الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة الشهيد حمّه لخضر - الوادي
UNIVERSITY OF ELOUED
كلية العلوم الدقيقة

مذكرة مقدمة لنيل شهادة ليسانس أكاديمي

قسم الإعلام الآلـــي

تخصص : نظم معلوماتية

الموضوع

التنبؤ بأمراض القلب بإستخدام طرق التعلم العميق

تحت إشراف :
د.ستو طرابلس

من إعداد :
تامة أسامة
دقموش يوسف
بـن عمر نـور الهدى

السنة الدراسية 2024/2023

شكر وتقدير

في هذه المناسبة، نتوجه بخالص الشكر والتقدير لأستاذتنا ستو طرابلسي، التي كانت بمثابة النور الهادي لنا طوال فترة إعداد هذه المذكرة. لم تكن مجرد مرشدة، بل كانت مصدر إلهام ودعم لا محدود، فكانت بحكمتها ومعرفتها السبب الرئيسي لتحقيق هذا الإنجاز. كلمات الشكر لا توفيها حقها، فقد كانت بحق العمود الفقري لهذا العمل، وبدون توجيهاتها لم نكن لنصل إلى ما وصلنا إليه اليوم.

كما نتوجه بأسمى عبارات الشكر والتقدير لكل من درسونا، ولكل أطارات الكلية بصفة عامة، ولعميد الكلية البروفيسور عبد الوهاب منصور بصفة خاصة، على دعمهم المستمر وتوجيهاتهم القيمة التي ساعدتنا في تحقيق هذا العمل.

كما نأمل أن تكون هذه الدراسة مصدر إلهام للباحثين والمطورين في المجال الطبي، وأن تسهم في تطوير أدوات وتقنيات جديدة تسهم في تحسين حياة المرضى وتقديم رعاية صحية أفضل.

الإهداء

نهدي هذه المذكرة إلى أبائنا الأعزاء، سواء كانوا معنا اليوم أو قد غادروا هذا الدنيا.
إلى أبي الحبيب، الذي كان دائماً قدوتي ومصدر قوتي، كان دعمك اللامتناهي ووجه الخير مشروط هو القوة الدافعة التي جعلتني أؤمن بنفسي وأواصل طريق العلم. كلمات الشكر لا يمكن أن تعبر عن مدى امتناني لك. فكنت النور الذي يضيء دربي والأمان الذي يحتضنني في كل لحظة. لولاك لما كنت هنا اليوم.

أما أمهاتنا الغاليات، فلهن الحنان الذي يروي أرواحنا والعطاء الذي لا ينضب. كانت تضحيتهن وصبرهن هما السر وراء نجاحنا. فكنت السند والداعم الأول لنا. قلوبكم الكبيرة وحكمكم الذي لا يعرف الحدود هو ما جعلنا نستطيع مواجهة كل التحديات. شكراً لكن على كل شيء. فأنتم البوصلة التي توجهنا نحو النجاح.

الملخص

تُعتبر أمراض القلب من أخطر الأمراض شيُوعًا، حيث يُعتبر التنبؤ بأمراض القلب من أكثر الأمور تعقيدًا وصُعوبة كما أنها مكلِّفة، وذلك بسبب كثرة التحليلات والمؤشرات التي يصعُب تحليلُها وتقييمها، بما في ذلك العمر، الجنس، مستوى الكولسترول، مستوى السكر، معدل ضربات القلب، ...ألخ. وبالتالي تواجه الأنظمة الصِحية تحديات كبيرة في تشخيص هذه الأمراض بالطُرق التقليدية.

مؤخرًا، أثبتت تقنيات الذكاء الإصطناعي، تحديدًا طُرق التعلم الآلي وطُرق التعلم العميق بِصفة خاصة، دورها الكبير في تحسين جودة الطب، خُصوصًا التنبؤ بالأمراض المختلفة مثل أمراض القلب والسُكري وأورام الدماغ ومرض الزهايمر وغيرها.

في هذه الدراسة قُمنا بتطوير نموذج تعلم عميق DNN لتحليل مؤشرات البيانات الصحية من أجل التنبؤ بمرض القلب، حيث قُمنا بتدريب النموذج على مجموعة بيانات كبيرة من أجل تقديم نتائج دقيقة. بالإضافة إلى ذلك تحققنا من آداء النموذج المُقترح مُقارنة بطرق التعلم الآلي الأكثر إستخداما في الدراسات الحديثة، فأظهرت التجارب المكثفة فعالية النموذج حيث حقق تحسنًا كبيرًا مقارنة بالأساليب السابقة.

كلمات مفتاحية

HD	مرض القلب
DD	الكشف عن المرض
ML	التعلم الالي
SL	التعلم الخاضع الإشراف
SVM	دعم شاحنات النقل
LR	الانحدار اللوجستي
KNN	ك-أقرب الجيران
ANN	الشبكات العصبية الاصطناعية
UL	تعليم غير خاضع للاشراف
SSL	التعلم شبه الخاضع للإشراف
DL	التعلم العميق
NN	الشبكات العصبية
DNN	الشبكات العصبية العميقة

Abstract

Heart disease is considered one of the most common and dangerous diseases. Predicting heart disease is considered one of the most complex, difficult and expensive matters, due to the large number of analyzes and indicators that are difficult to analyze and evaluate, including age, gender, cholesterol level, sugar level, heart rate, etc. ..etc. Therefore, health systems face great challenges in diagnosing these diseases using traditional methods.

Recently, artificial intelligence techniques, specifically machine learning methods and deep learning methods in particular, have proven their major role in improving the quality of medicine, especially predicting various diseases such as heart disease, diabetes, brain tumors, Alzheimer's disease, and others.

In this study, we developed a deep learning model (DNN) to analyze health data indicators in order to predict heart disease. We trained the model on a large data set in order to provide accurate results. In addition, we verified the performance of the proposed model compared to the machine learning methods most used in recent studies. Extensive experiments showed the effectiveness of the model, as it achieved a significant improvement compared to previous methods.

Keywords

HD	Heart disease
DD	Disease Detectoin
ML	Machine learning
SL	Supervised learning
SVM	Support Vector Machine
LR	Logistic regression
KNN	K-nearest nighbers
ANN	Artifical Neural networks
UL	Unsupervised learning
SSL	Semi Supervised learning
DL	Deep learning
NN	Neural networkss
DNN	Deep Neural networks

Résumé

Les maladies cardiaques sont considérées comme l'une des maladies les plus courantes et les plus dangereuses. La prévision des maladies cardiaques est considérée comme l'une des questions les plus complexes, les plus difficiles et les plus coûteuses, en raison du grand nombre d'analyses et d'indicateurs difficiles à analyser et à évaluer, notamment l'âge, le sexe, le taux de cholestérol, le taux de sucre, la fréquence cardiaque, etc. Les systèmes de santé sont donc confrontés à de grands défis pour diagnostiquer ces maladies à l'aide de méthodes traditionnelles.

Récemment, les techniques d'intelligence artificielle, en particulier les méthodes d'apprentissage automatique et les méthodes d'apprentissage profond, ont prouvé leur rôle majeur dans l'amélioration de la qualité de la médecine, notamment dans la prévision de diverses maladies telles que les maladies cardiaques, le diabète, les tumeurs cérébrales, la maladie d'Alzheimer, etc.

Dans cette étude, nous avons développé un modèle d'apprentissage profond (DNN) pour analyser les indicateurs de données de santé afin de prédire les maladies cardiaques. Nous avons formé le modèle sur un grand ensemble de données afin de fournir des résultats précis. De plus, nous avons vérifié les performances du modèle proposé par rapport aux méthodes d'apprentissage automatique les plus utilisées dans les études récentes. Des expériences approfondies ont montré l'efficacité du modèle, car il a permis d'obtenir une amélioration significative par rapport aux méthodes précédentes.

Mots Clés

HD	Maladie cardiaque
DD	Détection des maladies
ML	Apprentissage automatique
SL	Enseignement supervisé
SVM	Machine à vecteurs de support
LR	Régression logistique
KNN	k-Voisins les plus proches
ANN	Réseaux de neurones artificiels
UL	Apprentissage non supervisé
SSL	Apprentissage semi-supervisé
DL	L'apprentissage en profondeur
NN	Les réseaux de neurones
DNN	Réseaux de neurones profonds

الفهرس

قائمة الجداول ..... أ
قائمة الأشكال. ..... ب
المقدمة العامة: 1
الفصل الأول: 2
المقدمة: 2
1. تعريف القلب : 2
1.1. بنية القلب :..... 2
1.2. الدورة الدموية للقلب : 3
1.1.2. الدورة الدموية الكبرى: 3
2.1.2. الدورة الدموية الصغرى. 3
1.3. وظيفة القلب : 4
2. مرض القلب : 4
2.1. أسباب مرض القلب : 5
2.2. أنواع أمراض القلب و أعراضه: 6
3. تشخيص مرض القلب : 7
3.1. التشخيص بالطرق التقليدية : 7
1.3.1. تحاليل الدم :..... 7
2.3.1. مخطط كهربية القلب .(ECG) : 7
3.3.1. ممارسة اختبار الإجهاد : 7
4.3.1. مخطط صدى القلب (موجات فوق الصوتية) :..... 7
5.3.1. صوير الأوعية التاجية: 7
6.3.1. تصوير الأوعية الدموية المقطعي المحوسب للشريان التاجي(CCTA):. 7
3.2. تشخيص عن طريق تقنيات الذكاء الإصطناعي : 8
4. طرق العلاج من أمراض القلب : 11
الخاتمة: 11
الفصل الثاني: 12
المقدمة: 12
1. خوارزميات التعلم الآلي : 12
1.1. تعريف التعلم الآلي : 12
1.2. انواع التعلم الالي : 13

1.1.2. التعلم الخاضع للإشراف: 13
2.1.2. التعلم غير الخاضع للإشراف: 16
3.1.2. التعلم شبه الخاضع للإشراف: 16
2 خوارزميات التعلم العميق: 17
2.1 تعريف التعلم العميق: 17
2.2 طرق التعلم العميق: 17
1.2.2.الشبكات العصبية العميقة .(DNN) 17
2.2.2.الشبكات العصبية التلافيفية .(CNN) 19
3.2.2. الشبكات العصبية المتكررة.(RNN) 19
4.2.2.لشبكات العصبية طويلة المدى وقصيرة المدى(LSTM). 19
5.2.2. الشبكات التوليدية الخصومة(GAN). 19
2.3. مقارنة بين التعلم الآلي و التعلم العميق: 20
3. أعمال ذات صلة : 21
3.1. التعرف على مرض القلب عن طريق التعلم الآلي: 21
3.2. التعرف على مرض القلب عن طريق التعلم العميق: 23
الخاتمة: 25
الفصل الثالث: 26
المقدمة: 26
1. التصميم : 26
1.1. وصف البيانات : 26
1.2. معالجة البيانات : 27
1.3. التعلم و التدريب : 28
2. الشبكة العصبية العميقة المقترحة : 28
2.1. طبقة الإدخال : 28
2.2. الطبقات المخفية : 29
2.3. طبقة الإخراج : 29
2.4. دوال التنشيط : 30
2.5. الإختبار و التحقق : 30
3. التنفيذ : 31
3.1. بيئة التنفيذ : 31
1.3.1. أدوات التطوير: 31
2.3.1. خصائص الجهاز: 32
3.3.1. كيفية حساب الدقة : 32
3.2. خطوات التنفيذ : 33
1.3.2. تحميل وتحضير البيانات: 33
2.3.2. حويل البيانات النصية لتمثيل رقمي: 33

3.3.2. تقسيم البيانات إلى مجموعات التدريب والاختبار :. 34
4.3.2. إستخدام دوال التنشيط وتحديد الخسارة : 35
5.3.2. لتحسين و إختيار الخصائص: 36
6.3.2. تحليل الأداء: 37
3.3 التجارب: 37
1.3.3. تجارب التعلم الالي: 38
2.3.3. تجارب التعلم العميق: 38
3.4. النتائج : 41
4. إستنتاج : 41
5. واجهات النظام : 41
الخاتمة: 43
الخاتمة العامة: 44

قائمة الجداول

الفصل الثاني:

1.2 أهم الفروقات بين التعلم الآلي والتعلم العميق: 20
2.2 بعض الدراسات السابقة في استخدام التعلم الآلي التعرف على أمراض القلب: 22
3.2 بعض الدراسات السابقة في استخدام التعلم العميق التعرف على أمراض القلب: 23

قائمة الاشكال

الفصل الأول:
1.1 رسم تخطيطي لبنية القلب: 3
الفصل الثاني:
1.2 علاقة التعلم الآلي بالذكاء الاصطناعي: 13
2.2 LR : 14
3.2 SVM : 14
4.2 خوارزمية الجار الأقرب: 15
5.2 الشبكات العصبية الاصطناعية 16
6.2 علاقة التعلم العميق بالتعلم الالي:. 17
1.3 كيفية عمل الشبكات العصبية العميقة:. . . . . 18.
الفصل الثالث:
2.3 مخطط هندسة الشبكة العصبية العميقة المقترح لتشخيص أمراض القلب:.....29
3.3 مقارنة بين طريقة عمل دوال التنشيط: 30
4.3 تحميل البيانات من ملف CSV 33
5.3 تحويل البيانات الفئوية لتمثيل رقمي:. 34
6.3 تقسيم البيانات للتدريب و الاختبار:. 34
7.3 نموذج كيراس للتصنيف الثنائي: 35
8.3 تحسين النموذج:. 36
9.3 تطبيق تقنية pca لتقليل الأبعاد.... 36
10.3 نتيجة دقة النموذج: 37
11.3 الشكل النهائي لنموذج الشبكة العصبية العميقة:. 40
12.3 مقارنة بين نتائج خوارزميات التعلم الآلي و العميق: 41
13.3 واجهة برمجة الهاتف لنموذج التنبؤ بأمراض القلب: 43

المقدمة العامة

تُعد أمراض القلب من أكثر الأمراض المُزمنة في العالم و التي بسببها أصبحت الأنظمة الصحية تواجه تحديات كبيرة في تشخيصها ومتابعتها نظراً لأنها معقَّدة ومُكلفة ، بحيث تعتمد الطرق التقليدية في التشخيص على الفحوصات الطبية والإختبارات المخبرية التي تكون غالباً مُكلفة كما أنها تستغرق وقتاً طويلاً، وهذا يَحدُ من فعاليتها.

بينما تتضمن الحلول الأُخرى إستخدام تقنيات متقدمة مثل تخطيط كهربائية القلب ،(ECG) فحوصات الدم المتنوعة، وتقنيات التصوير الطبي، لكن هذه الحلول أيضا تواجه قيوداً مثل التكلفة العالية، الحاجة إلى أجهزة متخصصة، وضرورة المتابعة الطبية المستمرة.

مؤخرًا، شهدت مجالات الذكاء الإصطناعي والتعلم الآلي تطورات سريعة ومُذهلة، مما جعلها أدوات فعَّالة في تحسين جودة الطب والرعاية الصحية. وبالتالي يُعدُ تطبيق هذه التقنيات المُتقدمة بشكل خاص في التنبُؤ بالأمراض المزمنة مثل أمراض القلب، السكري،أورام الدماغ، الجلد، أورام الكبد، مرض الزهايمر وأمراض الكلى المزمنة وغيرها، خطوة مهمة جدًا و حاسمة من أجل تحسين جودة الرعاية الصحية لأن هذه الأمراض تتطلب تشخيصاً دقيقاً ومتابعة مستمرة.

تهدف هذه الدراسة إلى تطوير نموذج تعلم عميق يُمكنه تحليل مؤشرات البيانات الصحية وتقديم توقعات دقيقة من أجل التنبؤ بمرض القلب. مما يُساهم في تحسين النتائج الصحية وتقليل التكاليف المُرتبطة بالعلاج والمتابعة الطبية. بالإضافة إلى ذلك قُمنا بمُقارنة آداء النموذج المُقترح مع التعلم الآلي الأكثر إستخداما في الدراسات الحديثة.
سنقوم بتقسيم دراستنا الى ثلاثة فصول للقيام بهذا العمل :

الفصل الأول "مرض القلب": يتناول تعريف القلب وبنيته، ووظيفة القلب، أمراض القلب وأعراضها، وأساليب التشخيص التقليدية والحديثة.

الفصل الثاني "التعلم الآلي والتعلم العميق": سنعرض أساسيات التعلم الآلي والتعلم العميق، والفرق بينهما، وكيفية استخدامهما في تحليل البيانات والتنبؤ بالأمراض.

الفصل الثالث "التصميم والتنفيذ": سنشرح تصميم وتطوير النموذج المإقترح بإستخدام تقنيات التعلم العميق،ويشمل خُطوات مُعالجة البيانات، تدريب النموذج، وتقييم أدائه.

باب 1. الفصل الأول

مقدمة :

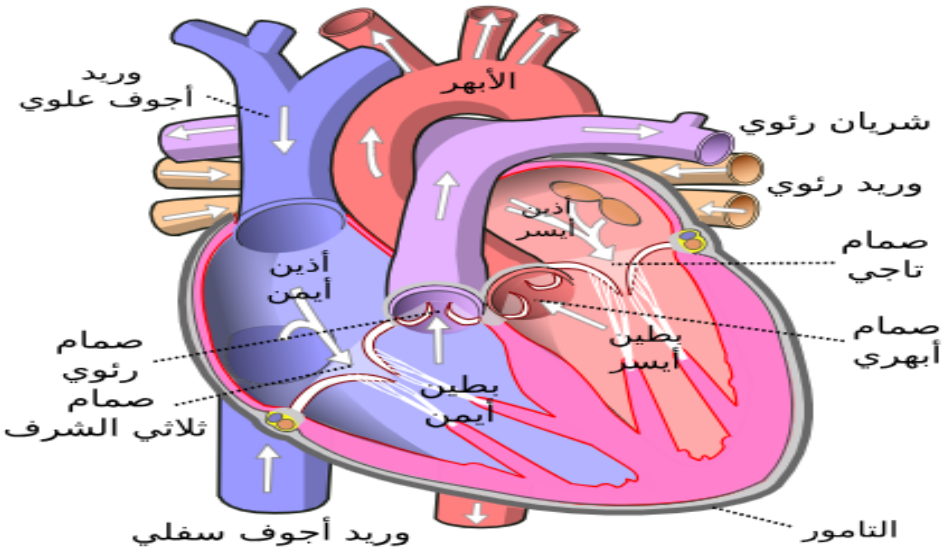
في هذا الفصل، سوف نتحدث عن مرض القلب ، بحيث سنتكلم من جهة على كل ما يخُص القلب من تعريف وشرح لِوظيفته الاساسية في الجسم، و سنتعرف على طبقاته الغشائية ، و دوره المهم في ضخ الدم ، و من جهة أخرى سنعرض مختلف الأمراض التي تؤثر على القلب و كيفية حدوثها ،وسنتعمق ايضا في العوامل المسببة لمرض القلب وأهم أعراضه ، و في الاخيرسننهي الفصل بدراسة طرق تشخيص مرض القلب وذلك إما بالطرق التقليدية أوعن طريق تقنيات الذكاء الاصطناعي.

1 تعريف القلب :

يعتقد الإنسان أن مكان القلب في جسم الإنسان على الجانب الأيسر من الصدر، ولكن في الحقيقة هذا الإعتقاد خاطئ حيث أن موقع القلب في وسط صدر جسم الإنسان، بالضبط بين الرئتين خلف عظمة القص (عضلة مسطحة تتواجد في منتصف الصدر)،وتميل قمته نحو اليسار قليلاً. القلب هو العضو الأساسي في جسم الإنسان حيث يمكن تعريفه بأنه مضخة عضلية بحجم قبضة اليد تضخ الدم إلى أنحاء الجسم عبر الأوعية الدموية، تعمل دون توقف وتنظم الإيقاع والنبضات. يزن القلب ما بين 230 إلى 350 جراماً تقريباً، ويضخ من 5 إلى 7 لتر في الدقيقة، و7600 لتر تقريباً في اليوم .[1]

1.1 بنية القلب :

يتكون القلب من أجزاء مختلة كل منها دور في الدورة الدموية: الأذينان والبطينان الأيمن والأيسر، بالإضافة إلى الصمامات التي تسمح بالدوران بين الأذينين والبطينين.يتكون قلب الانسان من ثلاث طبقات كما هو موضح في الشكل 1، الطبقة ألاولى هي الشغاف وهي الطبقة الداخلية الرقيقة في القلب، ثم تليها الطبقة الثانية و المسماة بعضلة القلب وهي الطبقة الوسطى و السميكة في القلب، وفي الاخير الطبقة الثالثة وهي الطبقة الخارجية التي تحيط بالقلب وهي النخاب. يتم تغذية القلب عن طريق الشرايين التاجية والشرايين المنعطفة. يدخل الشريان الأورطي والشرايين الرئوية إلى القلب، وكذلك الأوردة الرئوية والأجوف السفلي.[2]



شكل 1.1. **رسم تخطيطي لبنية القلب**

1.2 الدورة الدموية للقلب :

يقوم القلب بالتعاون مع الأوعية الدموية المنتشرة في جميع أنحاء الجسم، لضخ الدم و تزويد الخلايا بالأكسجين و المواد الضرورية من خلال ما يعرف بالدورة الدموية الصغرى و الدورة الدموية الكبرى .

1.1.2. الدورة الدموية الكبرى :

يعتبر الجزء الأيسر من القلب المسؤول الرئيسي عن الدورة الدموية الكبرى. التي تبدأ من عملية انقباض البطين الأيسر وفتح الصمام الأبهر؛ لضخ الدم المؤكسد الموجود بداخله إلى جميع أنحاء الجسم، من خلال الشريان الأبهر. و بعد ذلك يعود الدم غير المؤكسد القادم من جميع أعضاء الجسم إلى الأذين الأيمن عبر الوريد الأجوف العلوي والوريد الأجوف السفلي إلى القلب، ليستقر بعدها في البطين الأيمن عبر مروره من الصمام ثلاثي الشرف.

2.1.2. الدورة الدموية الصغرى :

يعتبر الجزء الأيمن من القلب المسؤول الرئيسي عن الدورة الدموية الصغرى. التي تتضمن عملية انقباض البطين الأيمن المحمل بالدم غير المؤكسد، ليفتح الصمام الرئوي، ويخرج الدم عبر الشريان الرئوي. فيصل الدم غير المؤكسد إلى الرئتين ويتأكسد هناك، بعد ذلك يعود الدم المؤكسد القادم من الرئتين إلى القلب. فيدخل إلى الأذين الأيسر عبر الأوردة الرئوية، ويستقر بالبطين الأيسر بعد مروره من خلال الصمام البطيني الأذيني، المعروف باسم الصمام المترالي أو الصمام ثنائي الشرف. وذلك لحدوث الدورة الدموية الكبرى مرة أخرى. [1]

1.3 وظيفة القلب :

ينبض القلب أكثر من 100 ألف مرة في اليوم ويعمل كمضخة تنقبض وتتوسع: بعد تلقي الدم الغني بالأكسجين، يعيده إلى جميع الأعضاء بشكل مستمر. تنقسم دورة القلب إلى مرحلتين: الانبساط الذي يمتلئ فيه القلب بالدم، والانقباض الذي ينقبض خلاله الدم لإخراج الدم إلى الأعضاء الأخرى. ثم يتوسع القلب قبل أن يتلقى تدفقًا جديدًا من الدم، ثم ينقبض مرة أخرى. وتشير التقديرات إلى أن قلب الشخص السليم ينبض ما بين 60 إلى 80 مرة في الدقيقة. ومع ذلك، فإن هذه الفكرة نسبية، فمن المرجح أن ينبض قلب الرياضيين بشكل أبطأ بكثير، وأحيانًا من 30 إلى 40 نبضة في الدقيقة. يتم تحديد معدل ضربات القلب (في الراحة أو النشاط) من خلال العقدة الجيبية، الموجودة في الجزء العلوي من الأذين الأيمن للقلب. يتم ري القلب عن طريق الشرايين التاجية والشرايين المنعطفة، المسؤولة عن تزويده بالأكسجين والمواد المغذية التي يحتاجها ليعمل بشكل طبيعي.[2]

2 مرض القلب :

يشير مصطلح مرض القلب إلى مجموعة متنوعة من الأمراض، التي تؤثر على وظيفة القلب، بحيث تصيب هذه الأمراض عضلة القلب، أو صماماته، أو الغشاء المحيط به، أو الشرايين والأوردة الرئيسية من القلب . تبدأ أمراض القلب بنوبات من الألم الحاد، نتيجة انسداد أحد الشرايين التي توصل الدم والأكسجين إلى القلب، وبالتالي يقل معدل الأكسجين الذي يصل إليه، أو قد يتوقف كليًا مسببًا الجلطات القلبية و الذبحات الصدرية، وغيرها من الأمراض المزمنة، والتي قد تشكل خطرًا على حياة المريض. ووفقًا لمراكز السيطرة على الأمراض والوقاية منها، فإن أمراض القلب هي السبب الرئيسي للوفاة في الولايات المتحدة .[1]

2.1 أسباب مرض القلب :

هناك العديد من أسباب مرض القلب، أكثرها شيوعا : إرتفاع ضغط الدم ، إرتفاع مستوى الكوليسترول ، التدخين ، إستهلاك الكحول ،السمنة، قلة النشاط البدني،سوء التغذية ، السكري، و التقدم في العمر، التوتر العصبي:

- **إرتفاع ضغط الدم :**

يؤثر ارتفاع ضغط الدم أيضًا على الأوعية الدموية، فيسبب زيادة في سمك جدرانها نتيجة لترسب الكوليسترول والدهون على جدر الأوعية من الداخل، وهو ما يرفع من معدلات الإصابة بالنوبات القلبية والسكتات الدماغية.[1]

- **إرتفاع نسبة الكوليسترول :**

إرتفاع مستويات الكولسترول في الدم يمكن أن يزيد من خطر تكوين اللويحات وتصلب الشرايين .[3]

- **التدخين :**

نعلم أن التدخين من العوامل الخطرة و الرئيسية للإصابة بمرض القلب و الأوعية الدموية ، بحيث يعمل التدخن في ترسب الدهون في الأوعية الدموية مما يؤدي إلى تضييقها ، كما أن تعاطيه يزيد من معدل الوفيات بأمراض القلب ، ولذلك فإن الإقلاع عن التدخين هو الطريقة الأكثر فعالية للوقاية من أمراض القلب و الأوعية الدموية.[3]

- **إستهلاك الكحول :**

أكدت جمعية القلب الأمريكية أن الإستهلاك القوي للكحول يؤدي إلى زيادة الدهون الثلاثية في الجسم، وزيادة التوتر الشرياني وزيادة خطر قصور القلب وإيقاع القلب.[3]

- **السمنة :**

تعتبر السمنة إحدى عوامل المسببة لأمراض القلب ، حيث أن زيادة مستويات الدهون في الجسم، وخاصة دهون البطن، يمكن أن تتسبب برفع مستويات الالتهاب في الجسم وتزيد من مقاومة الأنسولين، وهي عوامل خطرة يمكن أن تزيد من احتمالية حدوث تصلب الشرايين.[1]

- **قلة النشاط البدني :**

يؤدي جلوس الفرد فترات طويلة إلى الحدّ من قدرة العضلات على حرق الدهون، مما يزيد من احتمالية حدوث الانسداد والإصابة بمرض الشريان التاجي، والإصابة بأمراض القلب الوعائية.[1]

- **سوء التغذية :**

اتباع نظام غذائي غني بالدهون والملح والسكر والكوليسترول يمكن أن يساهم في تطور أمراض القلب.[2]

- **مرض السكري :**

يؤدي إرتفاع مستورى السكر في الدم إلى تلف الأوعية الدموية ، و بالتالي تؤدي المواد الدهنية بالتراكم بداخلها ، مما يؤدي إلى تصلبها ويحدث ما يسمى بتصلب الشرايين الذي يعمل على إعاقة تدفق الدم إلى القلب و الدماغ و حدوث نوبة قلبية أو سكتة دماغية قاتلة .[1]

- **التقدم في العمر :**

يزيد التقدم في السن من خطر تلف الشرايين وضيقها وضعف أو سماكة عضلة القلب.[2]

- **التوتر العصبي :**

يؤدي التوتر العصبي إلى ضيق شديد ينتج عنه تغيرات فيسيولوجية حادة مثل: ارتفاع في ضغط الدم و عدم إنتظام في ضربات القلب، وقد يحث أيضا تغيرات في ديناميكية حركة الدم على جدران الشرايين مما يؤدي إلى ترسيب الكوليستيرول و الدهنيات.[4]

2.2 أنواع مرض القلب و أعراضه:

هناك عدة أنواع من أمراض القلب كل منهم يختلف بإختلاف مسبباته، بالإضافة إلى ذلك، كل نوع له أعراض خاصة به، نذكر فما يلي أهمها :

● مرض القلب التاجي :

مرض القلب التاجي أو ما يسمى أيضًا بمرض الشريان التاجي، هو مرض يصيب الأوعية الدموية التي تغذي عضلة القلب.[5]

أعراضه :

- برودة الذراعين و الساقين
- الغثيان
- تنميل

● أمراض الوعائية الدماغية :

الأمراض الدماغية الوعائية أو السكتة الدماغية أو الحوادث المخية، هو مرض يصيب الأوعية الدموية التي تغذي الدماغ.[5]

أعراضه :

- صداع شديد و مستمر
- ضعف في الأطراف
- تشنجات العصبية

● مرض الشرايين المحيطية :

مرض الشرايين المحيطية، هو مرض يصيب الأوعية الدموية التي تغذي الذراعين و الساقين.[5]

أعراضه:

- تخدر و برودة في الساق
- تقلصات مؤلمة في عضلات الوركين أو الفخذين
- إنعدام النبض في الساقين أو القدمين

● أمراض القلب الروماتيزمية :

هو مرض التهابي، تلف عضلة القلب وصمامات القلب بسبب الحمى الروماتيزمية التي تسببها البكتيريا.[5]

أعراضه :

- طفح جلدي بارز أحمر اللون
- ضيق في التنفس

● **أمراض القلب الخلقية :**

العيوب الخلقية التي تؤثر على التطور الطبيعي ووظيفة القلب الناجمة عن عيوب في بنية القلب منذ الولادة.[4]

أعراضه :

- تورم الأطراف و إزرقاق الجلد
- إنتفاخ الساقين و القدمين
- عدم إنتظام ضربات القلب

3 تشخيص مرض القلب :

يمكن تشخيص أمراض القلب بالطرق التقليدية التي تشمل الفحوصات الطبية المباشرة ، ولكنها قد تكون مكلفة و تتطلب وقتاً طويلاً.مع تقدم التكنولوجيا ، دخلت تقنيات الذكاء الإصطناعي في تشخيص أمراض القلب لتحسين الدقة و الكفاءة مما يساهم في تحسين تجربة المريض و تقليل تكاليف الرعاية الصحية.

3.1 التشخيص بالطرق التقليدية :

1.3.1. تحاليل الدم :

عندما تتضرر عضلة القلب ، كما هو الحال في نوبة قلبية، يفرز الجسم مواد في دم. يمكن لاختبارات الدم قياس المواد وإظهار ما إذا كانت عضلة القلب قد تضررت ومدى تلفها. يتم إجراء اختبارات الدم أيضًا لقياس مستوى المواد الأخرى في الدم، مثل دهون الدم (مثل الكوليسترول والدهون الثلاثية) والفيتامينات والمعادن. يتم أخذ عينة من دم المريض من الوريد في ذراع. ثم يقوم المختبر باختباره وإرسال النتائج إلى طبيب مختص لمعاينتها.[6]

2.3.1. مخطط كهربية القلب (ECG):

يقرأ مخطط الكهربائي للقلب (ECG) النبضات الكهربائية للقلب. يظهر مدى نبض قلب. يتم وضع نقاط لزجة صغيرة وأسلاك سلكية على الصدر و الذراع والساقيين. يتم توصيل الخيوط بجهاز تخطيط القلب الذي يسجل النبضات الكهربائية ويطبعها على الورق. قد يستخدم الطبيب مخطط الكهربائي للقلب (ECG) لتشخيص نوبة قلبية أو عدم انتظام ضربات القلب.[6]

3.3.1. ممارسة اختبار الإجهاد :

اختبار الإجهاد، الذي يُطلق عليه أحيانًا اختبار "المطحنة" أو "التمرين"، هو نوع من تخطيط الكهربائي للقلب (ECG) يتم إجراؤه أثناء ممارسة التمارين الرياضية. يساعد طبيب على معرفة مدى جودة عمل القلب عندما يكون الشخص نشيطًا بدنيًا.[6]

4.3.1. مخطط صدى القلب (الموجات فوق الصوتية) :

مخطط صدى القلب هو اختبار شائع. يعطي صورة للقلب باستخدام الموجات فوق الصوتية، وهو نوع من الأشعة السينية. ويستخدم مسبارًا إما على الصدر أو أسفل المريء (الحلق). فهو يساعد الطبيب على التحقق مما إذا كانت هناك أي مشاكل في صمامات القلب وغرفه، ومعرفة مدى قوة ضخ القلب للدم.[6]

5.3.1. تصوير الأوعية التاجية :

يمكن إجراء تصوير الأوعية التاجية، والذي يسمى أحيانًا "قسطرة القلب"، بعد الإصابة بنوبة قلبية أو ذبحة صدرية. يتم وضع قسطرة (أنبوب صغير) في شريان في الفخذ أو الذراع أو الرسغ. يتم تحريك القسطرة لأعلى داخل الشريان حتى تصل إلى القلب. يتم حقن صبغة خاصة في الشرايين التاجية ويتم أخذ الأشعة السينية. تُظهر الأشعة السينية لطبيب مكان ومدى انسداد أو انسداد الشرايين التاجية. كما يظهر أيضًا مدى جودة ضخ القلب.[6]

6.3.1. تصوير الأوعية الدموية المقطعي المحوسب للشريان التاجي (CCTA) :

هذا نوع متخصص من التصوير المقطعي المحوسب (CT)، الذي يمكن استخدامه للمساعدة في تشخيص مرض الشريان التاجي. إنه اختبار غير جراحي للأشخاص الذين قد يعانون من أعراض قلبية غير عادية.[6]

3.2 تشخيص عن طريق تقنيات الذكاء الإصطناعي :

بالرغم من أنه يُمكن الكشف عن مرض القلب بإستخدام التقنيات التقليدية وذلك عن طريق الطبيب الذي يقوم بتحديد المرض و تشخيصه من خلال التحاليل و الإختبارات الطبية وغيرها، إلا أن هذه التقنيات المُعقدة تتطلب جُهداً، ووقتا طويلا من أجل التشخيص الدقيق، كما أنها مكلِّفة. لِهذا في السنوات الأخيرة نجِد أن تطور التكنولوجيا في مجال الذكاء الإصطناعي الذي يحاكي الذكاء البشري في معالجة البيانات و التعامل مع القضايا المُعقدة، أثبت فعاليته و قُدرته الكبيرة في العديد من المجالات، من بينها المجال الطِبي. فنجد أن إستخدام تقنيات الذكاء الإصطناعي في معالجة البيانات الطبية قد حسَّن من فعالية أنظمة التشخيص. فعلى سبيل المثال تُستخدام تقنيات الذكاء الإصطناعي، تحديدا التعلم الآلي والتعلم العميق لتشخيص مرض قصور القلب بِناءاً على تحليل الإشارات الكهربائية للقلب. كما أنه تم إستخدامه في معالجة وتحليل صور مخططات صدى القلب والأوعية الدموية. كما نجد أنه تم إستخدام هذه التقنيات، خصوصا التعلم العميق

لإكتشاف التشوُهات الهيكلية عن طريق صور مرضى القلب. كما أنه تم إستخدامه في تحليل بيانات وسجِلات المرضى وذلك من أجل تدريب نماذج قادرة على التنبؤ بمرض القلب، وهذا بناءًا على بعض المؤشرات المهمة والحاسمة والتي تختلف قِيَمُها بين الشخص السليم والشخص المريض بمرض القلب. فيما يلي سنشرح أهم المؤشرات الرقمية التي يتم من خلالها الكشف عن مرض القلب:

- **فئة العمر AgeCategory:**

مؤشر فئات العمر المُدرجة في البيانات تصنف للأشخاص ضمن مجموعات عمرية محددة تساعد في تحليل البيانات بناءً على العمر.

- **كتلة الجسم BMI :**

مؤشر الكتلة الجسم ، وهو مقياس لتقدير الدهون في الجسم إستناداً إلى الطول و الوزن.

- **الجنس Sex:**

يمثل الجنس.(Female = أنثى) (Male = ذكر)

- **التدخين Smoking:**

يشير ما إذا كان الشخص مدخناً أم لا.(Yes = مدخن , No = غير مدخن)

- **شرب الكحول AlcoholDrinking:**

الإستهلاك المفرط للكحول يؤدي إلى زيادة الدهون الثلاثية في الجسم ، التي تؤثر سلباً على صحة القلب . حيث تشير(Yes = يستهلك الكحول No, = لا يستهلك الكحول)

- **النشاط البدني PhysicalActivity:**

يشير إلى ما إذا كان الشخص يمارس النشاط البدني بانتظام.(Yes = يمارس النشاط البدني) (No =لا يمارس النشاط البدني)

- **مدة النوم SleepTime:**

يشير هذا المؤشر إلى متوسط عدد ساعات النوم في الليلة .

- **مرض السكري Diabetic:**

يصف هذا المؤشر إذا كان الفرد مصاباً بالسكري ، و يشمل :
- (Yes = مصاب بمرض السكري)
- (No = ليس مصاب بمرض السكري)
- لا، سكري حدودي(No, borderline diabetes)
- نعم، أثناء السكري(Yes, during pregnancy)

- **الجلطة الدماغية Stroke:**

يشير ما إذا كان الشخص قد اصيب بجلطة دماغية أم لا.(Yes = أصيب بجلطة دماغية , No = لم يصاب بجلطة دماغية)

- **الصحة العقلية MentalHealth:**

يشر إلى عدد الأيام في الشهر الأخير التي شعر فيه الشخص بأن صحته العقلية ليست جيدة.

- **الصحة الجسدية PhysicalHealth:**

يشر إلى عدد الأيام في الشهر الأخير التي شعر فيه الشخص بأن صحته الجسدية ليست جيدة.

- **الصحة العامة GenHealth:**

صحة القلب مرتبطة بشكل كبير بالصحة العامة للفرد ، حيث يشير هذا المؤشر إلى تقييم الحالة الصحية العامة للفرد . بحيث تم تصنيفها إلى :

- Excellent (ممتازة)
- good Very (جيدة جدًا)
- Good (جيدة)
- Fair (معقولة)
- Poor (سيئة)

- **أمراض الكلى KidneyDisease:**

يشير إلى ما إذا كان الشخص يعاني من مرض الكلى أم لا.(Yes = يعاني من مرض الكلى) (No = لا يعاني من مرض الكلى)

- **سرطان الجلد SkinCanser:**

يشير إلى ما إذا كان الشخص قد أصيب بسرطان الجلد أم لا.(Yes = أصيب بسرطان الجلد) (No = لم يصاب بسرطان الجلد)

- **الربو Asthma:**

يشير إلى ما إذا كان الشخص يعاني من الربو أم لا.(Yes = يعاني من الربو) (No = لا يعاني من الربو)

- **صعوبة المشي Diffwalking:**

مؤشر صعوبة المشي ، يشير إلى ما إذا كان الشخص يعاني من صعوبة في المشي أو صعود السلالم.(Yes = يعاني من صعوبة في المشي) (No = لا يعاني من صعوبة في المشي)

- **العرق Race:**

مؤشر العرق أو ما يسمى بالأصل العرقي ، حيث يتم تصنيف الأشخاص وفقا للعرق أو الجنسية ، مما يسمح بتحليل البيانات لمعرفة كيف تختلف الصحة بين الجنسيات المختلف . حيث الأعراق

المدرجة في البيانات تشمل :
- White (أبيض)
- Black (أسود)

● **مرض القلب HeartDisease:**

يشير ما إذا كان الشخص يعاني من مرض القلب أم لا.(Yes = مصاب بمرض القلب)(No = غير مصاب بمرض القلب)

لكن وقع إختيارنا على الـ11 خاصية الاولى فقط بناءً على أهميتها في التأثير على صحة القلب وإمكانية إستخدامها في التنبؤ بمرض القلب.

4 طرق العلاج من أمراض القلب :

يتم علاج الحالات البسيطة من أمراض القلب عن طريق الأدوية و التي من الممكن أن تكون كافية لإدارة الأعراض و منع تطور الحالة المرضية. و من بين هذه الأدوية مضادات التخثر مثل الأسبرين، وأدوية خفض الكوليسترول، ومثبطات الإنزيم المحول للأنجيوتنسين، ومثبطات الألدوستيرون، وغيرها من الأدوية . أما في الحالات المعقدة التي تشمل إنسداداً في الشرايين، أو فشلًا في عضلة القلب، أو إضطرابات نظم القلب الخطيرة، فقد يتطلب الأمر تدخلًا طبيًا أكثر تعقيدًا مثل العمليات الجراحية. تشمل هذه العمليات جراحة الشرايين التاجية، وزرع القلب، وإجراءات أخرى مثل تركيب أجهزة تنظيم ضربات القلب.[1]

خاتمة :

عرضنا خلال هذا الفصل حوصلة عامة حول مرض القلب، من تعريف وأسباب وأعراض وصولا إلى الأنواع الموجودة. كذلك تناولنا طرق الكشف عن المرض التقليدية وسلبياتها من جهة, ثم ذكرنا أهمية إستغلال تقنيات الذكاء الإصطناعي في الكشف عن المرض من جهة أُخرى. هذه الأخيرة سنتعمق فيها خلال الفصل الثاني، حيث سنشرح كل من التعلم الآلي والتعلم العميق، كما أننا سنعرض أهم الدراسات في التنبؤ بمرض القلب التي تعتمد على هذه التقنيات.

باب 2. الفصل الثاني

مقدمة :

خلال السنوات الاخيره اثبتت طرق التعلم الالي والتعلم العميق بصفه خاصه دورها الكبير في مختلف المجالات مثل التعرف على الصور التعرف على الصوت التعرف على الخط الكشف عن الكائنات داخل صوره او فيديو انظمه الامن التنبؤ بالكوارث والامراض الى اخره من خلال الدراسات الحديثه نجد ان طرق تعلم العميق حققت نسب نجاح عاليه جدا وذلك لقدرتها العاليه على استخراج خصائص دقيقه من المعطيات الخام وايضا القدره على تحليل البيانات ذات حجم كبير

قدمنا خلال الفصل السابق كل المفاهيم المتعلقه بالكشف عن مرض القلب حيث بدانا بشرح المفاهيم الاساسيه المختلفه المتعلقه بقلب الانسان ثم تطرقنا الى الحالات المرضيه للقلب من حيث الاسباب والانواع والاعراض وانهينا الفصل بدراسه طرق تشخيص مرض مرضى مرض القلب وذلك بالطرق التقليديه او عن طريق تقنيه الذكاء الاصطناعي

في هذا الفصل سنتعمق في طرق الكشف عن امراض القلب بتقنيات الحديثه حيث سنقدم احدث ما توصلت اليه التكنولوجيا في التعلم الالي والتعلم العام اولا سنقوم بتفصيل المفاهيم الاساسيه لتعلم الالي وذكر انواعها والطرق الموجوده ثم سنشرح الانتقال من التعلم الالي الى التعلم العميق بالاضافه الى ذلك مفهوم التعلم العميق ونذكر جميع الطرق الموجوده

1 خوارزميات التعلم الآلي :

لقد احدث مجال التعلم الالي ثوره في مجال الذكاء الاصطناعي من خلال دراسه تصميم وتطوير الخوارزميه التي تعطي الالة القدره على تعلم دون برمجه صريحه حيث تم استخدام هذه الخوارزميات في العديد من المجالات مثل المجالات الطبيه والتشخيصيه المجالات التجاريه وغيرها , حيث نجد ان استخدام هذه الطرق في مختلف المجالات يعطي نتائج مرضيه الى حد ما

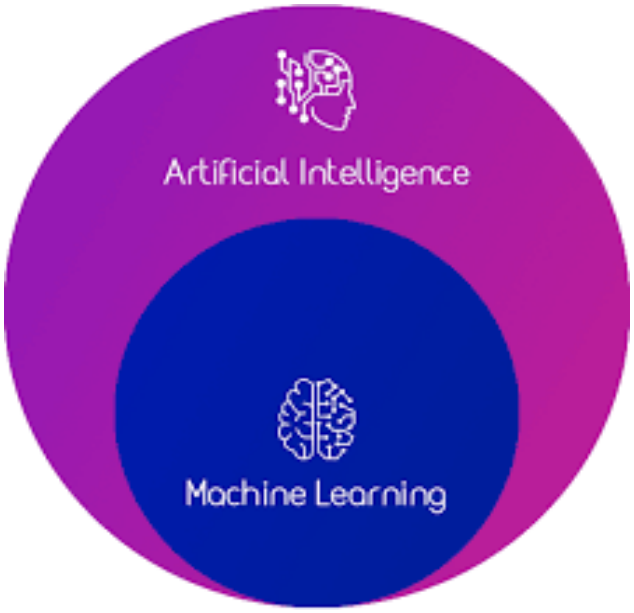
1.1 تعريف التعلم الآلي :

يُعد التعلم الآلي أو Machine learning بالإنجيلزية, رمزه المختصر ML ,فرعًا رئيسيًا من فروع الذكاء الإصطناعي (أنضر الشكل 1.2). يمكن تعريفه على أنه قدرة الآلة على تعلم إتخاذ القرار وذلك من خلال التدريب عن طريق مجموعة كبيرة من البيانات (بيانات التدريب). حيث نجد أن بيانات التدريب تختلف باختلاف مجال البحث فمثلا تكون عبارة عن عدد كبير من الصور في التعرف على الصور, النصوص في معالجة النصوص, الصوت في التعرف على الأصوات, قواعد البيانات أو جداول البيانات ...ألخ. وبالتالي ML هو علم يحاكي طريقة الإنسان في التعلم. فهو يهدف إلى جعل الآلة تتصرف وتحل مختلف المشاكل بالطريقة التي يقوم بها الإنسان الطبيعي. [7]

يكون التعلم الآلي عبرمرحلتين مهمتين متكاملتين, الأولى التعلم أو Training بالانجيلزية والثانية الإختبار وإتخاذ القرار أو Testing

1- التعلم : تبدأ مرحلة التعلم بمعالجة البيانات عن طريق تشفير البيانات و تحجيمها و استخراج الميزات الاكثر تاثيرا ، ثم يتم تقسيم البيانات الى قسمين بيانات التعلم و بيانات الاختبار ، حيث يتم تمرير بيانات التعلم مع الميزة المستهدفة ليقوم بتحليل البيانات و تحديد الانماط و العلاقات بينها ، ثم يقوم بتوقعات و يقارنها بالقيم الحقيقية ، ثم يقوم بتحسين اداءه عن طريق تقليل الفارق بين القيم الحقيقية و القيم الناتجة عن النموذج .

2- الاختبار : بعد انتهاء مرحلة التعلم يتم تمرير بيانات الاختبار للنموذج بدون الميزة المستهدفة ، ليقوم النموذج بتوقعات ثم يتم حساب الفرق بين توقعات النموذج للميزة المستهدفة بالميزة المستهدفة الاصلية لتقييم دقة النموذج ، حيث يمثل مجموع الفرق نسبة خطأ النموذج



شكل 1.2. **علاقة التعلم الآلي بالذكاء الاصطناعي**

1.2 انواع التعلم الالي :

يتم تحديد نوع التعلم الآلي وذلك بناءا على طريقة تلقي التعلم أثناء مرحلة تدريب النموذج. وبالتالي نجد ثلاثة أنواع التعلم الخاضع للاشراف ، التعلم غير الخاضع للاشراف و التعلم شبه الخاضع للاشراف .

1.1.2. التعلم الخاضع للإشراف :

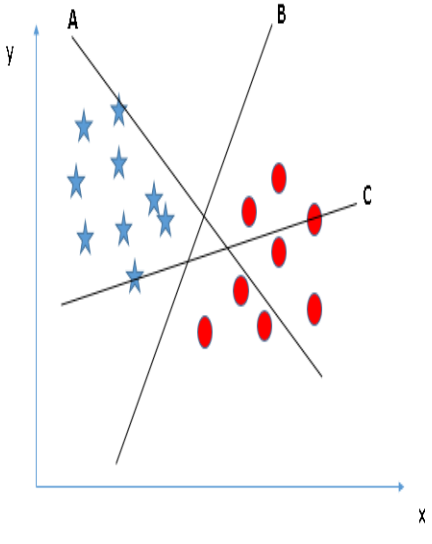
يُعدُ التعلم الخاضع للإشراف, (بالانجيلزية Supervised Learning) ,من أشهر أنواع التعلم الآلي حيث يكون فيها التدخُل البشري أثناء مرحة التدريب, فنجد أن التدريب موجها عن طريق مشرف أو مصمم, الذي يقوم بعملية إستخراج خصائص البيانات, ثم يرفقها بنوعها أي class, وبالتالي التدريب في هذه الحالة يكون عن طريق مدخلات مشكلة من أزواج (خصائص

البيانات, نوع المخرجة). الخصائص المستخرجة من البيانات يجب أن تتميز بالتباين العالي بين مختلف الأصناف,بينما يجب أن تتشابه بالنسبة لنفس الصنف. وهذا ينتج عليه الدقة العالية في التعلم,وبالتالي يُصبح النموذج قادرا على التنبؤ بعناصر جديدة.

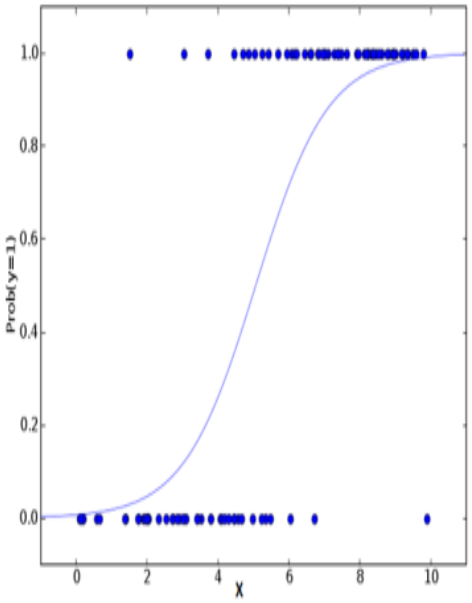
يتم استخدام العديد من الخوارزميات اثناء تطوير نماذج للتعلم الالي الخاضع للاشراف ، تختلف هذه الخوارزميات في طريقة عملها ، حيث تنقسم الى قسمين بناء على طريقة التنبؤ بالنتائج الى خوارزميات التصنيف التي تعتمد على تقسيم البيانات الى فئات حيث تمثل كل فئة صنف من النتائج ، وخوارزميات الانحدار التي تعتمد على ايجاد العلاقات الخطية و الغير خطية بين البيانات و النتائج ، نذكر منها الخوارزميات التالية :

- الة المتجه الداعم (Support Vector Machine) : خوارزمية آلة المتجه الداعم يرمز لها SVM هي خوارزمية تعلم آلي خاضع للإشراف تستخدم غالبا في مسائل التصنيف و التنبؤ ، تعتمد في طريقة عملها على ايجاد المستو الفائق (hyperplane) الذي يقسم البيانات الى عدة اصناف (انظر الشكل 2.2) في الفضاء ذو البعد n حيث يمثل n عدد السمات

- الانحدار اللوجستي (Logistic Regression) : خوارزمية الانحدار اللوجستي يرمز لها LR هي خوارزمية تعلم آلي خاضع للإشراف تستخدم غالبا في مسائل التصنيف و التنبؤ بالمتغيرات الفئوية ، تعتمد في طريقة عملها على استخدام الدالة اللوجيستية التي تقسم البيانات الى صنفين 0 و 1 (انظر الشكل 3.2) في الفضاء ذو البعد n حيث يمثل n عدد السمات

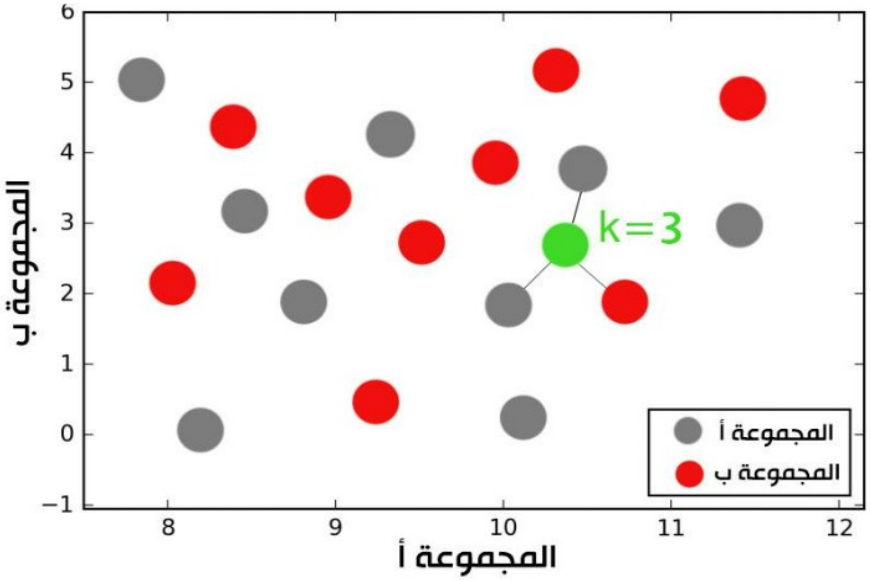


شكل 2.2. **SVM**



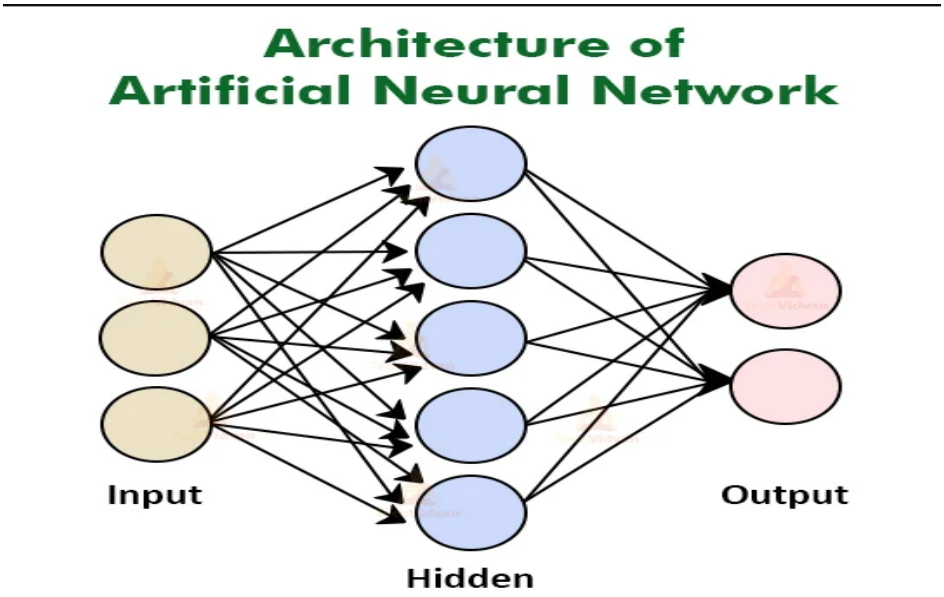
شكل 3.2. **LR**

- ك-أقرب الجيران (neighbors k-nearest) : خوارزمية ك-أقرب الجيران يرمز لها knn هي خوارزمية تعلم آلي خاضع للإشراف تستخدم غالبا في مسائل التصنيف و التنبؤ لها القدرة أيضا على التعامل مع البيانات الشاذة أو القيم المتطرفة بكفاءة عالية ، تعتمد في طريقة عملها على حساب المسافة الاقليدية بين النقاط في الفضاء ذو البعد n حيث يمثل n عدد الاصناف (انظر الشكل 4.2) ، حيث انه كلما قلت المسافة بين نقطتين زادت إحتمالية إنتماء النقطتين لنفس الفئة أي class ، و يشير الحرف k إلى عدد العينات التي سيتم تصنيف نقطة ما بناء على المسافات بينها و بين جيرانها الذين يبلغ عددهم k ، يعني إذا افترضنا أن k = 3 فإن الخوارزمية تقيس المسافة بين النقطة المستهدفة و أقرب ثلاث نقاط إليها .



شكل 4.2. **ك-أقرب الجيران**

- الشبكات العصبية الاصطناعية (networks neural Artificial) : الشبكات العصبية الاصطناعية يرمز لها ANN هي خوارزمية تعلم آلي خاضع للإشراف مستوحاة من عمل الدماغ البشري ، تتكون من ثلاث طبقات ، طبقة المدخلات و طبقة او طبقات مخفية و طبقة المخرجات ، تحتوي كل طبقة على مجموعة من العقد ،بحيث عُقد طبقة المدخلات هي تمثل البيانات المُدخلة، ثم يتم تمريرُها لعُقد الطبقة المخفية. يتم إستقبال هذه البيانات على مستوى كل عُقدة من عُقد الطبقة المخفية مضروبة في الأوزان، بحيث يتم جمعها وإدخالها إلى دالة التفعيل ، ثم يتم إرسال مُخرجات دالة التفعيل إلي الطبقة الموالية. تستقبل عُقد طبقة المُخرجات بنفس الطريقة البيانات ، إلا أن هذه الطبقة تتميز بوجود دالة تصنيف قادرة على تمييز مخرجات البيانات المُدخلة (الشكل 5.2) [7]



شكل 5.2. **الشبكات العصبية الاصطناعية**

2.1.2. التعلم غير الخاضع للإشراف :

يُعدُ التعلم غير الخاضع للإشراف ، (بالانجيلزية Unsupervised Learning) ، من أشهر أنواع التعلم الآلي حيث يغيب فيه التدخُل البشري أثناء مرحة التدريب ،فنجد أن التدريب يكون بشكل ألي فلا يكون موّجها ولا يحتاج لمُشرف، حيث أن الخوارزمية تقوم بإستنتاج نوع الفئات المُخرجة بشكل تلقائي من خلال بيانات التدريب غير مُرفقة بنوع المخرجة ، وبالتالي التدريب في هذه الحالة يكون عن طريق مدخلات بدون اي بيانات اخراج (بيانات الادخال).

يكمن الفرق بين التعلم الخاضع للاشراف و التعلم غير الخاضع للاشراف في مرحلة التدريب ، حيث نقوم تمرير خصائص البيانات و نوع المخرج في حالة التعلم الخاضع الاشراف ، بينما نمرر خصائص البيانات فقط في حالة التعلم غير الخاضع للاشراف و يتم استخلاص نوع المخرج عن طريق خوارزميات التجميع التي تقوم بتقسيم البيانات الى مجموعات تتشابه في ميزات معينة معا او خوارزميات قواعد الربط لاكتشاف العلاقات بين البيانات لإستخلاص النتائج .

3.1.2. التعلم شبه الخاضع للإشراف :

التعلم الالي شبه الخاضع للإشراف (بالانجليزية Semi-supervised learning) ، من اشهر انواع التعلم الالي اين يتم دمج اسلوب التعلم الالي الخاضع خاضع للإشراف (super-vised learning) ، حيث تستخدم خصائص البيانات و نوع المخرجة في تدريب الخوارزميات على النتبؤ بالمخرجات الصحيحة ، و الغير خاضع للإشراف (unsupervised learning) حيث تستخدم بيانات الادخال في تدريب الخوارزميات على اكتشاف علاقات و هياكل اكثر تعقدا في البيانات .

2 خوارزميات التعلم العميق :

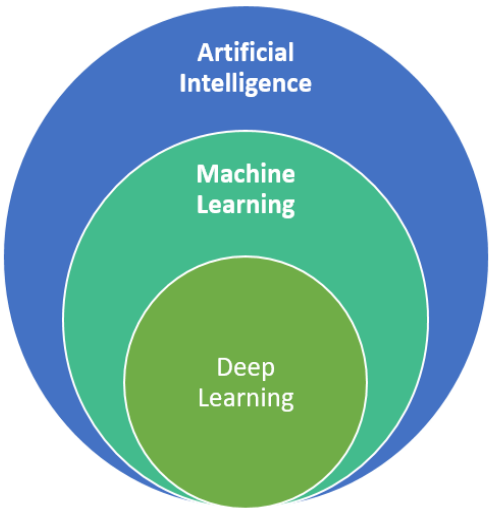
لقد أحدث مجال التعلم العميق تطور في العديد من جوانب الذكاء اصطناعي والحوسبة بشكل عام ، فنجد ان الجمع بين الشبكات العصبية الاصطناعية المعقدة وخوارزميات التعلم الآلي المتطورة حقق تقدم كبير في مجالات مثل رؤية الكمبيوتر و معالجة اللغة الطبيعية، والتعرف على الكلام ، و انظمة تشخيص الأمراض والتنبؤ بالمخاطر، والعديد من التطبيقات الأخرى.

2.1 تعريف التعلم العميق:

التعلم العميق أو Deep Learning بالإنجليزية، رمزه المختصر DL، هو مجموعة فرعية من طرق التعلم الآلي، وبالتالي هو بدوره فرع من فروع الذكاء الاصطناعي (انظر الشكل 6.2). يعتمد على حل المشكلات عن طريق الشبكات العصبية (neural networks). لهذا يطلق عليه ببساطة شبكة عصبية عميقة. بمعنى أدق، هو عبارة عن شبكة عصبية عادية ذات عدد كبير من الطبقات المخفية.

التعلم العميق يمكن أن يتعامل مع كميات ضخمة من البيانات ويستطيع استخراج الأنماط والخصائص المخفية التي لا يمكن للطرق التقليدية اكتشافها بسهولة. يستخدم التعلم العميق بشكل واسع في مجالات متعددة مثل معالجة الصور، التعرف على الصوت، معالجة اللغات الطبيعية، وغيرها من التطبيقات التي تتطلب تحليل بيانات معقدة.

الشبكات العصبية العميقة تتكون من طبقات متعددة من العُقد (nodes)، حيث تكون الطبقة الأولى طبقة الإدخال، تليها طبقات مخفية متعددة، وتنتهي بطبقة الإخراج. كل عقدة في الشبكة تقوم بعملية تحويل خطية أو غير خطية على البيانات التي تستقبلها من الطبقة السابقة، ثم تقوم بإرسال النتيجة إلى الطبقة التالية.



شكل 6.2. **علاقة التعلم العميق بالتعلم الالي**

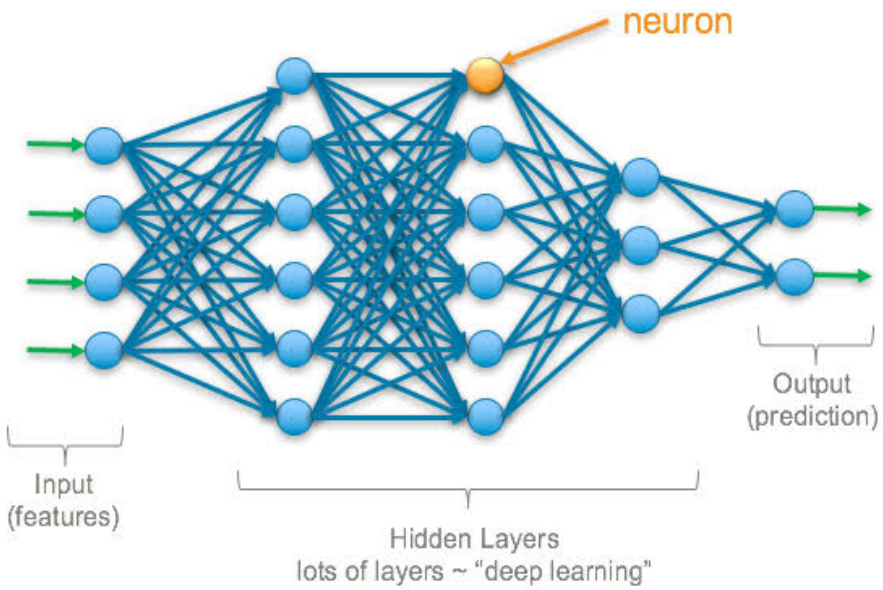
2.2 طرق التعلم العميق

1.2.2. الشبكات العصبية العميقة (DNN)

الشبكات العصبية العميقة (Deep Neural Networks - DNN) هي نوع من الشبكات العصبية الاصطناعية التي تتكون من طبقات متعددة من العقد العصبية. تستخدم DNN بشكل رئيسي في حل المشكلات المعقدة التي تتطلب اكتشاف الأنماط والتنبؤات الدقيقة.
تتكون DNN من ثلاث طبقات رئيسية:

- **طبقة الإدخال (Input Layer)**
  تستقبل البيانات الخام أو المدخلات.
- **الطبقات المخفية (Hidden Layers)**
  تتكون من عدة طبقات بين طبقة الإدخال وطبقة الإخراج. تعتمد كل عقدة على عمليات حسابية لتحويل البيانات.
- **طبقة الإخراج (Output Layer)**
  تنتج النتائج النهائية مثل التصنيف أو التنبؤ.

تتدرب DNN عبر عمليتي الانتشار الأمامي (Forward Propagation) والانتشار العكسي (Backward Propagation) لتحديث الأوزان بناءً على الخطأ الناتج في التنبؤ.
كما ان DNN أثبتت نجاحها في العديد من المجالا مثل تصنيف الصور، التعرف على الصوت، معالجة اللغات الطبيعية، والألعاب. والشكل التالي وضح طريقة عملها



شكل 7.2. **كيفية عمل الشبكات العصبية العميقة**

2.2.2. الشبكات العصبية التلافيفية (CNN)

الشبكات العصبية التلافيفية (Convolutional Neural Networks) هي نوع من الشبكات العصبية العميقة تستخدم بشكل رئيسي في معالجة الصور والتعرف على الأنماط البصرية. تعتمد هذه الخوارزمية على تطبيق مرشحات (filters) على الصور لاستخراج الميزات ذات الصلة مثل الحواف، الزوايا، والأنماط الأخرى في الصور. يتم استخدامها في العديد من التطبيقات مثل تصنيف الصور، اكتشاف الأشياء، ومعالجة الصور الطبية. تعتمد الخوارزمية على طبقات تلافيفية تليها طبقات تجميعية (pooling layers) ثم طبقات كاملة الترابط (fully connected layers) للوصول إلى النتيجة النهائية.[8]

من بين عيوبها، يمكن أن تكون الشبكات العصبية التلافيفية معقدة وتتطلب كميات كبيرة من البيانات والتدريب، مما يجعلها مكلفة من حيث الحسابات والوقت. كما أنها قد تعاني من مشاكل في التعرف على الأنماط إذا كانت البيانات المدربة عليها محدودة أو غير متنوعة بشكل كافٍ.

3.2.2. الشبكات العصبية المتكررة (RNN)

الشبكات العصبية المتكررة (Recurrent Neural Networks) هي نوع من الشبكات العصبية تستخدم لمعالجة البيانات التسلسلية مثل النصوص والكلام. تعتمد هذه الخوارزمية على استخدام مكونات متكررة يمكنها تذكر المعلومات السابقة في التسلسل من خلال وحدات الذاكرة الخاصة بها. تُستخدم الشبكات العصبية المتكررة في تطبيقات مثل الترجمة الآلية، تحليل المشاعر، والتعرف على الكلام. تعتمد الخوارزمية على توصيل الخلايا العصبية بطريقة تسمح للمعلومات بالانتقال من خطوة إلى أخرى عبر الزمن.

ومع ذلك، قد تواجه الشبكات العصبية المتكررة التقليدية صعوبة في تذكر المعلومات على مدى فترات طويلة وتعاني من مشكلة تلاشي التدرجات (vanishing gradients)، مما يجعل تدريبها أكثر تحديًا.

4.2.2. الشبكات العصبية طويلة المدى وقصيرة المدى (LSTM)

الشبكات العصبية طويلة المدى وقصيرة المدى (Long Short-Term Memory) هي نوع من الشبكات العصبية المتكررة (RNN) يمكنه تذكر المعلومات لفترات زمنية طويلة ويستخدم في معالجة النصوص والكلام. تعتمد هذه الخوارزمية على وحدات LSTM التي تحتوي على خلايا ذاكرة يمكنها الاحتفاظ بالمعلومات لفترة طويلة عن طريق بوابات تنظم تدفق المعلومات. يتم استخدامها في تطبيقات مثل الترجمة الآلية، التعرف على الكلام، وتوليد النصوص. توفر هذه الخوارزمية حلولًا لمشاكل النسيان والتداخل التي تواجه الشبكات العصبية المتكررة التقليدية.

لكن عيوب LSTM تشمل تعقيدها الحسابي الذي يتطلب موارد حاسوبية كبيرة وزمن تدريب طويل، مما قد يجعلها غير مناسبة لبعض التطبيقات التي تتطلب معالجة سريعة وفورية.

5.2.2. الشبكات التوليدية الخصومة (GAN)

الشبكات التوليدية الخصومة (Generative Adversarial Networks) تتكون من شبكتين عصبيتين تتنافسان مع بعضهما البعض لتوليد بيانات جديدة تبدو واقعية. الشبكة الأولى تُعرف

بالمولد (generator) الذي يحاول إنشاء بيانات مزيفة، بينما الشبكة الثانية تُعرف بالمميز (discriminator) الذي يحاول تمييز البيانات المزيفة عن البيانات الحقيقية. تُستخدم هذه الخوارزمية في تطبيقات مثل توليد الصور، تحسين الصور، وإنشاء بيانات تركيبية لأغراض التدريب. تتيح هذه الخوارزمية توليد محتوى جديد وإبداعي يعتمد على البيانات الحقيقية.

ومع ذلك، يمكن أن تكون تدريب GANs غير مستقر وصعب، وقد تتطلب ضبطًا دقيقًا للعديد من المعلمات لتحقيق توازن بين الشبكتين المولدة والمميزة. كما أنها قد تعاني من مشكلة التدرج المنحدر (mode collapse)، حيث يتم توليد أنواع محدودة من البيانات فقط.

2.3 مقارنة بين التعلم الآلي و التعلم العميق:

و في هذا الجدول سوف نستعرض اهم الفروقات بين التعلم الالي و التعلم العميق

جدول 1.2. اهم الفروقات بين التعلم الالي و التعلم العميق

الفرق	التعلم الآلي	التعلم العميق
تعريف	مجال فرعي من الذكاء الاصطناعي يسمح للحواسيب بالتعلم من البيانات دون برمجة صريحة	مجموعة فرعية من التعلم الآلي تستخدم الشبكات العصبية العميقة لمعالجة البيانات
الهياكل الأساسية	يعتمد على خوارزميات مثل الانحدار اللوجستي وأشجار القرار وK-نيرست نيبرز	يعتمد على الشبكات العصبية العميقة مع طبقات متعددة
حجم البيانات	يمكن أن يعمل بكفاءة على مجموعات بيانات صغيرة إلى متوسطة الحجم	يتطلب كميات كبيرة من البيانات لتحقيق أداء جيد
متطلبات الأجهزة	يمكن تشغيله على أجهزة حوسبة تقليدية	يتطلب أجهزة حوسبة قوية مثل وحدات معالجة الرسومات (GPU)
الأداء	جيد في مهام التحليل البسيطة والمتوسطة التعقيد	متفوق في معالجة المهام المعقدة مثل التعرف على الصور ومعالجة اللغات الطبيعية
السرعة	أسرع في التدريب على مجموعات البيانات الصغيرة	أبطأ في التدريب بسبب تعقيد الشبكات العصبية وعدد الطبقات

الفرق	التعلم الآلي	التعلم العميق
التفسير	أكثر قابلية للتفسير والتحليل	أقل قابلية للتفسير بسبب التعقيد الكبير في الطبقات العميقة
النطاق	يستخدم في تطبيقات متعددة مثل التصنيف والانحدار والتجميع	يستخدم بشكل رئيسي في تطبيقات معقدة مثل التعرف على الكلام والصور
الأمثلة	الانحدار اللوجستي، أشجار القرار، الغابات العشوائية	الشبكات العصبية العميقة، الشبكات العصبية التلافيفية، شبكات الذاكرة الطويلة القصيرة

3 أعمال ذات صلة :

بعد التطرق الى طرق التعلم الآلي والتعلم العميق, نحتاج أن نقوم بدراسة شاملة لما هو موجود من أبحاث تركتز على هذه الطرق من أجل الكشف عن مرض القلب. بحيث انقسمت الدراسات بين التي تستخدم التعلم الآلي في التعرف على المرض وأخرى تستخدم التعلم العميق.

3.1 التعرف على مرض القلب عن طريق التعلم الآلي:

في السنوات الأخيرة، شهد مجال التعلم الآلي تقدمًا هائلًا في تطبيقاته على الرعاية الصحية، وخاصة في التعرف على أمراض القلب. تعتمد هذه التقنيات على استخدام الخوارزميات والنماذج الرياضية لتحليل البيانات الكبيرة والمتنوعة المستخرجة من المرضى، مثل البيانات السريرية، نتائج الاختبارات المخبرية، وصور الأشعة. تتيح تقنيات التعلم الآلي إمكانية الكشف المبكر عن مرض القلب من خلال التعرف على الأنماط والعلاقات الخفية التي قد لا تكون واضحة للأطباء. من بين الأساليب الشائعة المستخدمة في هذا المجال نذكر الانحدار اللوجستي (Logistic Regression)، غابات القرار (Random Forest)، آلات المتجهات الداعمة (Sup-port Vector Machines - SVM)، بايز البسيط (Naive Bayes)، والتعزيز التدريجي (XGBoost). يُعدّ دمج هذه الأساليب المتقدمة مع البيانات الضخمة خطوة هامة نحو تحسين التشخيص، وتحديد العوامل الخطرة، وتقديم خطط علاجية مخصصة. توضح الدراسات السابقة في مجال التعلم الآلي للتعرف على أمراض القلب فعالية هذه التقنيات ودورها الكبير في تحسين آداء هذه التطبيقات .

فيما يلي جدول يلخص بعض الدراسات السابقة في استخدام التعلم الآلي للتعرف على أمراض القلب مرفوقاً برابط لكل دراسة:

جدول 2.2. بعض الدراسات السابقة في استخدام التعلم الآلي للتعرف على أمراض القلب

اسم الدراسة	الخوارزمية	الدقة	رابط الدراسة	قاعدة البيانات
Enhancing Disease Heart Prediction Accuracy through Machine Learning Techniques and Optimization	Voting Soft Ensemble	% 95	رابط الدراسة	Cleveland IEEE and Dataport disease heart datasets
Cardiovascular risk disease prediction using automated machine A learning: prospective of study UK 423,604 Biobank participants	AutoPrognosis	% 95	رابط الدراسة	data Using 423,604 on participants CVD without in baseline at Biobank UK
Early of prediction disease heart data with using analysis supervised with learning stochastic gradient boosting	HDP-DTRF	% 86	رابط الدراسة	The Cleveland Disease Heart dataset
Disease Heart Prediction Using Machine Learning Techniques	hybrid random a with forest model. linear	% 92	رابط الدراسة	. dataset UCI
Disease Heart Prediction Machine with Learning Approaches	Logistic regression (LR)	% 84	رابط الدراسة	Cleveland the Disease Heart database UCI at Repository

اسم الدراسة	الخوارزمية	الدقة	رابط الدراسة	قاعدة البيانات
Heart to Heart Application: Predicting Presence of Heart Diseases Using Machine Learning Techniques	SVM K-NN	% 87	رابط الدراسة	dataset UCI .
HEART FAILURE PREDICTION USING MACHINE LEARNING TECHNIQUES	SVM	% 85.2	رابط الدراسة	dataset UCI .

3.2 التعرف على مرض القلب عن طريق التعلم العميق:

يُعتبر التعلم العميق واحدًا من أهم الفروع في الذكاء الاصطناعي، وقد أحدث ثورة في كيفية التعامل مع البيانات المعقدة والمتنوعة، خاصة في مجال الطب. تعتمد تقنيات التعلم العميق على الشبكات العصبية العميقة التي تمتاز بقدرتها على معالجة كميات هائلة من البيانات واكتشاف الأنماط الدقيقة التي قد تكون غير مرئية للبشر. في مجال تشخيص أمراض القلب، يتم استخدام هذه التقنيات لتحليل البيانات الطبية مثل الصور الإشعاعية، تخطيط كهربية القلب، والبيانات السريرية لتحسين دقة التشخيص والتنبؤ بحالة المرضى. من بين الخوارزميات الشائعة في هذا المجال نذكر الشبكات العصبية التلافيفية (CNN)، الشبكات العصبية المتكررة (RNN)، الشبكات العصبية طويلة المدى وقصيرة المدى (LSTM)، الشبكات العصبية العميقة (DNN)، و SqueezeNet، والتي تساهم في تقديم حلول مبتكرة وفعالة في التشخيص الطبي. توضح الدراسات الحديثة كيف يمكن للتعلم العميق أن يعزز من فعالية تشخيص أمراض القلب ويدعم الأطباء في اتخاذ قرارات علاجية أكثر دقة.

فيما يلي جدول يلخص بعض الدراسات السابقة في استخدام التعلم العميق للتعرف على أمراض القلب مرفوقاً برابط لكل دراسة:

جدول 3.2. بعض الدراسات السابقة في استخدام التعلم العميق للتعرف على أمراض القلب

اسم الدراسة	الخوارزمية	الدقة	رابط الدراسة	قاعدة البيانات
Efficient Disease Heart Prediction Using Deep Hybrid Learning	DPA RNN+LSTM	% 99.2	رابط الدراسة	diagnosis HD and UCI failure heart clinical dataset.

اسم الدراسة	الخوارزمية	الدقة	رابط الدراسة	قاعدة البيانات
disease Heart prediction using based IoT framework improved and learning deep approach: Medical application	MDLSTM	% 98.01	رابط الدراسة	Hungarian disease heart dataset
Deep Novel Learning Architecture Heart for Disease Prediction using Convolutional Neural Network	CNN	% 87.5	رابط الدراسة	Cleveland . dataset
Feature Selection Dwarf and Mongoose Optimization Deep Enabled for Learning Disease Heart Detection	SqueezeNet	% 87.2	رابط الدراسة	disease heart dataset (Cleveland) Z- the ,and Sani Alizadeh . dataset
of Diagnosis cardiovascular using disease learning deep technique	CNN BiLSTM	% 94.5	رابط الدراسة	Disease Heart . dataset uci
Ensemble Learning on Based Deep Hybrid Learning Model Heart for Early Disease Prediction	CNN-LSTM and CNN-GRU	% 97.17	رابط الدراسة	Cleveland dataset

اسم الدراسة	الخوارزمية	الدقة	رابط الدراسة	قاعدة البيانات
deep Hybrid learning using model recurrent neural network gated and recurrent for unit disease heart prediction	RNN+GRU	% 97.68	رابط الدراسة	Cleveland dataset .

الخاتمة :

في هذا الفصل، استعرضنا كل ما يخص التعلم الآلي والتعلم العميق، حيث قمنا بتعريف كل منهما وشرح طرق عمل الخوارزميات الخاضعة وغير الخاضعة للإشراف، بالإضافة إلى الشبكات العصبية الاصطناعية. قمنا بمقارنة الخوارزميات بناءً على عدة عوامل مختلفة ، واستعرضنا الاعمال السابقة للتعلم الآلي والتعلم العميق.

في الفصل القادم، سنقوم بتصميم وتنفيذ نظام قادر على تشخيص مرض القلب وذلك بإستخدام وتدريب نموذج في التعلم العميق(DNN) ، سنتناول بالتفصيل خطوات تحميل البيانات ومعالجتها، وتدريب النموذج وتقييم أدائه، بالإضافة إلى عرض الأدوات التي استخدمناها في التطوير .

باب 3. الفصل الثالث

مقدمة :

قدمنا خلال الفصلين السابقين كل المفاهيم المتعلقة بالكشف عن مرض القلب، حيث بدأنا بِشرح المفاهيم الأساسية المتعلقة بمرض القلب، من تعريف وأنواع وأعرض ثم تطرقنا الى ذكر طرق تشخيص المرض وذلك إما بالطرق التقليدية أوعن طريق تقنيات الذكاء الاصطناعي. بعد ذلك تطرقنا بِشرح مفصل عن أساسيات التعلم الآلي والتعلم العميق، موضحين أهم الفروقات بينهما، بالإضافة إلى ذلك ذكرنا أهم الطرق الموجودة لكل منهما. وفي الأخير قدمنا أهم ما توصلت إليه الدراسات الحديثة للكشف عن مرض القلب.

في هذا الفصل،سنعرض النظام الخاص بنا للكشف عن مرض القلب بإستخدام طرق التعلم العميق، بحيث سنشرح بالتفصيل كيف قمنا ببناء النموذج المقترح وكيف قمنا بتدريبه وتقييمه. بالإضافة إلى ذلك سنعرض النتائج المتحصّل عليها ثم سنقوم بمُناقشتها.

1 التصميم :

تهدف دراستنا للكشف عن أمراض القلب وذلك بالإعتماد على طرق التعلم العميق، بحيث سنقوم بتدريب نموذج من نماذج الشبكات العصبية العميقة (DNN) . في مايلي سنشرح بالتفصيل جميع مراحل تطوير النموذج الخاص بنا، بداية من بيانات التدريب وصولا إلى تدريب النموذج، ثم التنبؤ بعناصر جديدة.

1.1 وصف البيانات :

ما يميز مجموعة بيانات (المؤشرات الشخصية لأمراض القلب) هو حداثتها حيث أنها أصدرت سنة 2020، مما يجعلها تعكس الواقع الحالي لصحة الأفراد وعاداتهم. تم تجميع مجموعة البيانات المؤشرات الشخصية لأمراض القلب لمساعدة الأبحاث القائمة على التعلم العميق للوصول إلى نتائج ذات دقة عالية.

تم استخدام هذه البيانات في العديد من الدراسات لتحليل العوامل المؤثرة في أمراض القلب وتطوير نماذج تنبؤية متقدمة.
حيث تم جمع هذه البيانات من قبل مركز السيطرة على الأمراض والوقاية منها (CDC) من خلال نظام مراقبة عوامل الخطر السلوكي (BRFSS)، والذي يجمع البيانات الصحية من خلال استطلاعات سنوية يتم إجراؤها على مستوى السكان في الولايات المتحدة.
تحتوي مجموعة البيانات [9] على معلومات تفصيلية حول الحالة الصحية، السلوكيات، والعوامل البيولوجية للأفراد، مما يوفر أساسًا قويًا لتحليل وتقدير مخاطر الإصابة بأمراض القلب.

تشمل المجموعة على 319,795 مشاركًا، و على 18 خاصية، حيث يتم توفير بيانات حول عوامل متعددة مثل العمر، الجنس، العرق، وزن الجسم، مستوى النشاط البدني، التدخين، استهلاك الكحول، إلى جانب معلومات أخرى قد تساهم في تطوير الأمراض القلبية الوعائية.

1.2 معالجة البيانات :

تُعتبر مرحلة معالجة البيانات مرحلة أولية وضرورية جدا من أجل تحديد النموذج وتدريبه. تمُرُ هذه المرحلة بمجموعة من الخطوات :

- **أولا : معالجة البيانات المفقودة :**

مشكلة البيانات المفقودة هي أول المشكلات التي واجهناها خلال دراستنا، فهي مشكلة شائعة بإمكانها أن تُعيق عملية التحليل إذا لم يتم علاجها بشكل مناسب. بالتالي لحل هذه المشكلة، إستخدمنا تقنيات متقدمة لتقدير القيم المفقودة وضمان الإستفادة الكاملة من المعطيات المتاحة، حيث قُمنا بإستدعاء مكتبة (Scikit-learn)، التي تُوفر طُرقًا مُتعددّة من أجل تعبئة القيم الغائبة. حيث أننا قُمنا بإستخدام أكثر الطرق شيوعًا، (mean)، والتي تقوم على إختيار متوسط القيم الموجودة وتُملأ به الفراغات الموجودة.

- **ثانيا : معالجة البيانات الفئوية :**

المشكلة التالية تكمن في البيانات الفئوية التي لا يمكن للنموذج التعامل معها بشكل مباشر. لمعالجة هذه المشكلة، تم ترميز السمات الفئوية باستخدام تقنية (One-Hot Encoding) لتحويل البيانات الفئوية (البيانات التي على شكل لفئات) إلى شكل يمكن للنموذج التعامل معه بشكل أفضل أي إلى بيانات رقمية. بالإضافة إلى ذلك، تم تطبيق ترميز السمات (Label Encoding) لتحويل القيم النصية إلى أرقام، مما يسهل على النموذج فهم ومعالجة هذه القيم.

- **ثالثا: توحيد مجال المُتغيرات :**

في هذه الخطوة يتم توحيد مجال قيم المتغيرات، لأن البيانات لدينا تتميز بإختلاف مجال القيم بين المتغيرات وهذا قد يؤدي إلى سوء أداء النموذج. لتوحيد البيانات (Scaling Feature) وضمان أن جميع المتغيرات لديها نفس المجال، قُمنا بإستخدام تقنيات مثل (Normalization) و (Standardization) ووقع اختيارنا على StandardScaler لأنه يناسب طبيعة البيانات التي تتضمن مؤشرات متنوعة، مما يساعد في تحسين أداء النموذج ودقته.

- **رابعا: تحديد عدد الخصائص الفعّالة :**

وأخيرا، نواجه مشكلة في تحديد عدد الخصائص في مجموعة البيانات، حيث أنه عند إستخدام الخصائص الفعّالة فقط، ينتج عنه تحسين أداء النموذج. وبالتالي، قُمنا بتجرُبة تقنيات تخفيض الأبعاد (Dimensionality Reduction). حيث توجد عدة تقنيات لتخفيض الأبعاد مثل تحليل

المكونات الأساسية (PCA)، التحليل التمييزي الخطي (LDA)، وتحليل التوزيع التلقائي (Autoencoder) حيث أن هذه التقنيات تساعد بشكل كبير على التقليل من عدد الخصائص المستخدمة مع الحفاظ على المعلومات الهامة والفعّالة. في هذه الدراسة، إستخدمنا طريقة تحليل المكونات الأساسية (PCA)، لأنها تُساعد على الحفاظ على أكبر قدر ممكن من المعلومات الفعّالة، وهذا ساعدنا في تحديد عدد الخصائص الفعّالة من جهة وتحسين دقة النموذج من جهة أُخرى.

1.3 التعلم و التدريب :

شبكتنا العصبية العميقة (DNN) تستخدم بنية تسلسلية تتكون من طبقات متعددة. يتم تدريب النموذج باستخدام البيانات التي تم تحضيرها، حيث يتعلم النموذج من خلال الأمثلة المقدمة له. نستخدم الخسارة الثنائية للتقاطع (Binary Cross-Entropy) كدالة خسارة لتقييم أداء النموذج في مهمة التصنيف الثنائي (مريض ، غير مريض). كما نستعمل مُحسّن آدم (Adam Optimizer) لتحديث الأوزان في الشبكة بهدف تقليل الخسارة.

2 الشبكة العصبية العميقة المقترحة :

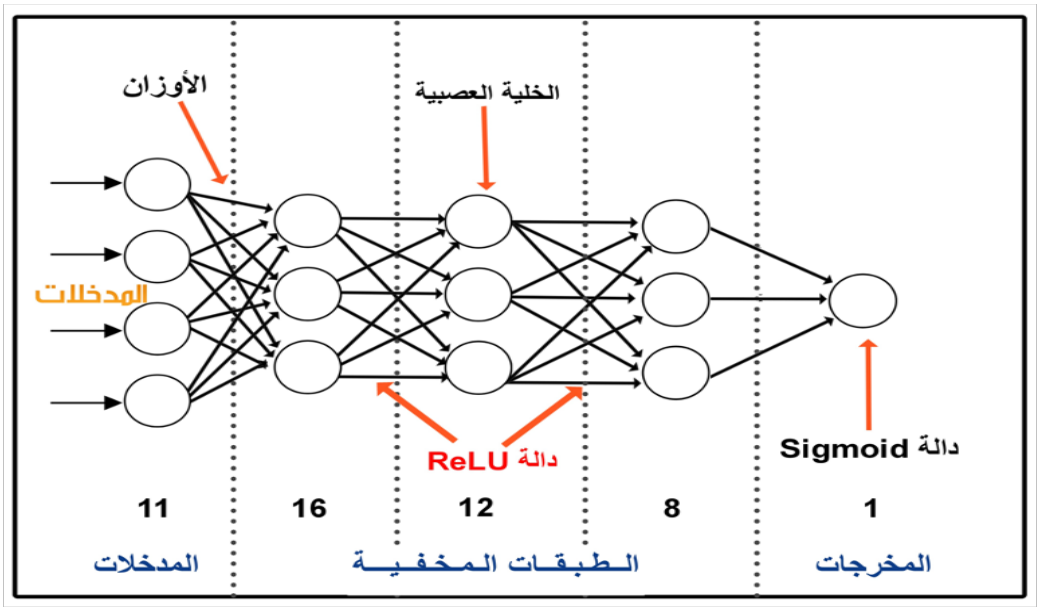
قُمنا بإنشاء نموذج الشبكة العصبية العميقة (DNN) لحل إشكالية التنبؤ بمرض القلب وذلك باستخدام تقنيات الذكاء الاصطناعي.

مرحلة بناء الشبكة تُعد مرحلة حاسمة ومهمة في دراستنا، بحيث تهدف هذه المرحلة إلى تحديد الهندسة العامة للنموذج. الشكل 1 يوضح مخطط النموذج المُقترح حيث يتكون من طبقة إدخال، تليها 3 طبقات مخفيه (16 عقدة،12 عقدة، و8 عُقد)، متبوعة بطبقة الخروج.

2.1 طبقة الإدخال :

تتكون طبقة الإدخال من 11 عقدة، كل عقدة تمثل خاصية صحية للمرضى مثل العمر، الجنس، والوضع الصحي العام، وغيرها. هذه الخصائص تم تحويلها إلى صيغة رقمية مناسبة بعد مرورها عبر مراحل المعالجة الأولية للبيانات. حيث تصبح البيانات بعد هذه المرحلة جاهزة لتدريب وتقييم النموذج المُقترح.

في طبقة الإدخال، يتم إرسال المعلومات المدخلة مباشرة إلى الطبقات المخفية دون أي تعديل بحيث كل عقدة في طبقة الإدخال ترتبط بجميع العقد في الطبقة المخفية الأولى، مما يُمكن الشبكة من تحليل العلاقات بين المتغيرات الداخلة بشكل أعمق ، كما يرتبط كل متغير في البيانات بعقدة مستقلة في طبقة الادخال بحيث تعكس قيمة المتغير للمدخلة.



شكل 1.3. **مخطط هندسة الشبكة العصبية العميقة المُقترح لتشخيص أمراض القلب**

2.2 الطبقات المخفية :

تحتوي شبكتنا على 3 طبقات مخفية، الأولى تحتوي على 16 عقدة، والثانية على 12 عقدة، بينما الثالثة تحتوي على 8 عقد. كل عقدة في الطبقات المخفية تقوم بحساب مجموع الإدخالات الموزونة من الطبقة السابقة، وبعدها تُطبق دالة الانتقال (ReLU) على هذا المجموع، مما يعدل القيم بطريقة غير خطية، ثم يتم تمرير الناتج إلى الطبقة التالية.

الطبقات المخفية مهمة لاكتشاف العلاقات المعقدة في البيانات باستخدام دوال التنشيط. و استخدامنا لدالة التنشيط ReLU لهذه الطبقات كان لأنها تساعد في الحفاظ على التدرج الصحيح خلال عملية التدريب ولتقليل مشكلة العقد الميتة (dead neurons).

2.3 طبقة الإخراج :

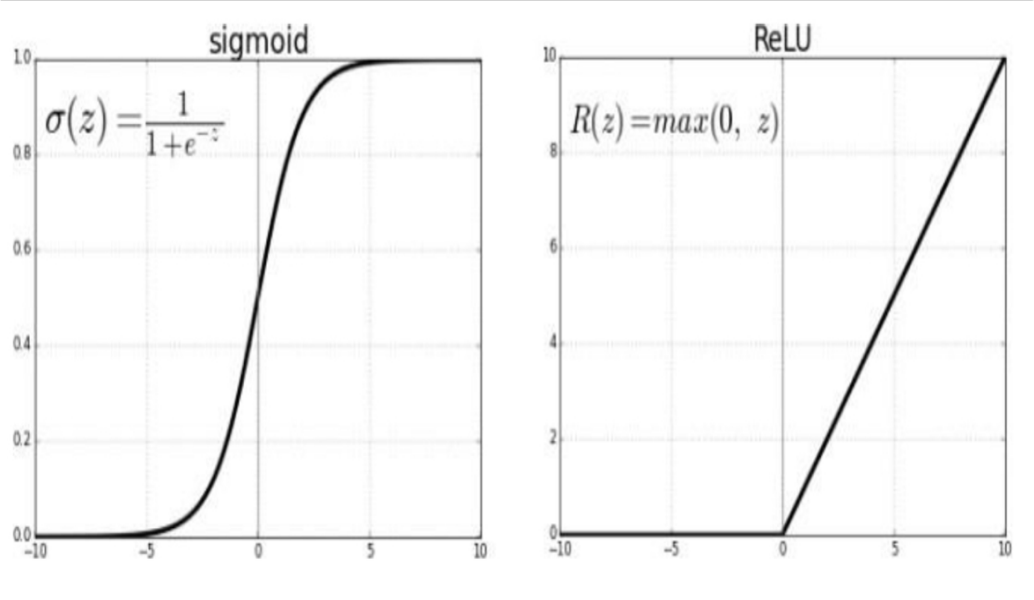
تحتوي طبقة الاخراج على عقدة واحدة. هذه العقدة تقوم بحساب مجموع مرجح للإدخالات القادمة من الطبقة السابقة، حيث يتم ضرب كل إدخال في وزن محدد.

بعد ذلك، يتم تطبيق دالة التنشيط (sigmoid) على هذا المجموع، مما يحول القيمة إلى نطاق بين 0 و 1. هذه القيمة تمثل احتمالية وجود المرض أو عدمه.

تتم هذه العمليات للعقدة الواحدة في الطبقة النهائية حيث يتم حساب مجموع مرجح للإدخالات ثم تطبيق دالة التنشيط للحصول على النتيجة النهائية .

2.4 دوال التنشيط :

بالمقارنة بين طريقة عمل دوال التنشيط إستعملنا دالة ReLU للطبقات المخفية حيث تعمل على تحسين عملية التدريب بحيث تمرر فقط القيم الإيجابية،بينما إستعملنا دالةSigmoid لإحتساب احتمالية وجود المرض حيث هذه الأخيرة تقوم بتحويل القيم في مجال ما بين 0 و 1. الشكل 2 يوضح طريقة عمل كل دالة من هذه الدوال.



شكل 2.3. **مقارنة بين طريقة عمل دوال التنشيط (Sigmoid/ReLU)**

2.5 الإختبار و التحقق :

بعد تدريب الشبكة العصبية، من الضروري اختبارها باستخدام بيانات لم تُستخدم أثناء عملية التدريب لتقييم دقة وكفاءة النموذج. يشمل ذلك تقسيم البيانات إلى مجموعتين: مجموعة تدريب ومجموعة اختبار. اخترنا تقسيم البيانات بحيث تكون 50% منها للتدريب و50% للاختبار. هذا التقسيم يسمح بتقييم موضوعي لأداء النموذج عند التعامل مع بيانات جديدة، مما يعطينا مؤشراً على مدى فعاليته في التطبيقات الحقيقية، أي قدرته على تقديم نتائج دقيقة عند استخدامه في مواقف وأوضاع واقعية خارج بيئة الاختبار أو التدريب.

3 التنفيذ :

بعد شرح كل مرحلة من مراحل النظام الخاص بنا، في هذه المرحلة سنقوم بتنفيذ جميع الخطوات من تحميل البيانات إلى غاية التنبؤ بعنصر جديد مريض أو غير مريض.وبالتالي سنقوم أولا بشرح كل من بيئة و خطوات التنفيذ و التجارب و النتائج

3.1 بيئة التنفيذ :

1.3.1. أدوات التطوير :

- **بايثون :**

تم تطوير لغة البرمجة Python بواسطة Guido van Rossum وأُطلقت لأول مرة في عام 1991. تتميز Python بكونها لغة سهلة القراءة والكتابة، مما يجعلها مثالية لكل من المبتدئين والمحترفين على حد سواء. بفضل مكتباتها المتقدمة مثل TensorFlow و Keras، استخدمنا Python لبناء وتدريب نموذج التعلم العميق الخاص بنا، لأنها تسهل إدارة النماذج العصبية المعقدة. Python تتضمن أيضًا مكتبات قوية أخرى مثل NumPy و Pandas للتعامل مع البيانات، و Matplotlib [10].

- **مكتبة (Pandas) :**

Pandas هي مكتبة في Python تم تطويرها بواسطة McKinney Wes في عام 2008. توفر Pandas هياكل بيانات قوية ووظائف تحليل مثل DataFrame، وهي مثالية للتعامل مع البيانات المنظمة والجداول الزمنية. تُستخدم Pandas على نطاق واسع في المهام التي تتطلب معالجة البيانات والتحليل الاستكشافي، وتشمل ميزات مثل الانضمام، الفرز، والتجميع السريع للبيانات. في مشروعنا، استخدمنا Pandas لأنها تمكننا من التعامل بكفاءة مع البيانات الكبيرة، مما يسهل إعدادها للتحليل والنمذجة، خصوصًا في مراحل ما قبل معالجة البيانات وتجهيزها لتدريب النموذج[10].

- **مكتبة (Scikit-learn) :**

Scikit-learn هي مكتبة في Python مخصصة للتعلم الآلي، تم تطويرها بمساهمة مجموعة من المطورين كجزء من مشروع SciPy. أُطلقت في عام 2007 حيث توفر واجهة برمجة تطبيقات سهلة الاستخدام لمختلف خوارزميات التعلم الآلي مثل الآلات الداعمة، الغابات العشوائية، و k-means. Scikit-learn, تدعم أيضًا أدوات للتقييم المتقاطع وتقنيات تقليل الأبعاد، مما يجعلها أداة مفيدة جدًا في البحث العلمي. في مشروعنا، استخدمنا Scikit-learn لأنها تُمكننا من التعامل بسهولة مع البيانات الكبيرة،وهذا ً يُسهل علينا إعداد البيانات للتحليل والنمذجة، خصوصا في مراحل ما قبل معالجة البيانات وتجهيزها لتدريب النموذج[10].

- **مكتبة (Keras & TensorFlow) :**

TensorFlow، تم تطويره بواسطة Google Brain Team، هو إطار عمل قوي تم إطلاقه في عام 2015 . يسمح بتنفيذ الحسابات عبر منصات متعددة (مثل الحواسيب المكتبية والخوادم) ويدعم بشكل خاص التعلم الآلي على نطاق واسع. Keras ، التي تم تطويرها بواسطة François Chollet و هي واجهة برمجية عالية المستوى لبناء وتدريب نماذج الشبكات العصبية ، تم تصميمها لتسهيل التجارب السريعة ويمكنها تشغيل TensorFlow كواجهة خلفية [10].

2.3.1. خصائص الجهاز :

تم تدريب وتقييم النموذج باستخدام جهاز MacBook Pro 2019 الذي يتميز بخصائص تقنية عالية تمكنه من التعامل مع العمليات الحسابية المعقدة بسرعة وكفاءة.

يحتوي الجهاز على معالج Intel i9 بثمانية أنوية (8 cores)، مما يوفر قدرة معالجة عالية لتسريع عملية التدريب.

بالإضافة إلى ذلك، يحتوي الجهاز على ذاكرة وصول عشوائي (RAM) بحجم 16 جيجا بايت، مما يتيح تخزين البيانات المؤقتة بكفاءة ودعم عمليات المعالجة الكبيرة.

للتخزين، يحتوي الجهاز على قرص SSD بسعة 1 تيرابايت، مما يوفر سرعة عالية في قراءة وكتابة البيانات، ويعزز أداء النظام بشكل عام.

يعمل الجهاز بنظام تشغيل macOS Sonoma، الذي يوفر بيئة مستقرة ومتقدمة لتطوير وتنفيذ نماذج التعلم العميق.

هذه الخصائص مجتمعة تساهم في توفير بيئة مناسبة لبناء النموذج وتدريبه بكفاءة عالية، مما يضمن أداءً ممتازًا وسرعة في إنجاز العمليات.

كما تم الاستعانة بمنصة Kaggle لتشغيل الكود وتحليل النتائج بكفاءة، مما ساعدنا في الوصول إلى أفضل نموذج.

3.3.1. كيفية حساب الدقة :

تم حساب الدقة من خلال تقييم أداء النموذج على مجموعة بيانات الاختبار، حيث يتم قياس نسبة التوقعات الصحيحة التي يقوم بها النموذج مقارنة بالتوقعات الفعلية. العلاقة الرياضية لحساب الدقة تُعبر عنها كالآتي:

$$\text{Accuracy} = \left(\frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \right) \times 100\%$$

هذه الصيغة توضح كيفية تحديد مدى فعالية النموذج في إصدار توقعات صحيحة بناءً على البيانات المُختبرة.

3.2 خطوات التنفيذ :

1.3.2. تحميل وتحضير البيانات :

تبدأ عملية بناء نموذج التعلم العميق بتحميل البيانات من الملف. في هذه الحالة، نستخدم مكتبة pandas لتحميل البيانات لأنها توفر واجهة سهلة وفعالة لمعالجة البيانات الجدولية. هنا نستعمل دالة read_csv لقراءة البيانات من الملف كما في الشكل 3.

```
data =
pd.read_csv('heart_2020_cleaned.csv')
```

شكل 3.3. **تحميل البيانات من ملف csv**

2.3.2. تحويل البيانات النصية لتمثيل رقمي :

الخطوة التالية هي تحويل البيانات الفئوية إلى صيغة رقمية باستخدام تقنية (one hot encoding)، والتي تساعد في تحويل البيانات النصية إلى صيغة رقمية يمكن للنموذج العمل معها بكفاءة أكبر. يتم ذلك باستخدام دالة get_dummies من مكتبة Pandas كما في الشكل 4 .

```
data = pd.get_dummies(data, columns=
['AgeCategory', 'Race',
'GenHealth'], drop_first=True)
label_encoder = LabelEncoder()
for column in data.columns:
    data[column] =
label_encoder.fit_transform(data[col
umn])
```

شكل 4.3. **تحويل البيانات الفئوية لتمثيل رقمي**

3.3.2. تقسيم البيانات إلى مجموعات التدريب والاختبار :

قمنا بتجريب عدة تقسيمات مختلفة للبيانات، مثل (30% - 70%) و (20% - 80%) وذلك بهدف معرفة التقسيم الأمثل لهذه البيانات الضخمة. بعد عدة تجارب، وجدنا أن تقسيم البيانات بنسبة 50% للتدريب و50% للاختبار هو الأنسب لتحقيق توازن جيد بين حجم بيانات التدريب و تقييم النموذج.

في هذا التقسيم، استخدمنا دالة train_test_split من مكتبة sklearn.model_selection لتقسيم البيانات بشكل عشوائي لمحاكاة الظروف الحقيقية، كما استخدمنا StandardScaler لتوحيد نطاق القيم.

حيث أن مجموعة 50% الأولى استُعملت لتدريب النموذج، لكي يتعلم من خلالها الأنماط و العلاقات المخفية بين البيانات، مما يمكنه من تعميم وتطبيق ما تعلمه على الـ 50% من بيانات الاختبار الجديدة التي لم يسبق له مشاهدتها من قبل شكل 5.

```
X_train_val, X_test, y_train_val,
y_test = train_test_split(X_pca, y,
test_size=0.5, random_state=42)
```

شكل 5.3. **تقسيم البيانات للتدريب و الإختبار**

4.3.2. إستخدام دوال التنشيط وتحديد الخسارة :

في النموذج، نستخدم عدة دوال تنشيط ودالة خسارة لتحقيق أهداف محددة:

- **دالة التنشيط ReLU**:

تُستخدم في الطبقات الداخلية لزيادة القدرة على التعامل مع العلاقات غير الخطية بين الخصائص. هذه الدالة تساعد في تعزيز التعقيد الحسابي للنموذج دون التأثير على سرعة التدريب.

- **دالة التنشيط sigmoid**:

تُستخدم في طبقة الخروج لتصنيف النتائج إلى فئتين. هذه الدالة مثالية للمشاكل ثنائية التصنيف حيث تحدد النتيجة إما صفر أو واحد، مما يُمكن النموذج من تحديد أحد الخيارين بدقة.

- **دالة الخسارة binary crossentropy**:

تُستخدم لحساب الفرق بين التوقعات والقيم الفعلية في مشاكل التصنيف الثنائي. هذه الدالة تعزز دقة النموذج في التنبؤ بأحد خيارين ممكنين وتساعد في تحسين الأداء عن طريق تقليل الخسائر خلال عملية التدريب.

هذه الدوال معاً تضمن تحقيق النموذج لأعلى مستويات الأداء والدقة في تصنيف البيانات الشكل6.

```
model = Sequential()
model.add(Input(shape(X_train_val.sha
pe[1],)))
model.add(Dense(12,
activation='relu'))
model.add(Dense(8,
activation='relu'))
model.add(Dense(1,
activation='sigmoid'))
```

شكل 6.3. **نموذج كيراس للتصنيف الثنائي**

5.3.2. التحسين و إختيار الخصائص :

قمنا بتحسين النموذج باستخدام دالة الخسارة binary crossentropy والمُحسِّن adam.

تُعد هذه الاختيارات شائعة لمشكلات التصنيف الثنائي نظراً لفعاليتها في تحسين دقة النموذج وتسريع عملية التدريب. بالإضافة إلى ذلك، قمنا بتطبيق تحليل المكونات الرئيسية (PCA) على البيانات المعيارية باستخدام دالة fit transform لتحويل البيانات وتقليل أبعادها. بفضل استخدام PCA، استطعنا تقليل عدد الخصائص إلى 11 خاصية، مما ساعدنا في تسهيل عملية التدريب وتحسين أداء النموذج.

قمنا بتجربة عدة قيم مختلفة لـ PCA، ولكن لم تتحسن الدقة بشكل كبير مع القيم الأخرى. واعتمدنا 11 مكوناً رئيسياً لأنها أعطت أفضل النتائج من حيث الأداء والدقة.

```
model.compile(loss='binary_crossentro
py', optimizer='adam', metrics=
['accuracy'])
```

شكل 7.3. **تحسين النموذج**

```
from sklearn.decomposition import PCA
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
pca = PCA(n_components=11)
X_pca = pca.fit_transform(X_scaled)
```

شكل 8.3. **تطبيق PCA لتقليل الأبعاد**

6.3.2. تحليل الأداء :

بعد تدريب النموذج على مجموعة البيانات، قمنا باستخدام مجموعة الاختبار لتقييم أداء النموذج. النموذج الذي تم بناؤه باستخدام TensorFlow و Keras (شكل 10)، تمكن من تحقيق نتائج جيدة جدا حيث بلغت نسبة دقة النموذج 91.58 % كما في شكل 10 .

```
loss, accuracy =
model.evaluate(X_test, y_test)
print(f'Accuracy on the original test
set: {accuracy*100:.2f}%')
```

شكل 9.3. **أداء النموذج**

```
Epoch 1/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 38s 2ms/step - accuracy: 0.9075 - loss: 0.2529
Epoch 2/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 35s 2ms/step - accuracy: 0.9150 - loss: 0.2344
Epoch 3/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 43s 2ms/step - accuracy: 0.9141 - loss: 0.2350
Epoch 4/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 42s 2ms/step - accuracy: 0.9150 - loss: 0.2336
Epoch 5/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 32s 2ms/step - accuracy: 0.9162 - loss: 0.2295
Epoch 6/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 39s 2ms/step - accuracy: 0.9157 - loss: 0.2292
Epoch 7/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 30s 2ms/step - accuracy: 0.9145 - loss: 0.2320
Epoch 8/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 30s 2ms/step - accuracy: 0.9164 - loss: 0.2286
Epoch 9/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 41s 2ms/step - accuracy: 0.9169 - loss: 0.2278
Epoch 10/10
15990/15990 ━━━━━━━━━━━━━━━━━━━━ 29s 2ms/step - accuracy: 0.9149 - loss: 0.2312
4997/4997 ━━━━━━━━━━━━━━━━━━━━ 7s 1ms/step - accuracy: 0.9137 - loss: 0.2353
Accuracy on the original test set: 91.58%
```

شكل 10.3. **نتيجة دقة النموذج**

3.3 التجارب :

في هذا القسم، سنتناول تقييم أداء عدة نماذج قمنا باستخدامها لتحديد مخاطر أمراض القلب، وهي خوارزمية الجار الأقرب (K-Nearest Neighbors - KNN)، الانحدار اللوجستي (Logistic Regression)، والشبكة العصبية العميقة (Deep Neural Network - DNN).

حيث قمنا بتطبيق كل من هذه الخوارزميات على نفس مجموعة البيانات، مما مكننا من إجراء مقارنة موضوعية بينها لفهم أي منها يوفر أفضل أداء في تشخيص وتقدير مخاطر هذه الأمراض.

1.3.3. تجارب التعلم الالي :

● الانحدار اللوجستي (Logistic Regression) :

الانحدار اللوجستي هو نموذج إحصائي يستخدم لتقدير احتمالات نتائج ثنائية (مثل الإصابة بمرض القلب أو عدم الإصابة به). يعتمد النموذج على دالة لوجستية لتحويل المخرجات الخطية إلى احتمالات تتراوح بين 0 و 1.

● خوارزمية الجار الأقرب (K-Nearest Neighbors):

خوارزمية الجار الأقرب هي خوارزمية تصنيف تعتمد على المسافة بين النقاط. يتم تصنيف النقاط الجديدة بناءً على الفئة الأكثر تكرارًا بين أقرب جيران لها. في مشروعنا، قمنا بتجربة عدة قيم لـ k[1,2,3,4] لتحديد عدد الجيران، ووجدنا أن القيمة $K = 2$ كانت الأفضل من حيث الدقة. هذه الخوارزمية بسيطة لكنها فعالة في كثير من الحالات، خصوصًا عند توفر بيانات ذات توزيع جيد.

2.3.3. تجارب التعلم العميق :

لقد قمنا بالعديد من التجارب على بياناتنا لتحديد النموذج الأمثل لشبكة DNN الخاصة بنا. الهدف من هذه التجارب هو تحسين دقة النموذج وتقليل نسبة الخطأ. قمنا بتجربة تكوينات متعددة على البيانات باستخدام تقنيات مختلفة وتوثيق النتائج لكل منها كما يلي:

● الأولى: استخدام 11 خاصية رئيسي :

في هذه التجربة، قمنا باستخدام 11 خاصية رئيسي بعد تطبيق PCA على البيانات. قمنا بتدريب النموذج باستخدام هذه البيانات وكانت النتائج كما يلي:

- دقة النموذج: 85%

- نسبة الخطأ: 15%

● التجربة الثانية: تعديل عدد الطبقات المخفية

في هذه الحالة، قُمنا بتجربة العديد من التشكيلات، مثل (32،256،256) و (10،100،205) و كانت الدقة على التوالي كما يلي:

- دقة التشكيلة (32،256،256): 91.57 %

- دقة التشكيلة (10،100،205): 91.53 %

لكننا في الاخير قمنا بتغيير عدد العقد في كل طبقة إلى 16 و12 و8 على التوالي مما أدى لتحسن دقة النموذج و النتائج كانت كالتالي:

- دقة النموذج: 88%

- نسبة الخطأ : 12%

- **التجربة الثالثة: استخدام تقنية RandomUnderSampler :**

جربنا استخدام تقنية RandomUnderSampler لتحقيق توازن أفضل في قيم البيانات قبل التدريب. النتائج كانت كالتالي:

- دقة النموذج: 90%

- نسبة الخطأ: 10%

- **التجربة النهائية: النموذج المقترح :**

بعد العديد من التجارب، وصلنا إلى النموذج النهائي لشبكة DNN الخاصة بنا والذي حقق أعلى دقة بنسبة 91.58 %. التكوين النهائي للنموذج هو كما يلي: - استخدام 11 خاصية رئيسي بعد تطبيق PCA - ثلاث طبقات مخفية بعدد عقد 16، 12، 8 على التوالي - استخدام دالة التنشيط ReLU في الطبقات المخفية ودالة Sigmoid في طبقة الإخراج

- **النموذج النهائي :**

الشكل 11 يوضح التكوين النهائي للنموذج الذي توصلنا إليه:

```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.decomposition import PCA
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Input
from imblearn.under_sampling import RandomUnderSampler
from imblearn.over_sampling import RandomOverSampler

data = pd.read_csv('/content/heart_2020_cleaned.csv')
data.dropna(inplace=True)
label_encoder = LabelEncoder()

data = pd.get_dummies(data, columns=['AgeCategory', 'Race', 'GenHealth'], drop_first=True)

for column in data.columns:
    data[column] = label_encoder.fit_transform(data[column])

X = data.drop('HeartDisease', axis=1)
y = data['HeartDisease']

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

pca = PCA(n_components=11)
X_pca = pca.fit_transform(X_scaled)

X_train_val, X_test, y_train_val, y_test = train_test_split(X_pca, y, test_size=0.5, random_state=42)

model = Sequential()
model.add(Input(shape=(X_train_val.shape[1],)))
model.add(Dense(16, activation='relu'))
model.add(Dense(12, activation='relu'))
model.add(Dense(8, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

model.fit(X_train_val, y_train_val, epochs=10, batch_size=10, validation_split=0.1)

loss, accuracy = model.evaluate(X_test, y_test)
print(f'Accuracy on the original test set: {accuracy*100:.2f}%')
```

شكل 11.3. **الشكل النهائي لنموذج شبكة DNN**

3.4 النتائج :

تم تقييم أداء كل النماذج الثلاثة باستخدام مجموعة البيانات نفسها وبنفس معالجة البيانات التي تشمل كل من الترميز (Encoding)، اختيار السمات (Feature Selection)، توحيد البيانات (Feature Scaling)، وتخفيض الأبعاد باستخدام تحليل المكونات الرئيسية (PCA). ونتيجة الدقة التي وصلت لها كل خوارزمية موضحة في الشكل 12.

اسم الخوارزمية	الدقة (Accuracy)
Deep Neural Network (DNN)	91.58%
Logistic Regression	91.25%
K-Nearest Neighbors (K=2)	90.00%

شكل 12.3. **مقارنة بين نتائج خوارزميات التعلم الآلي و العميق**

4 إستنتاج :

أظهرت النتائج أن الشبكة العصبية العميقة المقترحة (DNN) حققت أفضل أداء بدقة بلغت 91.58 %، تلاها الانحدار اللوجستي بدقة 91.25 %، ثم خوارزمية الجار الأقرب (KNN) بدقة 90.00 %.

يوضح هذا أن النماذج الأكثر تعقيدًا مثل DNN يمكن أن تقدم أداءً أفضل في التنبؤ بمخاطر أمراض القلب نظرًا لقدرتها على معالجة العلاقات غير الخطية بين المتغيرات.

5 واجهات النظام :

في هذا القسم، سنوضح كيفية استخدام واجهات النظام الخاص بنا للتنبؤ بمرض القلب. لقد قمنا بتصميم واجهة بسيطة وسهلة الاستخدام، يمكن من خلالها إدخال البيانات المطلوبة مثل العمر، الجنس، مستوى النشاط البدني، وغيرها من العوامل الصحية.

من مميزات هذه الواجهة إمكانية تشغيلها على جميع أنظمة التشغيل المتاحة، وذلك لأنه تم تطويرها باستخدام بيئة Flutter التي تُسهّل إنشاء واجهات متعددة المنصات. هذا يتيح للمستخدمين تجربة

موحدة وسلسة سواء كانوا يستخدمون أجهزة Android أو iOS.

تشمل خطوات تطوير واجهة برمجة الهاتف ما يلي:

1. **شاشة التشخيص**:

- تحتوي على خيارات لتشخيص أمراض القلب والسكري، وتقديم نصائح للمرضى.

- زر يأخذك إلى صفحة الأطباء.

- زر يأخذك إلى صفحة المطورين.

- زر للتحديثات.

2. **شاشة إدخال البيانات**:

- تحتوي على مجموعة من حقول الإدخال لجمع المعلومات الأساسية عن المستخدم لتشخيص حالته الصحية.

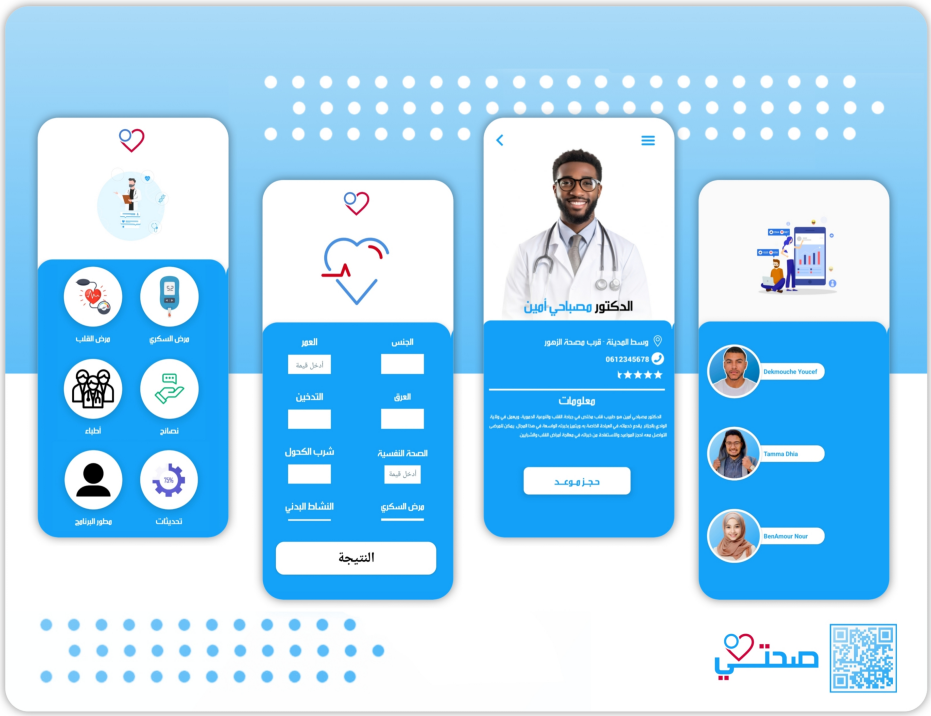
3. **شاشة الطبيب**:

- تعرض معلومات الطبيب وتتيح للمريض حجز موعد معه.

4. **شاشة المطورين**:

- تحتوي على معلومات مطوري البرنامج.

و الشكل 13 يوضح هاته واجهات.



شكل 13.3. **واجهة برمجة الهاتف لنموذج التنبؤ بأمراض القلب**

خاتمة :

في هذا الفصل قُمنا بعرض النظام الخاص بنا من أجل الكشف عن مرض القلب وذلك عن طريق نموذج التعلم العميق (DNN). حيث أننا قُمنا بشرح مفصّل لكل مرحلة من مراحل النظام، كما أننا تطرقنا لشرح عمليات التنفيذ لكل مرحلة. بالإضافة إلى ذلك شرحنا بالتفصيل كيفية بناء النموذج المُقترح وكيف قُمنا بتدريبه وتقييمه. وفي الأخير تحصلنا على النتائج، حيث حققت الشبكة العصبية العميقة المقترحة آداء 91.58 %.

كما أننا قُمنا بمقارنة النموذج مع أهم نماذج التعلم الآلي (LR , KNN) المتحصّلة على أحسن آداء في الدراسات السابقة. وكانت النتائج 90.00 % , 91.25 % على التوالي. وبالتالي نستنتج أن تقنيات التعلم العميق يُمكنها تحسين الآداء نظام التنبؤ بأمراض القلب مَقارنة بطرق التعلم الآلي.

الخاتمة العامة

تُشكِّل أمراض القلب تحديات كبيرة للنظم الصحية بسبب تعقيد تشخيصها ومتابعتها المستمرة، مما أدى إلى أن التنبؤ بالمرض بالإعتماد على الفحوصات التقليدية المُكلفة والتي تستغرق وقتًا طويلاً والتي تخلق إشكالية كبيرة تستدعي البحث عن حلول أكثر كفاءة وفعالية.

الأبحاث السابقة قدمت مجموعة من الحلول التي شملت تقنيات متقدمة مثل تخطيط كهربائية القلب ،(ECG) فحوصات الدم المتنوعة، والتصوير الطبي. ومع ذلك، كانت هذه الحلول محدودة بسبب تكاليفها العالية والحاجة إلى أجهزةمتخصصة، فضلاً عن المتابعة المستمرة.

مؤخرًا، أثبتت الدراسات الحديثة أن إستخدام تقنيات الذكاء الإصطناعي المختلفة، بما في ذلك التعلم الآلي والتعلم العميق، إلى تحسين جودة التنبؤات المرضية وزيادة فعالية الرعاية الصحية.

في هذا المشروع، تمكَّنا من تطوير نموذج تعلم عميق لتحليل مؤشرات البيانات الصِحية وتقديم توقعات دقيقة حول التنبؤ بمرض القلب. حقق النموذج نتائج مُبهرة عند إختباره على قاعدة بيانات ضخمة، حيث حسّن الأدآء مقارنة بالأعمال السابقة. بحيث توصَّلت الدراسة إلى أن استخدام تقنيات التعلم العميق يمكن أن يُوفر بديلاً موثوقًا للطرق التقليدية في التشخيص، كما أنه يُساهم في تحسين النتائج الصحية وتقليل التكاليف المُرتبطة بالعلاج.

وفي الأخير، نتائج هذه الدراسة تُؤكد على أهمية إستغلال التقنيات الحديثة، مثل الذكاء الإصطناعي والتعلم العميق، في تعزيز جودة الرعاية الصحية. كما أننا سنسعى في المستقبل إلى إضافة ميزة مسح التحاليل الطبية وقراءتها بإستخدام الذكاء الاصطناعي، بالإضافة إلى تطوير النموذج ليتمكن من التعرف على نوع المرض بدقة أكبر. سنعمل أيضًا على تحسين آداء النموذج وزيادة دقته. ومن بين الأفكار الأخرى التي يمكن تطبيقها، تطوير تطبيق يمكن المستخدمين من متابعة حالتهم الصحية بشكل دوري، وتقديم توصيات مخصصة بناءً على بياناتهم الصحية. كما يمكننا تحسين واجهة المستخدم لتكون أكثر تفاعلاً وسهولة في الاستخدام، مما يعزز من تجربة المستخدم ويجعل التطبيق أكثر فائدة.

المصادر

1 altibbi https://altibbi.com.**(16 January 2024)**

2 Mémoire de fin d'études de Master, Etude d'un Système Intelligent Pour Suivre les Patients Atteints de Maladies Cardiaques, UNIVERSITE ABDELHAMID IBN BADIS DE MOSTAGANEM. **(18 January 2024)**

3 Mémoire de fin d'études de Master, Prédiction des maladies cardiaques à l'aide des techniques d'apprentissage statistiques UNIVERSITE MOHAMED BOUDIAF - M'SILA.**(25 January 2024)**

4 D.Ayman Abou Majd , Maladies cardiaques et artères coronaires El Ckorouk.**(8 February 2024)**

5 Mémoire de fin d'études de Master, Système intelligent de prédiction des maladies cardiaques Université de 8 Mai 1945 – Guelma –.**(29 February 2024)**

6 healthywa , https://www.healthywa.wa.gov.au/.**(8 March 2024)**

7 Mitchell, T. M. (1997). *Machine Learning*. McGraw Hill.**(10 March 2024)**

8 ijeat , https://www.ijeat.org/wp-content/uploads/papers/v9i3/B3986129219.pdf.**(8 April 2024)**

9 Kaggle. (2020). Personal Key Indicators of Heart Disease. Retrieved from https://www.kaggle.com. **(7 April 2024)**

10 Python Software Foundation. (2021). Python. Retrieved from https://www.python.org.**(16 April 2024)**