Machine Learning - Initiation

QCM

Vous proposez à tous vos clients de répondre à un questionnaire de satisfaction en ligne. Quel biais va potentiellement se retrouver dans vos données récoltées ?

- 1. attrition
- 2. auto-sélection
- 3. sélection
- 4. mesure

Dans une base de donnée, vous avez des valeurs manquantes. Donnez une stratégie d'imputation pour une variable catégorielle.

Dans une base de donnée, vous avez des valeurs manquantes. Donnez une stratégie d'imputation pour une variable numérique.

Dans un problème de classification, quelle métrique est suffisante?

- 1. Précision
- 2. Rappel
- 3. Aire sous la courbe ROC
- 4. Aucune

Dans un problème de régression numérique, quel(s) algorithme(s) peut(peuvent)-être utilisé(s)?

- 1. régression polynomiale
- 2. Random Forest
- 3. SVM
- 4. Les 3

Dans un problème de régression logistique, quel(s) algorithme(s) peut(peuvent)-être utilisé(s) ?

- 1. régression polynomiale
- 2. Random Forest
- 3. SVM
- 4. Les 3

Quelle(s) condition(s) nous permet(tent) de nous assurer qu'une descente de gradient trouve les meilleurs paramètres d'un modèle (à un epsilon près) ?

- 1. Le théorème d'approximation universelle
- 2. La surface de l'erreur en fonction des paramètres est convexe
- 3. La variable de sortie est uniformément distribué
- 4. Le pas de gradient est plus petit que epsilon

Trouver l'assertion faussse :

- 1. Un neurone avec l'identité pour fonction d'activation est une régression linéaire
- 2. Un SVM revient à trouver un hyperplan séparateur des données en entrée
- 3. Les paramètres appris d'un SVM sont dépendants de l'initialisation
- 4. Les arbres de décision ont une tendance au surapprentissage

Vous avez un base de données de son. Quelle transformation peut vous permettre d'utiliser un SVM?

Sous quelle(s) condition(s) un dataset de vidéos peut être utilisé par des algorithmes de traitement de séquence ?

Expliquez rapidement comment utiliser les modèles de markov pour faire de la prédiction de classe.