1. Ce que j’ai vu sur internet : Variables prédictives : ensembles des variables utiles pour la prédiction ex : nbe de portes, nbe de km,…

Variables cible : variable sur laquelle on travaille, celle qui nous intéresse ex : prix d’une voiture

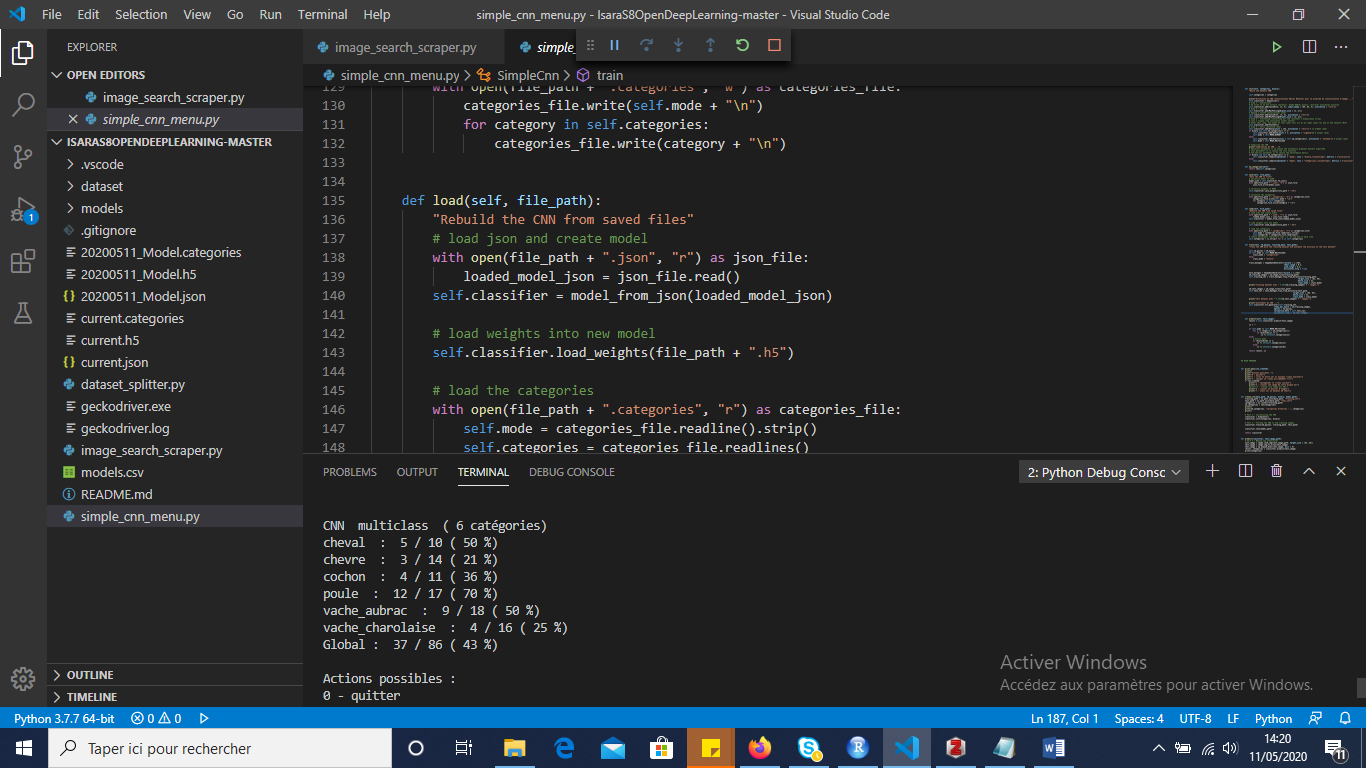
Notre réponse : prédictives = input (de x1 à xn) = training set (= un ensemble d'observation) et cible = output=y= ce que le pc renvoi

On a réalisé ce travail avec les animaux de la ferme =domaine et les catégories suivantes : vache charolaise, vache aubrac,...

1. L’IA nous sort un « raw set » ou il chope 100 image de chaque catégories correspondant à un domaine sur internet. Ensuite on trie manuellement pour enlever les images incohérentes. Toutes les images ne correspondent pas à ce qu'on aurait voulu. On e reçoit pas forcément l'animal vivant mais aussi des dessins et ce n'est pas ce qu'on souhaite apprendre à l'IA. LOrsque la photo comprenait une partie du corps de l'animal, plusieurs naimaux ou un animal avce un homme nous avons supprimées les images pour ne pas induire l'IA en erreur.Il nous reste à ce stade x images d'éléphants,....

Ensuite on sépare aléatoirement notre raw set en deux (résoudre le pb des 80% en demandant à Vivien)

1. Méthode pour l’apprentissage : on entraine l’ordi avec le « training-set » pour qu’il reconnaisse un éléphant en tant qu’éléphant par exemple.  
   On teste qu’il ait bien appris avec le test set pour voir combien d’images il reconnait comme étant un éléphant dans x images d’éléphant.
2. Observer et commenter la structure du fichier source :
3. Remarques :

CNN multiclass ( 6 catégories) avec 10 passes

cheval : 5 / 10 ( 50 %)

chevre : 3 / 14 ( 21 %)

cochon : 4 / 11 ( 36 %)

poule : 12 / 17 ( 70 %)

vache\_aubrac : 9 / 18 ( 50 %)

vache\_charolaise : 4 / 16 ( 25 %)

Global : 37 / 86 ( 43 %)

CNN multiclass ( 6 catégories) avec 25 passes

cheval : 5 / 10 ( 50 %)

chevre : 5 / 14 ( 35 %)

cochon : 4 / 11 ( 36 %)

poule : 11 / 17 ( 64 %)

vache\_aubrac : 8 / 18 ( 44 %)

vache\_charolaise : 5 / 16 ( 31 %)

Global : 38 / 86 ( 44 %)

* On a augmenté que de 1% en faisant + de 2x + de passes
* La solution a préférer est donc peut-être d’augmenter le corpus d’image de base (réponse en 6. Du coup)

1. On fait une ACP pour déterminer quels facteurs comptent le + pour avoir le meilleur %age d’apprentissage. Ensuite selon la réponse de l’ACP on peut tracer des graphiques du %age de réussite en fonction de la/des variables sélectionnées