Applied Mathematics & Information Sciences
An International Journal

An Optimal Strategy Model for Purchasing Initial Public Offerings in China

Dong Huang[1], *Xiaolong Wang*[1], *Jingjing Ma*[1] *and Ronggang Dou*[1]

[1]Computer Science and Technology Department, Harbin Institute of Technology Shenzhen graduate School, Shenzhen, Guangdong, China

Received: 15 Jul. 2012, Revised: 13 Oct. 2012, Accepted: 28 Oct. 2012
Published online: 1 May 2013

Abstract: An optimal recommend model for purchasing Chinese IPOs (Initial Public Offerings) based on artificial intelligence method is proposed here. In this paper, we focus on how to get the optimal income in a period of time. In order to avoid the problem of sparsity by separate method, combinations of classification, regression and maximum entropy are adopted to adjust the density of the filters in Double-layer filter model we propose. We improve the accuracy of the forecasting by structuring data in preprocessing, adjusting filters' density level by the entropy of maximum entropy, the probability of the classification and regression algorithm. The experimental results show that the strategy model we propose achieves better forecasting accuracy than the methods separately. Therefore, the Double-layer filter model not only can forecast precisely but also has much application value.

Keywords: Optimal strategy model, Initial Public Offerings, classification, regression, maximum entropy

1. Introduction

For the trading premium of the IPOs (Initial Public Offerings) on the first trading day, the purchasing of IPOs is always popular. People use many kinds of models and algorithms to predict the price although the price is uncertain [1]. BP Neural Networks [2] is one of the most popular tools in stock prediction, but there are problems with the noisy data and local optimum. Regression [3] is widely used in price prediction, but which factors should be taken and the factors should be adopted which expression are just speculate, so people use SVM (support vector machine) [4] to improve the precision and avoid the local optimum of neural networks. Each method has its limitation more or less. The performance of maximum entropy [5] in stock prediction is also not as well as anticipated because of the data sparse and the dependence of training data.

People forecast [6] the price of IPOs in order to get more revenue, but we recommend the IPOs for the investors in a period of time, make them get the income by sell-off on the first day of listing, then, add the income to the initial capital to take cycle investment on IPOs, so they can get optimal income and the income they get will be more than we predict the price directly, because the overall optimal is more practicable than the local optimal for investors.

We propose a new optimal model [7] for IPOs in China, which combined classification, regression and maximum entropy methods to avoid the disadvantage of the separate method. By the model, the IPOs we recommended arranged in an optimal path by time. The experiments show that the path of trend and the income we predict is very close to the ex post optimal path, and the model is used in our financial service platform for free now.

The rest of the paper is organized as follows. In Section 2 we present our optimal strategy model. The implementations of methods in the model are analyzed in Section 3. The comparative analysis is deferred to Section 4. In section 5 we present the experiments and results. We conclude our paper and suggest some possible future directions in Section 6.

2. Optimal Strategy Model for IPOs

Maximize the income is the goal of all investors, so getting the optimal income in a period of time is more important than in once, especially for IPOs. Freezing

[*] Corresponding author e-mail: donghuang2010@gmail.com

funds and uncertainty of listing date make the prediction difficult. The national policy and economic conditions also can affect the IPOs which issued but unlisted. For the reasons above, we propose Double-layer filter model as shown in Figure 1. The x-axis is IPOs issued time, and y-axis is the income we predict. IPOs are in the forms of ● ⬟ ☆.

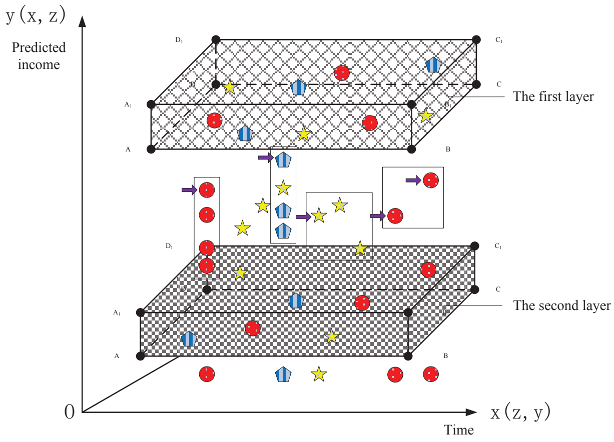


Figure 1 Double-layer filter model

The IPOs distribute according to the time sequence in the model. First, we preprocess IPOs issue announcement, extract features and make the data structured, and then filter primarily by the first layer. There are four kinds of forms arranged by the rule of purchase, which are shown in Figure 2.

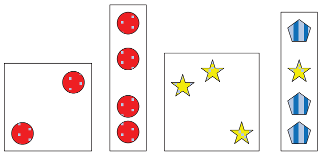


Figure 2 IPOs arrangement forms

①There are more than two days between the two IPOs issue.
②There are many IPOs are issued in one day.
③The IPOs are continuous in the same purchase cycle.
④It contains ② and ③ forms at the same time.

Second, we filter the IPOs that out of the first layer by the second layer which is denser than the first layer, and then get the optimal purchase path.

In Double-layer filter model, each layer use one method, so it is the combination of two methods from classification [8,9], regression [10,11] and maximum entropy [12].The important problem is when people predict the price, in order to get more accurate result, the adopted methods is always set more stringent in the layers, instead the result is not as good as expected. So in the model we proposed, the double-layer filter can overcome the disadvantage of separate method.

From the model we can see that we pick up the IPOs from the sequence that the arrows point to. For the third form in Figure 3 that the IPOs are continuous in the same purchase cycle, we do not choose the green star which we predict can get better income in first layer, but we take the yellow one arrow point to, because it performed better in the second layer and that consistent with the global optimum in cycle investment.

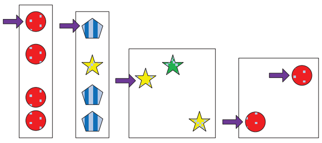


Figure 3 Optimal strategy for IPOs

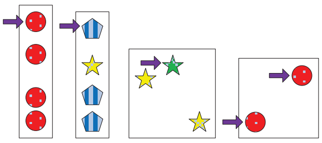


Figure 4 Comparison of strategy

If we choose the green star, the optimal road will be changed as is shown in Figure 4 because of freezing funds or insufficient balance. The income of IPOs (● + ⬟ + ☆ + ● + ●) in Figure 3 is more than the income of IPOs (● + ⬟ + ★ + ●) in Figure 4. In the model, we select the IPOs not only by considering the income we predict but also by the rules of IPOs subscription. So, we can get the global optimal strategy in a period of time as shown in Figure 3.

3. The implementation of methods in the model

Before applying the Double-layer filter model, we extract features from the IPOs issuance announcement in the form of PDF documents first, and then preprocess the data by K-means, clustering and χ^2.

3.1. Data preprocessing

We get the latest announcement by timing web crawler from financial official websites, then parse PDF document into text and extract features, thus converting the heterogeneous information into structured data. We select features by three algorithms: calculating the optimal weight by IIS (Improved Iterative Scaling) algorithm [13], considering the distance from the features to the center of the clusters after the clustering for features, and according to the rank of the probability by χ^2.

3.1.1. Weight of features

We calculate the optimal weight of each feature. The initial data: the empirical distribution is $\widetilde{p}$, the initial model is s_1, and features sequence is $f_1, f_2, ...f_n$. Additional condition

$$f_\gamma(g) = \sum_{i=0}^{n} f_i(g) \tag{1}$$

In Equation (1) $g \in G$ (training sample set). Implementation steps as follow:

i.Set initial iterations of the model $s^{(1)} = s_1$.
ii.$s^{(k)}\left[f_i e^{\lambda_i^{(k)} f_\gamma}\right] = \widetilde{p}[f_i]$, each value of i is unique for $\lambda_i^{(k)} \in (-\infty, \infty)$.
iii.$s^{(k+1)} = \lambda^{(k)} \circ s^{(k)}, k = k+1$.
iv.If $s^{(k)}$ is convergent, set $s^{(*)} = s^{(k)}$, then terminate the loop. Otherwise go to step ii.

After parsing the IPOs issuance announcement into text, we extract features from these documents by IIS algorithm.

3.1.2. Clustering

We define d_f is the distance from feature object f_i, which is belonged to the subclass to the center of the subclass cluster, and the center of the subclass is the average value of the whole subclass. So we calculate the minimum value of $\sum_i d_f$ cyclically by K-mean, until the function is convergent.

$$E_{obj} = \sum_{i=1}^{k} \sum_{O \in C_j} \left|O - avg_j\right|^2 \tag{2}$$

E_{obj} is the sum of error squares of all objects in the dataset, k is the number of clusters, C_j is the i-th cluster, O is the given object in the set, avg_j is the average value of C_j.

The K-means algorithm is as follow:

Input: the dataset include of N object (feature f_i), and the number of clusters you want to divide ($k < n$).

Output: k clusters.

i.Select k objects randomly as the default center from data set D.
ii.Clustering according to the average value of each object.
iii.Calculate the average value of each object again in the cluster.
iv.Repeat step ii and iii until no further changes in the clusters.

3.1.3. Select features

After calculating the weight of each feature by IIS algorithm and clustering for features, we get the dependence degree between f_i and C_j. We can calculate the valuation of f_i for C_j as follow:

$$x^2(f_i, cj) = \frac{\mu * (\mu_1 * \mu_4 - \mu_2 * \mu_3)^2}{(\mu_1 + \mu_2) * (\mu_3 + \mu_4) * (\mu_1 + \mu_3) * (\mu_2 + \mu_4)} \tag{3}$$

In Equation (3), $\mu = \sum_{i=1}^{4} \mu_i (n_i$ is the frequency between f_i and C_j). We can simplify the definition as:

$$x^2(f_i, cj) = \frac{(\mu_1 * \mu_4 - \mu_2 * \mu_3)^2}{(\mu_1 + \mu_2) * (\mu_3 + \mu_4)} \tag{4}$$

Since we want to select the features which can get the minimum value of $\sum_i d_f$, so the weight coefficient of $\mu_1 * \mu_4$ is bigger than that $\mu_1 * \mu_4$ for C_j.

$$\theta - x^2 = \theta * \frac{(\mu_1 * \mu_4 - \mu_2 * \mu_3)^2}{(\mu_1 + \mu_2) * (\mu_3 + \mu_4)} \tag{5}$$

$$\theta = \begin{cases} 1 & \text{if } (\mu_1 * \mu_4 - \mu_2 * \mu_3) > 0 \\ -1 & \text{if } (\mu_1 * \mu_4 - \mu_2 * \mu_3) \leq 0 \end{cases} \tag{6}$$

We only select the features that the value of $\theta - x^2$ is positive for C_j.

3.2. The methods of filters

In Double-layer filter model, we choose two methods from classification, regression, and maximum entropy as the filters. Different combination has different performance because of the methods' order and setting in the model.

(1) *Select the maximum entropy from distribution of constraint sets*

We set the categories finite set of predict event as Y, for each $y \in Y$ is affected and constrained by event X, the events $x \in X$ is known, the conditional probability of $y \in Y$ is $p(y|x)$ [14]. Training sample set is $G = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, feature set is $f = (f_1, f_2, ..., f_k)$, and the functional relationship $f_i(x, y)$ between feature f and x can be expressed by binary function as:

$$f_i(x,y) = \begin{cases} 1 & \text{if x,y satisfied the condition} \\ 0 & \text{if not} \end{cases} \tag{7}$$

Constraint equation is

$$\sum_{x,y} \widetilde{p}(x)p(x \mid y)f(x,y) = \sum_{x,y} p(x)p(x \mid y)f(x,y) \tag{8}$$

$\widetilde{p}(x)$ is the conditional probability in Equation (8), the set of constraints is

$$M = \{p \in P \mid p(x) = \widetilde{p}(x)\} \tag{9}$$

The constraint condition that all distributions should satisfied is

$$P = \{P \mid E_p f_i = E_p f_j, 1 \le i \le j\} \tag{10}$$

The model with maximum entropy is

$$p^* = \underset{p \in P}{\operatorname{argmax}} \sum_{x,y} p(y \mid x)\widetilde{p}(x) \log \frac{1}{p(y \mid x)} \tag{11}$$

We set the weight coefficient of feature f_i to λ_i, and then the exponential form of conditional probability that satisfied the maximum entropy is

$$p_\lambda(y \mid x) = \frac{1}{Z_\lambda} exp\left(\sum_i \lambda_i f_i(x,y)\right) \tag{12}$$

The normalized factor is

$$Z_\lambda = \sum_y exp\left(\sum \lambda_i f_i(x,y)\right) \tag{13}$$

(2) *Logistic Model Tree classification method*

LMT (Logistic Model Tree) is the combination of tree structure and logistic regression models, each leaf node is a logistic regression model. The accuracy of LMT is better than a single decision tree and logistic regression methods. Decision tree is constituted by N non-leaf nodes and T leaf nodes in the LMT.

S is the set of k gain data sample, n classification $C_i(i = 1,2,...,n)$, k_i is the number of sample in C_i, and then expected value of S is

$$I(i = 1,2,...,n) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{14}$$

$p_i = \frac{k_i}{k}$ in Equation(14).

There are q different features $f_i i = 1,2,...,q$, we divided IPOs object set S into disjoint subset $S_1, S_2, ..., S_q$ according to f_i, and each subset in the tree is represented by a leaf node. For each $t \in T$, logistic regression function is L_t, rather than a simple classification identification. The probability of the class members for each subset in the entire attribute set F by regression function L_t is

$$P_\gamma(G = j \mid X = x) = \frac{e^{I_j(x)}}{\sum_{k=1}^{T} e^{I_k(x)}} \tag{15}$$

x is the property value vector of an instance, and $I_j(x)$ is the function of the input parameters.

$$I_j(x) = a_0^j + \sum_{f \in f_t} \alpha_f^j \cdot f \tag{16}$$

If $f \notin f_t$, then $\alpha_f^j = 0$. We use LMT represent the model as

$$L(X) = \sum_{t \in T} L_t(x) \cdot I(x \in S_t) \tag{17}$$

If $x \in S_t$, the value of $I(x \in S_t)$ is 1, otherwise 0.

LMT classification process as follows:

i. Modeling and generating a new instance using the original data set.

ii. Create an original Logistic regression model with all the original data.

iii. There are two kinds of rules to determine how to divide branch for a property. For discrete properties, LMT generates a branch for each possible value. For numeric attribute, it generates two branches, for example, $X \le \alpha$ and $X > \alpha$. In order to determine the threshold α, we sort the instance set according to the value of X, and get an ordered instance set m, the set can be divided into two ordered subsets.

iv. LMT selects the best branch from candidate branch according to information gain value.

Set k_{ij} is the sample number of classification C_i in S_j, the entropy of subset divided by f_i is

$$E(A) = \sum_{j=1}^{q} \frac{s_{1j} + ... + s_{nj} I(s_{1j}, ...s_{nj})}{s} \tag{18}$$

The information gain ratio obtained by A branch is

$$GainRatio(A) = \frac{I(S_1, S_2, ..., S_m) - E(A)}{-\sum_{j=1}^{q} p_j \log_2 p_{(j)}} \tag{19}$$

(3) *Regression for income of IPOs*

We establish regression model [15] by the distribution of f (the features of IPOs) and corresponding variable r (the income of IPOs).

$$r = w_0 + w_1 f \tag{20}$$

We define the set that consists of prediction variable f and the corresponding response variable r is Q. The data pairs $(f_1, r_1), (f_2, r_2), ..., (f_{|Q|}, r_{|Q|})$ are in set Q. We use the following expression to estimate the regression coefficients:

$$w_1 = \frac{\sum_{i=1}^{|Q|}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{|Q|}(x_i - \bar{x})^2} \tag{21}$$

$$w_0 = \bar{y} - w_1\bar{x} \tag{22}$$

So the model we proposed performs better than the separate method and can get better income. That is not just simple combination of methods, because the combinations above can get better prediction by learning from each other and making up the disadvantage.

4. Comparative analysis

In order to evaluate the performance of the model we proposed, we get the ex post optimal income by Dijkstra algorithm and the average income.

(1)After the listing of IPO, we can get the optimal income by the optimal path, so the problem can be transformed into the SPP (Shortest Path Problem). We use the classic Dijkstra algorithm to get the ex post optimal income.

(2)The average income We calculate the average income by distributed subscription funds equally in one purchase cycle.

$$\bar{\Delta} S = \sum_{i=1}^{n}\sum_{j=1}^{k} F_{avg_i}\{\frac{S_P}{N_j}\} \tag{23}$$

The function F is to get the income by distributing the current subscription funds equally to the IPOs.

We can evaluate the income by the approaching ratio close to ex post optimal income CM_{LimR}:

$$CM_{LimR} = \frac{R_{act}}{R_{max} - R_{avg}} \tag{24}$$

The CM_{LimR} expresses the degree of predicted income close to the ex post optimal income based on average income.

5. Experiments and results

In our experiments, we consider the issuance, company's profitability, and overall market condition, so we extract 47 initial features including circulation ratio, current ratio, Price-earnings ratio, the market gain in the day before purchase day and so on.

5.1. Experiments

The five comparative tests consist of two parts: the prediction income of IPOs and the comparison with ex post optimal income and average income. The experiments data are 768 IPOs from June 29, 2009 to March 9, 2012 in China. We select 100 IPOs as test data by continuous time series. The calculation of income in each test according to the initial subscription funds is one million.

The feature f_i is in the form of numeric, so the range that the value can take is too wide for each feature. We clustering the features by K-means into two clusters C_1 and C_2. For each IPOs, it can be classified into Purchase (P) or Not Purchase $(\bar{P})$. We establish a matrix for $f_i(i = 1,2,...,n)$ as shown in Figure 5.

	C_1	C_2
P	a_{11}	a_{12}
$\bar{P}$	a_{21}	a_{22}

Figure 5 Optimal strategy for IPOs

a_{11} belongs to cluster C_1 after clustering for the numeric property of f_i and the corresponding stock property is P. a_{12} belongs to cluster C_2 and the corresponding stock property is P. a_{21} belongs to cluster C_1 and the corresponding stock property is $\bar{P}$. $a_2 2$ belongs to cluster C_2 and the corresponding stock property is $\bar{P}$. We calculate the value of $\theta - \chi^2$ by Equation (5) based on the result of clustering.

$$\theta - \chi^2 = \theta * \frac{(a_{11} * a_{22} - a_{12} * a_{21})^2}{(a_{11} + a_{12}) * (a_{21} + a_{22})} \tag{25}$$

We establish a table of property values $\theta - \chi^2$ for selecting the features from 47 initial features. At last, we get rid of online issue ratio, liquidity ratio, debt ratio and so on, and select 40 features according to the rank of the value.

After the data preprocessing, we take two methods from classification, regression, and maximum entropy. We select five combinations which performed better than the separated methods and other combinations as follow. In classification filer, we chose LMT algorithm as the classifier by which we can get better prediction than other tree classifiers.

①MaxEnt(Entropy & Prediction):
The first layer is MaxEnt's entropy, and the second layer is the probability of MaxEnt's prediction in the model.

②MaxEnt(Entropy) & Classification(LMT):
The first layer is MaxEnt's entropy, and the second layer is the probability of LMT's prediction.

③Classification(LMT & Prediction):
The first layer is LMT, and the second layer is the prediction of Classification.

④Classification(LMT) & MaxEnt:
The first layer is LMT, and the second layer is MaxEnt's entropy.
⑤Classification (LMT) & Linear Regression:
The first layer is LMT and the second layer is Linear Regression[1].

5.2. Results

The IPOs optimal incomes by Double-layer filter model predicted in our tests are shown in Table 1 to Table 5. The unit of the income is Yuan.

Table 1 Test1(10-01-22-10-05-11)

METHODS	NUMBER OF IPOs	INCOME	YIELD
M(C&E)	13	42212.123	4.22
M&C(P)	16	71013.226	7.10
C(L&P)	19	72181.340	7.22
C&M	19	70892.902	7.09
C(L)&LR(p)	20	80678.646	8.07

Number of IPOs is the number of IPOs in the optimal path we predict by the model. $Yield = \frac{Income}{Initial purchase funds} * 100\%$.

Table 2 Test2(10-09-01-10-12-20)

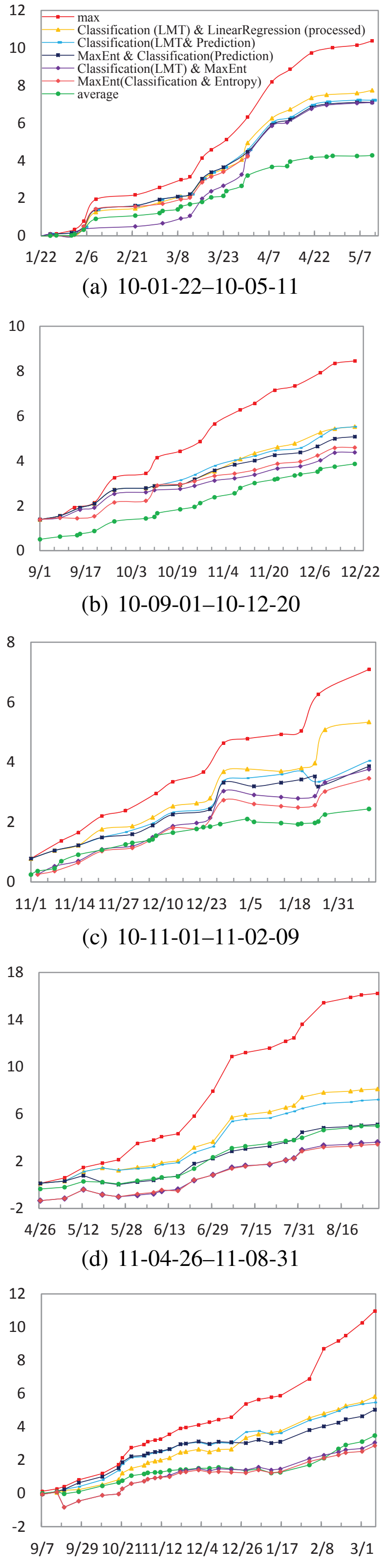
METHODS	NUMBER OF IPOs	INCOME	YIELD
M(C&E)	17	45952.161	4.60
M&C(P)	17	50813.149	5.08
C(L&P)	17	55225.275	5.52
C&M	17	48967.313	4.90
C(L)&LR(p)	17	55267.137	5.53

Table 3 Test3(10-11-01-11-02-09)

METHODS	NUMBER OF IPOs	INCOME	YIELD
M(C&E)	15	34551.7	3.46
M&C(P)	15	38597.117	3.86
C(L&P)	14	40421.331	4.04
C&M	16	37589.647	3.76
C(L)&LR(p)	16	53387.653	5.34

The comparison figures of prediction yield, ex post optimal yield and average yield are shown in Figure 6.

The ex post optimal incomes and average incomes are given in Table 6 and Table 7.



(a) 10-01-22–10-05-11

(b) 10-09-01–10-12-20

(c) 10-11-01–11-02-09

(d) 11-04-26–11-08-31

(e) 11-09-07–12-03-09

Figure 6 Comparison of IPOs income

[1] In the results of the Linear Regression, we select the IPOs which probabilities are greater than or equal to 0.8.

Table 4 Test4(11-04-26-11-08-31)

METHODS	NUMBER OF IPOs	INCOME	YIELD
M(C&E)	21	34261.423	3.43
M&C(P)	21	51152.028	5.12
C(L&P)	21	72183.659	7.22
C&M	21	36090.278	3.61
C(L)&LR(p)	21	81216.071	8.12

Table 5 Test5(11-09-07-12-03-09)

METHODS	NUMBER OF IPOs	INCOME	YIELD
M(C&E)	29	28697.984	2.87
M&C(P)	29	50386.717	5.04
C(L&P)	29	54681.738	5.47
C&M	29	30513.813	3.05
C(L)&LR(p)	29	58250.683	5.83

Table 6 Ex post optimal income

TEST	NUMBER OF IPOs	INCOME	YIELD
Test1	18	103791.565	10.38
Test2	17	84531.472	8.45
Test3	14	70956.553	7.10
Test4	22	162218.544	16.22
Test5	29	109756.282	10.98

Table 7 Average income of IPOs

TEST	NUMBER OF IPOs	INCOME	YIELD
Test1	100	42824.093	4.28
Test2	100	38664.197	3.87
Test3	100	24345.893	2.43
Test4	100	49898.788	4.99
Test5	100	34832.455	3.48

5.3. Analysis of results

From the experiments results and comparison figures we can see that the methods we take in the model can get more satisfactory income, and the trends we predicted are same to ex post optimal income. The ⑤ method (The first layer is LMT and the second layer is Linear Regression) can get better income of IPOs than other methods and the yield is more close to the ex post optimal yield. For example, in test 3 the ex post optimal income is 70.96 thousand Yuan and we can get 53.39 thousand Yuan (the detailed path is shown in Table 9) by our optimal strategy, the average income is 24.35 thousand Yuan, so the CM_{LimR} is 1.14, and the closeness the predicted income we can get to the ex post optimal income is 75.21% that is very close to the ex post optimal income.

Experiments also show that because of cycle investment and the freezing of purchase funds, the rate of income is not proportional to the number of IPOs in the optimal path.

6. Conclusion

In this paper, we propose the Double-layer filter model for Chinese IPOs. The optimal strategy model can improve the accuracy of the prediction for the income, and the combinations of methods can get more accurate prediction than the separate methods. Our IPOs recommended system is provided for free now in Haitianyuan financial service platform. (http://finance.haitianyuan.com/stock/newstock/).

However, the results also reveal that there is still room for improvement, especially in the selection of methods and features. Different combinations of features coordinate different combinations of methods can purify features and reduce noisy data. While, improving the adaptive of the prediction is the most important future work. That can make the prediction more accurate according to latest history data dynamically.

Acknowledgement

This work is supported in part by the National Natural Science Foundation of China (No. 61173075 and 60973076).

References

[1] M. Alghalith, Journal of Applied Economics **8**, 247-257 (2005).

[2] J. Wang, J. Wang, Z. Zhang and S. Guo, Expert Systems with Applications **38**, 14346-14355 (2011).

[3] C. Bouveyron and J. Jacques, Pattern Recognition Letters **31**, 2237-2247 (2010).

[4] O. L. Mangasarian and M. E. Thompson, Journal of Optimization Theory and Applications **131**, 315-325 (2006).

[5] S. Zhu, X. Ji, W. Xu, and Y. Gong, Proc. 28th annual international ACM SIGIR conference on Research and development in information retrieval, 274-281 (2005).

[6] R.R.P. Schumaker and H. Chen, ACM Transactions on Information Systems **27**, 12-19 (2009).

[7] C. Cecchetto and F. Dercole, Complexity in Engineering, 129-131 (2010).

[8] J.H. Min and C. Jeong, Expert Systems with Applications **36**, 5256-5263 (2009).

[9] E.L. Mencia, S. Park and J. Furnkranz, Neurocomputing **73**, 1164-1176 (2010).

[10] J. Qin, Y. Yao and J. Yang, International Journal of Digital Content Technology and its Applications **6**, 58-66 (2012).

[11] X. Jin and X. Hui, Journal of Convergence Information Technology **7**, 90-96 (2012).

[12] H. Hou, M. Wang and S. Ren, Journal of Convergence Information Technology **7**, 243-250 (2012).

[13] S.D. Pietra, V.D. Pietra and J. Lafferty, IEEE Transactions on Pattern Analysis and Machine Intelligence **19**, 380-393 (1997).

[14] A. Dukkipati, A.K. Yadav and M.N. Murty, Proc. 2010 International Conference on Pattern Recognition, 565-568 (2010).

[15] B. Wang, H. Huang and X. Wang, Neurocomputing **83**, 136-145 (2012).

Dong Huang is currently a PhD student in School of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School. Her research interests focus on natural language processing, machine learning,data mining, time series forecasting and the interdisciplinary research of Computer science and finance. The applications of her research are mainly in the financial field.

Xiaolong Wang received the B.E. degree in Computer Science from the Harbin Institute of Electrical Technology, China, the M.E. degree in Computer Architecture from Tianjin University, China, and the Ph.D. degree in Computer Science and Engineering from Harbin Institute of Technology in 1982, 1984, and 1989, respectively. He joined Harbin Institute of Technology as an Assistant Lecturer in 1984 and became an Associate Professor in 1990. He was a Senior Research Fellow in the Department of Computing, Hong Kong Polytechnic University from 1998 to 2000. Currently, he is a Professor of Computer Science at Harbin Institute of Technology. His research interest includes artificial intelligence, machine learning, computational linguistics, and Chinese information processing.