Veille du jour, qui fera suite a votre superbe journée d'hier : les services azure liés à la data et au machine learning :

- lesquels existent

- à quoi servent-ils

- Comment ils marchent en gros.

Veille du jour :

- quelles sont les grandes familles de méthode de ML non supervisés (Revisions)

- Quelles sont les grandes méthodes pour faire du clustering ? (Révisions)

- Quels sont les différentes méthodes pour faire de la sélection de variable ? Quelles méthodes existent sur azure ml studio ?

- Quelles méthodes / outils existent sur azure ML studio pour le traitement de texte ?

- Quelles différences entre Azure ml studio v1 et v2 ? Quelle différence avec Azure ML Service ?

Azure est une plateforme de cloud (serveur à distance) contenants de nombreux produits. Plusieurs sont en Data et ML :

- Cognitives services (luis, custom vision .ai, etc)

- Azure bot services

- bases de donnes

- azure datalake : infrastrcture pour le big data

- Hdinsight et cloudera traitement pour le big data

- Azure ml service : permet d'entraîner ses modèles sur le cloud

- Machine virtuelles : des servuer qu'on pop et on en fait ce qu'on veut

- Réviser les différences entre min-max scaler, Normalizer et standard scaler. Quelle boite dans azure ml studio permet de le faire ?

- Qu'est ce que SMOTE ? Quelle boite dans azure ml studio permet de l'utiliser ? Quelles sont les options de la boite ?

- Révision les métriques d'évaluation de modèles : Mean squared error, root mean squared error. Coefficient de détermination. Accuracy / precision / recall / F1-Score / AUC.

- Quelles boite permet d'afficher l'AUC d'un modèle ?

- Quel est l'intérêt de databricks, notamment pour l'utilisation de Spark. Avec quels langages peut-on l'utiliser ?

-- Quelles sont les principales méthodes pour faire de la selection de feature (feature selection) ?

- Quels sont les boites relatives au NLP ? Comment peut on retirer les stop words avec celles ci dans azure ml studio

- Quels sont les différentes méthodes de clustering disponible dans azure ml studio ?

Notions

- accuracy, precision, recall, f1 score

- matrice de confusion

- Courbe ROC

- l'influence du threshold sur le prédiction à partir de la sortie de la régression logisitique

Veille du matin (speciale DP100):

- Qu'est ce que DICE ?

- Quelles sont les différentes méthodes pour gérer les valeurs manquantes sur azure machine learning studio ?

- Quelle différence entre HDInsight Spark et azure databrick spark

- Qu'est ce que ONNX et NNEF ?

- Qu'est ce que Pytorch et Tensorflow, quelles différences avec Keras ? Qu'est ce que la différentiation automatique?

- Qu'est ce qu'une poisson régression ?

- Comment marche la boite Fine tune hyperparametre search de Azure ML studio

- Quelles boites permet de faire du NLP ? Que permettent-elles de faire ?

- Quelles sont les boites qui permettent de faire de la recommandation de contenu ?

@Apprenant qu'est ce que la Cohen kappa ? A quoi ça sert ?