

# Stratégie de Trading sur les Commodities Cacao

Robin SIODLAK, Margot BORDEL, Nicolas BIGEARD

April 13, 2025

## 1 Introduction

Ce projet vise à développer une stratégie de trading sur les indices de commodities cacao (BCOMCCTR Index) (c'est plus facile pour estimer notre performance totale) en utilisant des données météorologiques pour interpréter les changements de production des pays producteurs de cacao(FAOSTAT). Nous utilisons un réseau de neurones récurrents (RNN) pour générer des signaux d'achat/vente.

## 2 Étape 1 : Prétraitement et Fusion des Données

### 2.1 Importation et Prétraitement des Fichiers CSV

- Nous chargeons les fichiers CSV contenant l'historique des prix du cacao et les données de production annuelle, météorologiques et résultats de tests statistiques.
- Nous vérifions et convertissons la colonne « date » en format datetime.
- Nous traitons les incohérences (séparateurs, virgules dans les valeurs numériques, etc.).

### 2.2 Fusion des Données par Date

- Nous fusionnons les DataFrames sur la colonne « date ».
- Nous vérifions que le DataFrame fusionné contient toutes les informations désirées.

## **3 Étape 2 : Analyse Exploratoire et Ingénierie des Caractéristiques**

### **3.1 Exploration des Données**

- Nous visualisons les séries temporelles : prix du cacao, production annuelle, variables météo.
- Nous calculons et affichons des matrices de corrélation entre ces variables.

### **3.2 Création de Caractéristiques**

- Nous générons des features temporelles (lags, moyennes mobiles, indicateurs de tendance, saisonnalité).
- Nous créons des indicateurs basés sur les résultats statistiques.
- Nous normalisons et mettons à l'échelle les caractéristiques.

## **4 Étape 3 : Construction du Dataset d'Entraînement**

### **4.1 Construction du Dataset pour le RNN**

- Nous découpons la série temporelle fusionnée en séquences.
- Nous définissons les séries d'entrée (features) et la série cible (prix du cacao ou signal d'achat/vente).

### **4.2 Division du Dataset**

- Nous séparons le dataset en ensembles d'entraînement, de validation et de test.

## **5 Étape 4 : Conception et Entraînement du Modèle RNN**

### **5.1 Définition de l'Architecture du RNN**

- Nous choisissons un modèle RNN (LSTM ou GRU).
- Nous définissons le nombre de couches, de neurones, et la taille de la fenêtre d'entrée.
- Nous choisissons une fonction de perte et un optimiseur.

## 5.2 Entraînement du RNN

- Nous entraînons le modèle sur l'ensemble d'entraînement.
- Nous ajustons les hyperparamètres en fonction des performances.

## 6 Étape 5 : Génération de Signaux et Simulation de Trading

### 6.1 Génération de Signaux d'Achat/Vente

- Nous utilisons le modèle entraîné pour prédire les valeurs futures du prix du cacao ou générer des signaux.
- Nous comparons ces prédictions avec les valeurs réelles du marché.

### 6.2 Simulation de Stratégie de Trading

- Nous mettons en place une simulation de trading basée sur les signaux générés.
- Nous évaluons la performance de la stratégie sur l'ensemble de test.

## 7 Étape 6 : Itération et Optimisation

- Nous analysons les erreurs du modèle et identifions les points d'amélioration.
- Nous testons différentes architectures de RNN et affinons le feature engineering.
- Nous effectuons une validation croisée temporelle.

## 8 Stratégie de Trading et Rebalancement d'Indice

### 8.1 Stratégie de Trading

La stratégie de trading que nous effectuons repose sur l'utilisation d'un modèle RNN pour prédire les mouvements futurs des prix du cacao. Les étapes clés que nous suivons sont les suivantes :

- **Génération de Signaux d'Achat/Vente :**
  - Nous utilisons le modèle RNN pour prédire les prix futurs du cacao ou générer directement des signaux d'achat, de vente ou de maintien de position.

- Ces signaux sont basés sur les données historiques des prix, les données météorologiques, et d'autres caractéristiques pertinentes que nous avons intégrées.

- **Prise de Décision :**

- **Achat :** Si le modèle prédit une hausse significative des prix, nous générons un signal d'achat.
- **Vente :** Si le modèle prédit une baisse des prix, nous générons un signal de vente.
- **Neutre :** Si aucune tendance claire n'est prédite, nous maintenons la position.

- **Gestion des Stocks :**

- Nous permettons au modèle de détenir des stocks pour vendre à un meilleur prix ultérieurement, en prenant en compte les coûts de stockage et les risques associés.

## 8.2 Rebalancement d'Indice

Le rebalancement d'indice que nous effectuons permet de s'adapter aux changements saisonniers et aux conditions du marché. Voici comment nous l'intégrons :

- **Périodes de Rebalancement :**

- Nous identifions les dates spécifiques de rebalancement des indices de commodities cacao : Mars (H), Mai (K), Juillet (N), Septembre (U), et Décembre (Z).
- À ces dates, nous évaluons la performance actuelle et décidons s'il est nécessaire de réajuster les positions.

- **Processus de Rebalancement :**

- **Évaluation de la Performance :** Avant chaque date de rebalancement, nous évaluons la performance de la stratégie actuelle, y compris le calcul des profits/pertes et l'analyse des signaux générés.
- **Ajustement des Positions :** Nous ajustons les positions pour optimiser le portefeuille, en vendant des positions sous-performantes ou en achetant des positions prometteuses.
- **Coûts de Rebalancement :** Nous prenons en compte les frais associés au rebalancement, tels que les commissions de transaction et les coûts de gestion.

- **Optimisation Dynamique :**

- Nous utilisons des techniques d’optimisation pour déterminer les meilleures périodes de rebalancement en fonction des données historiques et des prévisions du modèle.
- Nous ajustons dynamiquement la fréquence et l’ampleur des rebalancements en fonction des conditions du marché et des performances passées.

### 8.3 Simulation et Évaluation

Pour tester l’efficacité de cette stratégie, nous effectuons des simulations de trading sur des données historiques :

- **Backtesting :**
  - Nous simulons la stratégie de trading, y compris les rebalancements d’indice, en utilisant des données historiques.
  - Nous évaluons les performances en termes de rendement, de volatilité, et de drawdown.
- **Analyse des Résultats :**
  - Nous analysons les résultats pour identifier les points forts et les faiblesses de la stratégie.
  - Nous ajustons les paramètres du modèle et la fréquence de rebalancement en fonction des résultats obtenus.
- **Validation Croisée Temporelle :**
  - Nous effectuons une validation croisée temporelle pour évaluer la robustesse de la stratégie sur différentes périodes.

## 9 Conclusion

L’approche consiste à fusionner les données historiques avec celles de la production et de la météo, explorer et créer des caractéristiques pertinentes, construire un dataset temporel pour entraîner un RNN, simuler et évaluer une stratégie de trading, et itérer pour améliorer la performance globale.