

CS Games 2015 - Université de Sherbrooke
Apprentissage automatique
Vincent Picard

La ville de Beerbrooke a récemment décidé d'installer un système de caméras intelligentes. Le système possède de nombreuses fonctionnalités, comme la détection de mouvement. Une caractéristique intéressante est la détection de chemin, qui récupère automatiquement la trajectoire 2D des piétons. La ville a reçu beaucoup de plaintes des résidents du centre-ville par rapport à des personnes en état d'ébriété ayant un mauvais comportement. Cela les a amenés à réfléchir à un système qui détecterait automatiquement les piétons ayant trop bu et les piétons sobres.

Vous devez élaborer un système qui distinguera les gens en état d'ébriété des gens sobres. Vous avez un échantillon de données de 500 chemins pour chaque type de chemin (bu, sobre). Votre système doit être en mesure de généraliser les données fournies. Une fois que le système est en place, vous pouvez vous attendre à avoir des demandes fréquentes. Le client (Beerbrooke) veut aussi avoir des coûts de fonctionnement minimales.

Instructions:

- Vous devez mettre en œuvre un algorithme d'apprentissage automatique qui sera en mesure d'identifier le chemin des personnes en état d'ébriété et sobres.
- Des solutions qui ne présentent pas un algorithme d'apprentissage automatique seront pénalisées (mais toujours regardées). Voir la définition de l'apprentissage de la machine:

“Machine learning is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs and using that to make predictions or decisions, rather than following only explicitly programmed instructions.” Wikipédia, Janvier 2015

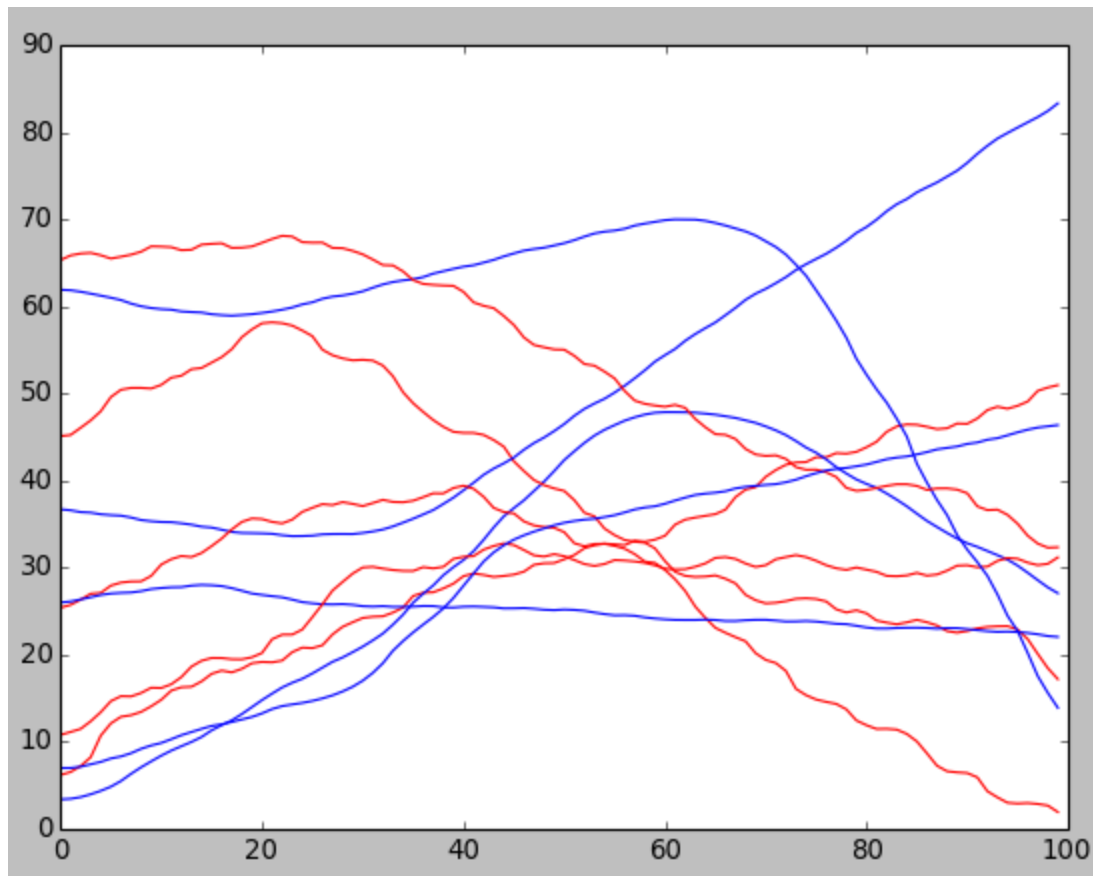
- Vous pouvez faire votre solution en utilisant Java ou Python.
- Vous devez compléter les fonctions *train* et *evaluate* dans la classe *Solution*. Voir Python/Solution.py ou Java/src/MachineLearningChallenge/Solution.java. La fonction *evaluate* sera appelée avec de nouvelles données pour évaluer l'efficacité de votre solution lors de la correction.
- Répondez aux questions sur les pages suivantes. Si vous avez besoin de plus d'espace, des feuilles supplémentaires seront données.
- Vos questions ne seront pas répondues.
- Vous avez trois heures pour terminer la compétition.

Nom (1) : _____

Nom (2) : _____

Équipe : _____

Voici un exemple de ce à quoi les données ressemblent :



Les chemins rouges sont les chemins des personnes en état d'ébriété. Les chemins bleus sont les chemins des personnes sobres. Votre algorithme recevra deux réseaux de 500 chemins dans la fonction *train*. Chaque chemin est une liste de 100 nombres à virgule flottante. Le nombre est la coordonnée y et la position dans la liste est la coordonnée x.

La fonction *evaluate* de votre algorithme doit être capable de catégoriser le chemin donné (une liste de 100 nombres à virgule flottante) en utilisant l'apprentissage automatique.

1. CHOIX D'ALGORITHME. Quel algorithme d'apprentissage automatique avez-vous choisi pour cette compétition? Pourquoi?

Si vous aviez plus de temps, quel aurait été votre choix? Auriez-vous essayé d'autres algorithmes d'apprentissage automatique? Lequel ou lesquels? Pourquoi?

2. UTILISATION DE L'ENSEMBLE DE DONNÉES. Expliquez précisément comment vous avez utilisé vos données pour l'apprentissage et le test de votre algorithme. Votre réponse doit expliquer la façon exacte dont vous avez utilisé les 1000 échantillons.

Toute transformation apportée aux données devrait aussi être expliquée. Quelle est l'entrée finale de votre algorithme?

3. *HYPER PARAMETERS* ET INITIALISATION. Expliquez exactement comment vous avez paramétré votre algorithme. Plus précisément, si votre algorithme a des *hyper parameters*, quelles valeurs avez-vous utilisées et pourquoi?

Votre algorithme a-t-il besoin d'initialisation? Quelles valeurs avez-vous choisies? Pourquoi?
