Multi-Task Learning for Stock Selection

Kunj Golwala
golwala.1@iitj.ac.in

Indian Institute of Technology Jodhpur,
Jodhpur,
Rajasthan,India

Artificial Neural Network (ANN) has been widely used previously for the prediction of financial markets since it incorporates technical analysis and the presence of non-linearities in the market only further justify the use of ANNs. These models can incorporate various types of explanatory variables: technical variables ,micro-economic stock-specific variables, and macro-economic variables. A series of experiments were performed on 9 years of data of 35 large capitalization companies of the Toronto Stock Exchange (TSE) to determine whether one should use the same network for all the stocks or to use different networks for each stock.

The learner takes (x,y) input-desired value pairs.A training criterion C is defined which is a function of desired values, y and outputs of the function f(x). The function f is parameterized by the parameters of the network and belongs to a Hypothesis space, H. The generalization in these models due to variance can be controlled by making the hypothesis space, H, as small as possible.Several neural networks are trained simultaneously,which helps in improving generalization performance.In multi-task learning,some or all hidden layers of neural networks are shared.This idea of multi-task learning is applied for stock selection and portfolio management.

For our experiment, we consider a universe of 35 risky assets (35 stocks from the TSE) and one risk-free asset (90-days Canadian treasury bills). The data is recorded every month for 9 years. Each month one can buy or sell some of their assets in a way so as to distribute their current worth between these assets. Five explanatory variables are chosen. Multiple experiments were carried out with different sets of initial weights.The training algorithm is based on the optimization of the neural network parameters with respect to a financial criterion. In the experiments, the ANN was trained for 120 epochs.Four sets of experiments with two different types of architectures, namely - 5-3-1 (5 inputs , a hidden layer of 3 units and 1 output) and 5-3-2-1 (5 inputs,first hidden layer of 3 units, second hidden layer of 2 units and 1 output). were performed.Four types of parameter sharing are done- sharing everything,sharing only the parameters of the first hidden layers, sharing only the output layer parameters, and not sharing anything.

From the observation tables and plots of the results,

Table 1: Comparative results for the 5-3-1 architecture: four types of sharing are compared with the buy-and-hold benchmark (see text).

	buy & hold	share all	share hidden	share output	no sharing
Average yearly return	8.3%	13%	23.4%	24.8%	22.8%
Standard deviation (monthly)	3.5%	4.3%	5.3%	5.3%	5.2%
Beta	1	1.07	1.30	1.26	1.26
Alpha (yearly)	0	9%	20.6%	21.8%	19.9%
t-statistic for alpha = 0	NA	11	14.9	15	14
Reward to variability	0.9%	9.6%	22.9%	24.7%	22.3%
Excess return above benchmark	0	4.7%	15.1%	16.4%	14.5%
Maximum drawdown	15.7%	13.3%	13.4%	13.3%	13.3%

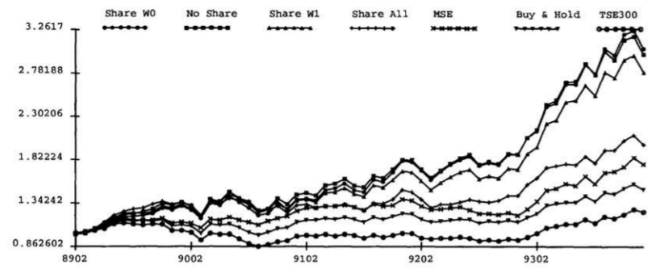

Figure 1: Evolution of total worth in the 5-year test period 02/89-01/94, for the 5-3-1 architecture, and different types of sharing. From top to bottom: sharing the hidden layer, no sharing across stocks, sharing the output layer, sharing everything, sharing everything with MSE training, Buy and Hold benchmark, TSE300 benchmark.

we obtain three main conclusions - 1)The returns from the stocks can be improved by not sharing some parameters (as observed in the case of 5-3-1 architecture.
2)Sharing some parameters yields better results than sharing no parameters at all.
3)The performance obtained using this model yields great returns (more than 14 percent yearly returns).In general, we conclude that multi-task learning is highly efficient for stock selection.