

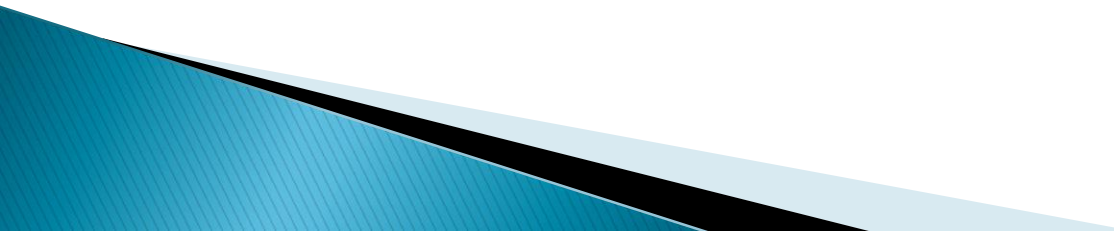
Mineure 6 – Informatique décisionnelle

Francis Dupé & Aubry Herve

Problématique

- ▶ Recherche de liens "me" entre différents réseaux sociaux
- ▶ Apprentissage via un TrainingSet
- ▶ Comparaison d'attributs :
 - Pseudos
 - Mail
 - Friends

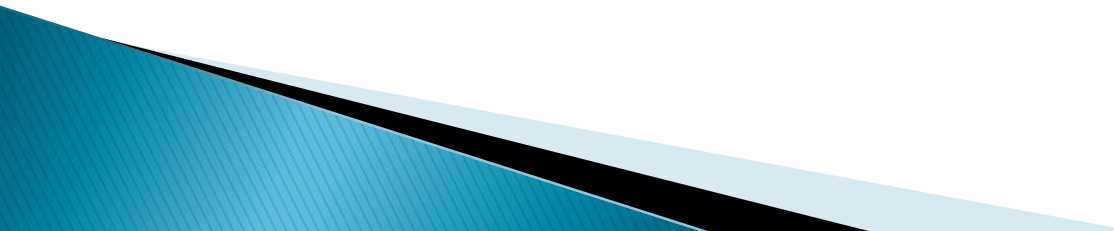
Statistiques

- ▶ 69,105 nodes
 - ▶ 2,990,738 "friends" relations
 - ▶ 501 "me" relations déjà dans la base
 - ▶ 86 "friends" en moyenne
 - ▶ 9874 "friends" max
- 

Données pour la classification

- ▶ Une classe : "me" {"yes", "no"}
- ▶ 6 Attributs :
 - Nicknames
 - Realname
 - Emails
 - Websites
 - profileLinks
 - Locations

Etablissement du Training set

- ▶ 50 paire de "yes" prises dans celles déjà incluses dans la base
 - ▶ 100 paire de "no" obtenu en croisant les paires de "yes" obtenues
 - ▶ Elimination des paires impliquant le même profil
 - ▶ Utilisation de la classe SMO de Weka pour entrainer la Support Vector Machine
- 

Algorithme de recherche

- ▶ Etablissement une liste de candidats potentiels
- ▶ Classement en utilisant le classifieur obtenu