FinTech Final Project Report

學號：R07922142 系級：資工所碩一 姓名:歐政鷹

在這次的Final Project中，我利用了課堂中學到的兩種指標：均線( Moving Average, MA )及RSI( Relative strength indicator ) 並利用Dynamic Programming及logistic regression模型，來作為我的預測模型中Baseline建立的方法。

我利用Dynamic Programming對Daily OHLCV中的Close price進行分析，找出現有資料中每天的最佳action，然後再利用每天當天過去14天的資料，透過logistic regression來預測出當天的action。由於Dynamic Programming是會用到未來的資料來作預測，因此以logistic regression所做出來的結果當然與Dynamic Programming的結果有著一定的差距，圖1顯示出logistic regression預測結果與baseline作比較的情況。所以我以以上的結果作為baseline，希望能透過改善及優化logistic regression model來得到更好的預測結果。

|  |
| --- |
|  |
| 圖1. logistic regression訓練曲線 |

在logistic regression的優化方面，我先以加入更多Feature作考慮，一方面可以觀察加入不同的feature是否會有更好的結果，另一方面可以讓之後在作feature selection時有更多的選擇。

我希望可以加入一些具判別性的feature，因此我以MA及RSI作為我的選擇。我先以所有資料為MA及RSI找出合適的長短期Window size，alpha及beta值，在尋找的過程中，為了避免出現overfitting的情況，所以回報率的多寡並不是尋找的目標，而是志在尋找出具有回報的參數組合。

在尋找出合適的參數組合後，我把其MA及RSI結果加入到feature中，再利用logistic regression作訓練。在得到初步的訓練結果後，再作features selection以找出最佳的features組合為目標。而在得到最佳的features組合後，我再透過ensemble，一方面可以提升model預測準確度，一方面可以減少單一model出現overfitting的情況。在ensemble的方法裡，不同的model是透過MA及RSI不同的參數組合訓練而成，不同的參數組合都是以可以得到正回報的前提下找出，因此我認為透過ensemble的方式可以讓model有更好的結果，圖2顯示出在經過features selection及ensemble後，logistic regression預測結果與baseline作比較的情況，可明顯看出結果有顯著的改善。

|  |
| --- |
|  |
| 圖2. logistic regression訓練曲線 |

最後，我利用上述mode來作為我final project的結果，希望能取得好的結果。

在我的model中，我是透過每天的action作為model好壞的比較依據，而沒有考慮到我的action最終得到的回報率的好壞。雖然不確定以考慮回報率的方式是否會有更好的結果，但我認為這是將來model要繼續改善時一個可考慮的方向，說不定可以讓model得到更好的結果。