par blocs :** Enfin, les histogrammes de chaque bloc sont normalisés pour réduire les effets de la variation de lumière dans l'image. Cela garantit que les features extraites sont robustes à la variation de la luminosité.
- **Le résultat final de l'algorithme HOG :** est un vecteur de descripteurs de l'apparence de l'image, qui peut être utilisé pour la reconnaissance d'images et la détection d'objets. Les algorithmes de reconnaissance d'images et de détection d'objets utilisant HOG peuvent être entraînés à reconnaître des objets spécifiques en utilisant des données d'entraînement annotées, et peuvent ensuite être utilisés pour détecter des objets similaires dans des images non annotées. [31]



2.4.2 Les principales caractéristiques de l'algorithme HOG :

- ❖ **Représentation de l'image :** L'algorithme HOG représente une image sous forme de descripteurs de gradient orienté.
- ❖ **Détection de bords :** HOG utilise des gradients orientés pour détecter les bords des objets dans une image.
- ❖ **Divisions en cellules :** L'image est divisée en petites cellules, et pour chaque cellule, l'histogramme des orientations des gradients est calculé.
- ❖ **Construction de blocs :** Les cellules sont regroupées en blocs plus grands, ce qui permet de capturer les relations spatiales entre les cellules.
- ❖ **Normalisation :** Pour gérer les variations de luminosité et de contraste, les histogrammes sont normalisés pour chaque bloc.

Figure 9: les étapes de HOG

- ❖ **Utilisation en reconnaissance d'objets :** L'algorithme HOG est souvent utilisé en conjonction avec des classificateurs tels que SVM (Support Vector Machine) pour la reconnaissance d'objets dans les images.

2.4.3 Les avantages de l'algorithme HOG :

L'algorithme HOG présente plusieurs avantages, notamment :

- ❖ **Robuste aux variations d'éclairage :**

L'algorithme HOG est capable de gérer les variations d'éclairage en normalisant les histogrammes pour chaque bloc, ce qui le rend moins sensible aux variations de luminosité et de contraste dans les images.

- ❖ **Performances élevées :**

L'algorithme HOG est considéré comme l'un des algorithmes les plus performants en matière de reconnaissance d'objets dans les images, en raison de sa capacité à capturer les caractéristiques de l'apparence des objets.

- **Rapide et efficace :**

HOG est un algorithme rapide et efficace en termes de consommation de ressources, ce qui le rend approprié pour une utilisation dans des applications en temps réel.

- ❖ **Facilité d'implémentation :**

L'algorithme HOG est relativement simple à implémenter, ce qui en fait un choix populaire pour les développeurs et les chercheurs.

- ❖ **Fonctionne avec des images en niveaux de gris :**

HOG peut être utilisé avec des images en niveaux de gris, ce qui peut réduire considérablement la quantité de données à traiter et accélérer les performances de l'algorithme.

- **Sélection des caractéristiques :** Dans cette étape, les caractéristiques les plus discriminantes sont sélectionnées à partir de l'ensemble complet de caractéristiques extraites. Cela peut être fait manuellement en utilisant une analyse de corrélation, une analyse en composantes principales (PCA) ou automatiquement à l'aide d'algorithmes de sélection de caractéristiques.

2.4.4 Les paramètres de HOG :

2.4.4.1 Orientation (nombre de bins d'orientation) :

L'orientation fait référence à la plage complète des angles possibles pour les gradients. Les gradients sont répartis dans des compartiments appelés bins. Le nombre de bins d'orientation détermine la précision de la représentation des orientations. Une valeur courante pour le nombre de bins d'orientation est de 9, ce qui divise la plage complète des angles en neuf bins.

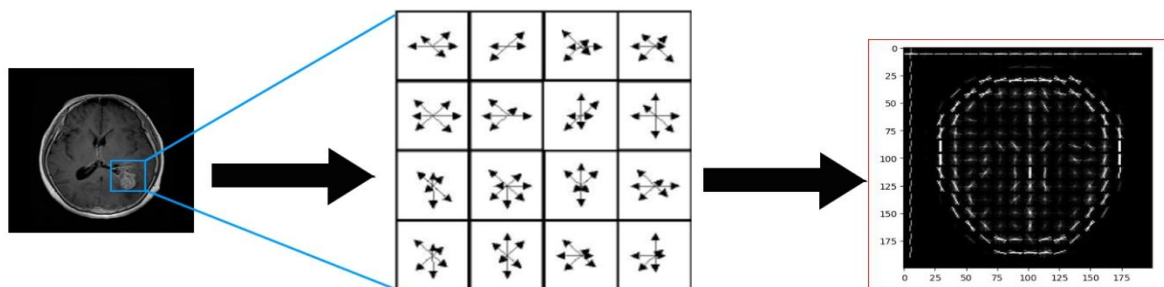


Figure 10: Le paramètre de l'orientation du HOG [31]

2.4.4.2 Cellules par bloc (cells per block) :

Les cellules sont les régions carrées sur lesquelles les gradients sont calculés. Les cellules par bloc font référence au nombre de cellules incluses dans un bloc rectangulaire. Un bloc est une région rectangulaire composée de plusieurs cellules adjacentes. Les cellules par bloc peuvent être réglées en fonction de la taille souhaitée pour les caractéristiques extraites. Une valeur courante pour les cellules par bloc est de 4x4, ce qui signifie que chaque bloc est composé de quatre cellules.[32]

2.4.4.3 Pixels par cellule (pixels per cell) :

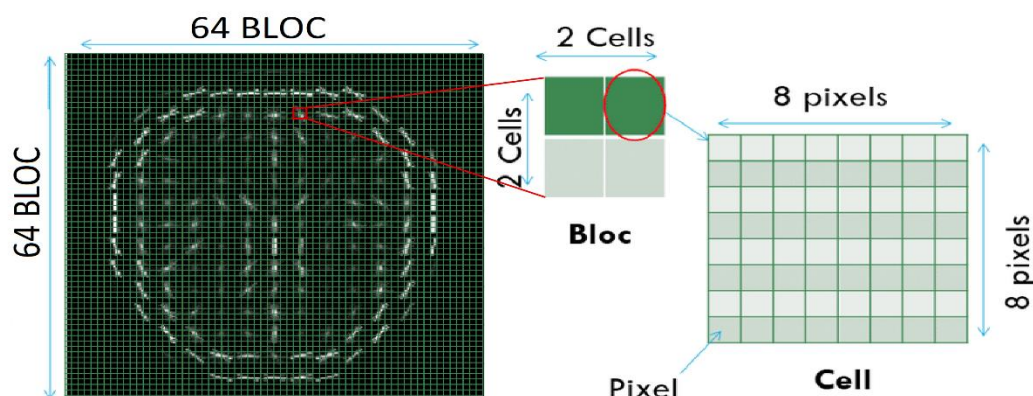
Les pixels par cellule font référence à la taille de la région carrée sur laquelle les gradients sont calculés. Une valeur typique pour les pixels par cellule est de 8x8, ce qui signifie que les gradients sont calculés sur une région de 8x8 pixels.[32]

Ces paramètres déterminent la granularité de la représentation des caractéristiques extraites par HOG. Ils peuvent être ajustés en fonction des spécificités de l'application et des objets à détecter pour obtenir de bons résultats en termes de précision et de performance.

Figure 11 : la distribution des pixels dans une cellule d'un bloc [32]

2.5 L'optimisation de HOG :

L'optimisation est conçue pour trouver les meilleurs réglages pour la technique d'extraction



de caractéristiques appelée HOG (Histogramme des gradients orientés), utilisée pour analyser des images et capturer des informations importantes sur la forme et la texture. L'objectif est de déterminer la combinaison optimale des paramètres de HOG qui permet d'obtenir les résultats les plus précis.

Le processus d'optimisation de HOG consiste à explorer différentes configurations pour trois paramètres clés : l'orientation, le nombre de cellules par bloc et le nombre de pixels par cellule. Ces réglages influencent la façon dont l'algorithme HOG traite l'image. Le code parcourt toutes les combinaisons possibles de ces valeurs de paramètres et teste chaque combinaison, pour chaque combinaison de paramètres, un ensemble d'images d'entraînement contenant des échantillons de tumeurs cérébrales et de tumeurs non cancéreuses est analysé. Les caractéristiques HOG sont extraites de chaque image, ce qui permet de représenter les caractéristiques uniques de l'image. Ces caractéristiques sont ensuite utilisées pour entraîner un modèle d'apprentissage automatique appelé un classifieur SVM, une fois l'entraînement terminé, le classifieur est évalué sur un ensemble distinct d'images de test. L'exactitude du classifieur est mesurée, et la meilleure exactitude obtenue jusqu'à présent ainsi que les valeurs de paramètres

correspondantes sont enregistrées. Le processus se poursuit jusqu'à ce que toutes les combinaisons aient été testées.

Enfin, la meilleure combinaison de valeurs de paramètres qui a donné la plus haute exactitude est affichée. Ces paramètres optimisés peuvent ensuite être utilisés dans des applications futures pour améliorer les performances de détection des tumeurs cérébrales.

En résumé, l'optimisation de HOG consiste à explorer différentes combinaisons de réglages, à extraire des caractéristiques d'image, à entraîner un classifieur et à évaluer son exactitude afin de trouver la meilleure combinaison de paramètres pour une détection précise des tumeurs cérébrales.

2.6 L'analyse en composantes principales du noyau (Kernel-PCA) :

Est une technique d'analyse de données qui permet de transformer des données non linéaires en un espace de dimension supérieure pour y détecter des relations linéaires cachées.

Contrairement à l'analyse en composantes principales classique (PCA), qui ne fonctionne qu'avec des données linéaires, le Kernel-PCA utilise des fonctions de noyau pour projeter les données non linéaires dans un espace de dimension supérieure. Cette technique permet de détecter des motifs dans les données qui ne seraient pas visibles avec l'analyse en composantes principales classique.

Le principe de base du Kernel-PCA est de trouver les vecteurs propres de la matrice de Gram, qui est une matrice symétrique définie positive obtenue en appliquant une fonction de noyau aux données d'entrée. Ces vecteurs propres sont utilisés pour projeter les données dans l'espace de dimension supérieure. Les coordonnées dans cet espace représentent les composantes principales du noyau.

Le choix de la fonction de noyau est crucial pour la performance du Kernel-PCA. Les fonctions de noyau les plus couramment utilisées sont la fonction gaussienne, la fonction polynomiale et la fonction de base radiale. Le choix de la fonction de noyau dépend des caractéristiques des données à analyser.[33]

➤ **Entraînement du modèle:** Après avoir extrait et sélectionné les caractéristiques, la prochaine étape consiste à entraîner un modèle de classification à partir des données

d'apprentissage. Les modèles couramment utilisés pour la classification d'images comprennent les SVM, les réseaux de neurones, les arbres de décision, les forêts aléatoires, etc.

- **Validation du modèle :** Dans cette étape, le modèle entraîné est évalué à l'aide de données de validation pour déterminer sa précision et sa capacité à généraliser les données. Des techniques telles que la validation croisée, la matrice de confusion, la courbe ROC (Receiver Operating Characteristic) sont utilisées pour évaluer la performance du modèle.
- **Application du modèle :** Après avoir validé le modèle, il peut être utilisé pour prédire des classes d'images inconnues. Les images inconnues sont prétraitées, les caractéristiques sont extraites, les caractéristiques sélectionnées sont alimentées dans le modèle et la classe prédite est renvoyée en sortie.

2.7 K-Nearest Neighbors (KNN ou méthode des k plus proches

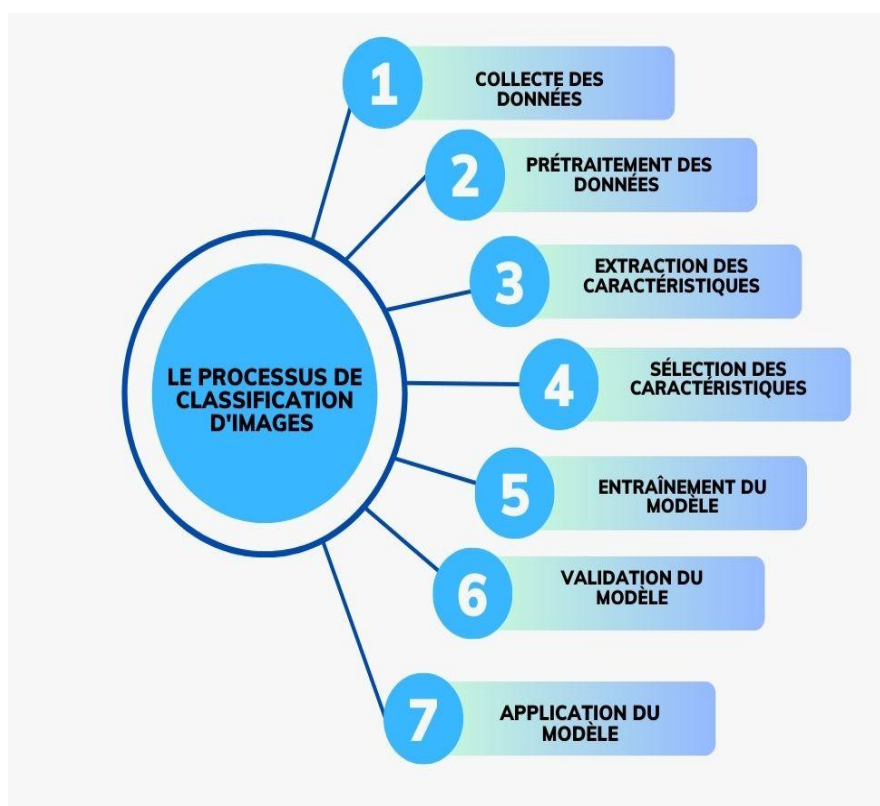


Figure 12 : le processus de classification d'images

voisins) :

L'algorithme des k plus proches voisins, également connu sous le nom de KNN ou k-NN, est un discriminant d'apprentissage supervisé non paramétrique, qui utilise la proximité pour effectuer des classifications ou des prédictions sur le regroupement d'un point de données individuel. Bien qu'il puisse être utilisé pour des problèmes de régression ou de classification, il est généralement utilisé comme algorithme de classification, en partant de l'hypothèse que des points similaires peuvent être trouvés les uns à côté des autres.

L'approche de classification KNN se base sur l'hypothèse que chaque cas de l'échantillon d'apprentissage est un vecteur aléatoire issu de R^n . Chaque point est décrit comme :

$x = [a_1(x), a_2(x), a_3(x), \dots, a_n(x)]$ correspond à la valeur I du n eme attribut. $a_n(x)$ peut être soit une variable quantitative soit une variable qualitative. Fin de déterminer la classe d'un point cible, chaque chacun des k points les plus proches de x_q procèdent à un vote. La classe de x_q correspond à la classe majoritaire. [34]

2.7.1 Le Principe général de KNN :

KNN ne construit pas de modèle prédictif comme il est le cas pour les algorithmes de régression. Cependant, il n'y a pas de phases d'apprentissage puisqu'il classe directement des points dont la classe est inconnue en fonction de leurs distances par rapport à des points appartenant à une classe connue auparavant. Cet algorithme est ainsi classé dans la catégorie de lazy Learning. [35]

L'algorithme KNN agit en deux étapes :

- **1ère étape :**

L'algorithme en premier lieu calcule la distance et stocke les k voisins plus proches pour chaque échantillon de l'ensemble d'apprentissages.

- **2ème étape :**

Ensuite pour un échantillon sans étiquette l'algorithme récupère les k voisins plus proches de celui-ci et prédit sa classe par vote.

2.7.2 Fonctionnement de KNN :

La méthode KNN fonctionne en comparant un point donné à ses plus proches voisins et en déterminant la classe à laquelle il appartient. Le principe de base de KNN est le suivant : pour chaque point donné, on calcule la distance entre ce point et tous les autres points dans le jeu de données. Les k plus proches voisins sont ensuite identifiés, et la classe à laquelle le point appartient est déterminée en fonction des classes des k plus proches voisins. La classe majoritaire parmi les k plus proches voisins est alors attribuée au point donné. KNN peut être appliqué à des jeux de données contenant des variables continues ou catégorielles. Il peut également être utilisé pour prédire des valeurs continues, telles que les prix des maisons ou les salaires, en se basant

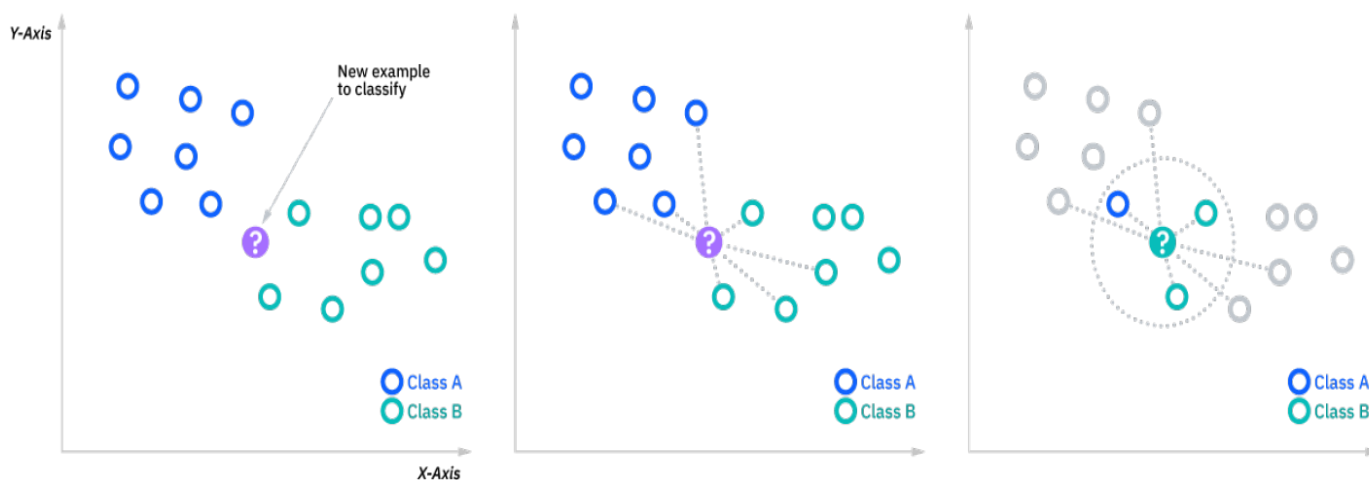


Figure 13 : le fonctionnement de KNN [36]

sur les valeurs connues d'autres points du jeu de données. [36]

2.7.3 Les avantages et les inconvénients de l'algorithme KNN :

Avantages :

- L'algorithme est simple et facile à mettre en œuvre.
- Aucune hypothèse sur les données (linéaires, affines,...)
- L'algorithme est polyvalent. Il peut être utilisé pour la classification, la régression.[37]

Inconvénients :

- L'algorithme devient beaucoup plus lent à mesure que le nombre d'exemples d'apprentissage augmente.
- Le choix de la méthode de calcul de la distance ainsi que le nombre de voisins K peut ne pas être évident
- L'étape de prédiction peut-être lente. La complexité est de l'ordre de $O(n)$ avec ($k \ll n$).
[37]

2.8 Machine à vecteurs de support (SVM) :

Les machines à vecteurs de support ou séparateurs à vaste marge (en anglais support-vector machine, SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. Les SVM sont une généralisation des classifieurs linéaires. Les SVM représentent **une famille d'algorithmes de Machine Learning** capables de résoudre un problème de discrimination, de régression mathématique ou de détection d'anomalie. Les premières machines à vecteurs de support ont été développées dans les années 1990 sur la base des concepts des informaticiens russes Vladimir Vapnik et Alexey Chervonenkis. Ils ont développé une nouvelle théorie statistique de l'apprentissage automatique connue sous le nom de la théorie de Vapnik-Chervonenkis. [38]

Très rapidement, leurs travaux ont suscité beaucoup d'engouement grâce à leur capacité à gérer des données de grandes dimensions. De plus, les SVM ne requièrent que peu de paramètres. Dans l'immense majorité des cas, l'utilisation du SVM ne nécessite que deux hyperparamètres : le choix de la technique de régularisation et celui du noyau.

Les séparateurs à vaste marge ont été développés dans les années 1990 à partir des considérations théoriques de Vladimir Vapnik sur le développement d'une théorie statistique de l'apprentissage : la théorie de Vapnik-Chervonenkis. Ils ont rapidement été adoptés pour leur capacité à travailler avec des données de grandes dimensions, le faible nombre d'hyperparamètres, leurs garanties théoriques, et leurs bons résultats en pratique.

Les SVM ont été appliqués à de très nombreux domaines (bio-informatique, recherche d'information, vision par ordinateur, finance¹...). Selon les données, la performance des machines à vecteurs de support est de même ordre, ou même supérieure, à celle d'un réseau de neurones ou d'un modèle de mélanges gaussiens

2.8.1 Fonctionnement de SVM :

Le Support Vector Machine (SVM) est une méthode d'apprentissage supervisé utilisée pour la classification et la régression. Il est basé sur le concept de séparation des données en classes distinctes par un hyperplan séparateur. [39]

Cette méthode est largement utilisée pour résoudre des problèmes de classification et de régression, car elle offre une précision élevée et une grande robustesse. Le SVM fonctionne en trouvant un hyperplan qui sépare les données en classes distinctes. L'hyperplan est déterminé en minimisant l'erreur de classification et en maximisant la marge entre les classes. La marge est la distance entre le plus proche point d'une classe à l'hyperplan séparateur. Une fois que l'hyperplan a été trouvé, il peut être utilisé pour classer les nouvelles données. Le SVM peut être appliquée à des données linéaires ou non-linéaires, selon le type de noyau utilisée. Les noyaux communs incluent le noyau linéaire, le noyau polynomial et le noyau radial (RBF). [40]

Le choix du bon type de noyau dépend du type de problème à résoudre et des caractères des données d'entrée. Le SVM peut également être appliquée à des problèmes multi-classes, où plusieurs classes doivent être distinguées par un seul modèle. Pour ce faire, plusieurs hyperplans

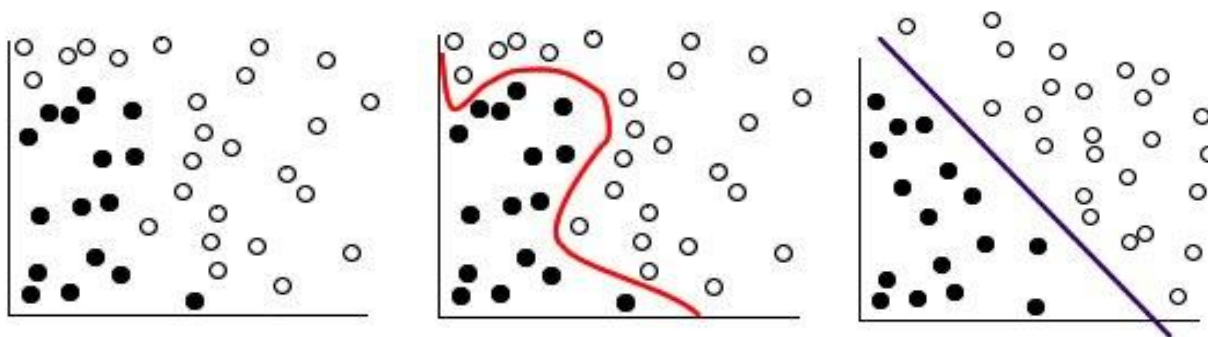


Figure 14 : fonctionnement de SVM

sont construits pour chaque classe afin de pouvoir distinguer chaque classe individuellement.

2.8.2 Notions de base (l'hyperplan, la marge, Les vecteurs de support, les classes) :

2.8.2.1 L'hyperplan :

Le rôle des machines à vecteur de support est de résoudre le problème ci-après : trouver une ligne qui assure **une séparation et une classification linéaire** de l'ensemble de données dont on dispose. Cette ligne est appelée l'hyperplan.

Plus les différents points de données ne sont éloignés de l'hyperplan, mieux c'est. Cela explique d'ailleurs pourquoi un algorithme de classification SVM cherchera toujours à placer la frontière aussi loin que possible des points noirs, mais également des points blancs. Cette frontière optimale qui est la plus éloignée de tous les points a une meilleure capacité de généralisation. On comprend ainsi mieux le nom donné aux points les plus proches de l'hyperplan à qui on doit également l'appellation de SVM : vecteur de support. Ils supportent d'une certaine manière la frontière.

Par ailleurs, les SVM ne s'appliquent pas uniquement à un espace à deux dimensions. Pour certains jeux de données par exemple, une séparation linéaire dans un espace à deux dimensions n'est pas envisageable. Afin de résoudre ce problème, le SVM cherche à mapper les données dans des dimensions de plus en plus grandes jusqu'à trouver l'hyperplan qui pourra les séparer de façon optimale. Le plus souvent, pour passer d'une dimension à une autre pour les données, on utilise **une fonction noyau ou kernel function** qui sert alors d'intermédiaire.

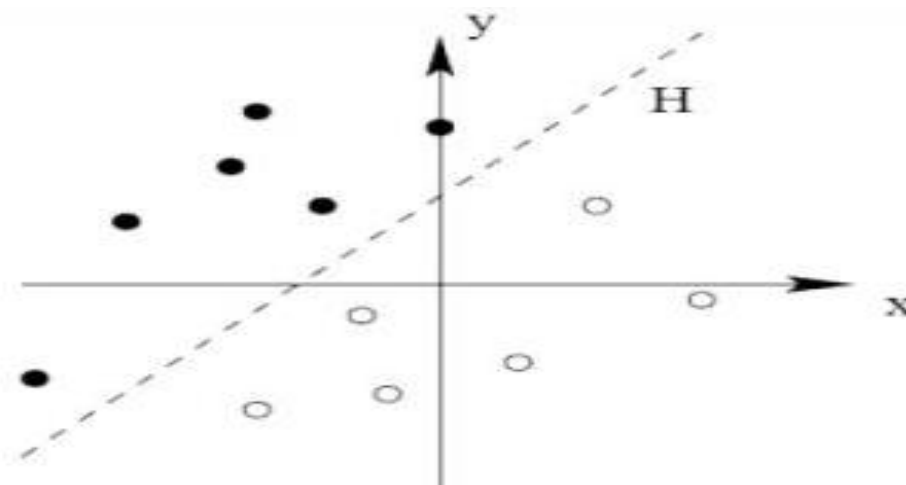


Figure 15 : kernel function

2.8.2.2 La marge :

La marge est la distance entre l'hyperplan et les points de données les plus proches de chaque classe. L'objectif de SVM est de maximiser la marge, c'est-à-dire de trouver l'hyperplan qui sépare les données en deux classes avec la plus grande distance entre l'hyperplan et les points de

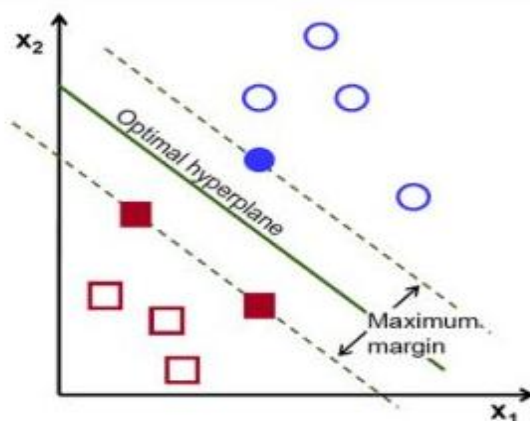


Figure 16 : la marge est la distance

données les plus proches. Cette marge maximale permet de mieux généraliser le modèle aux données inconnues.

2.8.2.3 Les vecteurs de support :

Les vecteurs de support déterminent les points les plus proches de la frontière. Dans un ensemble de données, si ceux-ci étaient supprimés, ils modifieraient la position de l'hyperplan en division. De ce fait, les vecteurs de support sont considérés comme des éléments cruciaux dans tout jeu de données.

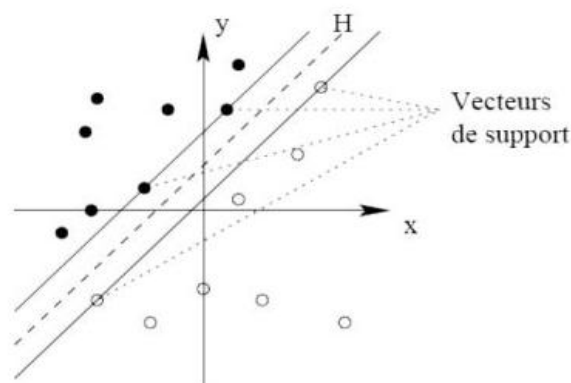


Figure 17 : les vecteurs de support

2.8.2.4 Les classes :

Les classes dans SVM sont les deux catégories de données que l'on cherche à séparer à l'aide de l'hyperplan. SVM est un algorithme de classification binaire, ce qui signifie qu'il peut séparer les données en deux classes seulement. Pour des problèmes de classification à plusieurs classes, des approches comme la classification en un contre tous ou la classification en un contre un peuvent être utilisées.

2.8.3 Le problème de discrimination linéaire :

Un problème de discrimination est dit linéairement séparable lorsqu'il existe une fonction de décision linéaire (appelé aussi séparateur linéaire), de la forme $D(x) = \text{signe } f(x)$ avec $f(x) = v \cdot x + a$, $v \in \mathbf{R}^p$ et $a \in \mathbf{R}$, classant correctement toutes les observations de l'ensemble d'apprentissage ($D(x_i) = y_i, i \in [1, n]$).

La fonction f est appelée fonction caractéristique. C'est un problème particulier qui semble très spécifique, mais qui permet d'introduire de manière pédagogique les principaux principes des SVM : marge, programmation quadratique, vecteur support, formulation duale et matrice de gram. Nous allons ensuite généraliser au cas des observations non séparables et non linéaires par l'introduction de variables d'écart et de noyaux. Ces différents types de problèmes sont illustrés figure 18.

A toute fonction de décision et donc aux fonction de décision linéaire ont peut associer une frontière de décision : $\Delta(v, a) = \{x \in \mathbb{R}^p \mid v \cdot x + a = 0\}$

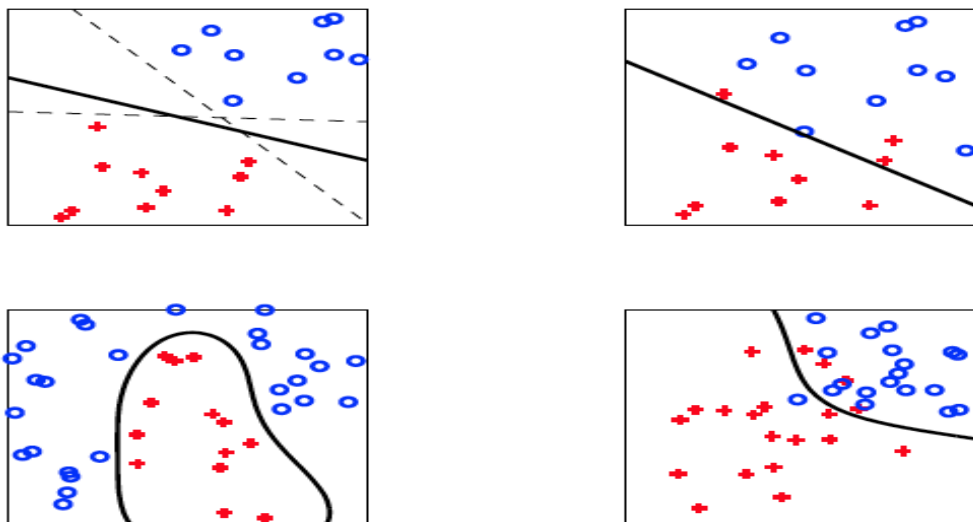
2.8.3.1 La marge d'un classifieur :

Pour un échantillon donné, il est possible d'associer deux marges à un même classifieur linéaire : sa marge géométrique et sa marge numérique. La marge géométrique m d'un échantillon linéairement séparable est donnée par la plus petite distance d'un point de l'échantillon à la frontière de décision. La marge numérique μ est donnée elle par la plus petite valeur de la fonction de décision, atteinte sur un point de l'échantillon. Leur définition mathématique est :

Figure 18 : les types de problèmes de discrimination binaire

<p>marge géométrique</p> $m = \min_{i \in [1, n]} \text{dist}(x_i, \Delta(v, a))$	<p>marge numérique</p> $\mu = \min_{i \in [1, n]} v^T x_i + a \dots \dots \dots (1)$
--	---

2.8.3.2 Cas séparable :



- Notion de marge :

Dans le cas séparable, on va considérer les points les plus près de l'hyperplan séparateur : vecteurs supports (support vectors).

Pour tout point de l'espace des exemples, la distance à l'hyperplan séparateur est donnée

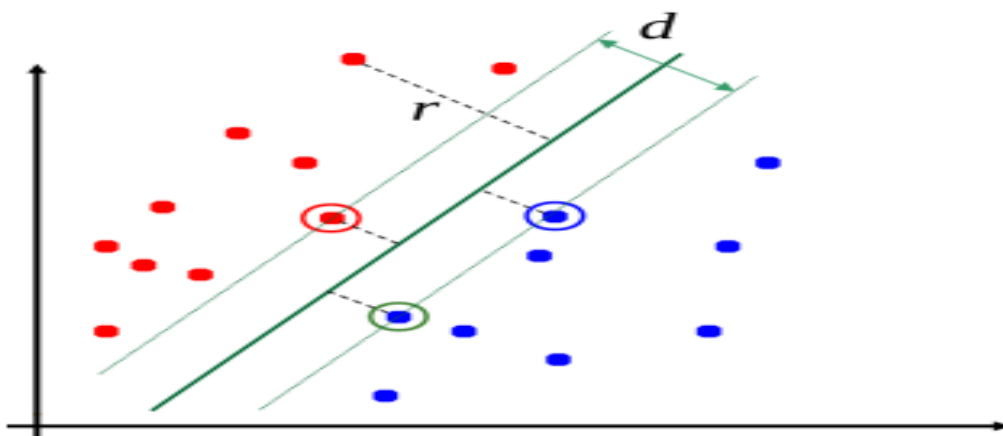


Figure 19 : Marges maximale pour séparer deux classes

Par :
$$r = \frac{|\mathbf{w} \cdot \mathbf{x} + b|}{\|\mathbf{w}\|} \dots\dots\dots(2)$$

• **Quantification de la marge :**

Pour limiter l'espace des possibles on considère que les points les plus proches sont situés sur les hyperplans canoniques donnés par : $\mathbf{w} \cdot \mathbf{x} + b = \pm 1$

Dans ce cas, la marge est :
$$d = \frac{2}{\|\mathbf{w}\|} \dots\dots\dots(3)$$

Les conditions d'une bonne classification sont :

$$\begin{cases} \mathbf{w} \cdot \mathbf{x} + b \geq 1, & \text{si } y_i = 1 \\ \mathbf{w} \cdot \mathbf{x} + b < -1, & \text{si } y_i = -1 \end{cases} \dots\dots\dots(4)$$

- Maximisation de la marge :

Le problème revient alors à trouver w et b tels que $d = \frac{2}{\|w\|}$ est maximale $\forall (x_i, y_i)$

Sous les contraintes :

$$\begin{cases} \mathbf{w} \cdot \mathbf{x} + \mathbf{b} \geq 1, & \text{si } y_i = 1 \\ \mathbf{w} \cdot \mathbf{x} + \mathbf{b} < 1, & \text{si } y_i = -1 \end{cases} \dots\dots\dots (6)$$

De manière équivalent, le problème peut s'écrire plus simplement comme la minimisation de :

$$\frac{1}{2} \| \mathbf{w} \|^2 \dots\dots\dots (7)$$

Sous les contraintes : $y_i(\mathbf{w} \cdot \mathbf{x} + \mathbf{b}) \geq 1, \forall i \in [1, N] \dots\dots\dots (8)$

Cette minimisation est possible sous les conditions dites de “**Karush-Kuhn-Tucker (KKT)**”

Soit le Lagrangien L : $\mathcal{L}(\mathbf{w}, \mathbf{b}, \lambda) = \frac{1}{2} \| \mathbf{w} \|^2 - \sum_{i=1}^N \lambda_i [y_i(\mathbf{w} \cdot \mathbf{x}_i + \mathbf{b}) - 1] \dots\dots (9)$

Les conditions de KKT sont alors :

$$\begin{aligned} \frac{\mathcal{L}}{\partial \mathbf{w}} = 0, \frac{\partial \mathcal{L}}{\partial \mathbf{b}} = 0, \frac{\partial \mathcal{L}}{\partial \lambda_j} \geq 0, \lambda_j \geq 0 \\ \lambda_i [y_i(\mathbf{w} \cdot \mathbf{x}_i + \mathbf{b}) - 1] = 0 \end{aligned} \dots\dots\dots (10)$$

Par ailleurs la dernière condition implique que pour tout point ne vérifiant pas $y_i(\mathbf{w} \cdot \mathbf{x}_i + \mathbf{b}) = 1$ le λ_i est nul.

Les points qui vérifient $y_i(\mathbf{w} \cdot \mathbf{x}_i + \mathbf{b}) = 1$, sont appelés “**vecteurs supports**”. Ce sont les points les plus près de la marge. Ils sont censés être peu nombreux par rapport à l'ensemble des exemples.

- Le problème dual :

Le problème s'exprime sous forme duale comme la minimisation de :

$$W(\lambda) = \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j (x_i \cdot x_j) \dots\dots\dots(11)$$

Fait partie des problèmes d'optimisation quadratique pour lesquels il existe de nombreux algorithmes de résolution

SMO : résolution analytique (par 2 points), gestion efficace de la mémoire, mais converge en un nombre d'étapes indéterminé

Simple SVM : facilite de la reprise à chaud, converge en moins d'étapes mais limitation mémoire

LA SVM : utilisation en ligne, résolution analytique mais solution sous optimale, plusieurs passes nécessaires pour les petites bases de donnée

2.8.3.3 Cas non-séparable :

Souvent il arrive que même si le problème est linéaire, les données sont affectées par un bruit (par ex. bruit de capteur) et les deux classes se retrouvent mélangées autour de l'hyperplan de séparation. Pour gérer ce type de problème on utilise une technique dite de marge souple, qui tolère les mauvais classements [41] :

- Rajouter des variables de relâchement des contraintes ξ_i ,
- Pénaliser ces relâchements dans la fonction objectif.

L'intuition est de modéliser les erreurs potentielles par des variables d'écart positives ξ_i associées aux observations $(x_i, y_i), i=1, \dots, n$

Si un point (x_i, y_i) vérifie la contrainte de marge $y_i(w^T x_i + b) \geq 1$ alors la variable d'écart (qui est une mesure du coût de l'erreur) est nulle.[42]

Nous avons donc deux situations :

- Pas d'erreur : $y_i(w^T x_i + b) \geq 1 \implies \xi_i = 0, \dots\dots\dots(12)$

▪ Erreur : $y_i(w^T x_i + b) < 1 \implies \xi_i = 1 - y_i(w^T x_i + b) > 0 \dots\dots(13)$

On associe à cette définition une fonction coût appelée « coût charnière » :

$$\xi_i = \max(0, 1 - y_i(w^T x_i + b)) \dots\dots\dots(14)$$

2.8.4 Le cas non linéaire

Dans le cas non-séparable, nous introduisons des variables d'écart et permettons que certains points soient mal classés. Dans le cas non linéaire, nous remplaçons le produit interne par un noyau $K(x_i, x_j)$ défini positif, où x_i et x_j sont les observations. La fonction noyau fournit une mesure de similarité entre deux observations, et définit implicitement une projection $\phi(x)$ des observations dans un espace de haute dimension, où elles peuvent être séparées par un hyperplan linéaire.

Cette technique de kernel trick est un aspect clé des SVM non linéaires et permet de traiter des problèmes de classification non linéaires en utilisant des fonctions noyau appropriées pour projeter les données dans un espace de dimension supérieure, où elles peuvent être séparées linéairement. [43]

2.8.4.1 L'astuce du noyau :

L'astuce du noyau est une technique pour ajuster des modèles non linéaires en transformant implicitement les observations dans un espace de grande dimension, où des modèles linéaires peuvent être utilisés. Cela est réalisé en utilisant une fonction noyau $k(x, x')$ qui calcule le produit scalaire entre les images de deux observations dans l'espace de grande dimension, sans avoir besoin de calculer explicitement la transformation.

L'astuce du noyau permet d'appliquer l'algorithme SVM à des problèmes non linéaires en projetant les données dans un espace de grande dimension de manière implicite, en utilisant une fonction noyau appropriée qui calcule la similarité entre les paires de données dans cet espace de grande dimension. Cette technique est utile car elle évite de devoir effectuer explicitement la transformation des données dans cet espace de grande dimension, ce qui peut être très coûteux en termes de calculs. [44]

2.8.4.2 Fonctions du noyau :

- Linéaire :

Le noyau linéaire calcule la similarité entre deux vecteurs d'entrée en effectuant simplement leur produit scalaire, sans effectuer de transformation dans un espace de caractéristiques de haute dimension. La formule mathématique pour le noyau linéaire est la suivante :

$$f(X) = w^t \cdot X + b \dots\dots\dots(15)$$

Dans cette équation, w est le vecteur de poids que vous souhaitez minimiser, X est les données que vous essayez de classer et b est le coefficient linéaire estimé à partir des données d'apprentissage. Cette équation définit la limite de décision renvoyée par la SVM.

Polynôme :

Le noyau polynomiale calcule la similarité entre deux vecteurs d'entrée en les projetant dans un espace de caractéristiques de plus haute dimension à l'aide d'une fonction polynomiale. La formule mathématique pour le noyau polynomiale est la suivante :

$$f(X1, X2) = (a + X1^t * X2)^b \dots\dots\dots(16)$$

Figure 20 : noyau linéaire [43]

C'est l'une des équations de noyau polynomiales les plus simples que vous puissiez utiliser.

SVC with polynomial (degree 3) kernel

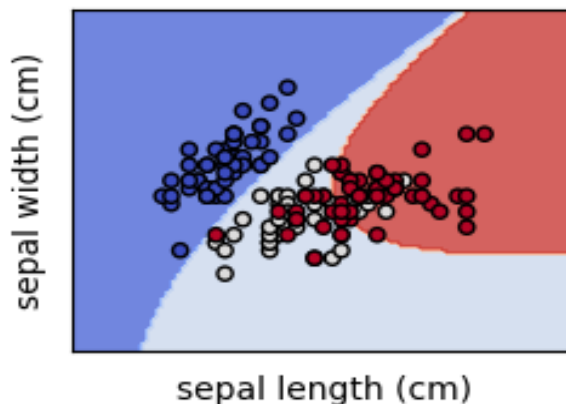
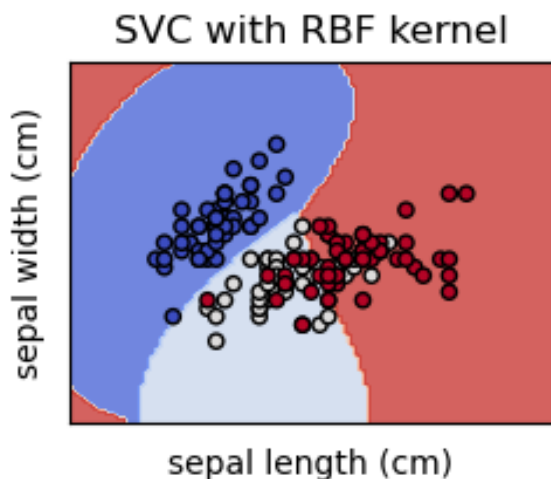


Figure 21 : noyau polynomial degré 3 [43]

Classification

$f(X1, X2)$ représente la limite de décision polynomiale qui séparera vos données. X1 et X2 représentent vos données



Radial-basis function (RBF):

Le noyau RBF calcule la similarité entre deux vecteurs d'entrée en les projetant dans un espace de caractéristiques de plus haute dimension à l'aide d'une fonction gaussienne. La formule mathématique pour le noyau RBF est la suivante :

$$f(X1, X2) = \exp(- \text{gamma} \cdot \| X1 - X2 \|^2) \dots \dots \dots (17)$$

Dans cette équation, gamma spécifie combien un seul point d'entraînement a sur les autres points de données qui l'entourent. $\|X1 - X2\|$ est le produit scalaire entre vos caractéristiques.

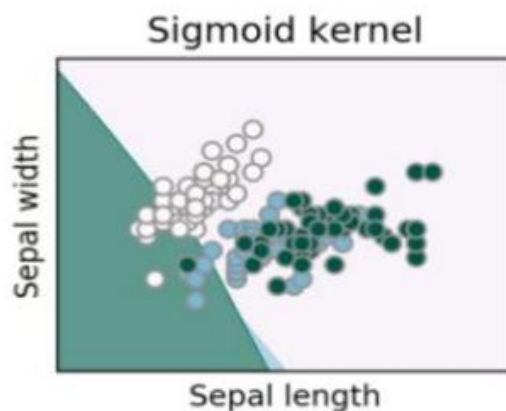
Figure 22 : radial basis function [43]

- **Noyau Tangent Hyperbolique- (sigmoïde) :**

Le noyau sigmoïd calcule la similarité entre deux vecteurs d'entrée en les projetant dans un espace de caractéristiques de plus haute dimension à l'aide d'une fonction sigmoïde. La formule mathématique pour le noyau sigmoïd est la suivante :

$$k(x, y) = \tanh(\alpha x^t y + c) \dots \dots \dots (18)$$

Il est intéressant de noter qu'un modèle SVM utilisant une fonction de noyau sigmoïde équivaut à un réseau de neurones perceptron à deux couches. Ce noyau était très populaire pour les machines à vecteurs de support en raison de son origine dans la théorie des réseaux de neurones. De plus, bien qu'il ne soit défini que positivement sous condition, il s'est avéré qu'il



fonctionnait bien dans la pratique. [45]

2.8.5 L'Utilité de SVM :

Le SVM est utilisé pour les problèmes de classification de texte telles que l'attribution de catégorie, la détection du spam ou encore l'analyse des sentiments. Ils sont également couramment utilisés pour les problèmes de reconnaissance d'image, particulièrement en reconnaissance de forme et en classification de couleur. SVM joue également un rôle essentiel dans de nombreux domaines de la reconnaissance manuscrite des symboles, tels que les services d'automatisation postale. [46]

2.8.6 Avantages des machines à vecteurs de support :

- Le SVM est très efficace pour moins de trois dimensions, ce qui peut être utilisé pour séparer les hyperplans de l'espace dans plus de quatre dimensions.
- Le SVM est robuste et peut être appliqué à des données non linéaires en utilisant des fonctions noyau.
- Le SVM est très efficace pour les données qui ont un grand nombre de variables et un petit nombre d'observations, ce qui fait partie de l'utilité pour les problèmes de classification et les complexes de régression.

Figure 23 : Noyau Tangent Hyperbolique [43]

Classification

- Le SVM est très efficace pour le prix dans le compte des données manquantes, car il peut généraliser ses résultats d'autant plus que certaines observations sont manquantes ou incomplètes.
- Le SVM est capable de générer des modèles prédictifs plus précis que d'autres algorithmes, ce qui le rend particulièrement utile pour les applications d'apprentissage automatique complexes telles que la reconnaissance faciale et vocale, l'analyse du langage naturel et la classification des images médicales.

2.8.7 Inconvénients des machines à vecteurs de support :

- Les algorithmes de SVM peuvent être très sensibles aux paramètres et aux choix initiaux, ce qui peut entraîner une mauvaise performance si ces paramètres ne sont pas correctement ajustés ou optimisés.
- Les algorithmes SVM prennent beaucoup de temps à entraîner, car ils doivent trouver l'hyperplan optimal parmi un grand nombre d'options possibles, ce qui peut prendre beaucoup de temps sur des jeux de données volumineux ou complexes.

Les algorithmes SVM ne s'adaptent pas aux donneurs jeux non linéaires ou non séparables, mais ils ne permettent pas une planification hyper-optimale optimale pour les types de dons donneurs.

2.9 L'optimisation par essais de particules (PSO) :

L'optimisation par essaims de particules (PSO - Particle Swarm Optimization en anglais) est une méthode d'optimisation stochastique inspirée par le comportement de certains essaims d'oiseaux et d'insectes. L'objectif de la PSO est de trouver la meilleure solution possible à un problème d'optimisation en cherchant dans l'espace des solutions possibles.

Le principe de base de la PSO est de faire évoluer un ensemble de "particules" dans l'espace des solutions, chacune représentant une solution potentielle. Chaque particule se déplace dans l'espace en fonction de sa propre expérience et de celle des autres particules de l'ensemble. À chaque itération, la position de chaque particule est mise à jour en fonction de sa position courante, de sa vitesse et de la meilleure position trouvée jusqu'à présent par la particule elle-même (meilleure solution personnelle) et par l'ensemble des particules (meilleure solution globale).

Le processus d'optimisation de la PSO peut être décrit en plusieurs étapes :

1. **Initialisation** : création de l'ensemble de particules, avec des positions et des vitesses initiales aléatoires.
2. **Évaluation** : chaque particule est évaluée en fonction de sa position dans l'espace des solutions.
3. **Mise à jour de la meilleure solution personnelle** : chaque particule garde en mémoire sa meilleure solution personnelle trouvée jusqu'à présent.
4. **Mise à jour de la meilleure solution globale** : l'ensemble des particules garde en mémoire la meilleure solution trouvée parmi toutes les particules.
5. **Mise à jour des positions et des vitesses des particules** : chaque particule calcule sa nouvelle position et sa nouvelle vitesse en fonction de sa position courante, de sa vitesse, de sa meilleure solution personnelle et de la meilleure solution globale.
6. **Critère d'arrêt** : le processus d'optimisation se poursuit tant que les critères d'arrêt ne sont pas satisfaits, par exemple le nombre d'itérations maximum ou la convergence des solutions.

La PSO est une méthode d'optimisation très populaire en raison de sa simplicité, de sa rapidité et de son efficacité pour de nombreux types de problèmes d'optimisation. Elle est souvent utilisée pour optimiser des fonctions non linéaires, continues et non convexes, ainsi que pour des problèmes de classification, de régression et de clustering .[47]

2.9.1 Les paramètres de PSO :

Les paramètres de l'optimisation par essaim de particules (PSO) sont des valeurs qui influencent le comportement de l'algorithme et doivent être choisis soigneusement pour obtenir de bons résultats d'optimisation. Voici une liste de quelques paramètres importants de la PSO :

1. **Nombre de particules** : le nombre de particules dans l'ensemble, qui détermine la diversité de l'ensemble et l'exploration de l'espace des solutions.
2. **Nombre d'itérations** : le nombre d'itérations de l'algorithme avant l'arrêt, qui détermine la durée de l'optimisation.
3. **Coefficient d'inertie** : un paramètre qui détermine la contribution de la vitesse actuelle de chaque particule à sa vitesse future. Une valeur élevée favorise l'exploration de

Classification

l'espace des solutions, tandis qu'une valeur faible favorise l'exploitation des zones les plus prometteuses.

4. **Coefficient de cognition** : un paramètre qui détermine l'importance de la meilleure solution personnelle de chaque particule dans le calcul de sa nouvelle position. Une valeur élevée favorise l'exploitation des solutions déjà connues, tandis qu'une valeur faible favorise l'exploration de nouvelles solutions.
5. **Coefficient de socialisation** : un paramètre qui détermine l'importance de la meilleure solution globale dans le calcul de la nouvelle position de chaque particule. Une valeur élevée favorise l'exploration de l'espace des solutions, tandis qu'une valeur faible favorise l'exploitation des solutions déjà connues.
6. **Vitesse maximale et minimale** : des bornes sur la vitesse de chaque particule pour éviter des sauts trop importants dans l'espace des solutions.
7. **Facteur d'accélération** : un paramètre qui détermine la quantité de changement de vitesse maximal autorisée à chaque itération, pour éviter des sauts trop importants dans l'espace des solutions.

Le choix des paramètres dépend du problème d'optimisation considéré, et doit être fait en fonction des connaissances a priori et des résultats empiriques pour trouver les valeurs les plus appropriées. [48]

2.9.2 Les avantages de L'optimisation par essaims de particules (PSO) :

- **Trouver un meilleur ensemble de paramètres :**

PSO peut aider à trouver les meilleurs paramètres pour SVM, tels que les paramètres de régularisation et les paramètres du noyau. Cela peut améliorer la performance de la classification et aider à éviter le surapprentissage.[49]

- **Éviter les minima locaux :**

Classification

PSO est une méthode d'optimisation globale qui peut aider à éviter de tomber dans des minima locaux lors de l'optimisation de SVM. Cela peut conduire à une meilleure performance de classification et à des modèles plus robustes.[50]

- **Réduire le temps de calcul :**

PSO peut être utilisé pour optimiser les paramètres de SVM plus efficacement que les méthodes d'optimisation traditionnelles telles que la recherche en grille. Cela peut réduire considérablement le temps de calcul nécessaire pour entraîner un modèle SVM.[51]

- **Flexibilité :**

PSO peut être appliqué à différentes formes de SVM, y compris les SVM linéaires et non linéaires. Cela permet une plus grande flexibilité lors de la modélisation de différents types de données.

- **Facilité d'utilisation :**

PSO est une méthode d'optimisation relativement simple à implémenter et à utiliser. Il ne nécessite pas de connaissances en optimisation avancées, ce qui le rend accessible à un large public.

2.9.3 SVM et PSO :

SVM-PSO est une méthode d'optimisation qui utilise une combinaison de l'algorithme d'optimisation par essaim de particules (PSO) et la technique de classification SVM (Support Vector Machine) pour optimiser les paramètres du SVM. L'objectif est d'améliorer la précision de la classification SVM en optimisant les paramètres de manière efficace et efficace.

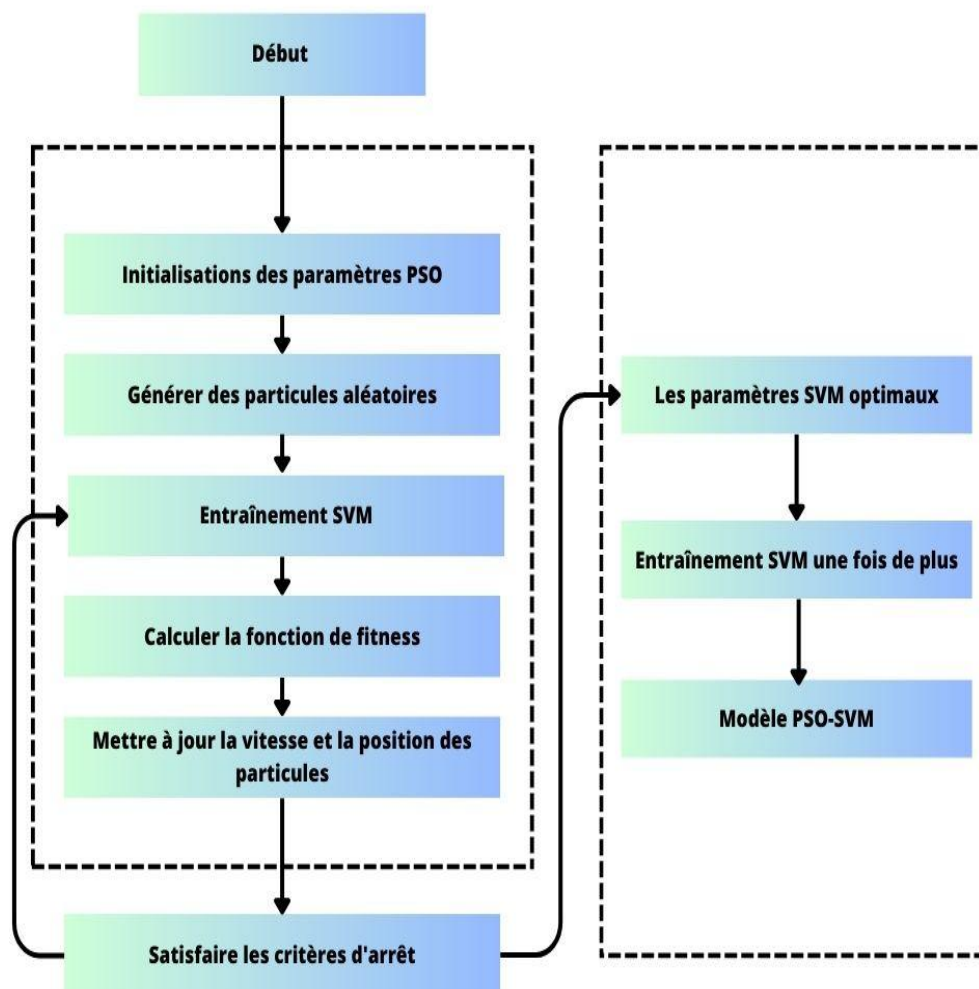


Figure 24 : La structure de fonctionnement de PSO avec le SVM

2.10 L'évaluation de performance :

L'évaluation de performance des SVM se fait généralement en utilisant des mesures telles que l'exactitude (accuracy), la précision (precision), le rappel (recall) et la F-mesure (F-measure) pour les problèmes de classification, et l'erreur quadratique moyenne (mean squared error, MSE) et le coefficient de détermination (coefficient of determination, R²) pour les problèmes de régression. [52]

Dans le cas de la classification binaire, l'exactitude mesure le pourcentage d'échantillons correctement classés par le modèle, tandis que la précision mesure le nombre de vrais positifs (échantillons correctement identifiés comme positifs) divisé par le nombre total de prédictions positives. Le rappel mesure la proportion de vrais positifs qui ont été correctement identifiés

Classification

parmi tous les échantillons positifs. La F-mesure combine la précision et le rappel pour fournir une mesure plus complète de la performance du modèle.

Pour les problèmes de classification multi-classes, les mesures d'évaluation sont souvent étendues pour prendre en compte toutes les classes. Par exemple, l'exactitude peut être calculée en considérant le nombre total de prédictions correctes pour toutes les classes.

Pour les problèmes de régression, l'erreur quadratique moyenne mesure la différence entre les valeurs prédites et les valeurs réelles, en prenant la moyenne des carrés de ces écarts. Le coefficient de détermination mesure la proportion de la variance totale de la variable cible qui est expliquée par le modèle.

Il est important de noter que le choix des métriques dépend du type de problème et des objectifs spécifiques de l'analyse. En outre, il est souvent recommandé d'utiliser plusieurs métriques pour avoir une vision complète de la performance du modèle. [53]

2.10.1 Matrice de confusion :

La matrice de confusion. Cette dernière correspond à un résumé tabulaire du nombre de prédictions correctes et non correctes, faites par le modèle. Dans cette matrice, chaque ligne correspond à une classe réelle et chaque colonne correspond à une classe estimée. [54]

Elle inclut les valeurs suivantes :

- **Vrais positifs (ou True Positive, TP)** soit lorsque la classe réelle et la classe estimée sont toutes les deux positives
- **Vrais négatifs (ou True Negative, TN)** soit lorsque la classe réelle et la classe estimée sont toutes les deux négatives
- **Faux positifs (ou False Positive, FP)** soit lorsque la classe réelle est négative mais que la classe estimée est positive. On appelle ceci une erreur de Type 1.
- **Faux négatifs (ou False Negative, FN)** soit lorsque la classe réelle est positive mais que la classe estimée est négative. On appelle ceci une erreur de Type 2.

Dans le cas d'une classification binaire, la matrice de confusion sera une matrice de 2 par 2, avec quatre valeurs, comme dans la figure suivante :

		TRUE CLASS	
		POSITIVE	NEGATIVE
PREDICTED CLASS	POSITIVE	<p>TP TRUE POSITIVE</p>	<p>FP FALSE POSITIVE</p>
	NEGATIVE	<p>FN FALSE NEGATIVE</p>	<p>TN TRUE NEGATIVE</p>

Figure 25 : matrice de confusion

- **Accuracy (Exactitude) :**

La accuracy correspond au nombre de prédictions correctes faites par le modèle. Elle représente le ratio entre le nombre de prédictions correctes et le nombre total de prédiction. Ceci peut être calculé en utilisant les valeurs de la matrice de confusion et en utilisant la formule suivante :

$$\text{Accuracy (Exactitude)} = \frac{TP + TN}{TP + FP + TN + FN}$$

Cette mesure est utilisée lorsque le nombre de True Positive et de True Negative sont les plus important.

- **Précision (ou Precision) :**

La précision est une mesure qui représente le nombre d'éléments correctement classés par le modèle. Elle est calculée en divisant le nombre de classifications positives correctes par le nombre total de prédictions positives effectuées par le modèle. La formule pour calculer la précision est la suivante :

$$\text{Précision} = \frac{TP}{TP + FP}$$

Cette mesure est utilisée lorsque le nombre de False Positive est le plus élevé.

- **Rappel (ou Recall)**

Le rappel détermine la proportion des valeurs positives qui ont été prédites avec précision. Cette mesure correspond donc au ratio entre le nombre de prédictions positives correctes et le nombre total de classifications de classe positive. On peut utiliser la formule suivante :

$$\text{Rappel} = \frac{TP}{TP + FN}$$

On utilise cette mesure lorsque le nombre de False Negative est le plus important.

- **Spécificité (ou Specificity) :**

La spécificité correspond au nombre de classes négatives prédites par le modèle. Cette mesure est déterminée par le ratio entre le nombre de prédictions négatives correctes et le nombre total de prédictions négatives. Elle peut se calculer de la manière suivante :

$$\text{Spécificité} = \frac{TN}{TN + FP}$$

- **Score F1 (ou F1 score) :**

Le score F1 est une mesure qui combine les mesures de rappel et de précision. Elle est utilisée lorsque la distinction entre ces deux mesures n'est pas claire ou lorsque les erreurs de type False Negative et False Positive sont d'une importance égale.

$$\text{F1 Score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

2.10.2 La surface sous la courbe ROC (AUC, Area Under the Curve) :

Est une mesure de performance couramment utilisée pour évaluer la capacité d'un modèle de classification binaire à distinguer les échantillons positifs des échantillons négatifs. Elle représente la surface entre la courbe ROC et l'axe diagonal du graphique, qui correspond à une performance aléatoire.

L'AUC mesure la capacité du modèle à classer correctement les échantillons positifs avant les échantillons négatifs, quelle que soit la valeur de seuil de classification. Elle varie de 0 à 1, où une valeur de 1 indique une performance de classification parfaite et une valeur de 0,5 indique une performance aléatoire. [55]

En d'autres termes, l'AUC représente la probabilité que le modèle classe correctement un échantillon positif au hasard par rapport à un échantillon négatif au hasard. Plus l'AUC est élevée, meilleure est la capacité du modèle à discriminer les échantillons positifs des échantillons négatifs.

Les règles pour évaluer la performance des classificateurs en utilisant la courbe ROC et l'AUC sont les suivantes :

- Plus l'AUC est proche de 1, meilleure est la performance du classificateur.
- Si l'AUC est supérieure à 0,5, le classificateur est meilleur que le hasard, c'est-à-dire qu'il est capable de distinguer les échantillons positifs des échantillons négatifs.
- Si l'AUC est égale à 0,5, le classificateur est équivalent à une décision aléatoire.
- Si l'AUC est inférieure à 0,5, le classificateur est pire que le hasard et doit être évité.

La position de la courbe ROC sur le graphique et sa forme peuvent fournir des informations supplémentaires sur la performance du classificateur. Une courbe qui se rapproche du coin supérieur gauche du graphique indique une meilleure performance, tandis qu'une courbe qui se rapproche de la ligne diagonale de 45 degrés indique une performance aléatoire.

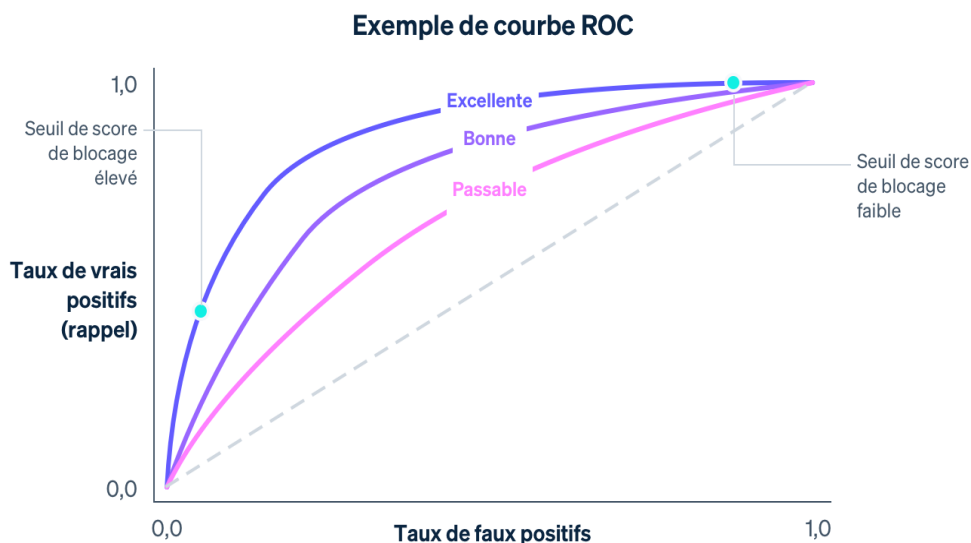


Figure 26 : exemple de courbe ROC (AUC) [55]

2.11 Conclusion :

Les machines à vecteurs de support (SVM) sont une technique populaire de classification et de régression en apprentissage automatique. Elles ont été largement utilisées dans une variété de domaines et présentent plusieurs avantages, tels qu'une grande précision de classification et une capacité à gérer des données non linéairement séparables. Cependant, il y a aussi quelques inconvénients, tels que la sensibilité aux valeurs aberrantes dans les données.

CHAPITRE 3

RÉSULTATS ET ANALYSE

3. Chapitre 3 : Résultats et Analyse

3.1 Introduction :

Dans le chapitre précédent, nous avons examiné en détail la classification et les méthodes associées. Ce chapitre met en œuvre et compare plusieurs modèles basés sur SVM et KNN en termes de précision. Nous explorons l'optimisation des paramètres HOG avec PSO pour améliorer la précision. Les outils, modèles et résultats sont présentés. Notre approche se concentre sur l'implémentation et la performance des modèles, ainsi que sur l'optimisation des paramètres pour une meilleure précision. En résumé, cette étude pratique des concepts de classification avec SVM et KNN évalue l'efficacité de PSO pour l'optimisation des paramètres HOG. Les résultats obtenus fourniront des informations utiles sur les performances des modèles dans le diagnostic des tumeurs cérébrales à partir d'images IRM.

3.2 Matériel :

Ce travail a été effectué sur un ordinateur personnel doté des caractéristiques suivantes :

- Processeur : Intel Core i5 de 7e génération
- Mémoire RAM : 8 Go
- Stockage : Un disque dur SSD de 128 Go et un disque dur HDD de 1 To
- Système d'exploitation : Windows 10 64 bits

3.3 Environnement de travail :

3.3.1 Python

Python est un langage de programmation interprété, orienté objet et de haut niveau doté d'une sémantique dynamique. Ses structures de données intégrées et de haut niveau, combinées au typage dynamique ainsi qu'à la liaison dynamique, l'ont rendu hyper attractif durant le développement rapide d'applications, ainsi que pour une utilisation comme langage de script pour connecter des composants entre eux. Citons qu'il est évident d'implémenter sous Python en accédant à des bibliothèques de haute qualité. En effet, il est omniprésent dans le domaine de

l'intelligence artificielle et de l'apprentissage automatique. prennent en charge Python et sont disponibles pour de plusieurs systèmes d'exploitation. Les bibliothèques les plus utilisées dans l'intelligence artificielle sont implémentées par Python , citons opencv et Scikit-learn , Pandas et



bien beaucoup d'autres.

Figure 27 : Logo python [56]

3.3.2 Jupyter Notebook :

Est un environnement interactif Jupyter Notebook est fréquemment utilisé pour la programmation, l'analyse de données et la visualisation. Cet outil en ligne ouvert permet la création et le partage de documents nommés "notebooks". Ces notebooks incluent du code exécutable, des représentations graphiques, des explications textuelles et d'autres contenus multimédias. Les notebooks sont structurés en cellules qui contiennent divers éléments tels que du code, du texte mis en forme, des équations mathématiques et des images. Les utilisateurs peuvent exécuter le code présent dans les cellules et observer les résultats instantanément affichés. Jupyter Notebook est compatible avec plusieurs langages de programmation tels que Python, R et Julia. Il est utilisé pour explorer des données, effectuer des apprentissages automatiques, réaliser des simulations numériques et présenter de façon interactive les résultats de recherches. Sa popularité est considérable au sein des milieux scientifiques et éducatifs en raison de sa simplicité d'utilisation, de sa flexibilité ainsi que de ses fonctionnalités favorisant la collaboration et le partage des connaissances.[57]



Figure 28: Logo jupyter [57]

3.3.3 Les bibliothèques

- NumPy.
- OpenCV.
- Matplotlib.
- Pandas.
- SciPy.
- Scikit-learn.
- Tensorflow.
- Scikit-image (skimage).
- Seaborn.

3.4 Les bases des données utilisées :

Nous avons utilisé deux bases de données qui ont été collectées à partir d'un site web appelé Kaggle, qui est une plateforme fournissant des ensembles de données téléchargés par une communauté de scientifiques et de chercheurs afin de permettre aux utilisateurs de participer à des compétitions d'apprentissage automatique.

3.4.1 Base des données 1 :

Les images IRM ont été réparties en 3002 images pour l'entraînement et 616 images pour les tests. Pour le dossier d'entraînement, nous avons deux dossiers pour les classes " **pituitary tumor** " et " **no tumor** ", chacun contenant 1501 images IRM. Pour le dossier de test, nous avons également deux dossiers pour les classes "tumeur hypophysaire" et "pas de tumeur", chacun contenant 308 images IRM.

3.4.2 Base des données 2 :

"Le deuxième dataset que nous avons utilisé est une base de données à classes multiples, qui nous avons téléchargé vers le site de kaggle dans cette [lien](#) qui se compose de 7023 images d'exams IRM, d'une taille de 164 Mo, réparties en 4 classes : "tumeur hypophysaire", "pas de tumeur", "gliome" et "méningiome". Les images IRM ont été divisées en 5712 images pour l'entraînement et 1311 images pour les tests.

Pour le dossier d'entraînement, nous avons quatre dossiers pour les classes :

- pituitary tumor (1457 images)
- no tumor (1595 images)
- glioma (1321 images)
- meningioma (1339 images)

3.5 Les Modèles et résultats :

3.5.1 Modèle 1 Extracteur HOG avec les Classificateurs SVM et KNN :

Ce modèle effectue une classification des images de tumeurs cérébrales en utilisant des modèles d'apprentissage automatique tels que les machines à vecteurs de support (SVM) et les plus proches voisins (KNN).

Le modèle commence par initialiser des listes vides pour stocker les données des images et les étiquettes correspondantes. Ensuite, il boucle sur chaque classe d'images de tumeurs cérébrales et lit les images à partir des répertoires correspondants. Pour chaque image,

il effectue les étapes suivantes :

1. Lecture de l'image en niveaux de gris.
2. Redimensionnement de l'image à une taille fixe de 200x200 pixels.
3. Extraction des caractéristiques de l'histogramme des gradients orientés (HOG) de l'image.
4. Rééchelonnement de l'intensité de l'image HOG.
5. Ajout de l'image HOG à la liste des données et de l'étiquette de classe correspondante à la liste des étiquettes.

Ensuite, les données et les étiquettes sont converties en tableaux NumPy. Le code analyse ensuite les classes uniques et leur répartition dans le jeu de données.

Les données d'image sont remodelées en un vecteur de caractéristiques unidimensionnel. Le jeu de données est divisé en ensembles d'apprentissage et de test en utilisant un pourcentage spécifié pour les tests.

Le modèle entraîne et évalue plusieurs modèles SVM avec différents noyaux (linéaire, RBF, polynomial, sigmoïde) et un classifieur KNN. Pour chaque modèle, il effectue les étapes suivantes :

1. Entraînement du modèle en utilisant les données d'apprentissage.
2. Prédiction des étiquettes pour les données de test.
3. Évaluation des performances du modèle à l'aide de mesures tel que la précision, le rappel, le score F1 et la matrice de confusion.
4. Affichage du rapport de classification avec les mesures de performance.

Enfin, ce modèle calcule la spécificité (taux de vrais négatifs) et visualise la matrice de confusion ainsi que la courbe ROC (Receiver Operating Characteristic).

Le modèle effectue donc les étapes suivantes : prétraitement des images, extraction des caractéristiques HOG, division des données en ensembles d'apprentissage et de test, entraînement et évaluation de plusieurs modèles SVM et d'un classifieur KNN, et affichage des performances et des visualisations associées.

Cela permet de classifier les images de tumeurs cérébrales en utilisant des modèles d'apprentissage automatique, ce qui peut aider à la détection et au diagnostic des tumeurs cérébrales.

3.5.2 Modèle 2 Extracteur HOG et Kernel-PCA avec les Classificateurs SVM et KNN:

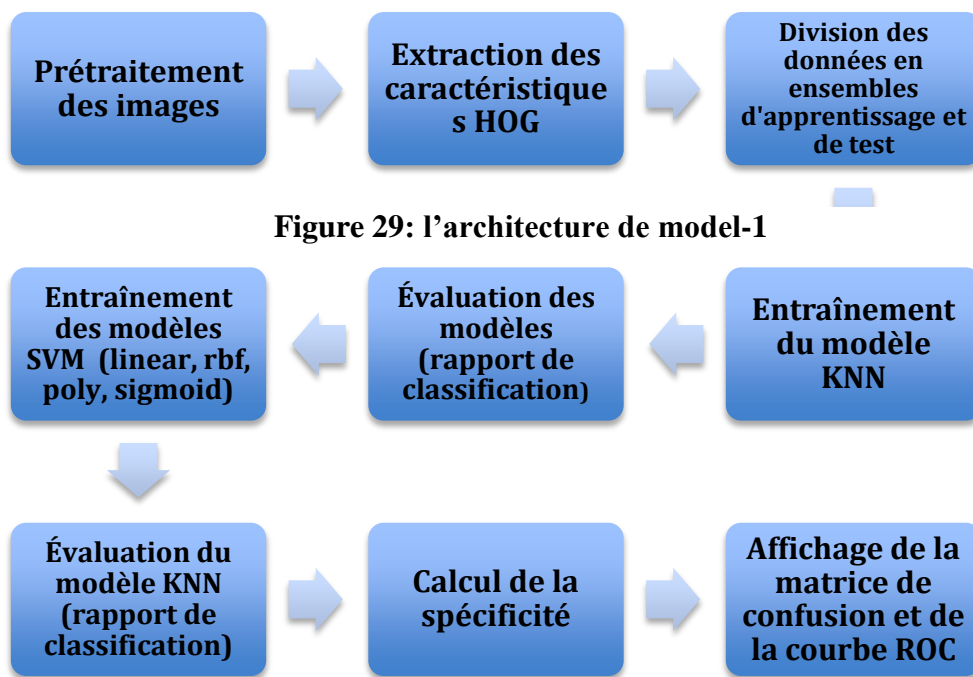


Figure 29: l'architecture de model-1

Analyse

Ce modèle effectue une classification des images de tumeurs cérébrales en utilisant des modèles d'apprentissage automatique tels que les machines à vecteurs de support (SVM) et les plus proches voisins (KNN). Le modèle a été modifié pour inclure l'algorithme Kernel PCA (Analyse en Composantes Principales à Noyau), qui permet de transformer les données d'image dans un espace de dimensions réduit en utilisant un noyau spécifié. après avoir redimensionné et extrait les caractéristiques HOG des images de tumeurs cérébrales, le modèle applique l'algorithme Kernel PCA pour réduire la dimensionnalité des données tout en conservant les informations discriminantes , cette étape supplémentaire permet de mieux représenter les images dans un espace de caractéristiques adapté à la tâche de classification , ensuite, le modèle divise les données transformées en ensembles d'apprentissage et de test, puis entraîne et évalue les modèles SVM (avec différents noyaux) et le classifieur KNN sur ces ensembles de données modifiés. Les

étapes d'entraînement, de prédiction et d'évaluation des performances restent les mêmes. l'ajout de l'algorithme Kernel PCA dans ce modèle permet de réduire la dimensionnalité des données d'image et potentiellement d'améliorer les performances de classification en fournissant une représentation plus discriminante des images de tumeurs cérébrales.

3.5.3 Modèle 3 Extracteur HOG avec Classificateurs SVM (optimisé par PSO) et KNN :

Ce modèle implémente l'algorithme d'optimisation par Essaim de Particules (PSO) pour optimiser les paramètres d'un modèle de Machine à Vecteurs de Support (SVM). PSO est un algorithme d'optimisation basé sur une population, inspiré du comportement des oiseaux en vol. Dans ce modèle, un groupe de particules représente des solutions potentielles, où la position de

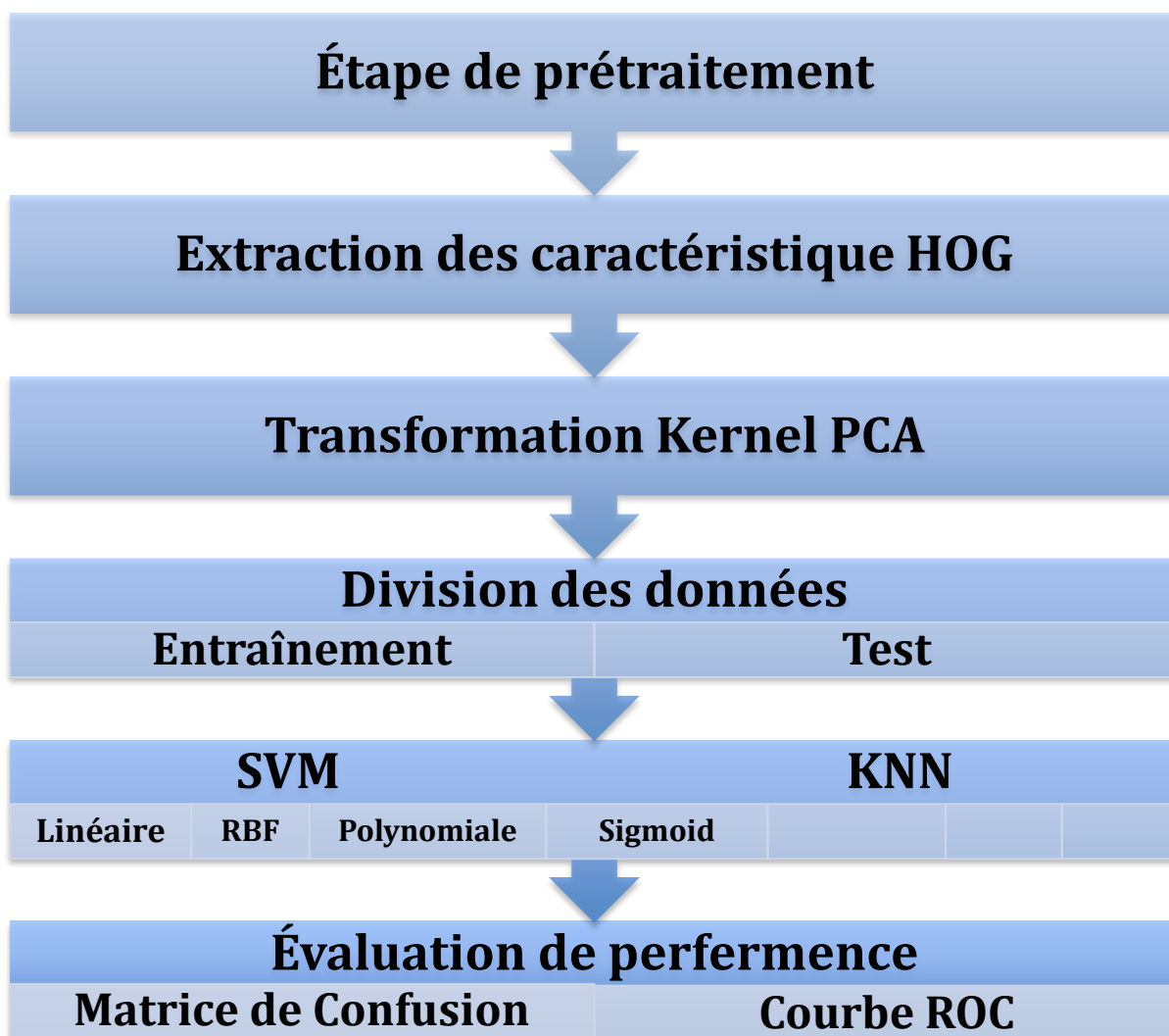


Figure 30: architecture de model-2

chaque particule correspond à un ensemble de paramètres SVM (gamma et C). L'objectif de l'algorithme est de trouver la meilleure combinaison de gamma et C qui maximise la précision du modèle SVM dans la classification des images de tumeurs cérébrales.

L'algorithme PSO débute en initialisant un ensemble de particules avec des positions aléatoires. Chaque particule évalue sa position en calculant la valeur de fitness, qui représente les performances du modèle SVM utilisant les valeurs gamma et C correspondantes. L'algorithme met ensuite à jour les meilleures positions (pbest) pour chaque particule et conserve la meilleure position globale (gbest) parmi toutes les particules ayant la plus faible valeur de fitness, à chaque itération, les particules ajustent leurs positions en fonction de leurs vitesses, qui sont influencées par leurs meilleures positions individuelles et la meilleure position globale trouvée jusqu'à présent. Ce processus d'exploration-exploitation se poursuit pendant un nombre spécifié d'itérations jusqu'à ce qu'une condition d'arrêt soit atteinte, telle qu'une erreur cible ou un nombre maximum d'itérations.

Le modèle renvoie la meilleure position trouvée, qui correspond aux valeurs optimales de gamma et C pour le modèle SVM, ainsi que la valeur de fitness (précision d'entraînement et de test ou erreur) obtenue à cette position.

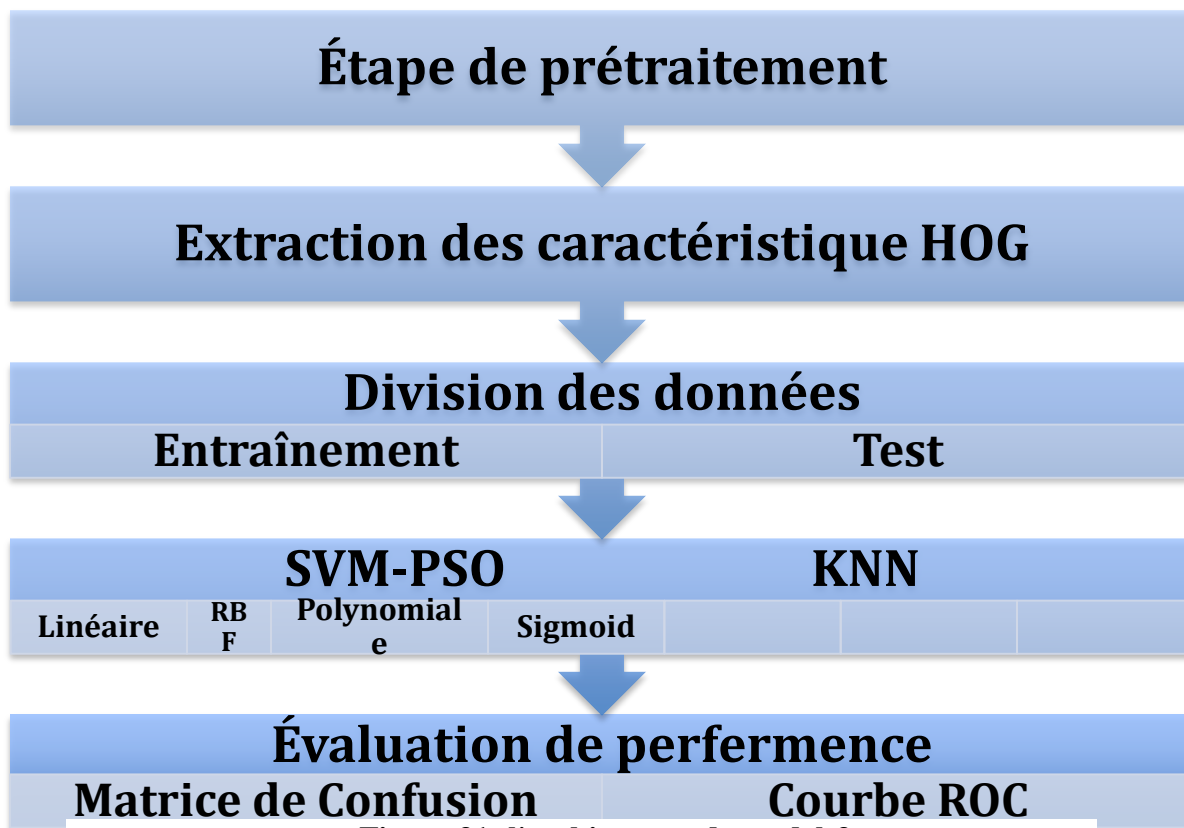


Figure 31 :l'architecture de model-3

3.5.4 Modèle 4 Extracteur HOG et Kernel-PCA avec Classificateurs SVM (optimisé par PSO) et KNN :

Dans ce modèle Nous avons ajouté l'algorithme Kernel PCA (Analyse en Composantes Principales à Noyau) pour réduire la dimensionnalité des données d'image tout en préservant les informations importantes. Cette étape permet de représenter de manière plus précise les images dans un espace de caractéristiques adapté à la classification. Ensuite, les données sont divisées en ensembles d'apprentissage et de test, et les modèles SVM (Machines à Vecteurs de Support) et KNN (K Plus Proches Voisins) sont entraînés et évalués. Pour améliorer les performances du SVM, nous avons appliqué l'algorithme d'optimisation par Essaim de Particules (PSO) pour rechercher les meilleurs paramètres gamma et C. Le PSO explore différentes combinaisons de paramètres en utilisant une population de particules, permettant ainsi d'obtenir une configuration optimale du SVM. En

utilisant cette approche, nous améliorons la précision de la classification des tumeurs cérébrales. Ce modèle combine différentes étapes pour obtenir une

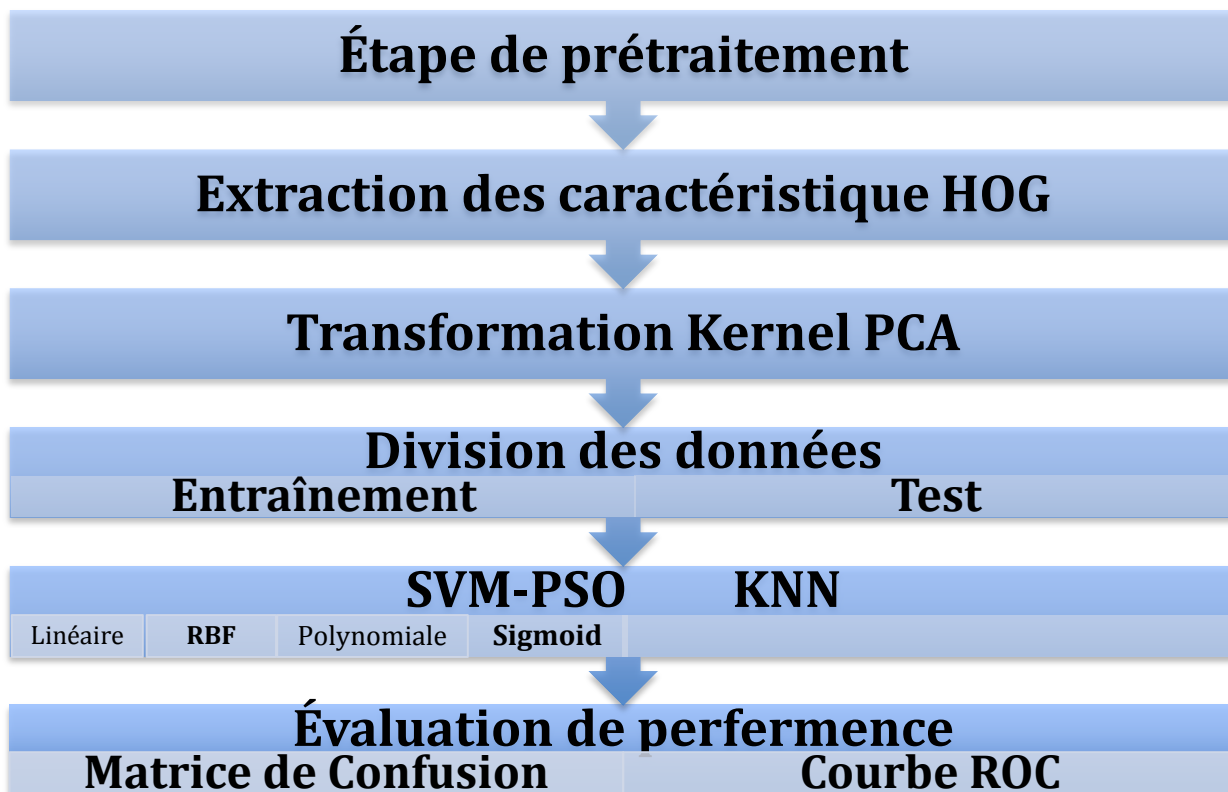


Figure 32: architecture de model-4

classification précise des images de tumeurs cérébrales.

3.5.5 Modèle 5 Extracteur HOG optimisé avec les Classificateurs SVM et KNN :

Ce modèle recherche les meilleurs paramètres pour l'extraction des caractéristiques HOG en testant différentes combinaisons d'orientation, de cellules par bloc et de pixels par cellule. Les caractéristiques HOG sont ensuite extraites des images et associées à leurs étiquettes correspondantes. Un classificateur SVM est ensuite entraîné avec un noyau RBF en utilisant un ensemble d'entraînement. L'exactitude du classificateur est évaluée sur un ensemble de test. Le modèle

évalue également d'autres classificateurs SVM avec différents noyaux (linéaire, polynomial, sigmoïde) en générant des rapports de classification, des matrices de confusion et des courbes ROC pour chacun. De plus, un classificateur des k plus proches voisins (KNN) est entraîné et évalué, avec un rapport de classification, une matrice de confusion et une courbe ROC. L'objectif est d'automatiser la recherche des paramètres optimaux pour les caractéristiques HOG et de sélectionner le

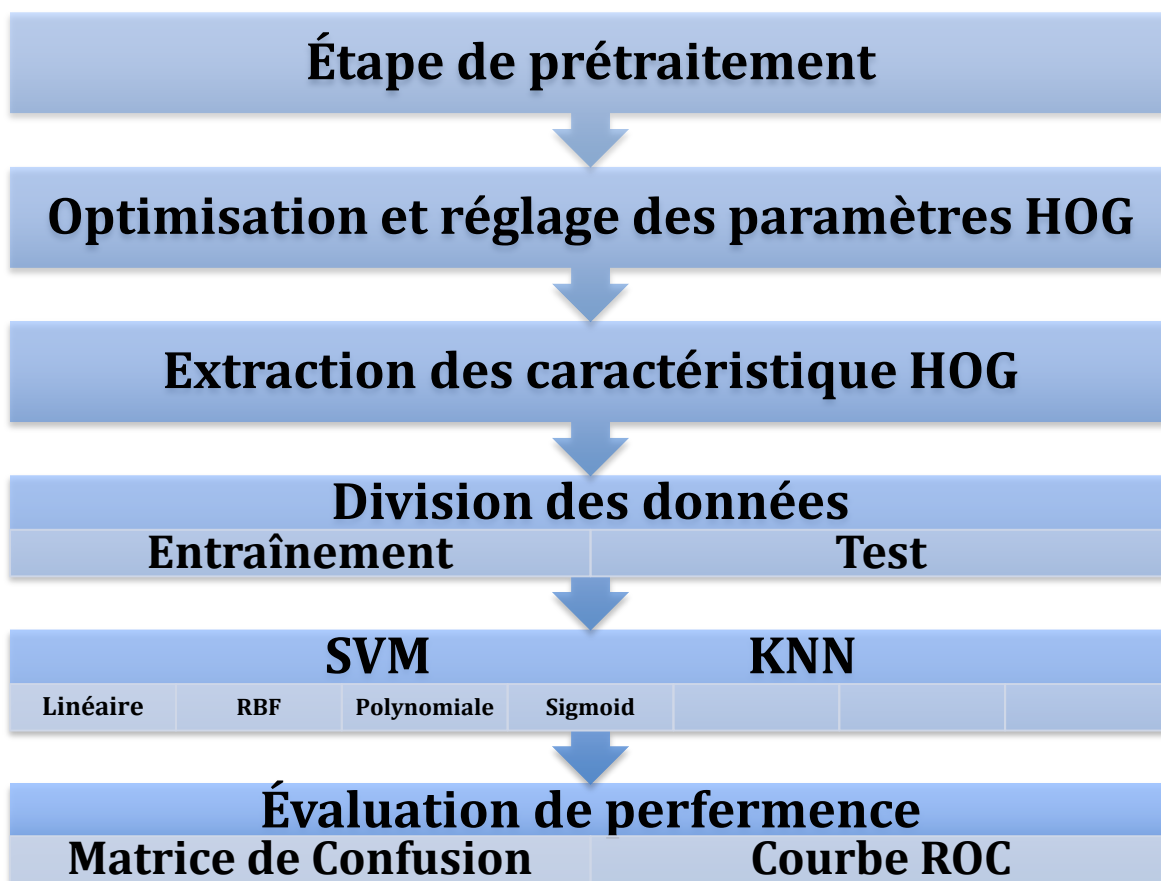


Figure 33: Architecture de model-5

meilleur classificateur pour la classification des tumeurs cérébrales. Les résultats et les visualisations générés aident à déterminer l'approche la plus efficace.

3.5.6 Modèle 6 Extracteur HOG optimisé avec Classificateurs (SVM optimisé par PSO) et KNN :

Le dernier modèle applique l'optimisation par essaim de particules (PSO) pour optimiser les paramètres d'un classifieur SVM (Support Vector Machine), en termes simplifiés, le modèle vise à trouver la meilleure combinaison de paramètres pour le classifieur SVM afin d'obtenir la meilleure précision possible dans la classification des images de tumeurs cérébrales. Les paramètres optimisés sont liés à la méthode d'extraction de caractéristiques Histogram of Oriented Gradients (HOG), qui est couramment utilisée pour les tâches de classification d'images. le modèle commence par importer les bibliothèques nécessaires et mettre en place l'ensemble des données contenant les images de tumeurs cérébrales. ensuite, il définit les plages de valeurs à explorer pour les paramètres HOG, tels que l'orientation, les cellules par bloc et les pixels par cellule. ensuite, le code itère sur toutes les combinaisons possibles de ces paramètres et effectue les étapes suivantes pour chaque combinaison :

- Extraction des caractéristiques HOG à partir des images de tumeurs cérébrales en utilisant la combinaison actuelle de paramètres.
- Segmentation l'ensemble de données en ensembles d'entraînement et de test.
- Entraînement d'un classifieur SVM à l'aide des données d'entraînement.
- Évaluation de la précision du classifieur sur les données de test.

Le modèle enregistre la meilleure précision obtenue jusqu'à présent et la combinaison de paramètres correspondante. Une fois toutes les combinaisons de paramètres évaluées, il affiche les meilleurs paramètres et leur précision correspondante , de plus, le code entraîne un classifieur SVM avec les meilleurs paramètres sur l'ensemble des données et évalue sa précision. Il génère également un rapport de classification, trace une matrice de confusion et affiche une courbe ROC pour le classifieur SVM , le modèle répète le même processus d'entraînement des classifieurs SVM avec différents noyaux (linéaire, polynomiale, fonction de base radiale et sigmoïde) et évalue leurs performances à l'aide de rapports de classification, de matrices de confusion et de courbes ROC. Enfin, le modèle entraîne et évalue un classifieur K-Nearest Neighbors (KNN) sur l'ensemble des données, génère un rapport de classification, trace une matrice de confusion et affiche une courbe ROC pour le classifieur KNN.

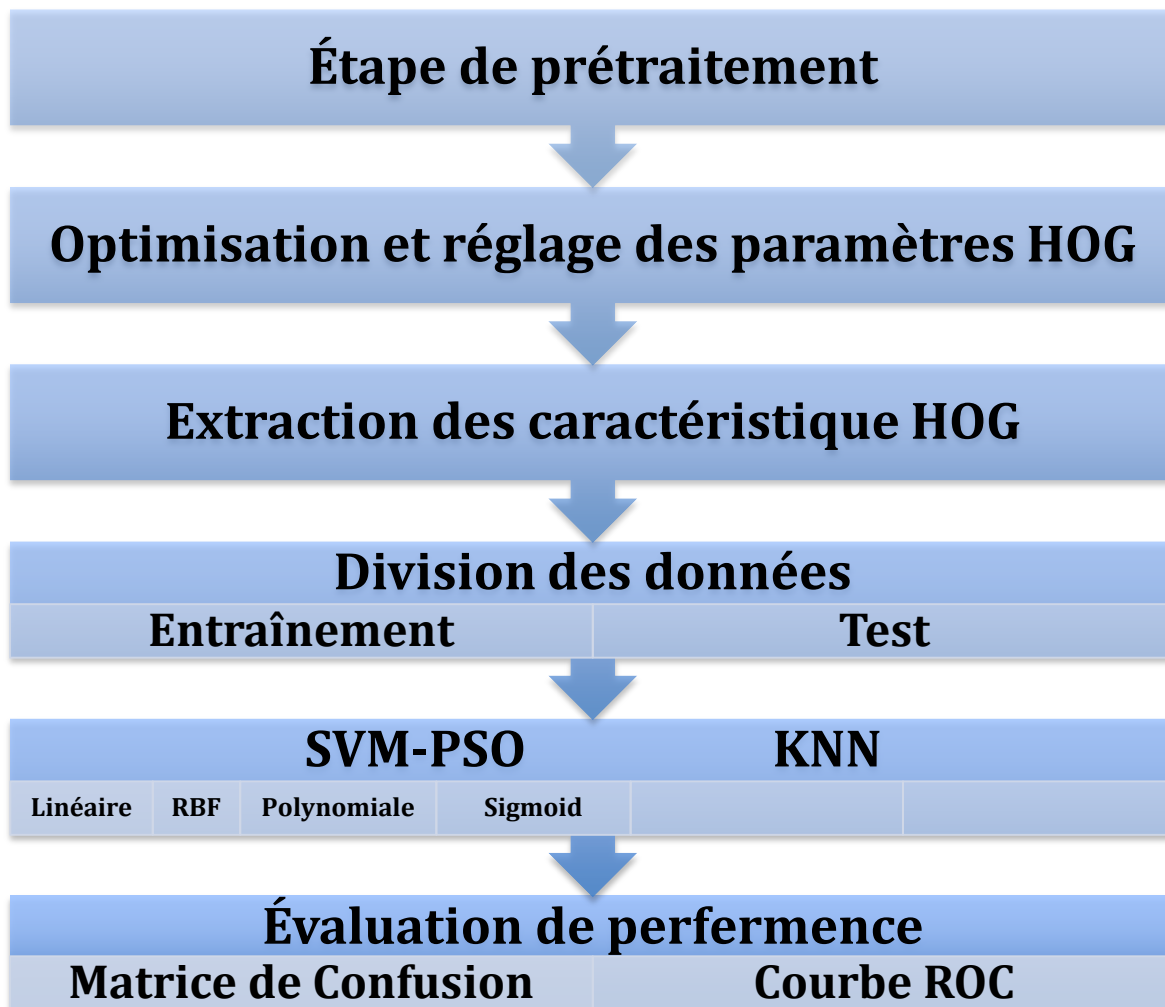


Figure 34: architecture de model-6

Tableau 2: Tableaux dès les paramètres de PSO

Paramètre	valeur
W	0.5
C1	0.5
C2	0.9

Tableau 3 : Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG pour dataset-1

N-itération	10
N-particule	100

3.6 Résultat :

3.6.1 Modèle 1 : Extracteur HOG avec les Classificateurs SVM et KNN :

Data 1 « Bas de données binaire » :

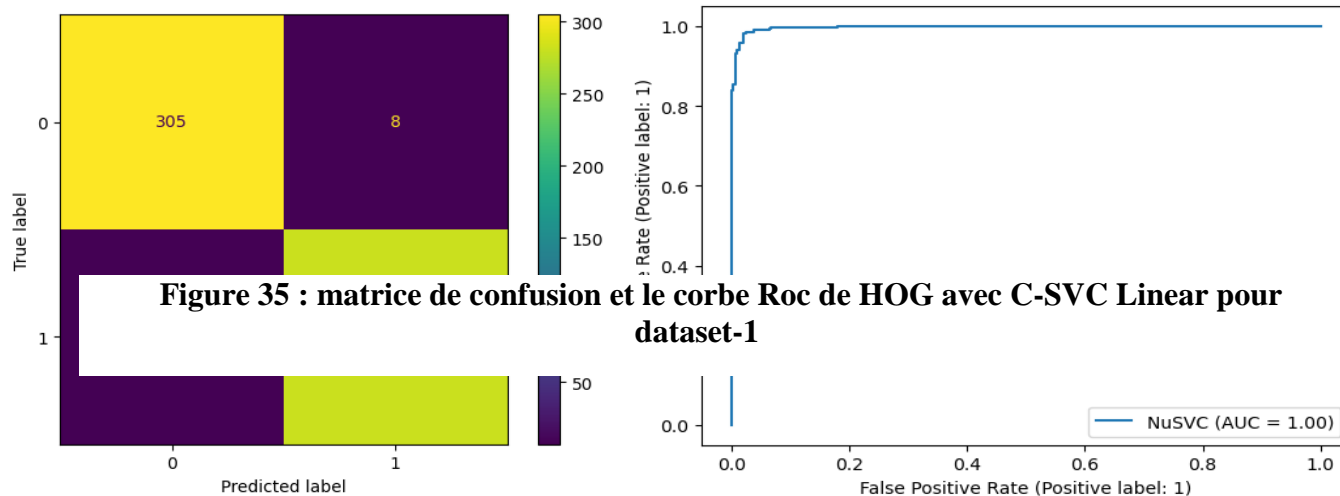


Figure 35 : matrice de confusion et le corbe Roc de HOG avec C-SVC Linear pour dataset-1

Data 2 « Bas de données multi class » :

Tableau 4: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG pour dataset-2

Classificateur	fonction	precision	sensitivity	f-mesure	Accuracy
linear-svc	linear	0.87799	0.877515	0.87773	0.87752
classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.8927	0.894138	0.8927	0.89414
linear-svc	linear	0.96506	0.965058	0.96506	0.96506
	poly	0.887489	0.887139	0.885631	0.887139
c-svc	linear	0.97505	0.975042	0.97504	0.97504
	rbf	0.873776	0.876640	0.874178	0.87664
c-svc	poly	0.93875	0.938436	0.93838	0.93844
	sigmoid	0.65224	0.654718	0.6482	0.65472
	rbf	0.93845	0.938436	0.93844	0.93844
	linear	0.80416	0.797025	0.79495	0.79703
	sigmoid	0.61683	0.617304	0.61572	0.6173
	poly	0.82262	0.812773	0.81129	0.81277
Nu-svc	linear	0.89019	0.890183	0.89015	0.89018
	rbf	0.82733	0.820647	0.81896	0.82065
	poly	0.90503	0.903494	0.90325	0.90349
Nu-svc	sigmoid	0.67459	0.670166	0.66305	0.67017
knn	rbf	0.90518	0.905158	0.90516	0.90516
	linear	0.88102	0.88189	0.87979	0.88189
	sigmoid	0.63192	0.63228	0.631	0.63228

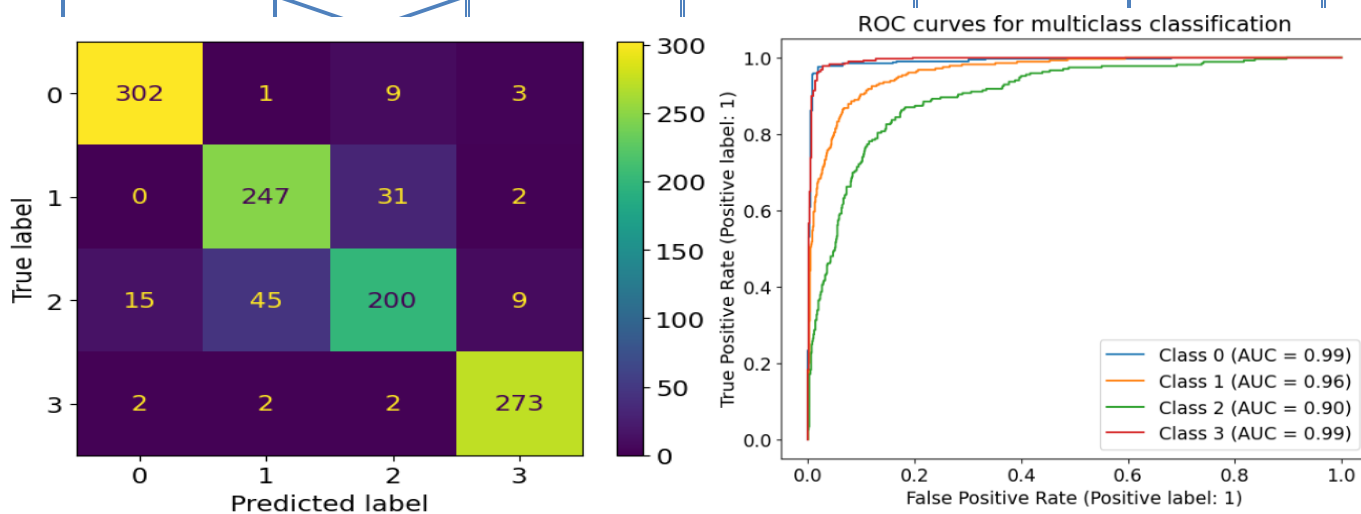


Figure 36 : matrice de confusion et le courbe Roc de HOG avec C-SVC linear pour dataset-2

3.6.2 Modèle 2 : Extracteur HOG et Kernel-PCA avec les Classificateurs SVM et KNN :

Data 1 « Bas de données binaire » :

Tableau 5: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG et Kernel-PCA pour dataset-1

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.9636	0.963394	0.963401	0.96339
c-svc	linear	0.973378	0.973378	0.973378	0.973378
	poly	0.980060	0.980033	0.980036	0.980033
	rbf	0.973406	0.973378	0.973381	0.973378
	sigmoid	0.965090	0.965058	0.965050	0.965058
Nu-svc	linear	0.975050	0.975042	0.975043	0.975042
	poly	0.970079	0.970050	0.970054	0.97005
	rbf	0.975043	0.975042	0.975040	0.975042
	sigmoid	0.930690	0.930116	0.930030	0.930116
Knn	X	0.963549	0.963394	0.963375	0.963394

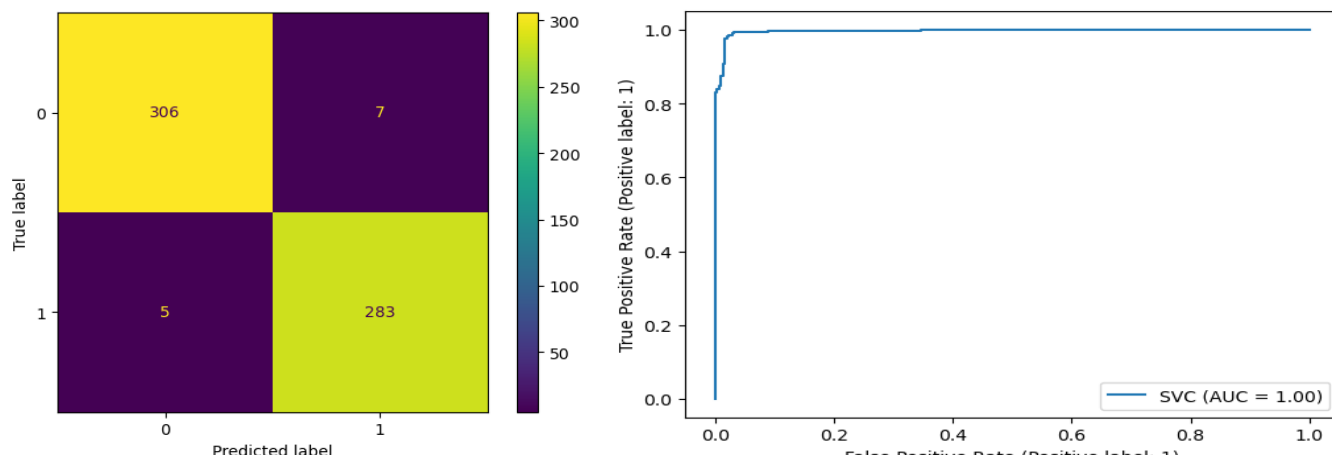


Figure 37 : matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC poly pour dataset-1

Data 2 « Bas de données multi class » :

Tableau 6: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG et Kernel-PCA pour dataset-2

classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.86051	0.861767	0.86106	0.86177
c-svc	linear	0.88325	0.884514	0.883247	0.88451
	poly	0.88504	0.886264	0.884615	0.88626
	rbf	0.90478	0.906387	0.905228	0.90639
	sigmoid	0.75961	0.76378	0.758411	0.76378
Nu-svc	linear	0.80405	0.797025	0.794854	0.79703
	poly	0.86368	0.861767	0.861302	0.86177
	rbf	0.83655	0.832896	0.830788	0.8329
	sigmoid	0.72739	0.727034	0.7201	0.72703
knn	X	0.88102	0.88189	0.879788	0.88189

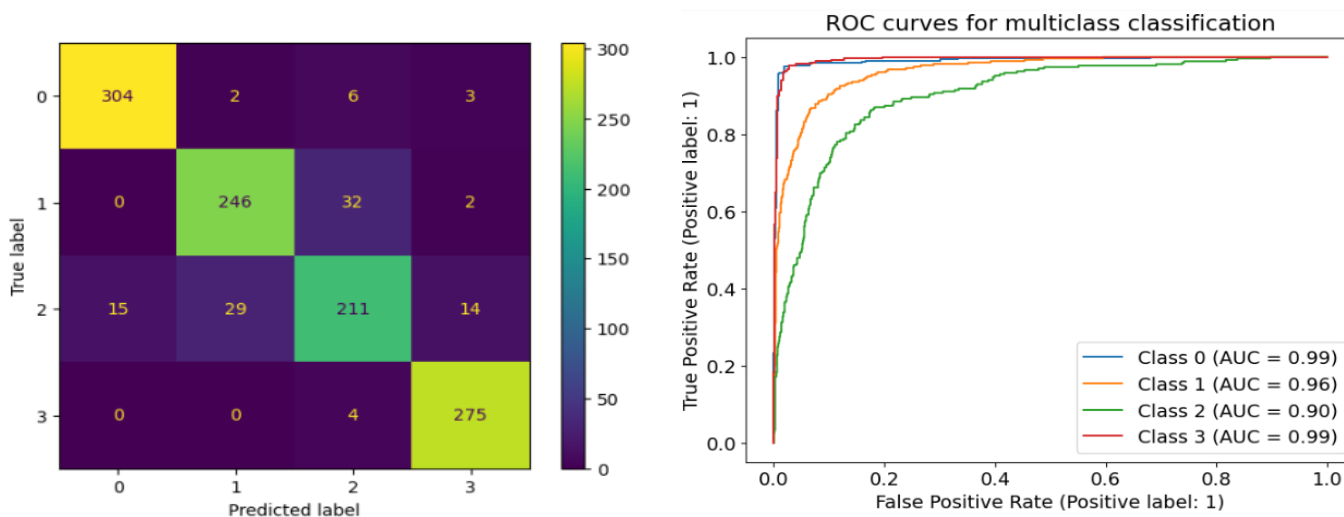


Figure 38: matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC RBF pour dataset-2

3.6.3 Modèle 3 : Extracteur HOG avec Classificateurs SVM (optimisé par PSO) KNN :


Les paramètres de PSO choisis pour le modèle :

Tableau 7: les Paramètres choisis par PSO pour le classificateur SVM .

Data	Svm	Noyou	C	Nu	Gamma	Degree	Coef0
Data1	C-svc	Linear	9.319722	/	/	/	/
		Poly	9.319722	/	8.200487	2	/
		Rbf	50.57815	/	0.002227	/	/
		Sigmoid	99.57815	/	0.09273	/	3
	Nu-svc	Linear	/	0.01	/	/	/
		Poly	/	0.093273	72.273	2	/
		Rbf	/	0.083273	0.00273	/	/
		Sigmoid	/	0.083273	0.00273	/	4
Data2	C-svc	Linear	9.410665	/	/	/	/
		Poly	3.4959	/	3.68132	2	/
		Rbf	50.28832	/	1.257801	/	/
		Sigmoid	16.8009	/	1.7377	/	1
	Nu-svc	Linear	/	0.09927	/	/	/
		Poly	/	0.093272	72.27995	2	/
		Rbf	/	0.083272	0.00273	/	/
		Sigmoid	/	0.832799	0.0028	/	0

Data 1 « Bas de données binaire » :

Tableau 8: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG pour dataset-1

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.97338	0.973378	0.97338	0.97338
c-svc	linear	0.97505	0.975042	0.97504	0.97504
	poly	0.98006	0.980033	0.98004	0.98003
	rbf	0.97008	0.97005	0.97005	0.97005
	sigmoid	0.5042	0.519135	0.40001	0.51914
Nu-svc	linear	0.96675	0.966722	0.96673	0.96672
	poly	0.98171	0.981697	0.9817	0.9817
	rbf	0.98171	0.96173	0.96174	0.96173
	sigmoid	0.60018	0.592346	0.58963	0.59235
Knn		0.91628	0.915141	0.91497	0.91514

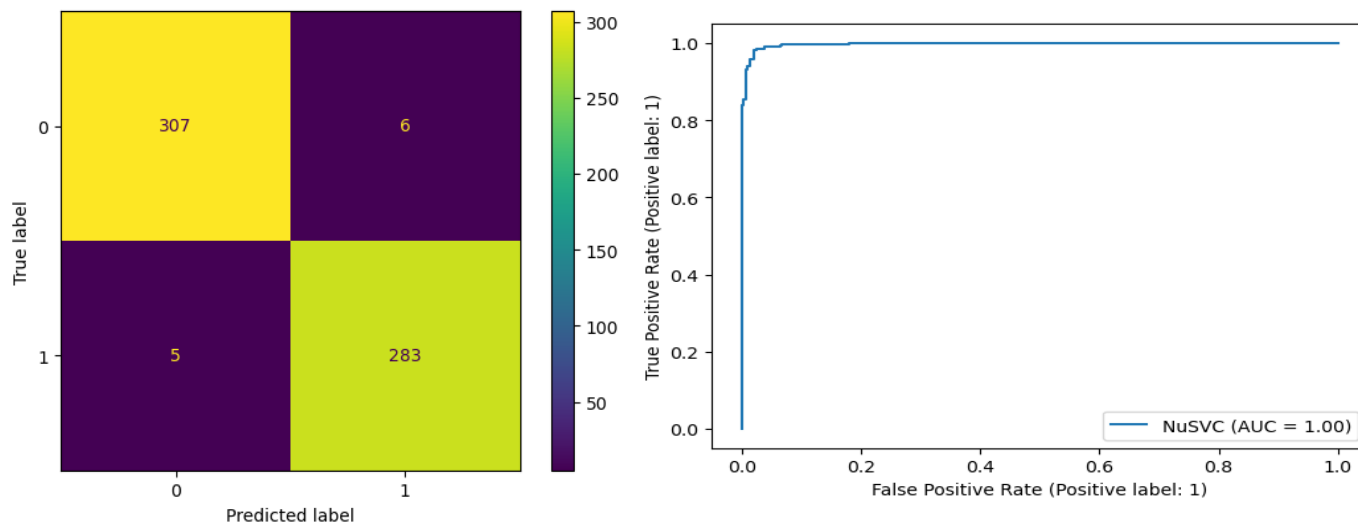
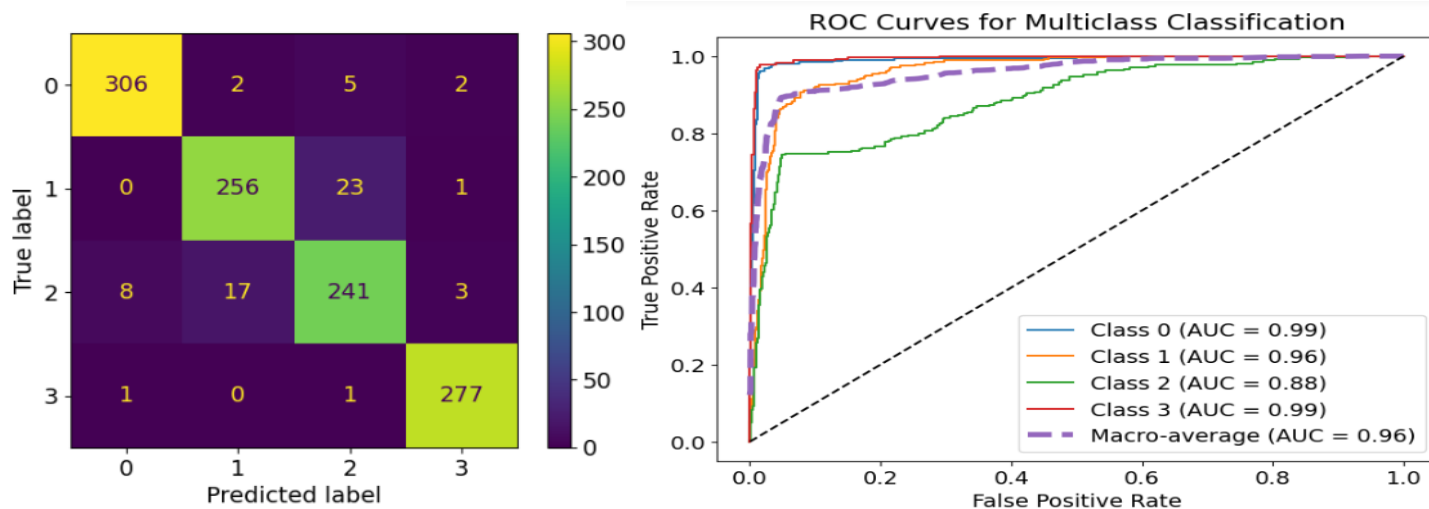



Figure 39 : matrice de confusion et le courbe Roc de HOG avec NU-SVC poly pour dataset-1

Data 2 « Bas de données multi class » :

Tableau 9: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG pour dataset-2

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.87799	0.877515	0.87773	0.87752
C-SVC	linear	0.8927	0.894138	0.8927	0.89414
	poly	0.93628	0.936133	0.93617	0.93613



	rbf	0.94475	0.944882	0.94478	0.94488
	sigmoid	0.76621	0.767349	0.76863	0.76935
Nu-svc	linear	0.88313	0.884514	0.88291	0.88451
	poly	0.93359	0.933508	0.93342	0.93351
	rbf	0.94289	0.943132	0.94297	0.94313
	sigmoid	0.88414	0.89007	0.88683	0.89007
knn		0.94107	0.941382	0.94108	0.94138

3.6.4 Modèle 4 : Extracteur HOG et Kernel-PCA avec Classificateurs SVM optimisé par PSO et KNN :

Les paramètres de PSO choisi pour le modèle :

Tableau 10: les paramètres choisis par PSO

Data	Svm	Noyou	C	Nu	Gamma	Degree	Coef0
Data1	C-svc	Linear	4.042562	/	/	/	/
		Poly	4.042562	/	78.48426	2	/
		Rbf	65.23465	/	26.0872	/	/
		Sigmoid	80.15635	/	2.571979	/	1
	Nu-svc	Linear	/	0.075635	/	/	/
		Poly	/	0.091606	78.48426	2	/
		Rbf	/	0.12724	0.988668	/	/
		Sigmoid	/	0.006922	0.69391	/	0
Data2	C-svc	Linear	4.042562	/	/	/	/
		Poly	4.042562	/	78.48426	2	/

		Rbf	65.23465	/	26.0872	/	/
		Sigmoid	80.15635	/	2.571979	/	1
	Nu-svc	Linear	/	0.075635	/	/	/
		Poly	/	0.091606	78.48426	2	/
		Rbf	/	0.12724	0.988668	/	/
		Sigmoid	/	0.006922	0.69391	/	0

Data 1 « Bas de données binaire » :

Tableau 11: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG et Kernel-PSA pour dataset-1

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.97338	0.973378	0.973378	0.97338
c-svc	linear	0.97338	0.973378	0.973378	0.97338
	poly	0.94665	0.945092	0.945092	0.94509
	rbf	0.95684	0.956739	0.956686	0.95674
	sigmoid	0.73371	0.733777	0.733737	0.73378
Nu-svc	linear	0.89188	0.891847	0.891802	0.89185

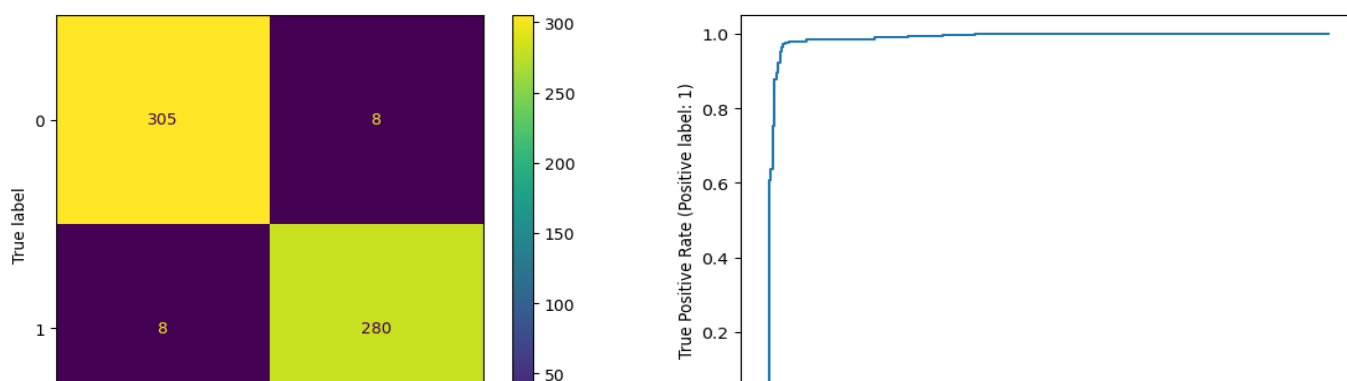


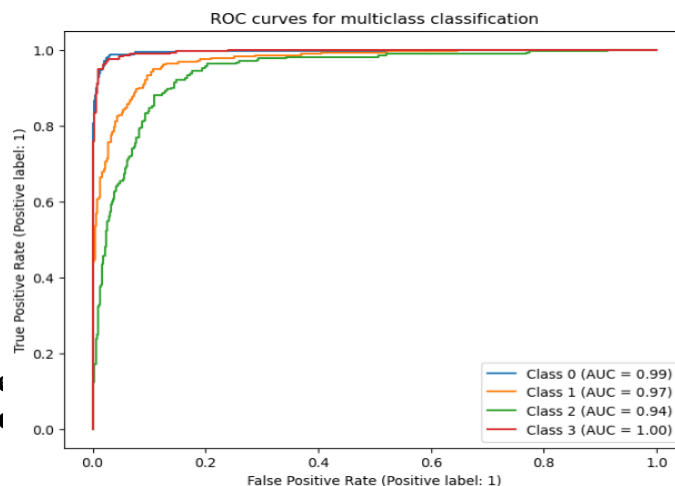
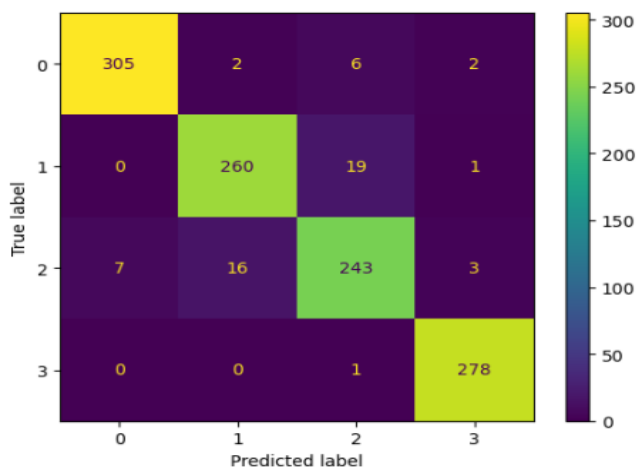
Figure 41 : matrice de confusion et le courbe Roc de HOG et Kernel-PCA avec C-SVC Linear pour dataset-1

	poly	0.90847	0.90183	0.90171	0.90183
	rbf	0.91039	0.91015	0.910178	0.91015
	sigmoid	0.78295	0.78203	0.782124	0.78203
Knn		0.91781	0.916805	0.916654	0.91681

Data 2 « Bas de données multi class » :

Tableau 12: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG et Kernel-PSA pour dataset-2

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.86051	0.861767	0.86106	0.86177
C-svc	linear	0.8927	0.894138	0.892703	0.89414
	poly	0.93628	0.936133	0.936166	0.93613
	rbf	0.950051	0.950131	0.950063	0.950131
	sigmoid	0.89448	0.895888	0.894671	0.89589
	Nu-svc	linear	0.8927	0.894138	0.892703
Nu-svc	poly	0.93359	0.933508	0.93342	0.93351
	rbf	0.89702	0.897638	0.89694	0.89764
	sigmoid	0.79196	0.79615	0.793272	0.79615
	Knn		0.94107	0.941382	0.941081



3.6.5 Modèle 5 : Extracteur HOG optimisé avec les Classificateurs SVM et KNN :

Data 1 « Bas de données binaire » :

Tableau 13: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG optimisé pour dataset-1

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.98006	0.980033	0.980035	0.98003
C-SVC	linear	0.96211	0.96173	0.961706	0.96173
	poly	0.9817	0.981697	0.981696	0.9817
	rbf	0.98006	0.980033	0.980035	0.98003
	sigmoid	0.77395	0.77371	0.773752	0.77371
Nu-svc	linear	0.89879	0.898502	0.898522	0.8985
	poly	0.92195	0.921797	0.92181	0.9218
	rbf	0.93366	0.933444	0.933456	0.93344
	sigmoid	0.76711	0.767055	0.767074	0.76706
Knn		0.97505	0.975042	0.97504	0.97504

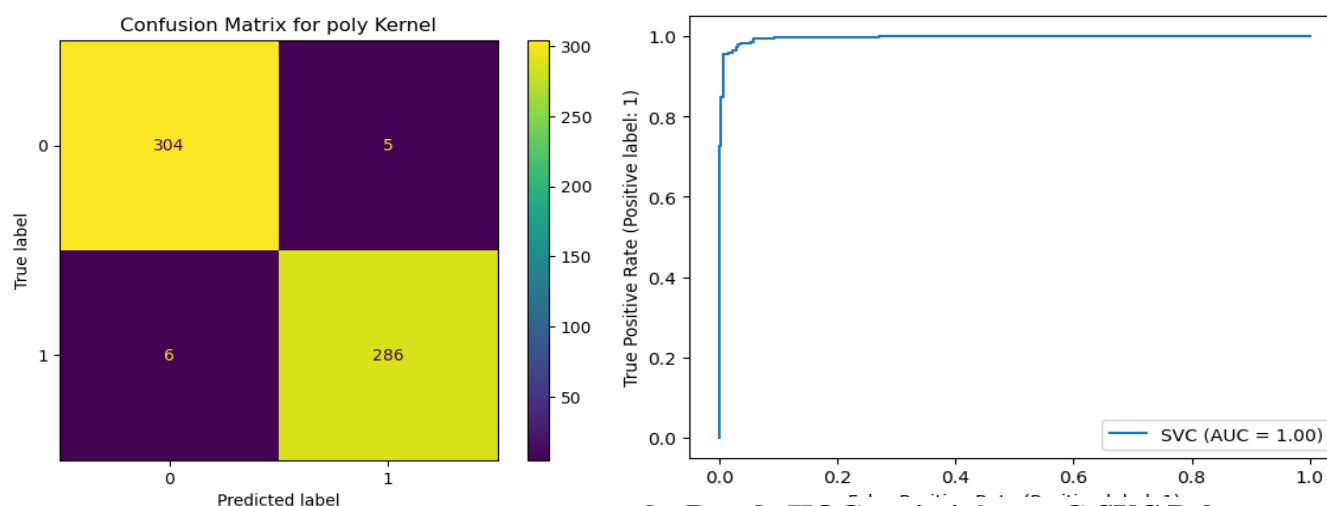



figure 45: matrice de comusion et le courbe Roc de HOG optimisé avec C-SVC Poly pour dataset-1

Data 2 « Bas de données multi class » :

Tableau 14: Tableau de comparaison entre les différents classificateurs de SVM et KNN avec l'extracteur HOG optimisé pour dataset-2

Classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.90551	0.907262	0.90606	0.90726
c-svc	linear	0.91004	0.909886	0.90989	0.90989
	poly	0.90463	0.904637	0.90385	0.90464
	rbf	0.90551	0.907262	0.90606	0.90726
	sigmoid	0.67198	0.674541	0.66795	0.67454
Nu-svc	linear	0.83553	0.831146	0.82932	0.83115
	poly	0.83749	0.828521	0.82846	0.82852
	rbf	0.84779	0.845144	0.84298	0.84514
	sigmoid	0.69162	0.690289	0.68423	0.69029
Knn		0.89548	0.895888	0.89423	0.89589

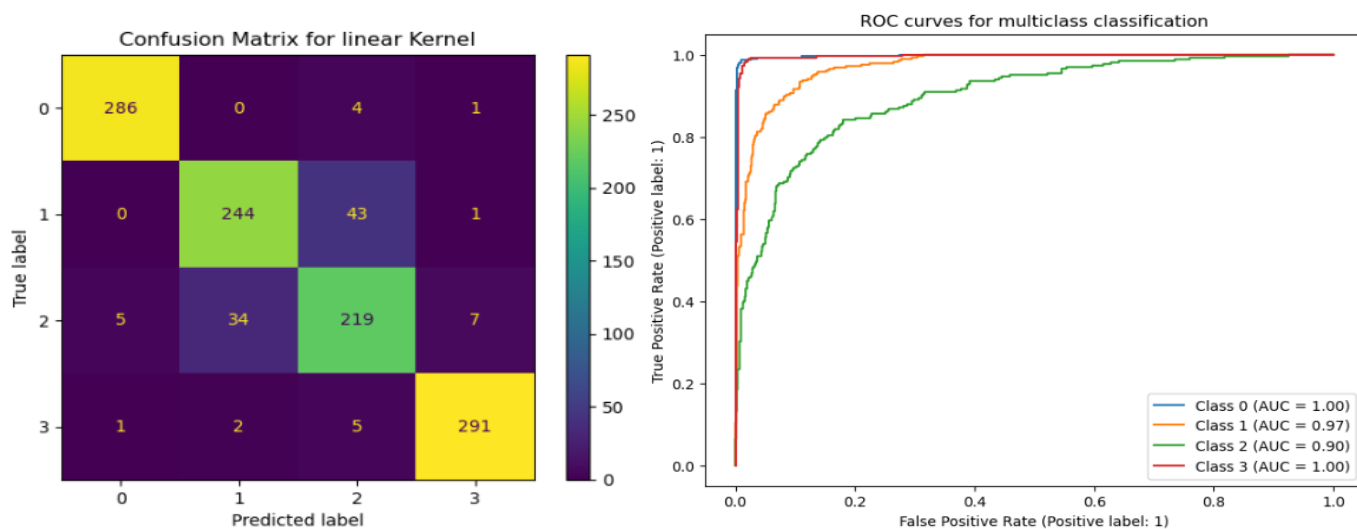


Figure 44 : matrice de confusion et le courbe Roc de HOG optimisé avec C-SVC Linear pour dataset-2

3.6.6 Modèle 6 : Extracteur HOG optimisé avec Classificateurs SVM optimisé par PSO et KNN :

Les parametres de PSO choisis pour le modele :

Tableau 15: les paramètres choisis par PSO

Data	Svm	Noyau	C	Nu	Gamma	Degree	Coef0
Data1	C-svc	Linear	1.578579	/	/	/	/
		Poly	1.578579	/	0.427865	2	/
		Rbf	1.578579	/	0.427865	/	/
		Sigmoid	1	/	0.427865	/	1
	Nu-svc	Linear	/	0.103488	/	/	/
		Poly	/	0.103488	0.1	4	/
		Rbf	/	0.1	0.01	/	/
		Sigmoid	/	0.1	0.001	/	3
Data2	C-svc	Linear	2.374327	/	/	/	/
		Poly	2.374327	/	78.07588	3	/
		Rbf	2.374327	/	78.07588	/	/
		Sigmoid	2.374327	/	78.07588	/	4
	Nu-svc	Linear	/	0.332497	/	/	/
		Poly	/	0.100045	35.46609	2	/
		Rbf	/	0.1	0.00273	/	/
		Sigmoid	/	0.303098	0.00273	/	0

Data 1 « Bas de données binaire » :

Tableau 16: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG optimisé pour dataset-1

classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.98006	0.980033	0.98004	0.98003
c-svc	linear	0.96211	0.96173	0.96171	0.96173
	poly	0.9817	0.981697	0.9817	0.9817
	rbf	0.90071	0.875208	0.87369	0.87521
	sigmoid	0.77395	0.77371	0.77375	0.77371
Nu-svc	linear	0.96703	0.966722	0.9667	0.96672
	poly	0.9817	0.981697	0.9817	0.9817
	rbf	0.98003	0.980033	0.98003	0.98003
	sigmoid	0.96356	0.963394	0.96338	0.96339
knn	X	0.97505	0.975042	0.97504	0.97504

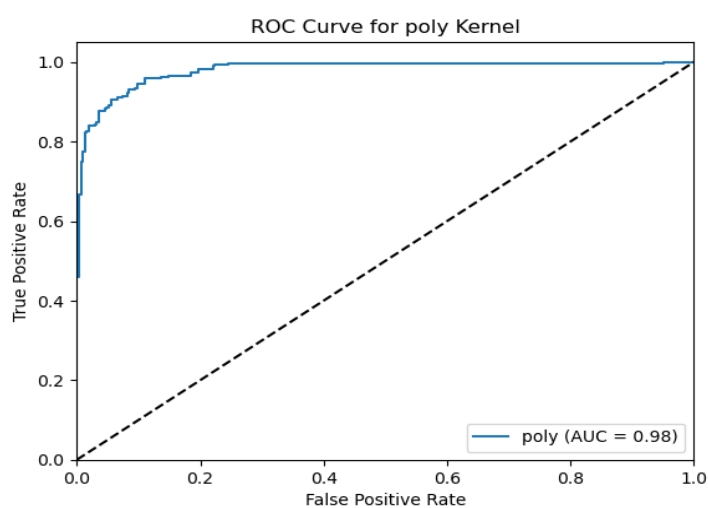
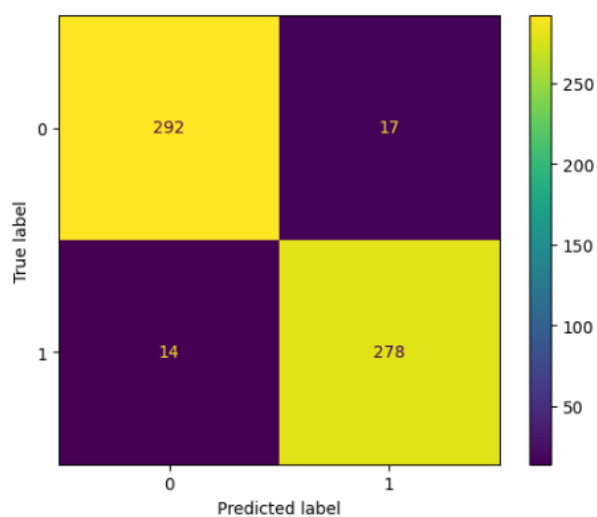


Figure 45 : matrice de confusion et le courbe Roc de HOG optimisé avec NU-SVC Poly pour dataset-1

Data 2 « Bas de données multi class » :

Tableau 17: Tableau de comparaison entre les différents classificateurs de SVM optimisé par PSO et KNN avec l'extracteur HOG optimisé pour dataset-2

classificateur	fonction	precision	sensitivity	f-mesure	accuracy
linear-svc	linear	0.90551	0.907262	0.906064	0.90726
c-svc	linear	0.91004	0.909886	0.909894	0.90989
	poly	0.94942	0.949256	0.949242	0.94926
	rbf	0.91121	0.91010	0.91103	0.91015
	sigmoid	0.69159	0.69105	0.69135	0.69129
Nu-svc	linear	0.88144	0.881015	0.880129	0.88102
	poly	0.94723	0.946632	0.946721	0.94663
	rbf	0.9394	0.939633	0.93931	0.93963
	sigmoid	0.69113	0.6893	0.69050	0.68929
knn	X	0.89548	0.895888	0.894232	0.89589

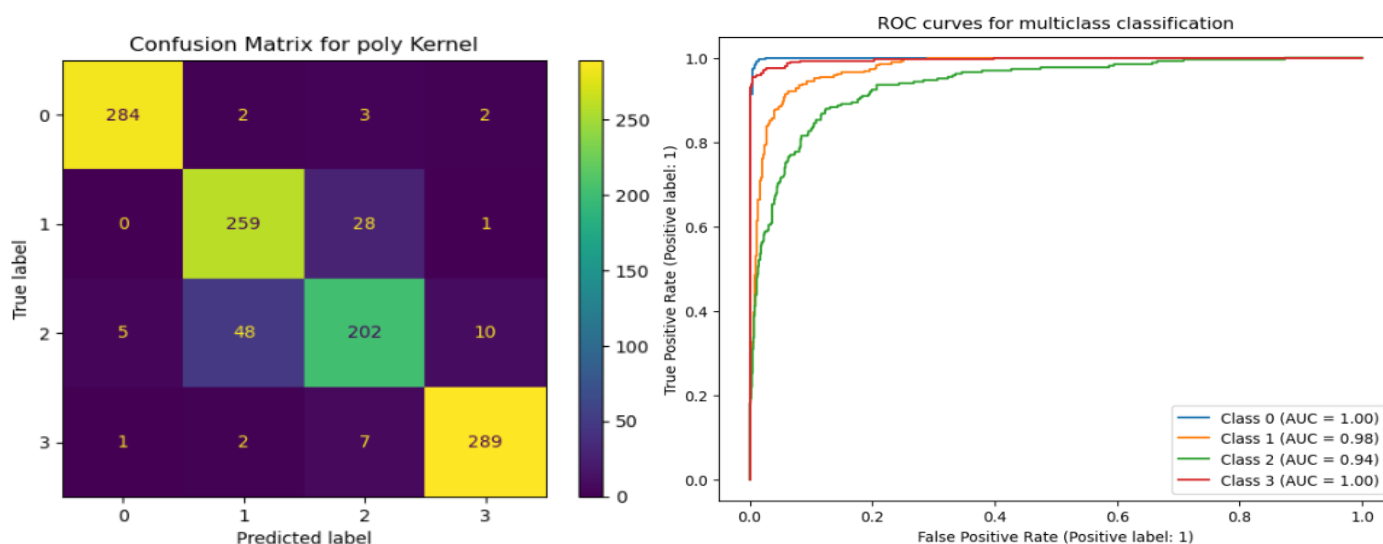


Figure 46: matrice de confusion et le courbe Roc de HOG optimisé avec C-SVC Poly pour dataset-1

3.7 Comparaison des résultats entre les six modèles :

Parmi les six modèles SVM différents, nous avons observé des niveaux de l'accuracy variables dans **data 1** et **data 2**. Dans **data 1**, les modèles **3**, **5** et **6** se sont démarqués en tant que meilleurs performeurs, atteignant tous une précision impressionnante de **98,17%**. Le modèle **2** suivait de près avec une accuracy de **98,0033%**. Les modèles **1** et **4** ont également affiché des précisions respectables de **97,504%** et **97,338%**, respectivement. En ce qui concerne l'ensemble **data 2**, c'est le modèle **4** qui s'est distingué en affichant la plus haute accuracy de **95,0131%**, démontrant ainsi sa performance solide. Le modèle **6** a obtenu la deuxième meilleure position avec une précision de **94,926%**, tandis que le modèle **3** a montré sa compétence avec une précision de **94,4882%**, lui valant ainsi la troisième meilleure position. Ces résultats soulignent l'importance de prendre en compte les caractéristiques spécifiques de chaque ensemble de données lors du choix du modèle SVM optimal.

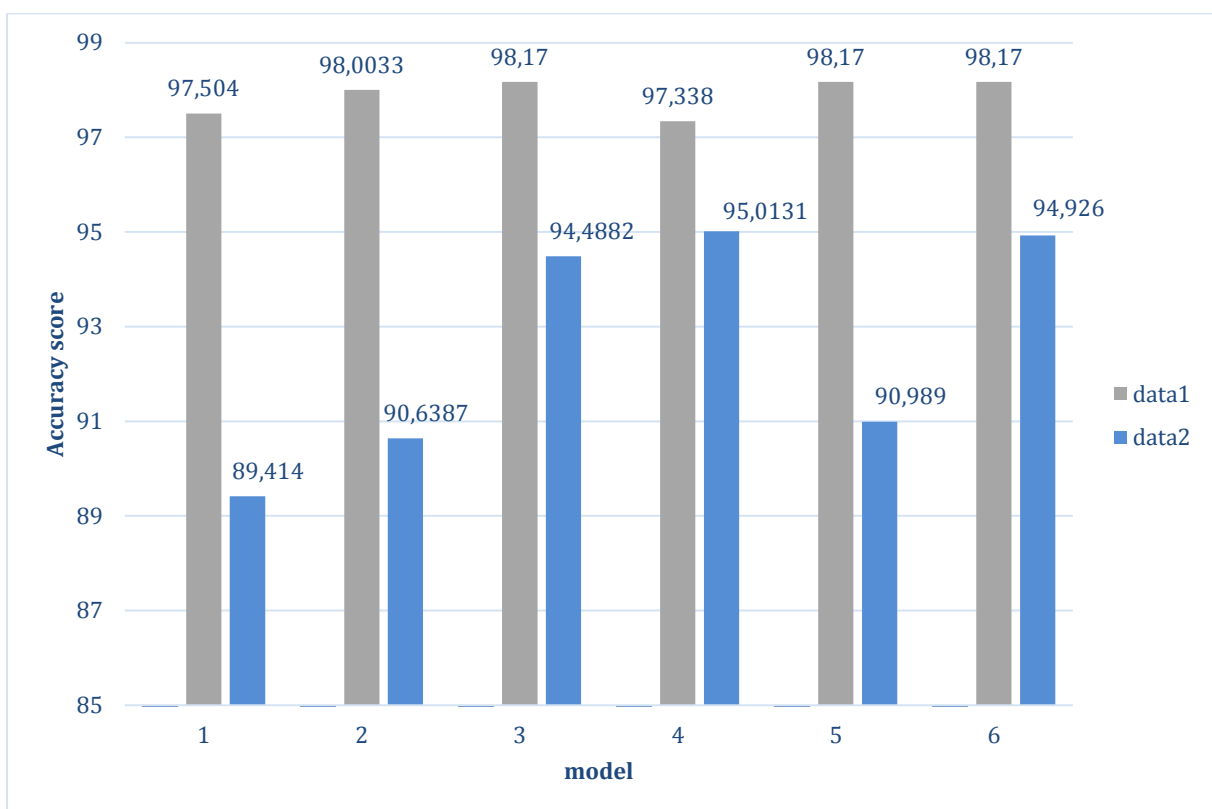


Figure 47 : Un diagramme à barres indique la meilleure accuracy entre les modèles

3.7.1 Comparaison des résultats entre les modèles de SVM et les modèles de SVM optimisé par PSO :

Dans les modèles SVM et les modèles SVM optimisés par PSO, comparons ceux qui ont obtenu les meilleurs résultats dans les bases de données data1 et data 2 :

Pour data 1 :

- Les modèles SVM les plus performants sont le modèle **3**, le modèle **5** et le modèle **6**, qui ont tous atteint une accuracy remarquable de **98,17%**.
- Le modèle SVM optimisé par PSO qui se rapproche le plus de ces performances est le modèle **2**, avec une accuracy de **98,0033%**.
- Les modèles SVM de base, le modèle **1** et le modèle **4**, ont également obtenu des accuracy respectables de **97,504%** et **97,338%** respectivement.

Pour data 2 :

- Le modèle SVM optimisé par PSO, le modèle **4**, s'est distingué en affichant l'accuracy la plus élevée de **95,0131%**.
- Le modèle SVM optimisé par PSO, le modèle **6**, a obtenu la deuxième meilleure accuracy avec **94,926%**.
- Le modèle SVM de base qui s'approche le plus de ces résultats est le modèle **3**, avec une accuracy de **94,4882%**.

Dans data 1, les modèles SVM ont globalement donné de très bons résultats, avec les modèles **3**, **5** et **6** se classant en tête. Dans data 2, ce sont les modèles SVM optimisés par PSO qui ont obtenu les meilleurs résultats, en particulier le modèle **4**. Cela indique que l'optimisation par PSO a contribué à améliorer les performances du SVM dans ce contexte spécifique.

Tableau 18 : Tableaux de comparaison entre les modèles avec SVM et les modèle SVM-PSO

SVM				
DATA	DATA 1		DATA 2	
Modèle	Fonction	Accuracy	Fonction	Accuracy
1	c-svc Linear	0.97504	c-svc Linear	0.89414
2	c-svc Poly	0.980033	c-svc Rbf	0.90639
5	c-svc Poly	0.9817	c-svc Linear	0.90989
SVM-PSO				
DATA	DATA 1		DATA 2	
Modèle	Fonction	Accuracy	Fonction	Accuracy
3	nu-svc Poly	0.9817	c-svc Rbf	0.94488
4	c-svc Linear	0.97338	c-svc Rbf	0.950131
6	c-svc Poly/ nu-svc Poly	0.9817	c-svc Poly	0.94926

3.7.2 Comparaison des résultats entre les modèles de HOG et les modèles de HOG optimisé :

Nous avons comparé entre les modèles qui utilisant les caractéristiques HOG simples (modèles 1, 2, 3 et 4) et les modèles utilisant les caractéristiques HOG optimisées (modèles 5 et 6) en termes de accuracy pour chaque base de données. Voici les résultats :

Pour data 1 :

- Les modèles avec les caractéristiques HOG simples ont obtenu l'accuracy suivantes : modèle 1 - **97,504 %** , modèle 2 - **98,0033 %** , modèle 3 - **98,17 %**, modèle 4 - **97,338 %**.
- Les modèles avec les caractéristiques HOG optimisées ont obtenu les précisions suivantes : modèle 5 - **98,17 %**, modèle 6 - **98,17 %**.

Pour data 2 :

- Les modèles avec les caractéristiques HOG simples ont obtenu les précisions suivantes : modèle 1 - **89,414 %**, modèle 2 - **90,6387 %**, modèle 3 - **94,4882 %**, modèle 4 - **95,0131 %**.

- Les modèles avec les caractéristiques HOG optimisées ont obtenu les précisions suivantes : modèle **5** - **90,989 %**, modèle **6** - **94,926 %**.

En conclusion, pour data 1, les modèles avec les caractéristiques HOG optimisées (modèles **5** et **6**) ont tous deux atteint une précision de **98,17 %**, les plaçant au premier rang. Pour data 2, le modèle avec les caractéristiques HOG simples (modèle **4**) a obtenu la meilleure précision de **95,0131 %**.

Ainsi, le modèle **5** avec les caractéristiques HOG optimisées est le meilleur pour le data 1, tandis que le modèle **4** avec les caractéristiques HOG simples est le meilleur pour le data 2. Ces modèles ont montré leur efficacité respective en fonction de la nature des données et des caractéristiques utilisées.

Tableau 19 : Tableaux de comparaison entre les modèles avec HOG et les modèle avec HOG optimisé

HOG				
DATA	DATA 1		DATA 2	
Modèle	Fonction	Accuracy	Fonction	Accuracy
1	c-svc Linear	0.97504	c-svc Linear	0.89414
2	c-svc Poly	0.980033	c-svc Rbf	0.90639
3	nu-svc Poly	0.9817	c-svc Rbf	0.94488
4	c-svc Linear	0.97338	c-svc Rbf	0.950131
HOG optimisé				
DATA	DATA 1		DATA 2	
Modèle	Fonction	Accuracy	Fonction	Accuracy
5	c-svc Poly	0.9817	c-svc Linear	0.90989
6	c-svc Poly/ nu-svc Poly	0.9817	c-svc Poly	0.94926

3.8 Conclusion :

Dans le précédent chapitre, nous avons examiné en détail la classification et les méthodes associées. Dans ce chapitre, nous mettons en œuvre et comparons plusieurs modèles basés sur les architectures d'apprentissage automatique SVM et KNN, en termes de précision. Nous explorons également l'impact de l'optimisation des paramètres HOG avec PSO sur SVM pour améliorer la précision. Nous présentons les outils utilisés, les modèles mis en place et les résultats obtenus. Notre approche se différencie de la conclusion précédente en mettant l'accent sur l'implémentation des modèles et leur performance, ainsi que sur l'optimisation des paramètres pour une meilleure précision. En résumé, notre étude pratique des concepts de classification en implémentant des modèles basés sur SVM et KNN, et évalue l'efficacité de PSO pour l'optimisation des paramètres HOG. Les résultats obtenus fourniront des informations utiles sur les performances des modèles dans le contexte du diagnostic des tumeurs cérébrales à partir d'images

IRM.

Conclusion Générale

Conclusion générale :

Notre recherche est axée sur l'application de l'apprentissage machine au diagnostic des tumeurs cérébrales humaines au moyen d'images IRM. L'imagerie par résonance magnétique est une forme d'imagerie médicale qui sert à obtenir des renseignements détaillés au sujet des organes et des tissus internes. Toutefois, l'interprétation d'images d'IRM complexes et la prise de décisions médicales peuvent s'avérer difficiles pour les médecins en raison du grand nombre de questions à analyser.

Dans notre étude, nous avons exploré différentes méthodes de classification supervisée, y compris les machines à vecteurs de support (SVM) et les k- plus proches voisins (KNN), pour améliorer la précision diagnostique des tumeurs cérébrales. Nous examinons également plusieurs méthodes d'extraction de caractéristiques, comme HOG et Kernel- PCA, et la segmentation avec K- signifie des méthodes pour améliorer la qualité des résultats.

En optimisant les paramètres SVM au moyen de l'algorithme d'optimisation PSO, nous avons atteint des performances supérieures en matière de classification des images IRM. Notre étude comparative de la SVM de base et optimisée démontre l'importance de l'optimisation des paramètres pour des résultats plus fiables et précis. À l'avenir, nous prévoyons élargir nos recherches en utilisant des ensembles de données plus vastes afin de créer des modèles plus puissants. Nous prévoyons également de déployer notre système dans un cadre clinique réel pour aider les professionnels de la santé à diagnostiquer avec précision les tumeurs cérébrales. En outre, nous souhaitons explorer davantage l'utilisation des algorithmes PSO pour sélectionner les fonctionnalités les plus pertinentes avant la classification.

En conclusion, nos recherches aident à améliorer les méthodes de diagnostic des tumeurs cérébrales en appliquant le machine learning aux images IRM. Nous espérons que notre travail offrira un soutien précieux aux professionnels de la santé et ouvrira de nouvelles opportunités pour améliorer les performances de la SVM dans le domaine médical.

Les Références :

- [01] Stadnik TW, Chaskis C, Michotte A. Imaging tutorial: differential diagnosis of bright lesions on diffusion-weighted MR images. *Radiographics*. 2003.
- [02] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*. Pages 221-248. 2017
- [03] Navneet Dalal ET Bill Triggs. "Histograms of Oriented Gradients for Human Detection" , pages 108-115 , 2005 .
- [04] Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., & White, L. E. *Neuroscience*. Sinauer Associates, Inc. 2012.
- [05] Bear, M. F., Connors, B. W., & Paradiso, M. A. *Neuroscience : Exploring the Brain*. Lippincott . 2016
- [07] , Ziad Chalboub : L'Intelligence Artificielle (IA) dans la santé , Ziad Chalboub manager de marketing digital de MCI
- [08] inskey ME, Weinand ME. Benign Brain Tumors. In: Winn HR, ed. *Youmans and Winn Neurological Surgery*. 7th ed. Philadelphia, pages 109-118 , 2017 .
- [09] Dr Daniel Gloaguen , Association médicale canadienne – Dec 13 , 2021 .
- [10] Steven A. Goldman, MD, PhD, Sana Biotechnology , Examen médical janv 2021.
- [11] Angelis LM. Brain tumors. Pages 114-123 , 2001 .
- [12] Ostrom QT, Gittleman H, Liao P , Primary Brain and Central Nervous System Tumors , Jun 2014
- [13] Dr Roux, Philippe Metellus, M.D., Ph.D., Clinique Clairval, Marseille, Johan Pallud, M.D., Ph.D., AP-HP Paris, Mars 2022.
- [14] Dr Sébastien Ronze , L'IRM (Imagerie par Résonance Magnétique) définition et déroulement et risques , Jan 01 , 2023
- [15] M. Zhou, J. Scott, B. Chaudhury, L. Hall, D. Goldgof, K.W. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert and R. Gatenby *American Journal of Neuroradiology* , 39 (2) 208-216 , Feb 2018 .
- [16] Thomas H. Berquist , Brain tumor imaging: a comprehensive review" - *American Journal of Roentgenology*, Jan 2015 - <https://www.ajronline.org/doi/full/10.2214/AJR.14.13619> .

- [17] Justin T Jordan, Elizabeth R Gerstner - Imaging of brain tumors: Recent advances and future directions - Neuroimaging Clinics of North America, Feb 2019 .
- [18] Zelaya, F. O., & Johnstone, Magnetic resonance imaging of the brain. In Essential Physics for Manual Medicine pages. 191-206. 2012.
- [19] LeCun, Y., Bengio, Y., & Hinton. Deep learning. Nature, 521(7553), pages 136-141 , 2015
- [20] Ian Witten , Eibe Frank , Data Mining: Practical Machine Learning Tools and Techniques Jun 2000.
- [21] Gareth James, Daniela Witten, Trevor Hastie et Robert Tibshirani , An Introduction to Statistical Learning , pages 205-209 , Feb 11 , 2013 .
- [22] Gareth James, Daniela Witten, Trevor Hastie et Robert Tibshirani , An Introduction to Statistical Learning , pages 221-227 , Feb 11 , 2013
- [23] Gareth James, Daniela Witten, Trevor Hastie et Robert Tibshirani , An Introduction to Statistical Learning , pages 238-242 Feb 11 , 2013
- [24] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." pages 436-444 , May 27 , 2015
- [25] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- [26] Kelleher, J. D., Mac Namee, B., & D'Arcy, Fundamentals of machine learning for predictive data analytics: algorithms, workflows, and case studies. Cambridge . April 2015
- [27] LeCun, Y., Bengio, Y., & Hinton, Deep learning. Nature, pages 445-464. 2015
- [28] Raschka , Python machine learning Vol. 1, pages. 1-12. Birmingham: Packt Publishing Ltd ., 2015
- [29] DR Aaron Kalb, Alation , Introduction aux algorithmes de classification , 09 sep 2014 .
- [30] K. Zhang, L. Zhang, and M. Patel, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 54, pages. 60-88, 2018.
- [31] Navneet Dalal ET Bill Triggs. "Histograms of Oriented Gradients for Human Detection" , pages 162-173 , 2005 .
- [32] Richard Szeliski , Computer Vision: Algorithms and Applications" , Aug 05 , 2010 .
- [33] Mika, Rätsch, Weston, Schölkopf et Müller , Kernel principal component analysis" , 2001
- [34] Wu, Y., Ianakiev, K., & Govindaraju, Improved k-nearest neighbor classification. Pattern recognition, 35(10), 2311-2318 , 2002.

- [35] Imane BENHMIDOU , Tuto Python & Scikit-learn : KNN (k-nearest neighbors). Nov 09 2022
- [36] .Alpaydin, E., Introduction to Machine Learning (3rd ed.), MIT Press, 2020.
- [37] Onel Harrison , Machine Learning Basics with the K-Nearest Neighbors Algorithm , Sep 10, 2018 .
- [38] Bernhard Schölkopf, Alexander J. Smola, Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond, 2002 .
- [39] Burges, C. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery , pages 121-167. 1998.
- [40] Cortes, C., Vapnik, V. Support-vector networks. Machine learning 20 (3), pp 273-297. , 1995
- [41] Picon A, Irusta U, Álvarez-Gila ,Machines à vecteurs supports , 2013 .
- [42] Steinwart, Christmann, Support Vector Machines, 2008.
- [43] Hastie, Tibshirani et Friedman "The Elements of Statistical Learning" (2e édition, page 413) , 2009
- [44] Hastie, Tibshirani et Friedman "The Elements of Statistical Learning" (2e édition, page 417) , 2009
- [45] Hastie, Tibshirani et Friedman , "The Elements of Statistical Learning" de (2e édition, pages 419-426) , 2009
- [46] Hausmane Issarane . "Les SVM, Support Vector Machine" . mars 01, 2019
- [47] Navneet Dalal ET Bill Triggs. "Histograms of Oriented Gradients for Human Detection" , pages 234-247 , 2005 .
- [48] Maurice Clerc , " Particle Swarm Optimization " pages 47-52 , Jan 01 , 2006 .
- [49] Wang, S., Han, Y., Li, Q., & Huang, Y. 2019. Parameter optimization of support vector machine based on particle swarm optimization for recognition of small targets. Journal of Applied Remote Sensing, 13(4), 046508. 2019
- [50] Zheng, Y., Cai, H., Li, H., & Feng, Z. Parameter optimization of SVM based on PSO algorithm for breast cancer diagnosis. BMC medical informatics and decision making, 19(S1), 24. 2019.

- [51] Liu, J., Chen, X., Huang, D., & Zheng, L. Multi-objective optimization of support vector machine parameters based on particle swarm optimization for image classification. *Journal of Ambient Intelligence and Humanized Computing*, 10(6), 2195-2204, 2019.
- [52] :F. Provost, T. Fawcett, and R. Kohavi. The Case Against Accuracy Estimation for Comparing Induction Algorithms. In *Proceedings of the Fifteenth International Conference on Machine Learning (ICML'98)*, pages 445-453, Madison, WI, 1998.
- [53] : Y. Bengio, J.-F. Paiement, P. Vincent, O. Delalleau, N. Le Roux, and M. Ouimet. Out-of-Sample Performance of Latent Variable Models with Multinomial Probit Regression Output. *Advances in Neural Information Processing Systems* 18, pages 89-96, 2006.
- [54] Amandine Allmang , LineData , Apprentissage Supervisé et Classification , May 03 , 2017 – LineData - <https://fr.linedata.com/apprentissage-supervise-et-classification>
- [55] Hand, D. J., & Till, R. J. A simple generalization of the area under the ROC curve for multiple class classification problems. *Machine learning*, 45(2), 171-186 , 2001.
- [56] Python Software Foundation. (Consulté le 26 juin 2023). Python Documentation. <https://www.python.org/doc/>.
- [57] Project Jupyter. (Consulté le 26 juin 2023). What is Jupyter ? <https://jupyter.org/>.