

# ESG Finance

## Machine Learning

### TP5

Binary options prediction

baptiste.robertie@gmail.com

21 avril 2022

## 1 Problématique

On se propose de développer un algorithme de trading qui puisse nous conseiller sur la position à prendre (achat ou vente) à un instant  $t$  dans le cadre de trading d'actions binaires : on souhaite prédire si le cours d'un indice va augmenter ou diminuer à  $t + 1$  avec l'objectif donc de maximiser nos gains pendant une période de temps finie.

Calcul de gain  $g_t \in \mathbb{R}$  pour les actions binaires :

- Vous pariez à l'instant  $t$  sur le fait que l'indice aura augmenté à  $t + 1$ . Si à  $t + 1$  la variation  $d_t$  est bien positive (le cours a bien augmenté entre  $t$  et  $t + 1$ ), vous gagnez cette différence de cours (gain positif,  $g_t = d_t$ ). Sinon, vous perdez cette différence de cours (gain négatif,  $g_t = -d_t$ ).
- Vous pariez la semaine  $t$  sur le fait que l'indice aura diminué à la semaine  $t + 1$ . Si à  $t + 1$  la variation est bien négative (le cours a bien diminué entre  $t$  et  $t + 1$ ), vous gagnez cette différence de cours (gain positif,  $g_t = d_t$ ). Sinon, vous perdez cette différence de cours (gain négatif,  $g_t = -d_t$ ).

Notez que votre algorithme peut se retenir de parier. On compte alors trois actions possible à chaque instant  $t$  :

- Pari à la hausse ( $g_t = \pm d_t$ ) ;
- Pari à la baisse ( $g_t = \pm d_t$ ) ;

— Aucun pari ( $g_t = 0$ )

L’objectif est donc de maximiser la somme des gains  $G = \sum_{t \in T} g_t$  sur une période  $T$  imposée. Bien sûr, les données associées à cette période ne pourront évidemment pas être utilisées pour l’entraînement de votre modèle.

## 2 Données

Vous disposez du cours de trois indices financiers (USGG2YR, SPX et VIX), chacun échantillonné à la semaine depuis 1995 (24 ans d’historique, soit près de 1300 semaines). Ces cours vous sont fournis via trois fichiers au format csv. Les cours de ces 3 indices sont tracés en Figure (1).

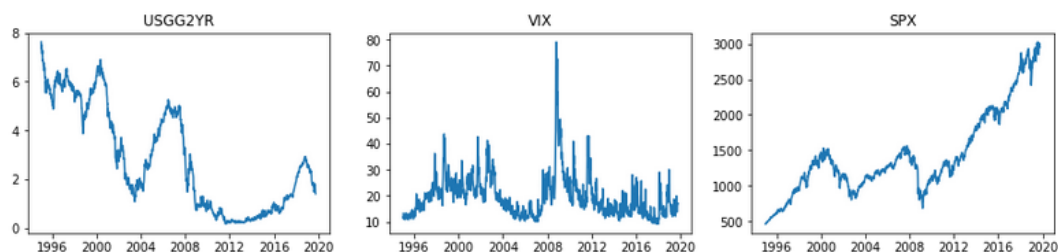


FIGURE 1 – Courbes

Vous utiliserez ces 3 indices pour parier sur le VIX. L’évaluation de votre modèle se fera sur les données publiées à partir de Janvier 2013, c’est à dire sur les 353 dernières semaines du jeu de données. Vous utiliserez alors pour l’apprentissage les données antérieur à 2013, allant de 1995 à Décembre 2012. Vous reporterez le gain total  $G$  de votre algorithme, sur ces 353 semaines, et le comparerez au gain total  $G_{rand}$  consistant à parier de façon totalement aléatoire chaque semaine.

G	$G_{rand}$
30,23 \$	-14,30 \$

TABLE 1 – Exemple de tableau attendu pour l’évaluation de votre modèle

## 3 Evaluation

Vous serez évalué sur votre capacité à formaliser théoriquement la problématique, à justifier vos choix de représentation des données, du modèle et des méthodes d’évaluation.