

Etalab

Quelques applications de la
data-science

Inspections

Exemple : Inspection des bâtiments à New-York, par Mike Flowers et une équipe de 5 data-scientists junior, 2011

Objectif : prévenir les incendies par des inspections sur 900.000 bâtiments.

	taux de contrôles positifs
contrôles assignés au hasard	13%
contrôles assignés suivant un score de risque	70%

NB : l'algorithme donne d'excellent résultats grâce aux données des permis de construire. Pourquoi ?

Médecine

Exemple : Puces à ADN

Objectif : Diagnostiquer le risque de cancer

Mesure le niveau d'expression de plusieurs milliers de gènes.



Publicité en ligne

Exemples : Google, Critéo, Facebook, ...

Objectif : Mieux cibler la publicité

Analyse du profil de l'utilisateur d'un site internet.

Gestion des stocks

Exemples : Entrepôts Amazon

Objectif : Réduire la latence et les stocks

Les prévisions de ventes enrichissent la gestion des stocks.

Finance

Exemples : Trading Haute Fréquence

Objectif : Anticiper les cours de valeurs boursières

Prédiction des cours à très court terme en fonction des transactions passées.

Sécurité informatique

Exemples : anti-spams

Objectif : Intercepter les emails frauduleux

Evaluation d'un score de suspicion en fonction d'un ensemble de spam connus.

Energie

Exemples : smart-grid

Objectif : Minimiser la production électrique à consommation constante.

Prévoir les pics de consommation électrique pour éviter d'allumer inutilement des centrales.

Transports

Exemples : voiture automatique

Objectif : faire une voiture sans conducteur, avec moins de risques d'accidents

Lutte contre la fraude

Exemples : fraude à l'assurance

Objectif : détecter les fraudes

Détecter les groupes d'individus liés à plusieurs sinistres.

Divertissement

Exemples : recommandations musicales Deezer, Netflix...

Objectif : proposer à un utilisateur des chansons / des films qu'il va aimer

Identifier d'autres utilisateurs ayant des goûts semblables.

Excellence opérationnelle

Exemples : process mining

Objectif : identifier les goulots d'étranglement,
détecter les anomalies

Exploiter les logs de production pour confronter les processus théorique à leur réalisation effective.

Gestion de l'information

Exemples : moteur de recommandation Google

Objectif : accéder facilement aux informations

Indexation de l'information, mesure de pertinence par rapport à une requête, mesure de centralité de ressources liées entre elles

Les familles d'applications

process mining
Décrire

diagnostic médical
Expliquer

gestion des stocks
lutte contre la fraude *publicité en ligne*
Prédire *anti-spams*
trading haute fréquence

smart grid
Prescrire

inspections
moteur de recherche
recommandations musicales

?

voiture automatique

Les familles d'applications

publicité en ligne

Prioriser

inspections

Classifier

anti-spams

lutte contre la fraude

Associer

diagnostic médical

moteur de recherche

Filtrer

recommandations musicales

?

smart grid

process mining

gestion des stocks

trading haute fréquence

voiture automatique

Les familles d'applications

