Documentation keras\_bitcoin.py

get\_unique\_words\_count : renvoi le nombre de mot unique différents dans un texte

Paramètre :

-text : texte dont on souhaite connaitre le nombre de mots différents

encode\_data : encode les textes d’une Dataframe en one\_hot encoding. Puis modifie les tableaux pour qu’ils soient tous à la même taille.

Renvoie :

- padded\_coded\_sentences : le texte encodé

- vocab\_length : le nombre de mots différent dans les textes

- max\_sentence\_size : la taille en nombre de mots du plus long texte

Paramètres :

-texts : les textes à encoder

get\_train\_test\_data : encode les textes avec la fonction encode\_data, et split la Dataframe en train/test

Paramètre :

- texts : textes à encoder

- labels : labels correspondant aux textes

Renvoie : train\_texts, train\_label, test\_texts, test\_labels, et les variables de la méthode encode\_data

get\_model : entraine et renvoi un modèle de réseaux de neurones Keras

Paramètres :

texts\_train : données d’entrainnement (textes encodées)

labels\_train : labels d’entrainnement

vocab\_length : nombre de mots différents

max\_sentence\_size : taille max des textes

epochs : nombre d’epochs pour le modèle

batch\_size : batch\_size pour le modèle

activations\_functions : liste des fonctions d’activations pour chaque couche de neurones

dropouts : tableau contenant les valeurs pour les couches drop\_out du modèle. Mettre valeurs < 0 pour sauter la couche et ne pas mettre de drop out

get\_data\_to\_predict : renvoie un tableau d’entier correspondant aux textes encodé.

Paramètres :

-vocab\_length : nombre de mots différents (défini lors de l’encodage des données d’entrainement)

- max\_sentence\_size : taille maximale des phrase (défini lors de l’encodage des données d’entrainement)

get\_predictions : renvoi les prediction pour les textes passé en paramètre.

Paramètres :

texts : textes dont on souhaite avoir la prédiction

model : modèle à utiliser pour prédire

-vocab\_length : nombre de mots différents (défini lors de l’encodage des données d’entrainement)

- max\_sentence\_size : taille maximale des phrase (défini lors de l’encodage des données d’entrainement)