

Evolutionary Rule-Based System for IPO Underpricing Prediction

David Quintana
Universidad Carlos III
Avda. Universidad, 30
Leganés 28911, Madrid, Spain
(+34) 91 624 8879
dquintan@inf.uc3m.es

Cristóbal Luque
Universidad Carlos III
Avda. Universidad, 30
Leganés 28911, Madrid, Spain
(+34) 91 624 8843
cluque@inf.uc3m.es

Pedro Isasi
Universidad Carlos III
Avda. Universidad, 30
Leganés 28911, Madrid, Spain
(+34) 91 624 9455
isasi@ia.u3m.es

ABSTRACT

Academic literature has documented for a long time the existence of important price gains in the first trading day of initial public offerings (IPOs).

Most of the empirical analysis that has been carried out to date to explain underpricing through the offering structure is based on multiple linear regression. The alternative that we suggest is a rule-based system defined by a genetic algorithm using a Michigan approach. The system offers significant advantages in two areas, 1) a higher predictive performance, and 2) robustness to outlier patterns. The importance of the latter should be emphasized since the non-trivial task of selecting the patterns to be excluded from the training sample severely affects the results.

We compare the predictions provided by the algorithm to those obtained from linear models frequently used in the IPO literature. The predictions are based on seven classic variables. The results suggest that there is a clear correlation between the selected variables and the initial return, therefore making possible to predict, to a certain extent, the closing price.

Categories and Subject Descriptors

I.2.6-Learning; J.4-Economics

General Terms

Algorithms

Keywords

Genetic algorithm, initial public offering, underpricing.

1. INTRODUCTION

Academic literature has documented for a long time the existence of abnormal first-day trading returns in initial public offerings

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO '05, June 25-29, 2005, Washington, DC, USA.

Copyright 2005 ACM 1-59593-010-8/05/0006...\$5.00.

(IPOs) that usually come as important price gains. That is, there is usually a big difference between the offering price and the closing price at the end of the first trading day. Ritter and Welch [29] report an average initial return of 18.8% on a sample of 6,249 US IPOs that took place between 1980 and 2001¹.

This phenomenon has been puzzling researchers in Financial Economics for more than thirty years and has been the subject of a vast amount of academic work. Many theories have been postulated to offer explanations and it is still a very active field. The mentioned paper by Ritter and Welch is good reference to the state of the art.

The use of Genetic Algorithms in financial prediction is hardly new. Many pieces of research have explored the suitability of this technique in areas such as trading [2], portfolio optimization [19] or bankruptcy prediction [32] but little has been done in IPO research.

Most of the empirical analysis that has been carried out to date in order to explain underpricing through variables related to the offering, is based on multiple linear regression. Other than the seminal work of Jain and Nag [17] trying to predict first-day returns using Artificial Neural Networks, very few efforts have been done using artificial intelligence. We suggest new approach based on genetic algorithms that will be compared to the traditional instrument in terms of predictive performance.

The rest of the paper will have the following structure: In Section 2, we introduce the explanatory variables and describe the data. Section 3 deals with the methodology and Section 4 will be used to report the results of our empirical analysis. Finally, Section 5 covers the summary and conclusions.

2. VARIABLES AND DATA

As a starting point, we provide a formal definition of the phenomenon we have been referring to as IPO underpricing.

We define IPO underpricing as the percentage change of the share price from the offer to the closing price on first day of trading minus the return on the appropriate index or:

$$R_i = \left(\frac{Pc_i - Po_i}{Po_i} \right) - \left(\frac{Mc_i - Mo_i}{Mo_i} \right)$$

¹ Equally weighted average first-day return measured from the offer price to the first CRSP-listed crossing price.

Where R_i is the adjusted first day return for stock i ; PO_i is the offering price for stock i ; P_c is the closing price for stock i ; Mo_i is the closing for the broad market index of the market where the stock i was floated for the day before the IPO and MC_i is the closing for the broad market index of the market where the stock i was floated on the day of the IPO².

Once the target variable has been clearly described, we introduce the independent variables.

2.1 Variables

As we have already mentioned in the introduction, the amount of literature regarding IPO underpricing is quite remarkable. This fact makes the initial number of potential explanatory variables very high. However, there seems to be a number of them concerning the structure of the offerings that show up very often.

These variables, which are about to be succinctly described, are the following: Underwriter prestige, price range width, price adjustment, offer price, retained stock offer size and relation to tech sector.

2.1.1 Underwriter prestige (PRESTIGE)

This is one of the factors whose influence has been studied the most, Hammond [26] [27] [8]. Our approach will be similar to the one suggested by Balbers et al [4]. The prestige of financial advisors will be modeled by a dummy variable. In case there is a prestigious underwriter in the role of lead manager of co-lead, the value will be one and zero otherwise. The label of “prestigious” will be granted to the financial advisors who were consistently among the 25 who underwrote more capital in the US during the years of the study.

2.1.2 Price range width (RANGE)

This variable represents the width of the non-binding reference price range offered to potential customers during the roadshow. This width can be interpreted as a sign of uncertainty regarding the real value of the company and a therefore, as a factor that could influence the initial return. Following Hanley [13], the representation to be used will be the difference between the maximum and minimum price divided by the minimum price.

2.1.3 Price adjustment (P_{-ADJ})

Benveniste and Spindt [5] and Hanley [13] suggest the relation between the final offer price and the mentioned price range might also be interpreted as sign of uncertainty. They state that this effect might be captured by the following expression.

$$P_{-ADJ} = \frac{|P_f - P_e|}{P_e}$$

Where P_f is the final offer price and P_e is the expected price defined as the middle point of the price range.

² The indexes used in the analysis were: S&P 500, NASDAQ Composite and AMEX Composite.

2.1.4 Offering price (PRICE)

The final offering price has been found to be a relevant variable not only as a part of the previous indicator, but on its own. Studies like [9] or [7], among others, support this idea.

2.1.5 Retained stock (RETAINED)

The influence of the capital retained by initial investors at the time of the IPO has been traditionally understood to have a signal the quality of the stock [20], [12] or [1]. Since we lack the breakdown of primary and secondary shares, we will proxy this variable through the ratio number of shares sold at the IPO divided by post-offering number of shares minus the number of shares sold at the IPO.

2.1.6 Offering size (LSIZE)

This variable is defined as the logarithm of the offering size in millions of dollars excluding the over-allotment option. Studies like [23], [14] and [16] support the need to include it in the models.

2.1.7 Technology (TECH)

The reason why we suggest a specific variable to control whether the industrial activities of a company are related to tech sector is the fact that they tend to show a higher underpricing. This fact is usually modeled by a dummy that equals one for tech companies [33] [21].

Our labeling criterion is based on IPO Data System’s definition. This company publishes every year a report with the list of tech companies taken public based on US Standard Industry Codes. Hence, we will consider an IPO to be “tech” if it is in the list.

2.2 Data

The sample consists of 1,040 companies taken public between April 1996 and November 1999 in US stock markets. This includes AMEX, NASDAQ and NYSE IPOs and excludes American Depositary Receipts; closed-end funds; real state investment trusts and unit offerings.

Our primary data source was Hoovers Online. The information was completed with IPO profiles and reports provided by IPO Data Systems. Index information was obtained from NASD (NASDAQ and AMEX composites) and Bloomberg (S&P 500).

The main descriptive statistics (mean, median, maximum, minimum and standard deviation) for the sample of 1,040 IPOs considered in the analysis are reported in table 1.

Table 1. Descriptive statistics

	Mean	Median	Max	Min	Std. Dev.
RETURN	0.26	0.10	6.06	-0.32	0.52
LSIZE	1.54	1.56	3.36	0.48	0.43
P_ADJ	0.10	0.08	3.00	0.00	0.13
PRESTIGE	0.57	1.00	1.00	0.00	0.49
PRICE	11.89	12.00	34.00	3.50	4.81
RANGE	0.16	0.16	1.33	0.00	0.09
RETAINED	0.50	0.40	7.84	0.06	0.46
TECH	0.29	0.00	1.00	0.00	0.45

3. METHODOLOGY

The aim of this paper is to show a rule-based algorithm tool that could be very useful in IPO research. To prove it, we will compare its forecasting ability to the one achieved by the linear regression, the standard tool in the literature.

In order to do that, we will break the sample in two subsamples. The first one, made of 840 IPOs, will be used as training set for the rule-based system. We will also fit the linear model on the very same sample. These two models will then be used to predict the rest of the original set of data.

The system has the ability to anticipate which observations could lead to potentially inaccurate forecasts. Should this be the case, the system will not provide any estimate for the initial return.

The accuracy of the predictions will be assessed comparing the normalized mean square error of the predictions of the models.

3.1 Rule-Based System

Genetic Algorithms [15][11][3][25] have already been used in prediction problems, mainly related to time series [28][20][22]. Following Packard's [28] prediction approach for dynamical systems, we introduce a rule-based system based on genetic algorithms.

3.1.1 Model Description

The system generates rules defined in the input variable space in order to make prediction in the output variable space. In this work we use the previously defined set of independent variables (LSIZE, P_ADJ, PRESTIGE, PRICE, RANGE, RETAINED and TECH) and RETURN as target variable. The process starts with the standardization of these values, which are forced within the 0 and 1 range. At this point, we bring in the rules. These are asserts such as "if variable 1 is smaller than 0.9 and bigger than 0.13, variable 2 is smaller than 0.81 and bigger than 0.2, variable 4 is smaller than 0.94 and bigger than 0.01, and variable 7 is smaller than 9.9 and bigger than 0.45, then the prediction (for the output variable) will be 0.73, with an expected error of 0.03. It could be expressed as

IF (0.13<v1<0.9) AND (0.2<v2<0.81) AND (0.01<v4<0.94)
AND (0.45<v7<0.99) THEN prediction = 0.73 +/- 0.03.

This information is encoded in an individual as follows:

(0.13, 0.9, 0.2, 0.81, dc, dc, 0.01, 0.94, dc, dc, dc, dc, 0.45, 0.99, 0.73, 0.03)

where *dc* means "don't care". Figure 1 illustrates an individual. The genetic algorithm paradigm might be useful to evolve new rules such as the one we just mentioned. Two individuals can generate an offspring. This offspring inherits each gene from one parent, as shown in figure 2. A gene is a pair (LL,UL), where LL is the lower limit for a variable, and UL is the upper limit for the same variable. In other words, the offspring receives a gene from each parent with equal probability for each variable. The offspring doesn't inherit parent's predictions and errors. Following the example above:

Parent 1: (0.13, 0.9, 0.2, 0.81, dc, dc, 0.01, 0.94, dc, dc, dc, dc, 0.45, 9.9, 0.73, 0.03)

Parent 2: (0.14, 0.77, 0.13, 0.85, 0.11, 0.22, dc, dc, 0.32, 0.51, 0.31, 0.66, 0.14, 0.27, 0.57, 0.06)

Offspring: (0.13, 0.9, 0.13, 0.85, dc, dc, dc, dc, 0.32, 0.51, dc, dc, 0.45, 9.9, ?, ?)

Obviously, the offspring's "prediction" and "error" are not assigned (and therefore are represented as "?"). Once generated, an offspring may suffer a gene mutation. This event takes place with a 10% probability for each new individual. Should the individual be selected, one of its genes will be altered. In table 2, the mutations over the gene (LL,UL) are showed. R(x,y) means a random value between x and y with $x < y$, and W means the 10% of the width of the interval defined by (LL,UL), in other words, $W=0.1 (UL-LL)$. Obviously, the transformation of (LL',UL') into (LL,UL) must fit the condition ($0 \leq LL' < UL' \leq 0$). Figure 3 summarizes these alternatives.

Table 2. Mutations

Name	Transformation
New random value	(LL,UL) → (R(0,1), R(0,1))
Null condition	(LL,UL) → (dc, dc)
Enlarge	(LL,UL) → (LL-R(0,1)W, UL+R(0,1)W)
Shrink	(LL,UL) → (LL+R(0,1)W, UL-R(0,1)W)
Move up	(LL,UL) → (LL+cW, UL+cW) c = R(0,1)
Move down	(LL,UL) → (LL-cW, UL-cW) c = R(0,1)

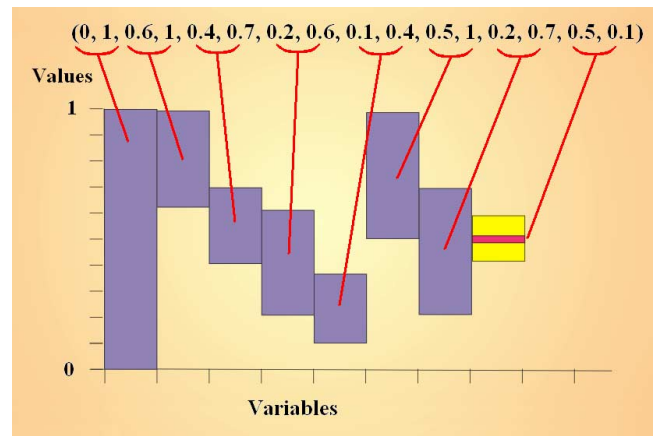


Figure 1. Graphical representation of an individual

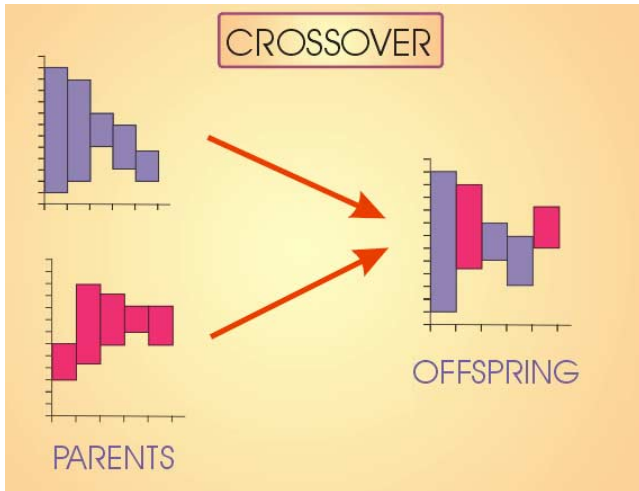


Figure 2. Crossover process

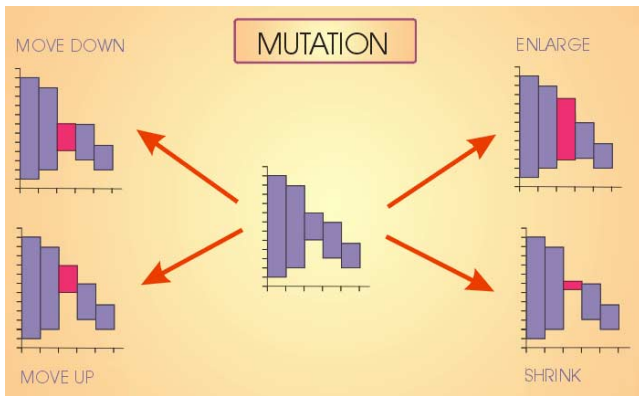


Figure 3. Mutation process

3.1.2 Fitness

The computation of fitness follows the idea suggested by Packard [28], by which good individuals are characterized by uniform values for the dependant variable across matching patterns. The assessment of these individuals (rules), does not take into account the forecasts, which are made at a later stage. Any individual has two associated parameters, a prediction and an error measure. Let C be an individual and i a training pattern. If the value of the variables for pattern i meet the requirements of C , we will say $C(i)=true$, otherwise $C(i)=false$. The prediction assigned to the rule will be derived by averaging the target variable of the patterns from the training sample that fulfill the requirements specified by C . The error magnitude, E , is the maximum absolute difference between the predicted value, p_i , and the real value, v_i , of the matching training patterns divided by the number of patterns in the training sample that meet rule C .

$$E = \text{Max}_i \{ |p_i - v_i| / C(i) \}$$

The model favors individuals with the ability to predict the highest number of patterns with the lowest possible error. The specific fitness function that we use is:

IF $((N_C > 1) \text{ AND } (e < EMAX))$ THEN $\text{fitness} = (N_C * 100) - C$
 ELSE $\text{fitness} = f_{min}$

where C is the individual, and N_C is the number of training patterns that match C (in other words, $N_C = \# \{ i / C(i) \}$). $EMAX$ is a constant for punishing individuals with a variance greater than its own value, and f_{min} is the lowest value assigned to the individual when the rule is not fitted.

3.1.3 Evolution of simple rules

Within the two main approaches to system design in Evolutionary Computation, Michigan and Pittsburgh, the second has traditionally been more widely used [34][18][10]. In this approach a solution is defined by an individual, and the population is made of a number of different alternatives. Michigan approach [6], however, defines the solution by the whole population. Each individual is just a piece of the solution, or the solution to a part of the domain.

The use of genetic algorithms in IPOs faces a major difficulty, namely, the presence of a number of outliers that are hard to characterize. In these cases, there are unusually high differences between the offering price and the closing price at the end of the first trading day or very extreme values of the independent variables. When using the Pittsburgh approach, each individual tries to predict the whole set of data, and usually only standard behaviors are characterized. Should this be the case, unexpected behaviors are not considered, which leads to poor predictions.

This factor justifies the choice of a Michigan approach, since it results in each individual being specialized in a section of the input space and, therefore, providing forecasts only for those patterns within its range of capability. Since the population as a whole is used like a set of rules to predict most of the data set, it is possible to evolve individuals that are specialized in those unusual cases, while others evolve focused on the bulk of the data with a regular behavior.

An additional benefit of encoding the solution in the whole population is the reduction in the size of individuals. This accelerates their evaluation, which is crucial in some domains.

We apply Michigan's approach selecting only two parents by three rounds trial each generation to produce a single offspring, following a Steady-State strategy This is subsequently used to replace the closest individual in terms of phenotypic distance. That is, we find the individual whose prediction is nearest to the offspring's prediction, and replace it by the offspring if and only if the offspring fitness is better than the individual's fitness. If this doesn't happen, there isn't any change this generation. That strategy generates a diverse population, in which each individual makes a prediction that is different to the rest instead of genetic clones of the best individual, produced by the standard genetic algorithm method. Finally, after each execution of the model (500.000 generations), we store in a file, that we call "rule pool", individuals that predicts more than 20 points in the training set (and not only one as we would do in a standard genetic algorithm model). This process is repeated many times and a file with a set of individuals is generated. Some individuals predict the same patterns of the test set, so the final prediction is the mean of all predictions (we must remember that, possibly, not all the individuals could make a prediction for a pattern).

The generation of the initial population hasn't been explained yet. We divide the prediction range (i.e., the interval [0,1]) in 100 subintervals that are 0.01 wide. We create an individual for each prediction interval. For example, for interval [0.34, 0.35] we

search for all the patterns i in the training set such that the predicting value $p(i)$ is in the $[0.34,0.35]$ interval, and we create an individual with an upper and lower limit for every input variable. For each input variable x_j the upper limit must be the minimum number U such that for each pattern i with predicting value $p(i)$ in the $[0.34,0.35]$ interval, the value of x_j is less than U . Accordingly, we select the lower limit for x_j as the maximum L such that for each pattern i with predicting value $p(i)$ in the $[0.34,0.35]$ interval, the value of x_j is greater than L . We repeat the process for each variable and interval to generate the initial population of 100 individuals and so we can assure that there is an individual in each interval of prediction.

3.2 Regression

The basic linear model that we use is the standard cross-sectional ordinary least squares regression, which will be initially fitted to the 840 training sample. However, given the sensitivity of the model to the presence of outliers, a robust alternative, least trimmed squares regression [30], will also be considered. We will use three trimming constants (424, 600 and 800), being the first one the maximum breakdown point as suggested by Rousseeuw and Leroy [31].

The pattern selection criterion used in the test sample deserves an explanation. Given that the rule-based system will not provide initial return forecasts for a number of IPOs that are expected to be hard to predict, we need a filtering criterion to be used with the regression in order to allow for meaningful model comparisons. The solution that we suggest is based on the standard forecast errors.

This measure is an indicator of the expected accuracy of the predictions made. High variability of forecasts leads to higher potential prediction errors and, therefore, lowers performance. Hence our choice of discarding from the test sample those IPOs that present the highest standard forecast errors. The number of patterns to be excluded is the number of IPOs discarded by the rule system.

4. EMPIRICAL ANALYSIS

4.1 Rule-Based System

The data was subject to several experiments using the described model in order to carry out the prediction task. The first step was computing the best value for the constant EMAX, referred in section 3.1.2. The lower value it has, the lower prediction error we get, but as a side effect, we simultaneously reduce the number of patterns that we are able to forecast. The results of the experiments done to identify the optimal value of EMAX, as are shown in table 3. We used a set of 840 patterns for these experiments. For each value of EMAX, we ran the model 5 times (adding the rules to the “rules pool” each time), and the best equilibrium between percent of prediction and prediction error was attained with a value of 0.15. Once this value was selected, more thorough experiments were carried out (the system was ran 25 times). This time, the prediction percentage increased to 90%, the normalized mean square error was 0.433 and the mean absolute error (for the normalized between 0 and 1 output) was 0.055. Cross-validation analysis was done to test the robustness of this prediction rule search scheme. We selected 200 new patterns from the training set to the validation test, and added the previous

200 validation patterns to the training set. The results are reported in table 4.

Table 3. EMAX experiments

EMAX value	NMSE	Percent of prediction
0.05	0.47638	32.5%
0.1	0.44031	61.5%
0.15	0.39520	80.5%
0.2	0.63399	84.0%
0.25	0.44392	85.0%
0.35	0.61332	91.5%
0.45	0.70942	95.5%
0.5	0.88462	89.5%
0.6	0.89787	93.5%

Table 4. Cross-validation

Segment in data for validation	NMSE	Percent of prediction
[840,1040]	0.43267	90%
[0,200]	0.43631	88.5%
[200,400]	0.38950	87.5%
[400,600]	0.45244	89.5%

4.2 Regression

The linear models were fitted on the training set and used to forecast the 200 pattern test set. Table 5, reports the performance of the different models used.

Table 5. Cross-sectional regressions

Linear model	Trimming constant	NMSE Test	NMSE Test-20
OLS	-	0.92302	0.77238
LTS	424	0.98268	0.87663
LTS	600	1.03793	0.96238
LTS	800	0.94482	0.88741

The best model offered a normalized mean square error of 0.923 for the whole test sample. All of the suggested explanatory variables were significant at the 5% conventional level.

As we mentioned before, we allowed the regression to discard the same number of patterns that was discarded by the rule system (20). As a result, the normalized mean square error goes down to 0.772, which is a clear improvement, but still underperforms the rule system.

5. SUMMARY AND CONCLUSIONS

In this paper we have discussed an implementation of a genetic algorithm to predict the underpricing of IPOs. The predictions are based on seven variables identified by literature review and their accuracy has been compared to the forecasts provided by a set of linear regressions.

In this domain, we have to deal with outliers that make prediction particularly difficult. In order to overcome this problem, we suggest a rule system based on a genetic algorithm with a Michigan approach. The system identifies local rules to be used

with patterns that show similar behavior so that we get local predictions instead of general rules.

Following this scheme, we obtain predictions for 90% of the test sample. The system doesn't find correlation between the patterns in the remaining 10% that we cannot forecast. In order to make the predictive performance comparable, we discard from the forecasts made by the regressions the 10% that are more likely to be the worst.

The results show significant differences in favor of the rule system, which beats the regression by 49% in terms of normalized mean square prediction error. This fact suggests that IPO research would benefit from the use of this tool and we understand that so would any prediction effort in domains where the above mentioned factor is applicable.

6. ACKNOWLEDGMENTS

This article has been financed by the Spanish founded research MCyT project TRACER, Ref: TIC2002-04498-C05-04M.

7. REFERENCES

- [1] Aggarwal, R. K., Krigman, L., and Womack, K. L. Strategic IPO underpricing, information momentum, and lockup expiration selling. *Journal of Financial Economics*, 66, 1 (2002), 105-137.
- [2] Allen, F., and Karjalainen, R.. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51 (1999), 245-272.
- [3] Bäck, T., Schwefel, H.P., Evolutionary Algorithms: Some Very Old Strategies for Optimization and Adaptation. In *Perret-Gallix* (1992), pp. 247-254.
- [4] Balvers, R. J., McDonald, B., and Miller, R. E. Underpricing of new issues and the choice of auditor as a signal of investment banker reputation. *Accounting Review*, 63, 4 (1988), 605-622.
- [5] Benveniste, L. M., and Spindt, P. A. How investment bankers determine the offer price and allocation of new issues. *Journal of Financial Economics*, 24 (1989), 343-362.
- [6] Booker, L.B., Goldberg, D.E., Holland, J.H., Classifier Systems and Genetic Algorithms. *Artificial Intelligence*, 40 (1989), pp. 235-282.
- [7] Brennan, M., and Hughes, P. Stock prices and the supply of information. *Journal of Finance*, 46 (1991), 1665-1691.
- [8] Carter, R. B., and Manaster, S. Initial public offering and underwriter reputation. *Journal of Finance*, 45 (1990), 1045-1067.
- [9] Chalk, A., and Peavy, J. Initial public offerings, daily returns, offering types and the price effect. *Financial Analysts Journal*, 43 (1987), 65-69.
- [10] De Jong, K.A., Spears, W.M., Gordon, F.D., Using Genetic Algorithms for Concept Learning. *Machine Learning* 13 (1993), pp. 198-228.
- [11] Fogel, D.B., An introduction to simulated evolutionary optimization. *IEEE transactions on neural networks*, 5, 1, (Jan. 1994).
- [12] Grinblatt, M., and Hwang, C.Y. Signaling and the pricing of new issues. *Journal of Finance*, 44 (1989), 393-420.
- [13] Hanley, K. W. The underpricing of initial public offerings and the partial adjustment phenomenon. *Journal of Financial Economics*, 34, 2 (1993), 231-250.
- [14] Hansen, R. S., and Torregrosa, P. Underwriter compensation and corporate monitoring. *Journal of Finance*, 47, 4 (1992), 1537-1555.
- [15] Holland, J.H., *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Mi, 1975.
- [16] Jain, B. A., and Kini, O. On investment banker monitoring in the new issues market". *Journal of Banking and Finance*, 23 (1999), 49-84.
- [17] Jain, B. A., and Nag, B. N. Artificial neural network models for pricing initial public offerings. *Decision Sciences*, 26, 3, (1995), 283-299.
- [18] Janikow, C.Z., A Knowledge Intensive Genetic Algorithm for Supervised Learning. *Machine Learning* 13 (1993), pp. 189-228.
- [19] Korczak, J., Lipinski, P., and Roger, P. Evolution Strategy in Portfolio Optimization. In *Artificial Evolution*, ed. P. Collet. Lecture Notes in Computer Science, 2310, Springer, 2002, 156-167.
- [20] Leland, H., and Pyle, D. Informational asymmetries, financial structure and financial intermediation. *Journal of Finance*, 32 (1977), 371-387.
- [21] Loughran, T., and Ritter, J. R. Why has IPO underpricing changed over time?. *Financial Management*, 33 (2004), 5-37.
- [22] Luque, C., Isasi P., Hernández, J.C., Forecasting Time Series by means of Evolutionary Algorithms. *Proceedings of Parallel Problem Solving from Nature - PPSN VIII*, vol 3242 (2004), pp. 1058--1067
- [23] Megginson, W. L., and Weiss, K. A. Venture capitalist certification in initial public offerings. *Journal of Finance*, 46, 3 (1991), 799-903.
- [24] Meyer, T.P., Packard, N. H., Local Forecasting of High-Dimensional Chaotic Dynamics. *Nonlinear modeling and forecasting*. (1990); editors, Martin Casdagli, Stephen Eubank, pp. 249-263.
- [25] Mitchell, M., *An introduction to Genetic Algorithms*. MIT Press (1996), Cambridge, MA, 1996, pp. 55-65.
- [26] Neuberger, B. M., and Hammond, C. T. A study of underwriters' experience with unseasoned new issues. *Journal of Financial and Quantitative Analysis*, 9, 2 (1974), 165-177.
- [27] Neuberger, B. M., and La Chapelle, C. A. Unseasoned new issue price performance on three tiers: 1975-1980. *Financial Management*, 12, 3 (1983), 23-28.
- [28] Packard, N. H. A genetic learning algorithm for the analysis of complex data. *complex systems* 4, no 5 (1990), pp. 543-572.

- [29] Ritter, J. R., and Welch, I. A review of IPO activity, pricing, and allocations. *Journal of Finance*, 57, 4 (2002), 1795-1828.
- [30] Rousseeuw, P. J. Multivariate estimation with high breakdown point. *Mathematical Statistics and Applications, Vol. B.* ed. W. Grossmann, G. Pflug, I. Vincze, and W. Werty, Reidel, Dordrecht, Netherlands, 1985, 283-297.
- [31] Rousseeuw, P. J., and Leroy, A. M. *Robust regression and outlier detection*, John Wiley and Sons, New York, 1987.
- [32] Shin, K. A genetic algorithm application in bankruptcy prediction modeling. *International Journal of Expert Systems with Applications*, 23 (2002), 321-28.
- [33] Smart, S. B., and Zutter C. J. Control as a motivation for underpricing: a comparison of dual and single-class IPOs. *Journal of Financial Economics*, 69 (2003), 85-110.
- [34] Smith, S.F., A Learning System Based on Genetic Adaptive Algorithms. *Ph.D. Thesis*, University of Pittsburgh (1980).