**Stratégie : Machine Learning Investment Strategy**

La stratégie Machine Learning Investment Strategy repose sur un processus structuré qui comprend plusieurs étapes : la collecte des données, l'ingénierie des features, l'entraînement des modèles, l'évaluation et l'exécution des transactions. Voici une description détaillée du déroulement de cette stratégie :

**1. Collecte et Préparation des Données :**

* + On commence par récupérer les prix historiques des actifs via yfinance, en remontant suffisamment loin pour obtenir un volume de données pertinent.
  + Les données brutes contiennent des prix d’ouverture, de clôture, des volumes échangés et d’autres informations financières.
  + Après récupération, on applique un nettoyage en remplissant les valeurs manquantes et en s’assurant que les index temporels sont bien alignés.

**2. Création des Indicateurs Techniques :**

* + On construit des indicateurs permettant de capturer les tendances des prix :
    - Retours quotidiens (Returns) pour mesurer la variation des prix.
    - Moyenne Mobile 20 jours (SMA\_20) pour identifier les tendances.
    - Indice RSI (Relative Strength Index) pour mesurer la force relative des actifs.
    - Ratio de Volume (Volume\_Ratio) pour analyser les fluctuations de volume.
  + Ces indicateurs sont calculés à partir des prix passés et serviront de variables explicatives pour le modèle.

**3. Construction des Features et de la Variable Cible :**

* + Pour que le modèle puisse prédire efficacement, on crée des features retardées :
    - Par exemple, pour chaque jour, on enregistre les valeurs des indicateurs sur les 6 jours précédents (lookback\_period).
  + La variable cible (Target) est définie comme une classification binaire :
    - 1 si le prix de clôture de J+1 est supérieur à celui de J.
    - 0 si le prix de clôture de J+1 est inférieur ou égal à celui de J.
  + On élimine ensuite les valeurs manquantes et normalise les données pour éviter des biais dus aux écarts d’échelle entre variables.

**4. Séparation des Données pour l'Entraînement et le Test :**

* + On divise l’ensemble des données en deux parties :
    - Données d’entraînement (avant start\_date) : utilisées pour construire le modèle.
    - Données de test (après start\_date jusqu’à end\_date) : utilisées pour valider la stratégie.
  + Une partie des données d'entraînement est également réservée pour la validation interne afin de choisir le meilleur modèle.

**5. Sélection et Entraînement du Modèle :**

* + Plusieurs modèles de Machine Learning sont testés :
    - Random Forest : modèle basé sur des arbres de décision pour capturer des relations complexes.
    - SVM (Support Vector Machine) : algorithme efficace pour la classification non linéaire.
    - KNN (K-Nearest Neighbors) : méthode basée sur la similarité des observations.
    - Gradient Boosting et AdaBoost : méthodes d’ensemble pour renforcer les performances des arbres de décision.
    - Régression Logistique : modèle simple et interprétable pour les prévisions.
  + Chaque modèle est entraîné sur l’échantillon d’entraînement, puis évalué sur un échantillon de validation.
  + Le modèle ayant la meilleure précision est sélectionné pour exécuter les transactions.

**6. Prédiction et Exécution des Transactions :**

* + Une fois le modèle entraîné, il est utilisé pour prédire chaque jour si le prix va monter (1) ou baisser (0).
  + À chaque nouvelle journée après start\_date :
    - Les dernières valeurs des indicateurs sont récupérées.
    - Ces valeurs sont transformées en features retardées comme pendant l’entraînement.
    - Le modèle utilise ces features pour prédire le mouvement du prix pour J+1.
    - Décision de trading :
      * Si le modèle prédit 1, on achète l’actif au prix d’ouverture.
      * Si le modèle prédit 0, on vend l’actif.
      * Le capital est ajusté en fonction du gain ou de la perte réalisée.

**7. Suivi de la Performance et Gestion du Risque :**

* + Tout au long de l’exécution, on suit l’évolution du capital et on mesure des indicateurs de performance :
    - Gain total et performance annualisée.
    - Nombre de transactions effectuées.
    - Métriques de risque comme la volatilité historique et le Value at Risk (VaR).
  + L’objectif est de s’assurer que la stratégie est rentable et qu’elle ne prend pas un risque excessif.

**8. Analyse des Résultats et Ajustements :**

* + Une fois la période d’investissement terminée, on analyse les résultats :
    - Quels modèles ont le mieux fonctionné ?
    - Quels indicateurs étaient les plus pertinents ?
    - Faut-il ajuster les paramètres comme lookback\_period ou le modèle utilisé ?
  + En fonction des conclusions, la stratégie peut être optimisée pour améliorer la robustesse des prévisions.

Ainsi, la stratégie Machine Learning Investment Strategy suit un cycle bien défini, partant de l’entraînement sur des données historiques jusqu’à la mise en œuvre des prédictions pour exécuter des transactions en temps réel. Son efficacité dépend du choix du modèle, de la qualité des données et de la stabilité des patterns exploités.

A diagram of a strategy

AI-generated content may be incorrect.