# Introducción

motivación y objetivos.

Objetivos:

1. Entender el Deep RL y lo necesario para implementarlo.
2. Implementar técnicas del libro de lopez de prado machine learning advances in finance
3. Utilizarlas para diseñar una estrategia de trading sencilla

# Estado del arte

Descripción general sobre el trading, series temporales financieras y sus métricas; el aprendizaje por refuerzo y los tipos de algoritmos, indicadores y métricas de trading.

## Financial time series

Descripción de las propiedades de las series temporales financieras, diferencias entre los retornos simples y los logarítmicos.

Propiedades de cada uno..

Explicar los tipos de datos que se van a utilizar en el proyecto.

## Deep RL

Descripcion general del RL, mítico esquema de agente y mundo.

El objetivo del agente es maximizar su retorno acumulado

### La política

Policy-based

Value-based

### Ecuación de Bellman

### Métodos monte carlo vs métodos de diferencias temporales

Esto aplica en DRL?¿?¿

Backetsting

Mastering backtesting for algo trading quedan los dos últimos capítulos.

Position sizing – vol targeting and meta labeling

Position sizing with machine learning (meta-labeling) 3.35h

Como el espacio de acciones del proyecto va ser discreto, yo me limitaría a utilizar y explicar solo los algoritmos discretos. Por ejemplo, DQN, A2C, PPO y también RecurrentPPO.

# Imbalance bars

Explicación de las imabalance bars del libro

Aplicación de la metodología.

Testeo de la normalidad de los log-retornos obtenidos

Conclusiones de la metodología.

Tabla

Descripción generada automáticamente

### Entorno gymnasium

Explicación de la librería y los métodos más importantes. Explicación de la implementación realizada. (observaciones, estados.. relacionarlos con la teoría del estado del arte)

Explicar las limitaciones del entorno: incluye costes de transación, no incluye liquidez de mercado ni slippage. Pensar en incluir turbulence index (Risk aversion for market crash: the turbulence index is introduced. When

turbulence index is high it indicates sudden sell offs that may yield a stock crash. If

turbulence index surpasses the set threshold all shares are sold and trading is

halted.)

explicación de las funciones de recompensa implementadas.

Differential sharpe ratio.

Tras la explicación del entorno, mostrar resultados de los algoritmos con las distintas funciones de recompensa.

Después mostrar resultados del mejor modelo con la mejor métrica out of saample con pyfolio.

# Meta labeling (position sizing + vol targeting)

### Proceso de Meta-Labeling

1. **Generación de Señales**: El modelo primario genera señales de trading (por ejemplo, comprar o vender un activo).
2. **Asignación de Meta-Labels**: Cada señal generada por el modelo primario se etiqueta como correcta o incorrecta en base al rendimiento observado. Esto implica observar si la señal resultó en una ganancia o pérdida después de un cierto período.
3. **Entrenamiento del Meta-Modelo**: El meta-modelo se entrena utilizando estas etiquetas (meta-labels) y otras características adicionales que podrían influir en la precisión de la señal. Estas características pueden incluir indicadores técnicos, datos de sentimiento, volatilidad, entre otros.
4. **Predicción de la Calidad de Señales**: Una vez entrenado, el meta-modelo se utiliza para predecir la probabilidad de que las futuras señales del modelo primario sean correctas. Sólo se ejecutan las señales con alta probabilidad de éxito según el meta-modelo.

Nuestro primer odelo será el mejor modelo de aprendizaje por refuerzo. Modelo que se utilizará para hacer las predicciones de compra/venta y labelar las señales.

Después,se utilizará un modelo de hidden markov models para sumar indicadores adicionales de regiones, y con ello entrenar un segundo modelo.

A las señales del modelo refinado se les aplicara un position-sizing mediante vol targeting para obtener el resultado final.

# Bibliografía

Stable baselines: <https://stable-baselines3.readthedocs.io/en/master/guide/algos.html>

Pp0-implementation details - https://ppo-details.cleanrl.dev//2021/11/05/ppo-implementation-details/