

Enseignant(s)

**VIDAL Nicolas**

Email(s)

[nvidal@myges.fr](mailto:nvidal@myges.fr)

## 4A IBD - Deep Learning - Kaggle Dataset Analysis with Keras

### 1 Matières, formations et groupes

Matière liée au projet :

Formations : -

Nombre d'étudiant  
par groupe :

**4**

Règles de constitution des groupes: **Imposé**

Charge de travail  
estimée par étudiant : **30,00 h**

### 2 Sujet(s) du projet

Type de sujet : **Imposé**

### 3 Détails du projet

**Objectif du projet (à la fin du projet les étudiants sauront réaliser un...)**

Utiliser le Framework Tensorflow 2.0 pour composer des modèles issus du Deep Learning et obtenir des résultats proches de l'état de l'art sur différents Datasets. Un soin tout particulier sera à apporter à l'étude des hyper paramètres des modèles étudiés.

### Descriptif détaillé

Après avoir installé et configuré Tensorflow 2.0 sur leur machine, les étudiants devront procéder à l'étude de différents modèles sur des Datasets en faisant varier les différents hyper paramètres de ces derniers :

- \* Modèles précédents
  - Modèle Linéaire
  - Perceptron Multicouches
- \* Nouveaux modèles
  - Conv Net(s)
  - ResNets / HighwayNets
  - RNN(s)

Il sera fortement conseillé aux étudiants de s'appuyer sur Tensorboard pour visualiser comparer et retranscrire les performances de ces modèles.

Il sera important de faire apparaître pour chaque dataset, pour chaque modèle :  
L'influence de tous les hyperparamètres des modèles (structure, fonctions d'activation, ...), ainsi que des paramètres des algorithmes d'apprentissages (learning rate, momentum, ...)

### Ouvrages de référence (livres, articles, revues, sites web...)

<http://www.deeplearningbook.org/>  
<https://www.tensorflow.org/>  
<https://keras.io/>

### Outils informatiques à installer

Keras / Tensorflow 2.0 / Jupyter

## 4 Livrables et étapes de suivi

1	Etape intermédiaire	Etude du dataset CIFAR-10 <a href="https://www.cs.toronto.edu/~kriz/cifar.html">https://www.cs.toronto.edu/~kriz/cifar.html</a>  Rendu : notebook Jupyter + pdf / Présentation	mardi 21/01/2020 23h59
2	Etape intermédiaire	Soutenance, Rendu de l'étape 1	mardi 26/02/2019 11h00
3	Etape intermédiaire	Etude des datasets Kaggle  Rendu : notebook Jupyter + pdf + Présentation powerpoint	jeudi 27/06/2019 23h59

4	Rendu final	<p>Etude du dataset 'Mystère'</p> <p>Téléchargez le Train Set sur :  <a href="http://www.greenkumquat.com/dataset/">http://www.greenkumquat.com/dataset/</a></p> <p>Et décompressez le. (attention, plus de 160Go à la décompression)</p> <p>Attention, votre modèle devra comporter 2 entrées nommées input_1 et input_2.  Et proposer 2 sorties, nommées output_1 et output_2.</p> <p>La couche correspondant à output_1 devra être une couche d'Activation simple avec une fonction d'activation softmax.  ex: output_1 = Activation(activation=softmax, name="output_1")(XXX1)</p> <p>La couche correspondant à output_2 devra être une couche Dense avec une fonction d'activation sigmoid.  ex: output_2 = Dense(1, activation=sigmoid, name="output_2")(XXX2)</p> <p>Il est fortement recommandé de privilégier la minimisation du loss de l'output_1 :  ex: loss_weights={'output_1': 1.0, 'output_2': 0.001})</p> <p>Rendu : modèle keras + notebook Jupyter + pdf +  Présentation powerpoint</p>	<p><b>mardi</b>  <b>16/07/2019</b>  <b>23h59</b></p>
---	-------------	--	--

5	Soutenance	
Durée de présentation par groupe :	<b>20 min</b>	Audience : <b>A huis clos</b>
Type de présentation :	<b>Présentation / PowerPoint - Démonstration</b>	
Précisions :		