**Projet de Vie Artificielle – Perceptron multicouches multi-classes probabiliste**

**Introduction**

Le but de ce projet est de comprendre le fonctionnement de réseaux neuronaux simples, et de voir leur capacité de prédiction et de reconnaissance de différents types de données. Dans notre cas nous étudierons 2 ensembles de données : MNIST (Chiffres) et les Silhouettes Cal101.

1. **Présentation du modèle**

L’architecture de base utilisée est un réseau à 3 couches :

* 784 neurones d’entrées
* 1000 neurones intermédiaire
* Activation Tangente Hyperbolique
* 10 neurones de sorties (ou 101 dans le cas de Cal101)
* Activation Softmax
* Entropie Croisée

Nous étudierons ensuite diverses modifications pour comprendre la dynamique du réseau, ou afin d’aider à la visualisation. Par exemple en changeant l’activation, la fonction de perte ou le nombre de neurones intermédiaires.

L’architecture du code adoptée est inspiré de la librarie Keras en python, c’est à dire que chaque type de « couche » possède une classe associée qui calcule indépendamment la propagation avant et arrière. Ces couches incluent les activations, la fonction softmax, les fonctions de pertes etc…

Toutes les couches se trouvent dans le package « layers ». Le package layers.activations contient les activations, le package layers.flat contient les couches « classiques » telle que la transformation affine, l’entropie croisée, la normalisation par batch (Batchnorm) etc…