



Tech4Tchad Machine Learning

Prédire la qualité du vin

Travaux Pratiques

Auteur : Groupe N°1

1. ALLARABAYE Julie
2. LAGRE GABBA BERTRAND
3. MAHAMAT Saleh Yacoub
4. MBAYAM Hippolyte Djegomdé

Plan de travail :

1. Lecture des données
2. Analyse Exploratoire et Visualisation des données
3. . Traitement des Valeurs aberrantes.
4. Entrainement du Modèle
5. Validation et test du modèle

INTRODUCTION

Le machine learning gagne de plus en plus de terrain dans la vie active. Des analyses sont faites pour automatiser certaines prédictions. Ce travail entre dans la vision.

METHODOLOGIE

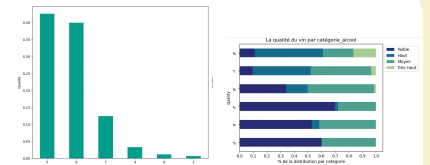
Les données sont téléchargées sur un site dont le lien est en biographie. Nous avons fait la lecture et avons procédé à la visualisation. Cette dernière nous a conduit au traitement de valeurs aberrantes. Le dataset obtenu à la fin a permis de fixer un modèle et de l'entraîner.

PRÉREQUIS :

Il faut installer les modules suivants:
1. Anaconda, jupyter notebook
2. Bibliothèque de langage python

RÉSULTATS

Nous avons divisé le feature quality en des catégories qui sont [5, 6, 7, 4, 8, 3]. En utilisant le modèle Tree Décision, nous obtenons une précision de 0.99 % sur la catégorie 6. La même précision est obtenue avec le modèle Random Forest. Donc nous concluons que les vins de la catégorie 6 sont de bonne qualité.



CONCLUSION

Après entraînement du modèle et après le test, nous pouvons dire que les vins de catégorie 6 sont de bonne qualité. Cela est confirmé par le modèle Random Forest qui s'est montré performant.

BIBLIOGRAPHIE

1. Cours de l'enseignant
2. <https://inside-machinelearning.com/skewness-et-kurtosis/>
3. <https://pandas.pydata.org/docs/reference/api/pandas>
4. <https://www.kaggle.com/learn>

[Click here Github page](#)