Dominic MICHAUD  
Dominic JOBIN  
Groupe 0001

**Reconnaissance de caractères à l’aide de réseau de neurones artificiels**  
  
Travail présenté à M. François Bertrand  
Exploration des nouvelles technologies  
420-501-SF

Département de l’informatique  
Cégep de Sainte-Foy  
Vendredi le 1er septembre 2017

Table des matières

[Sujet 2](#_Toc491433702)

[Enjeux 2](#_Toc491433703)

[Objectifs 3](#_Toc491433704)

[Expérimentation 3](#_Toc491433705)

[Limites 3](#_Toc491433706)

[Incertitudes 3](#_Toc491433707)

[Ressources 3](#_Toc491433708)

[Planification 3](#_Toc491433709)

# Sujet

Dans le cours d’exploration des nouvelles technologies, nous rechercherons comment développer un réseau de neurones artificiels (« *artificial neural network* » ou *ANN*) qui sera en mesure de détecter des caractères (chiffres et lettres) dessinées à l’écran. Le but de la recherche est de comprendre le fonctionnement derrière les technologies d’apprentissage automatique (« *machine learning »*) en exploitant un outil de reconnaissance optique de caractères.

Dans le contexte plus général de l’intelligence artificiel, l’enjeu principal est souvent d’automatiser des tâches qui sont faciles pour un humain mais difficiles pour un ordinateur. La reconnaissance optique de caractère (« *optical character recognition* » ou *OCR*) est un bon exemple de tâche qui est difficile à programmer de façon algorithmique.

La reconnaissance de caractères a une application importante au niveau de l’archivage de données, par exemple lorsque les documents sont seulement disponibles sur papier. Il est facile de numériser un document mais difficile d’extraire l’information sous forme de texte pour la traiter par la suite, à moins de transcrire le texte manuellement.

Ce projet fera le lien avec des connaissances que nous avons acquises dans nos cours de programmation et d’algèbre linéaire et nous permettra de les mettre en pratique concrètement.

# Enjeux

Pour arriver à la réussite du projet, beaucoup de connaissances seront nécessaires à acquérir. Le plus difficile sera le fonctionnement derrière les réseaux de neurones artificiels et les détails qui les entourent qui relèvent de l’algèbre linéaire et de calcul avancé. Heureusement, il existe plusieurs librairies dans le language Python qui nous faciliteront la tâche.

L’intelligence artificielle est un sujet qui est de plus en plus d’actualité dans des domaines qui jusqu’à maintenant étaient crus comme étant réservés aux humains, par exemple la conduite d’automobiles. Les réseaux de neurones artificiels sont souvent à la base des systèmes d’intelligence artificielle complexes. La motivation de notre recherche est donc de pouvoir expliquer le lien entre la reconnaissance de caractère et les technologies plus poussées qui ont été développées dans les dernières années par les compagnies comme Google et IBM.

Notre outil de démonstration se limitera à la reconnaissance individuelle de caractères, mais il serait possible de l’utiliser pour transcrire correctement des documents numérisés pour en extraire le texte. Il y a plusieurs domaines où il est utile d’extraire du texte contenu dans une image, par exemple pour archiver les données de façon plus compacte ou pour l’envoyer dans un système de synthèse vocale pour aider les personnes avec des problèmes de vision.

# Objectifs

* Créer un réseau de neurones artificiels
* Utiliser une collection d’images de caractères pour entraîner le réseau pour qu’il apprenne à détecter les caractères avec un taux de réussite d’au moins 95%
* Développer une interface graphique simple qui permettra de dessiner un caractère à l’écran
* À partir du réseau qui aura été entraîné, afficher sous forme de texte le caractère qui a été dessiné à l’écran et le niveau de certitude en pourcentage

# Expérimentation

# Limites

# Incertitudes

# Ressources

# Planification