



ITESO

Universidad Jesuita
de Guadalajara

Deep Learning Microestructura y Sistemas de Trading

Guillermo Aguilar Ochoa
Daniela Aguilar Castaño
Diego Echeverria Villaseñor
Jorge Varela Martinez
Mateo Vereas Dorantes

1. Carga de Datos

En esta sección, se cargan dos conjuntos de datos: `aapl_5m_train.csv` y `aapl_5m_validation.csv`, que representan los datos de entrenamiento y validación respectivamente. Estos datos probablemente contienen información histórica de precios de acciones de Apple en intervalos de 5 minutos.

2. Ingeniería de Características

Se calculan dos medias móviles (MA5 y MA10) para cada conjunto de datos. Estas medias móviles se utilizan como características para predecir si el precio de cierre siguiente será mayor que el precio de cierre actual. Se crea una variable objetivo (Target) que es 1 si el precio de cierre siguiente es mayor y 0 en caso contrario. Luego, se eliminan los valores NA para mantener la integridad de los datos.

3. Modelado

Se implementan cuatro modelos diferentes para la predicción:

Regresión Logística: Un modelo básico para clasificación binaria.

SVC (Support Vector Classifier): Un modelo más complejo que puede capturar relaciones no lineales.

XGBoost: Un modelo basado en árboles de decisión, conocido por su rendimiento en muchas tareas de clasificación.

Red Neuronal MLP (Multi-Layer Perceptron): Se construye y entrena una red neuronal utilizando TensorFlow, con dos capas ocultas.

4. Optimización

Se realiza una optimización de hiperparámetros para el modelo XGBoost utilizando la función minimize de SciPy. Se buscan los mejores valores para gamma y reg_alpha para maximizar la precisión del modelo.

5. Combinación de Modelos

Se combina la salida de los cuatro modelos para tomar una decisión final de compra o venta. Si la mayoría de los modelos predicen un aumento en el precio, se genera una señal de compra; de lo contrario, una señal de venta.

6. Backtesting

Se simula una estrategia de trading utilizando los datos de entrenamiento. Se inicia con un capital inicial y se compra o vende una acción basándose en las señales combinadas de los modelos. Se aplican reglas de stop-loss y take-profit para gestionar el riesgo.

7. Selección y Validación de la Estrategia

Se repite el proceso de backtesting en el conjunto de datos de validación para evaluar el rendimiento de la estrategia en un conjunto de datos no visto anteriormente.

8. Resultados y Conclusiones

Se visualizan los resultados del backtesting en ambos conjuntos de datos para comparar el rendimiento. Se espera que esta visualización muestre cómo cambia el valor del portafolio con el tiempo y proporcione una idea de la efectividad de la estrategia de trading.

Resultados:

Los resultados específicos dependerán de los datos y de cómo se comportaron los modelos en esos datos. Idealmente, se esperaría que la estrategia generara ganancias consistentes o al menos superara un benchmark relevante, como el rendimiento del mercado en general.

Procedimientos:

El procedimiento general involucra la preparación de datos, la construcción y entrenamiento de modelos, la optimización de hiperparámetros, la combinación de señales de diferentes modelos, y la evaluación de la estrategia de trading a través del backtesting. Este enfoque integral permite no solo predecir los movimientos del mercado sino también aplicar estas predicciones en una estrategia de trading simulada.