

專題題目：高頻交易深度學習平台加速器實作

指導老師：馬席彬教授

組別：A42

組員姓名：呂依凡

Abstract

隨著科技的發展，電腦的計算速度愈來愈快，隨著電腦計算速度的提升，深度學習的相關應用也愈來愈廣泛，高頻交易就是其中一個例子。利用深度學習做為高頻交易主要的演算法，主要是因為可以影響股票市場的因素過多也過於複雜的關係，深度學習可以在複雜的股票市場中找到資料間的關係。

因為高頻交易是一場比速度的競賽，若只用通用處理器來計算決策會太慢，所以需要利用硬體做加速，本次專題是利用FPGA加速。因為受限於FPGA的記憶體容量的關係，所以不能只用FPGA來完成整個高頻交易的過程，需要另一台電腦進行較不花時間的運算，並將佔用絕大多數時間的計算傳給FPGA加速。Vitis整合了C和C++，可以讓軟體工程師也享受到硬體快速的優點，這個專題是要利用FPGA和Vitis加速機器學習的過程，藉此熟悉Vitis的介面。

Implementation

先利用python來驗證KNN分類Fashion-MNIST的結果是正確的，再寫成C++的形式，再改成給硬體加速的版本，圖（一）是硬體加速的結構。Fashion-MNIST的每一筆資料都是 $28 * 28$ 個pixels的照片再加上這筆資料所屬的類別，將待測圖分類時，需要一套比較的程序，首先將每一個pixel都當作一個特徵，KNN會藉由比對和training data的特徵相似度，來將待測資料歸類，比較的方式是利用待測物每一個pixel的與每一筆training data的每一個pixel取差的平方和，藉由這個計算方式算出待測資料與每一筆training data的距離，並取出距離最近的若干個training data的labels並進行多數決。本份報告中，KNN的算法是利用19個鄰近的training data進行多數決。



圖（一）

Result

到目前為止的進度而言，有成果的部分只做到利用python驗證KNN的結果，從圖（二）可以看到KNN所計算出來的準確度是96%，在10000筆資料中，有368筆待測資料被分錯，此方法可以辨識待測圖片到達96%的精準度，而執行KNN演算法所花的CPU Time為17分30秒，之後會繼續做完。

```
Misclassified instances: 368
Accuracy: 0.96
CPU times: user 17min 30s, sys: 173 ms, total: 17min 30s
Wall time: 17min 32s
```

圖(二)

Conclusion

高頻交易是深度學習的一個應用面，高頻交易需要根據快速變動的市場進行決策，並且得比別人做的決策還要快，才可以搶得先機，而因為深度學習是個非線性的模型，所以比其他方式更能適應多變的市場，並在眾多資料裡找出關聯性。

硬體的優勢就是可以算的比一般軟體的程式要快速，而Vitis也提供了介面，讓只會寫軟體程式的工程師也可以享受到硬體的優勢，但是受限於FPGA板本身的運算速度和資料傳遞速度的關係，需要讓可以控制運算及資料傳遞的主程式進行良好的資料傳遞分配，也因為如此，所以需要好好了解FPGA板的特性和整個硬體加速的架構。

在用python驗證KNN的時候，得到了96%的準確度，在10000筆資料中只有368筆資料辨識錯誤，可知KNN很適合拿來辨識Fashion-MNIST的圖。