

Dossier : Modélisation et réseaux de neurones formels

Table des matières

Introduction.....	1
Présentation des projets.....	1
Projet 1- Réseau de neurones artificiel : prédire les chances de survit en cas de crise cardiaque.....	1
Projet 2 - Réseau de neurones convolutif : Déterminer les types d'un pokemon à partir d'une photo.....	1
Réflexion générale.....	1
Conclusion.....	1

Introduction

1) Le réseau de neurones artificiel

Haykin¹ va définir le réseau de neurones formel comme étant « une combinaison parallèle d'unité de traitement simple qui peut acquérir des connaissances de l'environnement à travers un processus d'apprentissage et stocker les connaissances dans ses connexions. ». Ce processus d'apprentissage est inspiré du mécanisme d'un réseau de neurones biologique.

Chaque unité va posséder des entrées auxquelles vont être attribuées des poids. A partir de ces entrées, l'unité va calculer la somme pondérée de ces poids, puis une fonction d'activation va déterminer comment cette unité va propager l'information.

Tableau 1. Similarités d'un neurone biologique et d'une unité²

Neurons	Processing Elements (PEs)
Synapses	Weights
Dendrites	Summing Function
Cell Body	Activation Function
Axon	Output
Threshold value	Bias

Les unités vont se combiner dans un réseau de neurones artificiel composé d'une succession de couches d'unités. Chaque couche va prendre comme entrées les sorties de la couche précédente.

Tableau 2. Similarités d'un réseau de neurones biologique et d'un réseau de neurones artificiel

¹ S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New Jersey, 1999.

² Guresen, Erkam, and Gulgun Kayakutlu. "Definition of Artificial Neural Networks with Comparison to Other Networks." *Procedia Computer Science*, vol. 3, Jan. 2011, pp. 426–33

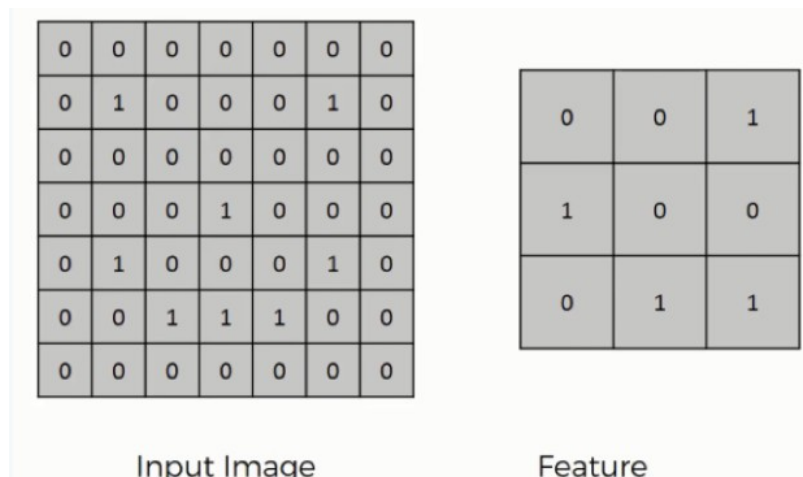
Biological Neural Networks	Artificial Neural Networks
Stimulus	Input
Receptors	Input Layer
Neural Net	Processing Layer(s)
Neuron	Processing Element
Effectors	Output Layer
Response	Output and an entry

En apprentissage profond, les poids d'entrées peuvent être attribués de manière aléatoire sur une base de donnée. A la suite de la propagation de l'information (sortie) sera calculé le gradient de l'erreur. Ce gradient correspond à la différence entre la sortie obtenue et la sortie recherchée.

Il va y ensuite avoir apprentissage par rétropropagation du gradient qui va venir modifier les poids des entrées. Chaque nouvelle itération permet un ajustement des poids.

2) Le réseau de neurones convolutif

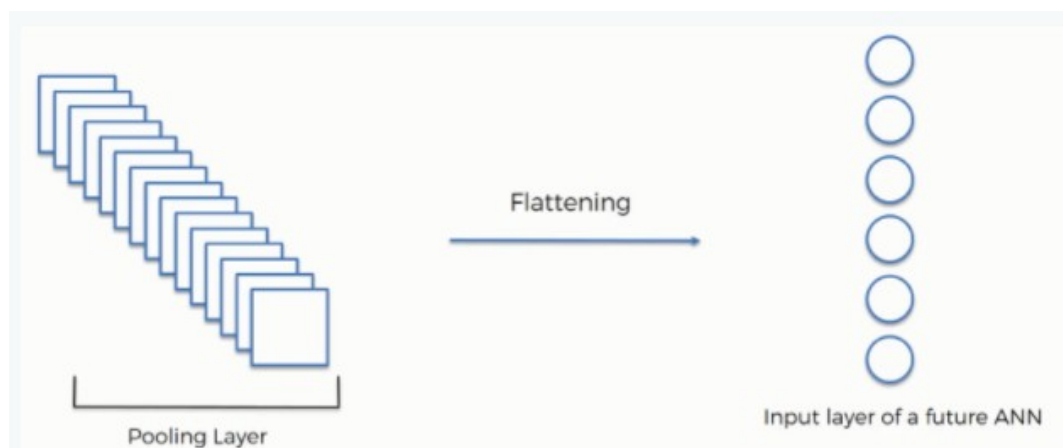
Le réseau de neurones convolutif est une forme de réseau de neurones qui permet d'analyser une image en une succession de 4 étapes. Chaque unité aura le même poids d'entrée et vont traiter une partie de l'image, cette partie de l'image sera légèrement décalée pour chaque unité et peuvent se chevaucher entre elles, c'est ce qu'on appelle un produit de convolution.



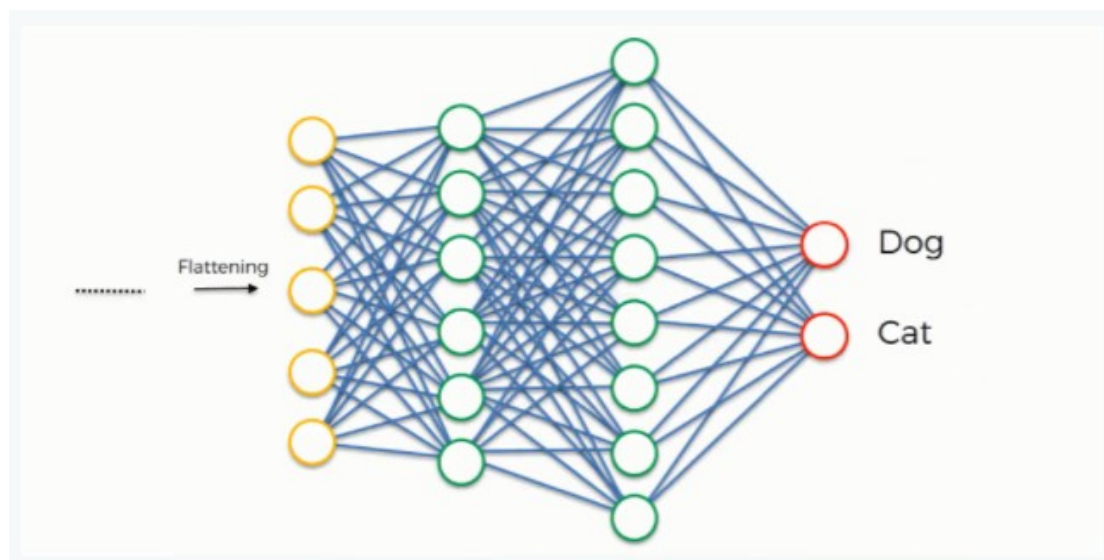
La couche suivante va s'occuper de mettre en commun les sorties.



Dans le but de pouvoir traiter ces données sous formes d'entrée, on va effectuer une couche d'aplatissement.



La dernière couche va permettre l'analyse profonde des vecteurs obtenus avec la méthode d'un réseau de neurones artificiel.



Projet 1- Réseau de neurones artificiel : prédire les probabilité de survie en cas de crise cardiaque

1. Présentation du sujet

La mort à la suite d'une crise cardiaque est la première cause de décès dans le monde. Le jeu de donnée³ provient du site kaggle.com et recueil 12 facteurs chez 300 personnes qui ont subi une crise cardiaque ainsi que si cela a engendré la mort ou si ils ont survécu.

L'objectif du réseau de neurones était de prédire selon 11 facteurs (le facteur « temps » ayant été retiré car il est impossible à récupérer dans un but prédictif en situation réelle) le risque de décès si une personne subit une crise cardiaque.

2. L'architecture utilisée

³ Davide Chicco, Giuseppe Jurman: Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Medical Informatics and Decision Making 20, 16 (2020)

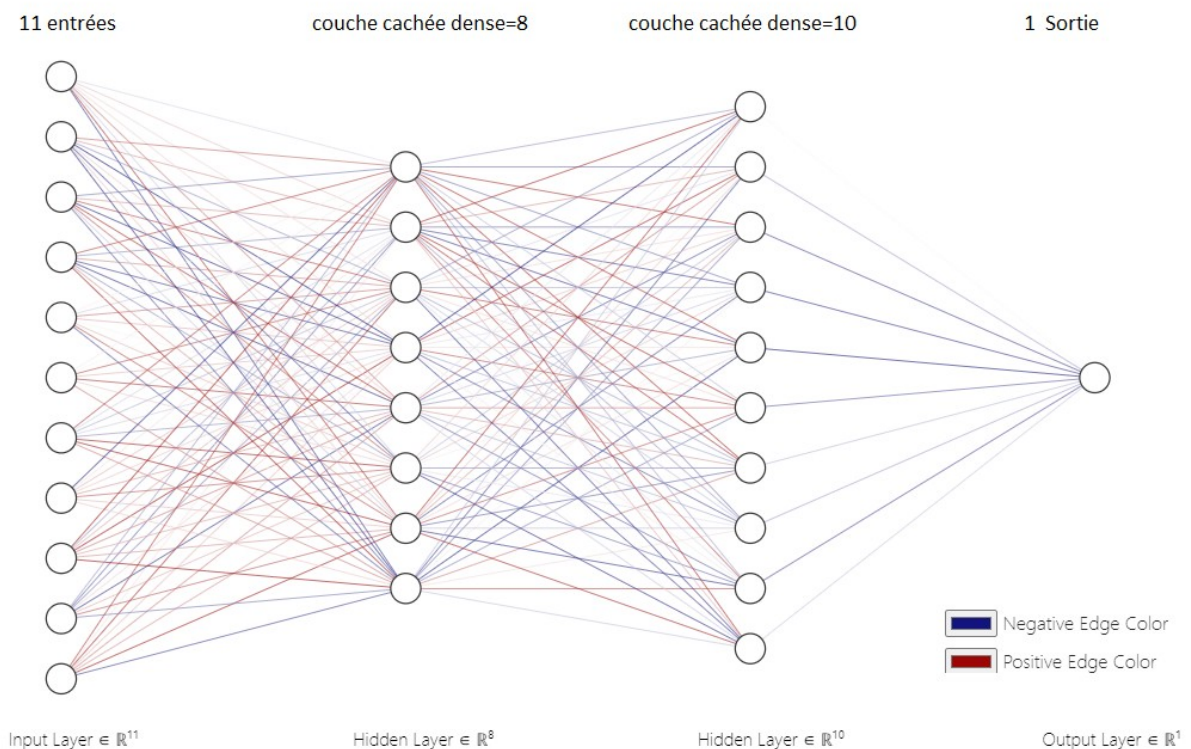


Tableau 3. Architecture finale du réseau

Layer (type)	Output Shape	Param #
dense_4828 (Dense)	(None, 8)	96
dense_4829 (Dense)	(None, 10)	90
dense_4830 (Dense)	(None, 1)	11
Total params: 197		
Trainable params: 197		
Non-trainable params: 0		

Le réseau est composé de onze catégories en entrée avec des poids aléatoire, deux couches cachées et une sortie qui prédit en probabilité le risque de survit en cas de crise cardiaque.

3. Procédure

Pour avoir le meilleur résultat possible, des hyperparamètres ont été utilisée faisant varier le nombre d'itérations, le type d'optimizer et le nombres de sorties des couches cachées :

- En faisant varier le nombres d'itérations, on cherche à déterminer jusqu'à quand le réseau continue d'apprendre. Il a été choisi des mesures relativement espacées pour garantir qu'il y est une différence.
- l'optimizer : deux optimizers robustes avec une logique différents ont été choisi : RMSpropre et Adam. RMSpropre divise le gradient par la racine de sa moyenne alors que Adam utilise une méthode de décente du gradient

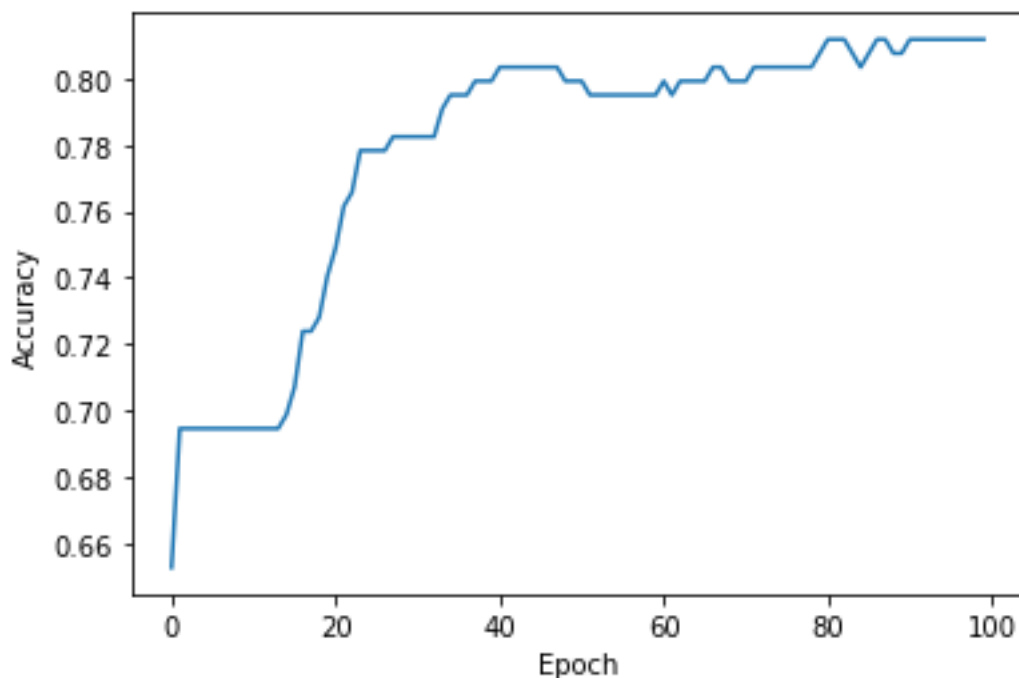
Pour éviter le surapprentissage, les données ont été décomposé en deux parties : une pour l'entraînement et une pour le test.

Tableau 4. Résultat à la fonction *bestparam*

Key	Type	Size	Value
algo_gradient	str	1	adam
batch_size	int	1	25
epochs	int	1	100
nb1	int	1	8
nb2	int	1	10

Je n'ai pas réussi à extraire et sauvegarder les données à la suite des hyperparamètres, ci-dessous les résultats obtenus à la suite de l'entraînement en utilisant les recommandations de *bestparam* :

```
In [70]: score=network.evaluate(X_train,
y_train,verbose=0)
...: print("%s: %.2f%%" % (network.metrics_names[1],
score[1]*100))
accuracy: 81.17%
```



Le réseau est capable de prédire correctement à 81,17 % si une personne va décéder en cas d'attaque cardiaque sur le jeu de test.

On peut également imaginer faire de nouvelles prédictions en condition réelle :

```
...: print(new_prediction)
Le patient a 44.85 pourcent de chance de s'en sortir en
cas d'attaque cardiaque
[[0.55154353]]
```

4. Réflexion

Cet outils peut s'avérer utile d'un point de vue préventif, on peut imaginer qu'à la suite d'un diagnostic, si les prédictions sont mauvaises, une attention ou une prise en charge particulière pour ces personnes. A la suite du suivi si les prédictions s'améliorent cela permettrait de prévenir le risque de décès à la suite d'une attaque cardiaque.

On peut également imaginer des questions éthiques quant à la connaissances de ces informations. D'un point de vue médical si le dossier est informatisé et qu'une personne arrive en réanimation à suite d'une attaque cardiaque, si la personne à une probabilité quasiment nulle de s'en sortir, les médecins pourraient décider de concentrer les efforts sur quelqu'un d'autre qui à plus de chance de s'en sortir. Il pourrait également y avoir une prophétie autoréalisatrice de la part des médecins qui vont modifier inconsciemment leur comportement et dans le cas présent, moins s'investir pour sauver la personne.

D'un point de vue économique, une compagnie d'assurance en possession de ces informations pourrait augmenter le prix de l'assurance d'une personne à risque.

Projet 2 - Réseau de neurones convolutif : Déterminer les types d'un pokemon à partir d'une photo

1. Présentation du sujet

Le jeu vidéo est l'activité qui génère le plus d'argent dans l'industrie du divertissement et pokemon est la licence la plus rentable de tout les temps. On peut donc imaginer voir arriver un jour de l'intelligence artificiel dans les jeux pokemon. Le jeu de données provient du site Kaggle.com et recense les 151 pokemon de première génération. Chaque pokemon peut avoir deux types parmi 14, l'objectif du réseau est de prédire quels sont les 2 types d'un pokemon de n'importe quelle génération (plus de 900 pokemon au total) à partir d'une image. Pour les besoins du projet, j'ai moi même classé les 151 pokemon en 14 catégories.

```
class_mode = 'categorical')
Found 11524 images belonging to 14
classes.
Found 3459 images belonging to 14 classes.

In [5]: train_generator.class_indices
Out[5]:
{'type_combat': 0,
 'type_dragon': 1,
 'type_eau': 2,
 'type_electrique': 3,
 'type_fantôme': 4,
 'type_feu': 5,
 'type_insecte': 6,
 'type_normal': 7,
 'type_plante': 8,
 'type_poison': 9,
 'type_psycho': 10,
 'type_roche': 11,
 'type_sol': 12,
 'type_vol': 13}

In [6]:
```


2. L'architecture utilisée

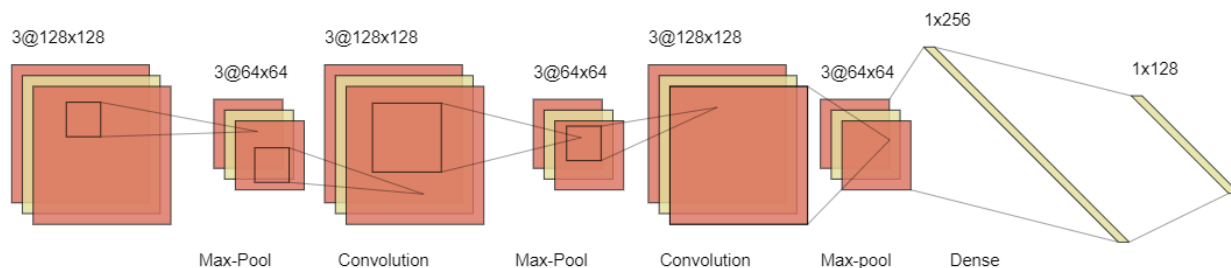


Tableau 5. Architecture finale du réseau

```
Model: "sequential_2"
```

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 126, 126, 32)	896
max_pooling2d_4 (MaxPooling2)	(None, 63, 63, 32)	0
conv2d_5 (Conv2D)	(None, 61, 61, 64)	18496
max_pooling2d_5 (MaxPooling2)	(None, 30, 30, 64)	0
conv2d_6 (Conv2D)	(None, 28, 28, 128)	73856
max_pooling2d_6 (MaxPooling2)	(None, 14, 14, 128)	0
flatten_2 (Flatten)	(None, 25088)	0
dense_3 (Dense)	(None, 128)	3211392
dense_4 (Dense)	(None, 14)	1806
Total params: 3,306,446		
Trainable params: 3,306,446		

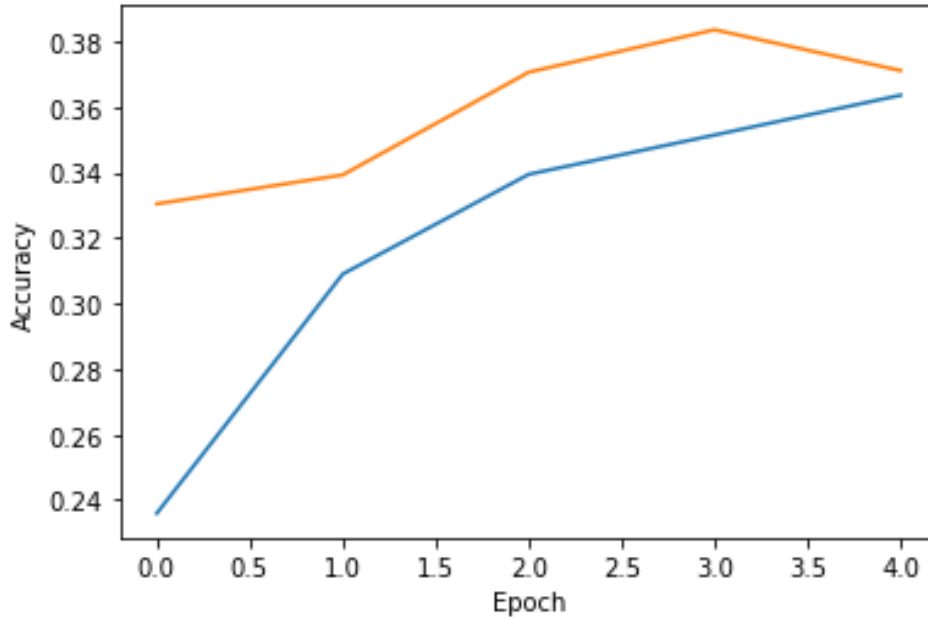
Le jeu de données est composé de 14000 images réparties en 14 catégories et divisées en un jeu d'entraînement et un jeu de test sur la base de 80/20. L'architecture est composée de 3 couches de convolution et 3 couches de pooling inspirée du modèle VGG-16.

3. Procédure

Pour ce réseau, deux architectures ont été proposées car utiliser les hyperparamètres n'aurait pas été raisonnable pour un réseau convolutif. Pour éviter le surapprentissage, la fonction ImageDataGenerator pour déformer l'image a été utilisée. Ci-dessous un comparatif des résultats obtenus :

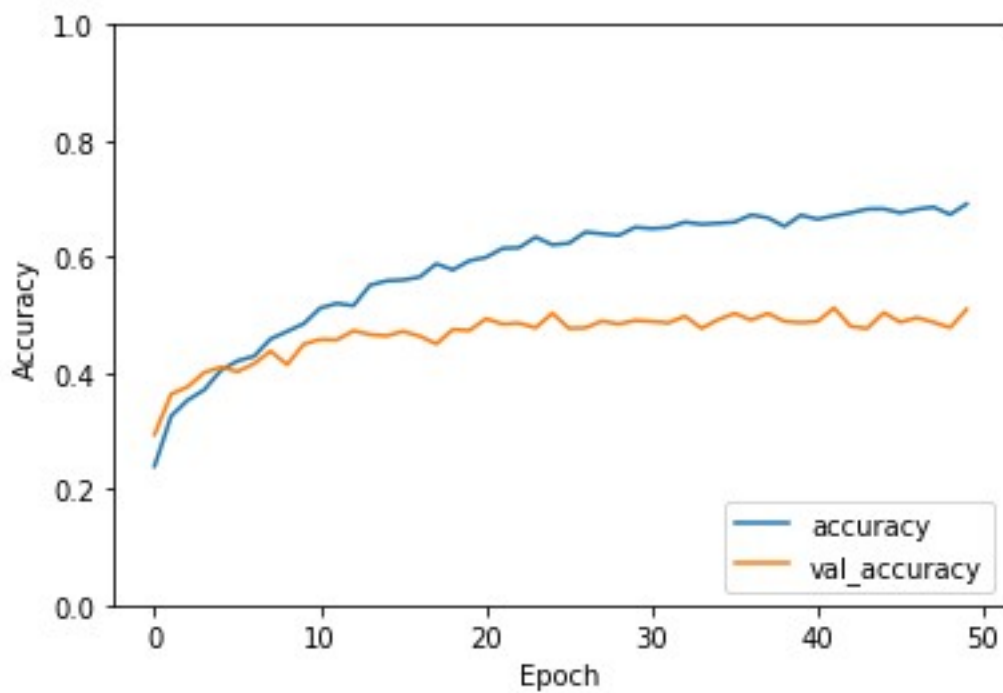
Score de prédiction avec le modèle vu en TP2

```
Epoch 5/5  
250/250 [=====] -  
110s 442ms/step - loss: 1.8280 - accuracy:  
0.3721 - val_loss: 1.6313 - val_accuracy:  
0.3869
```



Score de prédiction avec le modèle inspiré du VGG-16

```
Epoch 50/50  
250/250 [=====] -  
171s 683ms/step - loss: 0.6320 - accuracy:  
0.6911 - val_loss: 1.0292 - val_accuracy:  
0.5094
```



Sur la première architecture le réseau est capable de prédire à 37 % quel sera le type du pokemon dans le jeu de test alors que dans la deuxième architecture la prédiction monte à 69 %. On peut donc imaginer que ces scores seront d'autant plus bas dans une condition réelle. Cependant ces résultats pour autant de catégories et les premiers tests effectués sont encourageants.

Exemple



```
In [102]: test_image('data/test.jpg')  
ton pokemon est de type plante et de type poison
```

4. Réflexion

L'utilité de ce réseau serait dans un objectif de divertissement. Actuellement il existe un nombre fini de pokemon, il serait donc plus simple d'aller chercher de quels types est le pokemon qui nous intéresse. Il est probable que dans quelques années, le jeu pokemon évolue vers une intelligence artificielle capable d'inventer une infinité de pokemon, comme on peut le voir déjà dans certains réseaux de neurones, on pourrait donc avoir un intérêt à définir les types d'un pokemon en fonction de son apparence physique.

Tant que cela reste un monde tiré de l'imaginaire, la prédiction de catégorie ne semble pas poser de problème. En 2015, la reconnaissance d'image de Google avait catégoriser des personnes de couleur noire comme « gorille », cela peut donc avoir des répercussions sociales importantes, sans solution Google avait retiré les catégorie « singes » de leur réseau de neurones.

Réflexion générale

L'utilisation d'algorithmes et d'apprentissage profond ont permis d'améliorer nos conditions de vie dans de multiples domaines : santé, transport, divertissement. Ces répercussions positifs sont à mettre en balance avec de nombreuses questions et problèmes liés à ces intelligences artificielles.

Les jeux de données

Ces données sont à l'origine des sélections humaines. En cherchant à récupérer des données issues de nos vies, nos interactions, nous allons créer des machines qui risquent de reproduire les inégalités de notre société. Par exemple, Tinder utilise un mélange de technologies pour recommander des rencontres entre deux personnes notamment l'outil Recognition développée par Amazon. Il a été montré que Tinder allait suggérer des rencontres de personnes de la même origines ethniques, d'un même niveau social pour les hommes mais aussi des personnes d'un niveau social plus élevé pour les femmes. La machine n'est pas fondamentalement raciste ou sexiste mais est à l'image de notre société.

Il y a également un besoin important de contrôle humain sur la validité de ces données, beaucoup de ces vérifications sont effectuées sous la forme de « travail du clic » dans les pays émergents mais aussi locaux. Il

il y a un paradoxe sur des technologies qui ont pour but d'améliorer nos conditions de vies en créant des micro-travaux.⁴

Les conséquences sur nos libertés

Un algorithme qui nous recommande d'adopter un comportement, basé sur nos préférences reste un nudge. Il est intéressant de se demander si l'ensemble de ces recommandations n'est pas une atteinte à notre liberté. Par exemple, les réseaux sociaux utilisent la recommandation de contenus ce qui peut renforcer nos idées et former une « bulle des filtres ».⁵

Conclusion

Améliorer nos conditions de vie pour l'intérêt de qui ?

« On ne peut pas arrêter le progrès » et ce n'est pas le but, cependant il faut être vigilants sur les réelles intentions derrière. Bien sûr dans une société capitaliste on peut comprendre qu'il y ait des enjeux économiques, le risque est cependant de conduire à des nudges⁶ qui vont contre notre intérêt, par exemple choisir telle direction car il y a des lieux de consommation.

Il faut faire également attention au confort apporté par ces outils, il y a presque 100 ans A. Huxley décrivait un monde futuriste dans lequel les individus étaient emprisonnés par le plaisir⁷. Ce conformisme par le plaisir est un réel risque aujourd'hui, quand on connaît la puissance du renforcement par la récompense, il est plus difficile d'en sortir qu'un contrôle par la punition.

Je pense donc qu'il est une bonne chose de chercher à faire fonctionner ou inventer de nouvelles technologies qui ont un impact positif sur nos vies. Cependant je pense qu'il est tout aussi important de se demander en parallèle les questions suivantes :

Quels sont les intérêts pour l'individu et pour le concepteur ? Est-ce que ces intérêts rentrent en conflit ?

Dans une logique coût/bénéfice, quels sont les avantages et quels sont les risques ? Quelles mesures mettre en place pour limiter ces risques ?

L'ensemble de ces questions doivent être étudiées avant le déploiement des technologies et ce par des comités d'éthiques reconnus.

⁴ « En attendant les robots. Enquête sur le travail du clic. » Antonio Casilli

⁵ Pariser, Eli. *The Filter Bubble: What The Internet Is Hiding From You*. 2011.

⁶ Bovens, Luc. « The Ethics of Nudge. » *Preference Change: Approaches from Philosophy, Economics and Psychology*, 2009

⁷ « Le meilleur des Mondes - Aldous Huxley. »