

ABSTRACT

- Se aplicó Deep Q-Network con una red neuronal convolucional, el cual toma imágenes de gráficas de mercado como entrada para hacer predicciones de mercado a nivel mundial.
- El modelo se entrenó con data del mercado de US y se probó con con data del mercado de 31 países en un periodo de 12 años.
- Resultados:
 - Precios a futuro pueden ser predecidos incluso si el modelo fue entrenado y probado con data de distintos países.
 - El modelo puede ser entrenado en mercados relativamente grandes, y ser probado en mercados más pequeños.
 - IA basada en modelos de predicción de precios del mercado puede ser usada en mercado relativamente pequeños, incluso si estos no cuentan con suficiente cantidad de data para su entrenamiento

INTRODUCCIÓN

- Predecir precios futuros de las acciones del mercado siempre ha sido un tópico controversial debido al “*efficient market hypothesis*”
- Varios tipos estudios se concentraron en demostrar la rentabilidad de los mercados: análisis técnico, descubrimiento de señales no convencionales mediante la web y construcción de IA basada en modelos que utilizan machine learning.

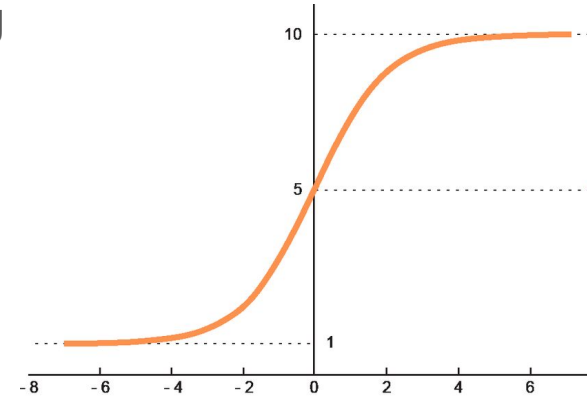
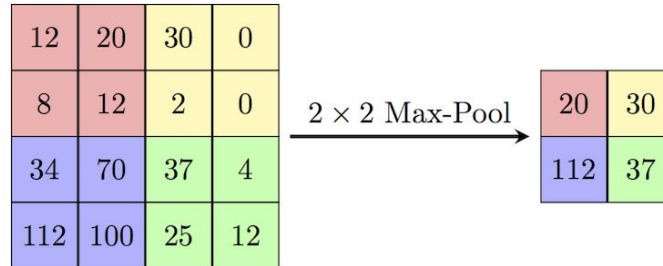
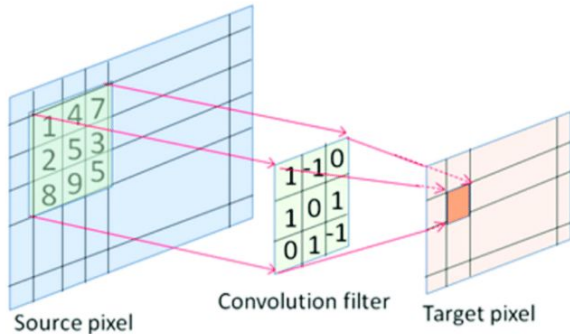
INTRODUCCIÓN

- Redes neuronales y aprendizaje reforzado son de los métodos de machine learning que más se utilizan actualmente.
- Se utilizó el framework Deep-Q-Network en vez del aprendizaje supervisado tradicional para entrenar el modelo. Como se trabaja con predicciones en los precios de acciones, el asignar valores binarios no es suficiente.

BACKGROUND

- Red neuronal convolucional (CNN):

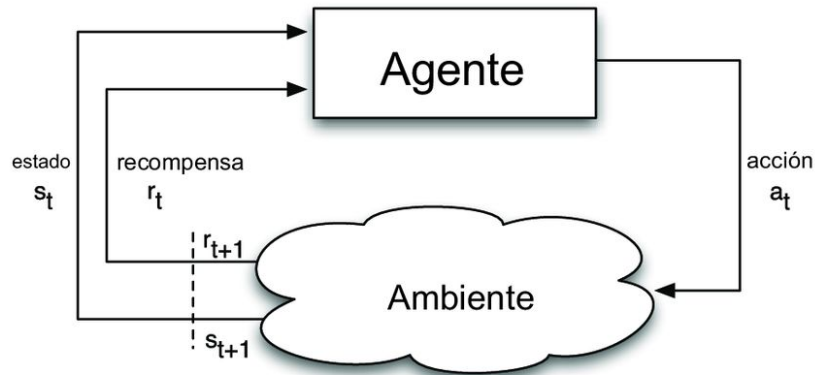
Usualmente toma imágenes 2D con 3 canales de colores como entrada. Se pasa por múltiples *capas escondidas*, consistiendo de capas convolucionales seguidas por capas de *nonlinearity* y de *pooling*. Las últimas capas escondidas están totalmente conectada y son usadas para funciones softmax. En este caso se utiliza como una función de aproximación para Q-learning



BACKGROUND

- Q-learning:

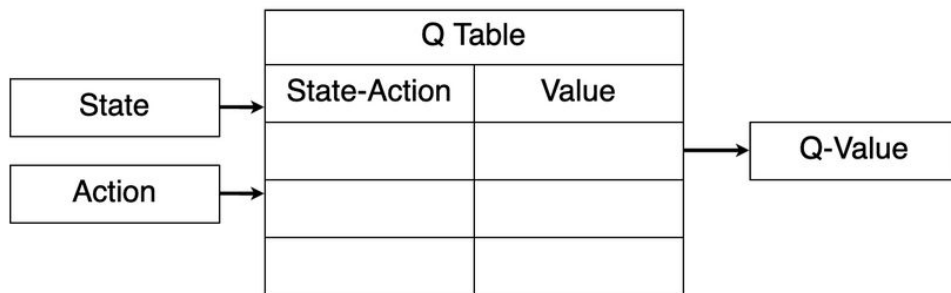
Algoritmo de aprendizaje basado en valores y se centra en la optimización de la función de valor según el entorno o el problema. La Q en el Q-learning representa la calidad con la que el modelo encuentra su próxima acción mejorando la calidad. El proceso puede ser automático y sencillo. El modelo almacena todos los valores en una tabla, que es la Tabla Q.



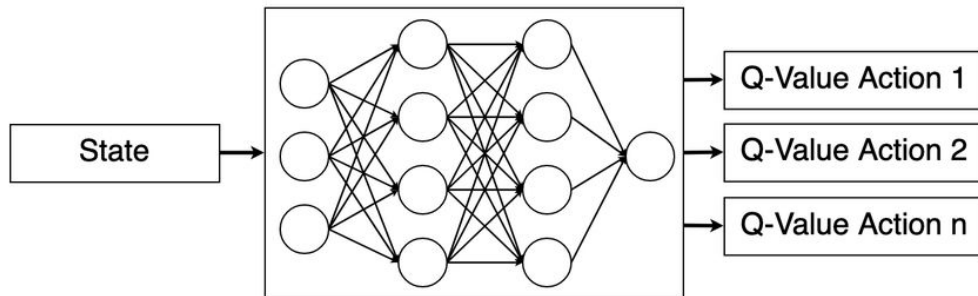
1. El agente “percibe” el ambiente y obtiene un estado.
2. En base al estado actual, el agente toma una acción en base a las estrategia del ambiente.
3. El agente obtiene una *reward* del ambiente y actualiza su estrategia.
4. Después de actuar, el ambiente se actualiza y se obtiene un sig. estado.
5. Se repite 1-4

BACKGROUND

Q-Learning



Deep Q-Learning



MÉTODO

Se muestra como la red neuronal lee la entrada y da de salida valores de acción para compañías individuales. El término “valor de la acción” se refiere a la recompensa acumulativa esperada de una acción.

Muestra La arquitectura de la red neuronal , en el ejemplo se ilustra una imagen de una gráfica de $W \times W$ en un tiempo T y $T+1$.

Por ejemplo, si $W=8$, como se muestra en la figura, la red la lee de entrada como una matriz de 8×8 , con todos los elementos llenados con un 0 o un 1, una sola columna en la matriz representa un solo día.

Los elementos en negro representan un 1 y los en blanco un 0, la parte superior de la matriz representa el valor relativo del valor de cierre y la media baja representa el valor relativo del volumen. Las dos líneas en medio de la grafica estan vacias para ayudar a la red a distinguir entre valor y volumen.

Arquitectura

- Toma imágenes de $32 \times 32 \times 1$ como entrada.
- Cuenta con 6 capas ocultas ($H=6$), las primeras 4 son capas convolucionales seguidas por una capa ReLu, y las últimas dos son capas totalmente conectadas.
- Después de la 2da y 4ta capa se encuentra una capa de max pooling con un filtro de 2×2 .
- Las últimas 2 capas son totalmente conectadas una de 2048×32 y otra de 32×3 .

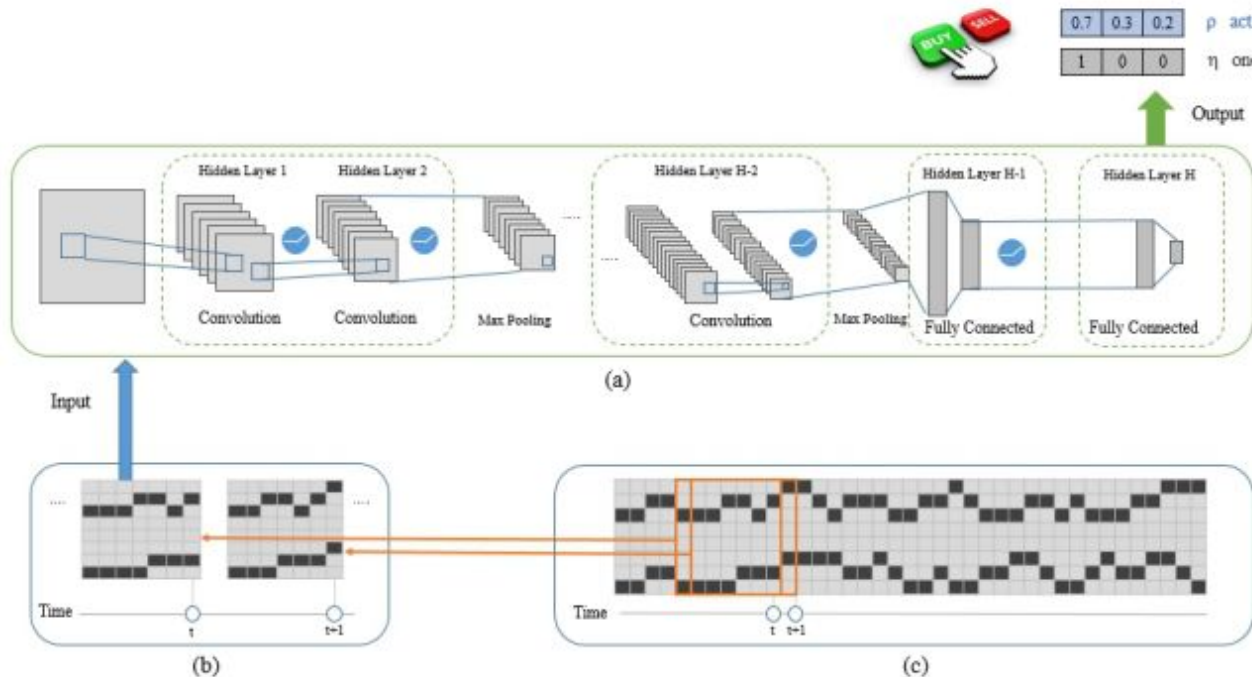


FIGURE 1. Overview of how our CNN reads an input chart of a single company at a specific time point (time t) and outputs the two vectors ρ and η .

DESCRIPCIÓN DE LA DATA

- **Com#:** Numero de compañías en cada país.
- **N y NL:** Número de compañías que se utilizaron para el experimento
- **Data#:** Número total de data utilizado.
- **Avg:** Promedio diario en retorno (en porcentaje) de la compra y sostener portafolios de un periodo dado.
- **STD:** Desviación estándar de los retornos diarios.
- **Excess Rate:** Porcentaje de data con un valor absoluto del escalar.

Proceso de entrenamiento

El entrenamiento consiste en $N \times T$ data points, donde $N = 1500$ aprox y $T = 1000$ (días hábiles en 4 años, el cual será el 80% del dataset para el entrenamiento).

TABLE 2. List of hyperparameters mentioned in the paper. Hyperparameter optimization is done using 20% of the training set.

| Hyperparameter | Description | Value |
|------------------|----------------------------------------------|-----------|
| $maxiter$ | The maximum number of iterations | 5,000,000 |
| $learning\ rate$ | The learning rate used by Adam optimizer | 0.00001 |
| ϵ_m | The minimum value of ϵ | 0.1 |
| W | Horizontal and vertical size of input matrix | 32 |
| M | The capacity of the memory buffer | 1,000 |
| B | The update interval of parameters θ | 10 |
| C | The update interval of parameters θ^* | 1,000 |
| P | The transaction penalty while training | 0.05 |
| γ | The discount factor | 0.99 |
| β | The batch size | 32 |