* بسم الله الرحمان الرحيم
* Bonjour Madame Monsieur les membres de jurés;
* En vue de l’obtention du diplôme de master, nous avons l’honneur de vous présenter notre travail qui porte sur l’Aide au diagnostic cardiaque basé sur ECG
* Nous nous présentons je suis xxxxx et mon camarade xxxxx
* Voici le sommaire de notre plan de présentation;
* Nous commençons par une petite introduction après nous allons aborder les principes de l’ECG ainsi que son acquisition
* Puis nous entamons une petite discussion sur les maladies cardiaques et leur impact sur la vie de l’être humain,
* Ensuite on aborde les différents méthodes d’apprentissage et par conséquent les réseaux de neurones.
* Et au finale on vous exposera les résultats obtenus de notre application.
* Et on conclus avec une conclusion générale qui englobe notre travaille
* L’IA, est en train de changer le monde, dans la médecine elle permet et facilite le suivi des patients et la manipulation des robots chirurgiens ainsi que les traitements personnalisés, prothèses intelligentes ou surveillance épidémiologique : l'intelligence artificielle a un énorme potentiel dans le domaine de la santé. Les progrès de la puissance de calcul ont permis d'analyser rapidement de grandes quantités de données avec cohérence et précision.
* Les maladies cardiovasculaires (MCV), malgré les progrès considérables réalisés en matière de diagnostic et de traitement, représentent toujours la principale cause de morbidité et de mortalité dans le monde. Afin d'améliorer et d'optimiser les résultats des MCV, les techniques d'intelligence artificielle ont le potentiel de changer radicalement la façon dont nous pratiquons la cardiologie, en nous offrant de nouveaux outils pour interpréter les données et prendre des décisions cliniques.
* Les techniques d'IA telles que l'apprentissage automatique et l'apprentissage profond peuvent également améliorer les connaissances médicales en raison de l'augmentation du volume et de la complexité des données, débloquant ainsi des informations cliniquement pertinentes.
* **Problimatique**:
* L’interprétation des ECG reste une tâche difficile, même pour les médecins surtout s’ils doivent interpréter en urgence ; Le problème étudier ici c’est comment peut-ont utilisé DL pour l’aide au diagnostic de ce dernier ? autrement, est ce que le Deep Learning peut faire une classification correcte au différents ECG acquis ?

**L’ECG : qu’est-ce qu’un ECG ?**

* Un equipement qui permet une représentation complète de l'activité électrique du cœur à la surface du corps humain, et il est largement appliqué dans le diagnostic clinique des pathologies cardiovasculaires depuis son développement au XXe siècle., il peut être utilisé de manière fiable comme mesure pour surveiller la fonctionnalité du système cardiovasculaire.)
* Les signaux ECG ont été largement utilisés pour détecter les maladies cardiaques en raison de leur simplicité et de leur caractère non invasif, Puisqu’ils reflètent l'activité électrique du cœur. Ainsi, les troubles du rythme cardiaque ou les altérations de la forme d'onde de l'ECG sont des preuves de problèmes cardiovasculaires sous-jacents, tels que les arythmies.

**L’enregistrement :**

* A l’intérieur d’un groupe de fibres cardiaques, les différences de potentiel que produit la dépolarisation de chacune d’entre-elles vont s’additionner pour atteindre une amplitude telle qu’elles peuvent être saisies à la surface du corps, par des électrodes cutanées.

Les électrodes se divise en deux catégories

* Électrodes précordiales : Six électrodes unipolaires sont disposées sur le thorax, sur un plan horizontal, chacune à un endroit bien précis, afin d’obtenir des tracés qui restent fait comparables d’un enregistrement a l’autre.
* Électrodes périphériques : Au nombre de quatre, marquées de couleurs conventionnelles (rouge, noir, vert et jaune), les électrodes périphériques sont placées au poignet droit (électrode rouge), au poignet gauche (électrode jaune), à la cheville droite (électrode noire) et à la cheville gauche (électrode verte).

**Les maladies :**

* Les coronaropathies (touchent les vaisseaux sanguins qui alimentent le muscle cardiaque) et l’insuffisance cardiaque sont les maladies cardiovasculaires les plus fréquentes. Les maladies cérébrovasculaires touchent les vaisseaux sanguins qui alimentent le cerveau et les maladies vasculaires périphériques touchent les vaisseaux sanguins qui alimentent les bras et les jambes.

Nous passons à la partie qui traite notre problème de classification tous d’abord parlons de quelques méthodes de classifications :

**ML :**

* L'apprentissage automatique est un sous-domaine de IA et la technique la plus ancienne et la plus simple. Il est basé sur un algorithme qui s'adapte au système basé sur le feedback humain. La mise en œuvre de cette technologie signifie l'existence de données organisées. Ensuite, fournissez des données structurées et classifiées au système afin qu'il puisse comprendre comment classer de nouvelles données similaires. En fonction de la base de cette classification, le système exécutera l'action programmée. Par exemple, il peut identifier si les données numérique qui proviennent de l’ECG sont normal ou bien il y a une arythmie, puis trier les fréquences cardiaque dans la classe correspondante.

**DL** :

* Le deep learning ou apprentissage profond est un type d'intelligence artificielle dérivé du machine learning où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter à la lettre des règles prédéterminées.
* Le deep Learning s'appuie sur un réseau de neurones artificiels s'inspirant du cerveau humain. Ce réseau est composé de dizaines voire de centaines de « couches » de neurones.

**Type d’apprentissage :**

Il existe trois types d’apprentissage nous allons discuter de seulement deux:

* Supervisé : consiste à former un modèle d'apprentissage automatique à l'aide de données étiquetées. Cela signifie que les données dont on dispose sont déjà associées à la bonne classification.
* Non supervisé : consiste à former un modèle avec des données non étiquetées. Cela signifie que le modèle doit trouver ses propres caractéristiques et faire des prédictions en fonction de la façon dont il classe les données.

Nous allons choisir deux algorithmes supervisé et un seul non supervisé

* + Support vector machines « SVM »: est utilisées pour la classification, la régression et la détection des valeurs aberrantes ; Les SVM se distinguent des autres algorithmes de classification par la manière dont ils choisissent la limite de décision qui maximise la distance entre les points de données les plus proches de toutes les classes.
  + Le graphe représente une séparation linéaire il existe aussi plusieurs type de séparateur selon la nature des données
  + Naive Bayes: Les classifieurs Naive Bayes sont une collection d'algorithmes de classification basés sur le théorème de Bayes ; l’algorithme nous permet de prédire une classe, étant donné un ensemble de caractéristiques utilisant des probabilités.
  + Avec la formule de probabilite suivante :
  + K-Means : (k-moyennes) est un algorithme de clustering non supervisé conçu pour partitionner des données non étiquetées en un certain nombre (c'est-à-dire le " K ") de groupes distincts. En d'autres termes, les k-means trouvent des observations qui partagent des caractéristiques importantes et les classent ensemble dans des clusters

**Difficultés rencontrées dans ML :**

nous allons aborder certains des principaux défis que nous pourrions rencontrer lors du développement de notre modèle d'apprentissage automatique.

* Collecte des données: Les données jouent un rôle essentiel dans tous les cas d'utilisation.
* Mauvaise qualité des données: Nous ne voulons pas que notre système fasse de fausses prédictions. La qualité des données est donc très importante pour obtenir des résultats précis.
* Surapprentissage des données : Le modèle est bien performant, faisant des prédictions probables sur l'ensemble de données d'entraînement, mais il n'est pas bien généralisé.
* Sous-apprentissage des données: Le sous-apprentissage, qui est l'opposé du sur-apprentissage, se produit généralement lorsque le modèle est trop simple pour comprendre la structure de base des données.

NN :

* Ils sont conçus pour reconnaître des modèles et ils nous aident à regrouper et à classifier et regrouper les données étiquetées et les données non étiquetées.
* Les réseaux de neurones sont parfois décrits en fonction de leur profondeur,   
  notamment du nombre de couches qu'ils possèdent entre l'entrée et la sortie,   
  ou des couches dites cachées du modèle. C'est pourquoi le terme "réseau de neurones" est presque synonyme d'apprentissage profond

Les principaux types de réseaux de neurones artificiels sont les suivants :

* perceptrons multicouches : est un type de réseau neuronal formel organisé en plusieurs couches au sein desquelles une information circule de la couche d'entrée vers la couche de sortie uniquement ; il s'agit donc d'un réseau à propagation directe (feedforward).
* RNN : Réseaux de neurone récurant Ils enregistrent la sortie des nœuds de traitement et réinjectent le résultat dans le modèle. C'est ainsi que l'on dit que le modèle apprend à prédire le résultat d'une couche ; Ce réseau de neurones commence par la même propagation frontale qu'un réseau à action directe, mais il se souvient ensuite de toutes les informations traitées afin de les réutiliser à l'avenir ; Ce type de réseau ANN est fréquemment utilisé dans les conversions texte-parole.
* CNN : Ce modèle de calcul de réseau de neurones utilise une variante des perceptrons multicouches et contient une ou plusieurs couches convolutionnelles qui peuvent être entièrement connectées ou regroupées. Ces couches convolutionnelles créent des cartes de caractéristiques qui enregistrent une région de l'image, laquelle est ensuite divisée en rectangles et envoyée à l'analyse non linéaire. Le modèle CNN est particulièrement populaire dans le domaine de la reconnaissance d'images.
* DNN : En général, un Deep Neural Network a une couche d’entrée, une couche de sortie et au moins une couche entre les deux. Plus le nombre de couches est élevé, plus un réseau est dit ” profond “.

Les outils qui nous

* Outils utilisés : Nous avons utiliser pour cette experomentation le language python baser sur le notebook Jupyter qui peut gèrer des codes Python basé sur une application web exécutable
* Nous avons travailler avec La base de données sur les arythmies du MIT-BIH, complétée en 1980, a une des premières bases de données introduites dans ce domaine. Cette base de données fournit une série de tests pour l'évaluation des détecteurs d'arythmie.
* Notre objectif est de classer les signaux ECG de notre base d’apprentissage qui contient 87 554 signales en cinq classes et de faire l’evaluation avec la base de test qui contien 21 891 signales elle aussi classé en cinq classes.

Voici une représentation graphique de chaque classes des différents battements

\*

Passant maintenent a notre application et le modèle que nous avons proposé !

* Nous avons créé notre modèle d’apprentissage simple PMC basé sur DNN qui se compose de 3 couches cachées

Les principales caractéristiques du modèle sont les suivantes :

* La couche d'entrée est constituée de 50 nœuds la fonction d'activation 'sigmoid'.
  + Nous avons utiliser sigmoid dans la couche d’entre car elle permis de prédir la probabilité de sortie car elle existe entre 0 et 1.
* Les trois couches cachées que nous avons utilisées ont la même fonction d’activation tanh et la seule défiance se trouve au nombre de neurones.
* Tanh représente a peu prés la meme chose que sigmoid elle existe entre -1 et 1 se qui permis d’avoir des résultat plus performant
* La quatrième et dernière couche est composée de 5 neurones c’est-à-dire 5 classes et de la fonction d'activation 'softmax'.
* La fonction Softmax distribue une probabilité entre les classes avec une somme égale a 1.

Résultat :

Comme on peut le voir nous avons utiliser les mesure d’évaluation basé sur loss et accuracy

La phase apprentissage de notre modèle c’est effectue sur 100 époques avec un batch size de 32

Chaque époque prenais a peu près 5 Sec

* La métrique qui été utilisé pour calculer loss c’est **categorical crossentropy**
* Car nous avon utiliser la fonction **d’activation softmax** dans la couche de sortie.
* Batch\_size du test a été augmenté à 100 pour reduire le temps d’éxécusion

Pour une meilleur représontation des résultat voici les graphe suivat :

**Le premie représent l accuracy** : la courbe bleu désigne validation et l‘autre désigne celle de l’entrainement

De meme pour le deuxéme graphe qui représente Loss

Le modèle que nous avons adobté nous a permi d’avoir **97.8% exactitude** ‘accuracy’ et **8.7% taux d’erreurs** ;

**Nous remarquons que dans le deuxéme graphe** la courbe de validation commance a divergé de celle d’entrainement mais pas d’une grande marge c’est-à-dire que notre apprentissage ne tond pas vers surapprentissage

On voix maintenant la matrice de confusion aussi connu sous le nom de matrice d’erreur.

Chaque ligne de la matrice représente les échantillons d’une classe prédite tandis que chaque colonne représent les échantillons d’une classe réel

Elle va nous permetre de encors mieux visualisé nos résultat car la diagonal de cette matrice représente les true positif c’est-à-dire les échantillons bien classé, au-dessus de cette ligne en retrouve les faux positifs c’est-à-dire ceux qui ont été classé dans cette classe mais n’appartient pas à celle-ci ,et au dessous en retrouve les faux négatif

Conclusion et :

Les réseaux neuronaux profonds ont été les plus appropriés pour la réalisation de notre projet.

Perspective :

Et nous aspirons au future de pouvoir crée un programme qui nous permettra de relie et concrétisé notre programme avec un équipement ECG et faire une classification en temps réel.

Donc, nous vous remercions d’avoir consacré de votre temps pour juger notre modeste travail