

Proposition de projet

GIF-7005: Introduction à l'apprentissage machine

Équipe 10: Stéphane Caron, Philippe Blais, Philippe Blouin-Leclerc, Samuel Lévesque et William Bourget

26 octobre 2018

Description du projet

Dans le cadre du projet à réaliser dans le cours GIF-7005, notre équipe avons choisi la problématique proposée par l'entreprise Coveo. Le problème consiste à bâtir un modèle d'apprentissage supervisée qui, à partir d'une nouvelle recherche, retourne un ensemble d'au plus 5 documents pertinents à un utilisateur donné.

Jeu de données

Pour entraîner notre modèle, nous aurons accès à des données portant sur différentes recherches réalisées par des utilisateurs. Nous pourrions donc utiliser certaines informations générales sur l'utilisateur et sur la recherche effectuée, par exemple:

- le langage
- le pays de provenance
- la ville
- la requête effectuée
- ...

De plus, nous pourrions accès aux clics effectués par l'utilisateur en réponse à sa recherche. Ces clics nous donnent l'information sur les documents choisis par l'utilisateur. Cela sera d'ailleurs ce que nous tenterons de modéliser avec notre algorithme d'apprentissage.

Méthodologie anticipée

Nous pensons utilisé l'approche proposée par (Hastie, Tibshirani, and Friedman 2001) sur ...

Références

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2001. *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.