



جامعة هواري بومدين للعلوم والتكنولوجيا

Université des Sciences et de la Technologie Houari Boumediene

Faculté de Génie Electrique

Département Télécommunications

Master Réseaux et Télécommunications

TP1 : Initiation et prise en main du langage Python (Modules : NumPy, matplotlib, Pandas, fichiers CSV, etc)

Objectif : Cette première séance de TP opte pour la familiarisation avec le langage de programmation Python. Nous abordons l'installation de Python, l'intégration des librairies, l'importation de modules, et la manipulation des fichiers.

Manipulation

Mettons en pratique ce que nous avons appris jusqu'à présent en cours (Modules : NumPy, matplotlib, Pandas, fichiers CSV, etc) !

Afin de valider cette partie, il vous est demandé d'accomplir les tâches suivantes :

1. Phase de prétraitement.

Après avoir importé l'ensemble de données à l'aide de Pandas, commencez par afficher l'entête de l'ensemble de données, puis quelques informations générales sur les colonnes et les valeurs de données.

Dans l'étape suivante, appliquez ce que vous avez appris dans le cours pour prétraiter vos données : recherchez les informations manquantes et remplacez-les par les valeurs appropriées.

Vérifiez que vos données sont prêtes à être traitées.

2. Phase de visualisation :

Commençons maintenant par la phase de visualisation des données :

Commencez par étudier la distribution des fonctionnalités les plus importantes en fonction de votre compréhension de l'ensemble de données et du problème.

Visualisez la corrélation entre le sexe et l'âge dans un graphique de votre choix. La figure visualisée devrait nous donner des déductions évidentes concernant l'importance de l'âge et du sexe dans la survie des individus.

Choisissez deux autres caractéristiques et étudiez leur impact sur la survie des individus.

Jetez un œil à cette fonction:

```
def plot_correlation_map( df ):
    corr = df.corr()
    s , ax = plt.subplots( figsize =( 12 , 10 ) )
    cmap = sns.diverging_palette( 220 , 10 , as_cmap = True )
    s = sns.heatmap(
        corr,
        cmap = cmap,
        square=True,
        cbar_kws={ 'shrink' : .9 },
        ax=ax,
        annot = True,
        annot_kws = { 'fontsize' : 12 }
    )
```

Essayez d'exécuter la fonction, analysez ce qu'elle fait exactement et ses utilités et enfin, rédigez un paragraphe décrivant votre analyse (le paragraphe doit être en anglais).

Module : Programmation orientée objets en Python
Chargé de cours : M.B. Abidine
Section : RTB₁
2023-2024

Utilisez la fonction `groupby` combinée avec `mean()` pour afficher la relation entre 'Pclass' et 'survived'

Comme dernière étape, essayez de supprimer les colonnes inutiles telles que les noms

En fait, nous avons supprimé la colonne 'Names' car nous l'avons jugée inutile. Eh bien, ce n'est pas le cas car nous pouvons créer une fonctionnalité très utile à partir des noms.

Votre tâche consiste maintenant à créer une nouvelle colonne appelée Titre qui contient le titre approprié pour chaque individu (indice : extrayez le titre de la colonne Names).

Visualisez la corrélation entre le titre et d'autres caractéristiques (par exemple, sex, fare, Age...)

Comme vous l'avez vu, il existe de nombreux titres qui rendent cette fonctionnalité très peu pratique, essayons de regrouper ces titres en titres plus utiles.

Astuce : utilisez le dictionnaire suivant

```
Title_Dictionary = {
    "Capt": "Officer",
    "Col": "Officer",
    "Major": "Officer",
    "Dr": "Officer",
    "Rev": "Officer",
    "Jonkheer": "Royalty",
    "Don": "Royalty",
    "Sir": "Royalty",
    "Lady": "Royalty",
    "the Countess": "Royalty",
    "Dona": "Royalty",
    "Mme": "Miss",
    "Mlle": "Miss",
    "Miss": "Miss",
    "Ms": "Mrs",
    "Mr": "Mrs",
    "Mrs": "Mrs",
    "Master": "Master"
}
```

Visualisons maintenant les mêmes corrélations en utilisant ces nouveaux titres.

Enfin, utilisez les colonnes Parch et SibSp pour créer une fonctionnalité plus utile, appelons-la FamilySize.

Utilisez ce que vous avez appris jusqu'à présent pour déterminer si cette fonctionnalité est utile ou non.

Instructions

Implémentez votre projet dans votre compte GitHub !