Rankings of financial analysts as means to profits

Artur Aiguzhinov[1,2], Carlos Soares[2,3], and Ana Paula Serra[1]

[1] FEP & CEF.UP, University of Porto
[2] INESC TEC
[3] FEUP, University of Porto
artur.aguzhinov@inesctec.pt, csoares@fe.up.pt, aserra@fep.up.pt

Abstract. Financial analysts are evaluated based on the value they create for those who follow their recommendations and some institutions use these evaluations to rank the analysts. The prediction of the most accurate analysts is typically modeled in terms of individual analyst' characteristics. The disadvantage of this approach is that these data are hard to collect and often unreliable. In this paper, we follow a different approach in which we characterize the general behavior of the rankings of analysts based upon state variables rather than individual analyst's characteristics. We extend an existing adaptation of the naive Bayes algorithm for label ranking with two functions: 1) dealing with numerical attributes; and 2) dealing with a time series of label ranking data. The results show that it is possible to accurately model the relation between the selected attributes and the rankings of analysts. Additionally, we develop a trading strategy that combines the predicted rankings with the Black-Litterman model to form optimal portfolios. This strategy applied to US stocks generates higher returns than the benchmark (S&P500).

1 Introduction

In recent years, some institutions were very successful selling the rankings of analysts based on their relative performance. For example, Thomson Reuters publishes the StarMine rankings of the financial analysts on an annual basis identifying the top. The Institutional Investors magazine and Bloomberg have been publishing and selling these rankings for decades and these attract investor attention and broad media coverage. Aside from personal acknowledgment among the peers, it is still arguable if rankings of financial analysts provide valuable information to market participants and help them in selecting which analysts to follow.

Following analysts' recommendations, on average, brings value to investors [15]. Hence, following the recommendations of the top analysts should result in a profitable trading strategy. Since analysts do not make recommendations frequently, at any given moment in time, an investor may only have recommendations from analysts other than the top ones. Given that, identifying the top analysts ahead of time is beneficial for an investor. In this paper, we propose a method to predict the rankings of the analysts and use these rankings to develop a successful trading strategy.

We address the problem of rankings of analysts as a label ranking (LR) problem. Many different algorithms have been adapted to deal with LR such as: naive Bayes [1], decision-trees [5], k-nn [4,5]. However, none of these algorithms is prepared to deal with time series of rankings, which is an important characteristic of our problem. It is expected that the ranking of analysts on a given period is not independent from the ranking in the previous period. Thus, some adaptation of existing LR algorithms is required to solve this problem.

Once we have predicted rankings, we apply a trading strategy that works within the framework of the Black-Litterman (BL) model [2]. The model admits a Bayesian setting and allows to transform stock views into optimal portfolio weights. We use analysts' target prices to obtain expected returns. Using the predicted rankings, we compute analysts' views for a particular stock. These views are the input for the BL model. The resulting portfolio maximizes the Sharpe ratio [12]. We use S&P500 as a proxy for market returns. We show that 1) our LR model outperforms other forecasting models; 2) the resulting trading strategy generates superior returns.

The contributions of our paper are the following. We are able to adapt the existing LR algorithm and apply it to a real world problem of predicting the rankings of financial analysts. Using the predicted rankings as inputs, we design a profitable trading strategy based on the BL model

The paper is organized as follows: Section 2 reviews the rankings of the analysts in the finance literature; Section 3 formalizes the label ranking problem and introduces the adaptation of the algorithm to deal with time series of the rankings; Section 4 outlines the trading strategy that uses the predicted rankings; Section 5 describes the datasets used for the experiments; Section 6 analyzes the results; and Section 7 concludes.

2 Ranking of Financial Analysts

In the finance literature there has been a long debate over whether financial analysts produce valuable advice. Some argue that following the advice of financial analysts, translated as recommendations of buying, holding, or selling a particular stock, does not yield abnormal returns, i.e., returns that are above the required return to compensate for risk [8]. If financial markets are efficient then any information regarding a stock would be reflected in its current price; hence, it would be impossible to generate abnormal returns based upon publicly available information. This is the Efficient Market Hypothesis (EMH).

Yet there are information-gathering costs and the information is not immediately reflected on prices [9]. As such, prices could not reflect all the available information because if that was the case, those who spent resources to collect and analyze information would not receive a compensation for it.

For market participants, rankings could be useful because they signal the top analysts. Evidence shows that market response to analysts' recommendations is stronger when they are issued by analysts with good forecasting tracking record [11]. Yet the value of these rankings for investors is arguable as they are ex-post and a good analyst in one year does not necessarily make equally good recommendations in the following year [7]. However, if we know the ranking of analysts ahead of time then it would be possible to create a successful trading strategy based upon that information. If we can, with reasonable accuracy, predict the rankings we can follow the recommendations of the analysts that are expected to be at the top and, in presence of contradictory recommendations, take the rank of the corresponding analysts into account.

3 Label ranking algorithm

The classical formalization of a label ranking problem is the following [13]. Let $\mathcal{X} = \{\mathcal{V}_1, \ldots, \mathcal{V}_m\}$ be an instance space of variables, such that $\mathcal{V}_a = \{v_{a,1}, \ldots, v_{a,n_a}\}$ is the domain of nominal variable a. Also, let $\mathcal{L} = \{\lambda_1, \ldots, \lambda_k\}$ be a set of labels, and $\mathcal{Y} = \Pi_{\mathcal{L}}$ be the output space of all possible total orders over $\mathcal{L}$ defined on the permutation space Π. The goal of a label ranking algorithm is to learn a mapping $h : \mathcal{X} \to \mathcal{Y}$, where h is chosen from a given hypothesis space $\mathcal{H}$, such that a predefined loss function $\ell : \mathcal{H} \times \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ is minimized. The algorithm learns h from a training set $\mathcal{T} = \{x_i, y_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{X} \times \mathcal{Y}$ of n examples, where $x_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,m}\} \in \mathcal{X}$ and $y_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,k}\} \in \mathcal{Y}$. With time-dependent problem in rankings, we replace the i index with t; that is $y_t = \{y_{t,1}, y_{t,2}, \ldots, y_{t,k}\}$ is the ranking of k labels at time t described by $x_t = \{x_{t,1}, x_{t,2}, \ldots, x_{t,m}\}$ at time t.

Consider an example of a time-dependent ranking problem presented in Table 1. In this example, we have three brokers ($k = 3$), four independent variables ($m = 4$) and a period of 7 quarters. Our goal is to predict the rankings for period t, given the values of independent variables and rankings known up to period $t-1$; that is, to predict the ranking for time $t = 7$, we use $n = 6$ ($t \in \{1 \ldots 6\}$) examples to train the ranking model.

Table 1. Example of label ranking problem

Period	$\mathcal{V}_1$	$\mathcal{V}_2$	$\mathcal{V}_3$	$\mathcal{V}_4$	Ranks		
					Alex	Brown	Credit
1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	1	2	3
2	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	2	3	1
3	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	1	2	3
4	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	3	2	1
5	$x_{5,1}$	$x_{5,2}$	$x_{5,3}$	$x_{5,4}$	3	2	1
6	$x_{6,1}$	$x_{6,2}$	$x_{6,3}$	$x_{6,4}$	2	1	3
7	$x_{7,1}$	$x_{7,2}$	$x_{7,3}$	$x_{7,4}$	1	2	3

3.1 Naive Bayes algorithm for label ranking

The naive Bayes for label ranking (NBLR) will output the ranking with the higher $P_{LR}(y|x)$ value [1]:

$$\begin{aligned} \hat{y} &= \underset{y \in \Pi_{\mathcal{L}}}{\arg\max}\, P_{LR}(y|x) = \\ &= \underset{y \in \Pi_{\mathcal{L}}}{\arg\max}\, P_{LR}(y) \prod_{i=1}^{m} P_{LR}(x_i|y) \end{aligned} \tag{1}$$

where $P_{LR}(y)$ is the prior label ranking probability of ranking $y \in Y$ based on the similarity between rankings obtained from the Spearman ranking correlation Equation (2):

$$\rho(y, y_i) = 1 - \frac{6 \sum_{j=1}^{k} (y - y_{i,j})^2}{k^3 - k} \tag{2}$$

$$P_{LR}(y) = \frac{\sum_{i=1}^{n} \rho(y, y_i)}{n} \tag{3}$$

Similarity and probability are different concepts; however, a connection as been established between probabilities and the general Euclidean distance measure [14]. It states that maximizing the likelihood is equivalent to minimizing the distance (i.e., maximizing the similarity) in a Euclidean space.

$P_{LR}(x_i|y)$ in Equation (1) is the conditional label ranking probability of a nominal variable x of attribute a, (v_a):

$$P_{LR}(x_i|y) = \frac{\sum_{i:x_a=v_a} \rho(y, y_i)}{|\{i : x_a = v_a\}|} \tag{4}$$

The predicted ranking for an example x_i is the one that will receive the maximum posterior label ranking probability $P_{LR}(y|x_i)$.

Continuous independent variables In its most basic form, the naive Bayes algorithm cannot deal with continuous attributes. The same happens with its adaptation for label ranking [1]. However, there are versions of the naive Bayes algorithm for classification that support continuous variables [3]. The authors modify the conditional label ranking probability by utilizing the Gaussian distribution of the independent variables. We apply the same approach in defining the conditional probability of label rankings:

$$P_{LR}(x|y) = \frac{1}{\sqrt{2\pi}\sigma(x|y)} e^{\frac{(x-\mu(x|y))^2}{2\sigma^2(x|y)}} \tag{5}$$

where $\mu(x|y)$ and $\sigma^2(x|y)$ weighted mean and weighted variance, defined as follows:

$$\mu(x|y) = \frac{\sum_{i=1}^{n} \rho(y, y_i)x}{\sum_{i=1}^{n} \rho(y, y_i)} \tag{6}$$

$$\sigma^2(x|y) = \frac{\sum_{i=1}^{n} \rho(y, y_i)[x - \mu(x|y)]^2}{\sum_{i=1}^{n} \rho(y, y_i)} \tag{7}$$

3.2 Time series of rankings

The time dependent label ranking (TDLR) problem takes the intertemporal dependence between the rankings into account. That is, rankings that are similar to the most recent ones are more likely to appear. To capture this, we propose the weighted TDLR prior probability:

$$P_{TDLR}(y_t) = \frac{\sum_{t=1}^{n} w_t \rho(y, y_t)}{\sum_{t=1}^{n} w_t} \tag{8}$$

where $w = \{w_1, \ldots, w_n\} \rightarrow \mathbf{w}$ is the vector of weights calculated from the exponential function $\mathbf{w} = b^{\frac{1-\{t\}_1^n}{t}}$. Parameter $b \in \{1 \ldots \infty\}$ sets the degree of the "memory" for the past rankings, i.e., the larger b, the more weight is given to the last known ranking (i.e, at $t-1$) and the weight diminishes to the rankings known at $t = 1$.

As for the conditional label ranking probability, the equation for the weighted mean (Equation (5)) becomes:

$$\mu(x_t|y_t) = \frac{\sum_{t=1}^{n} w_t \rho(y, y_t) x_t}{\sum_{t=1}^{n} \rho(y, y_t)} \tag{9}$$

and σ:

$$\sigma^2(x_t|y_t) = \frac{\sum_{i=1}^{n} w_t \rho(y, y_t)[x_t - \mu(x_t|y)]^2}{\sum_{i=1}^{n} \rho(y, y_t)} \tag{10}$$

4 Trading Strategy

4.1 Independent variables

Several studies try to analyze factors that affect the performance of the analysts [6,10]. However, most of these papers look at the individual characteristics of analysts such as their job experience, their affiliation, education background, industry specializations. These variables are very important to characterize the relative performance of the analysts in general. Yet, our goal is to predict the rankings of the analysts over a series of quarters. We assume that the variation in rankings is due to the different ability of the analysts to interpret the informational environment (e.g., whether the market is bull or bear). We, thus, use variables that describe this environment. We select variables based on different levels of information: analyst-specific (analysts' dispersion; analysts' information asymmetry; analysts' uncertainty), stock-specific (stock return volatility, Book-to-Market ratio; accruals; Debt-to-Equity ratio), industry-specific (Sector index volatility) and general economy (interest rate; Gross National Product; inflation rate; S&P 500 volatility).

Given the time series of the rankings and independent variables, we also need to capture the dynamics of independent variables from one time period to another; that is, to find signals that affect brokers' forecasts accuracy. We propose the following methods of dynamics:

- `last`—no dynamics of x: $x_{t,m} = x_{t-1,m}$;
- `diff`—first-difference of x: $x_{\Delta t,m} = x_{t,m} - x_{t-1,m}$;
- `random`—an unobserved component of time series decomposition of x: $x_{\Delta t,m} = T(t) + S(t) + \epsilon(t)$, where $T(t)$- trend, $S(t)$ - seasonal part and $\epsilon(t)$ - random part of time series decomposition.
- `roll.sd`—moving 8 quarters standard deviation of x [16]:

$$\mu_{t(8),m} = \frac{1}{8}\sum_{j=0}^{7} x_{t-j,m} \qquad \sigma^2_{t(8),m} = \frac{1}{7}\sum_{j=0}^{7} \left(x_{t-j,m} - \mu_{t(8),m}\right)^2 \tag{11}$$

Each of these methods produces a different set of attributes. By using the algorithm on each one of them separately, we get different rankings. By evaluating them, we can get an idea of which one is the most informative.

4.2 Strategy setup

The Black-Litterman model [2] is a tool for active portfolio management. The objective of the model is to estimate expected returns and optimally allocate the stocks in a mean-variance setting, i.e., maximize the Sharpe ratio.

The BL model has established notations for the views part of the model and we use the same notations in this paper. The views are made of: Q—the expected stock return; Ω—the confidence of Q. For the market inputs, the model requires a vector of equilibrium returns.

The trading strategy is applied as follows:

1. For each stock s, at the beginning of quarter q, we predict the rankings of all analysts that we expect to be at the end of the quarter q;
2. Based on these predicted rankings and analysts' price targets, we define $Q_{q,s}$ and $\Omega_{q,s}$;
3. Using market information available at the last day of quarter $q-1$, we obtain the market inputs;
4. Apply BL model to get optimized portfolio weights and buy/sell stocks accordingly;

To measure the performance of our portfolio, we compare it to the baseline which is the market portfolio (S&P500). We compare the relative performance of our portfolio using the Sharpe ratio:

$$SR = \frac{r_p - r_f}{\sigma_p} \tag{12}$$

where r_p is the portfolio quarterly return and r_f is the risk-free rate; σ_p is the standard deviation of the portfolio returns.

5 Data and experimental setup

To implement the trading strategy, we focus on the S&P500 stocks. Given that we base stock views on the analysts' price target information, the period of the strategy experiment runs from the first quarter of 2001 until the last quarter of 2009. We get price target from ThomsonReuters. The list of S&P constituents and stock daily prices data is from DataStream as well as the market capitalization data. The total number of brokers in price target dataset includes 158 brokers covering 448 stocks all of which at some point in time were part of the S&P 500. Given the fact that analysts issue price targets annually, we assume that an analyst keeps her price target forecast valid for one calendar year until it either is revised or expire.

5.1 Target rankings

We build the target rankings of analysts based on the Proportional Mean Absolute Forecast Error (PMAFE) that measures the accuracy of a forecasted price target ξ. First, we define the forecast error (Δ) as an absolute value of the difference between actual price ξ_s and the price target made by an analyst k ($\hat{\xi}_{k,s}$):

$$\Delta_{t,k,s} = |\xi_{t,s} - \hat{\xi}_{t,k,s}| \tag{13}$$

Then, we calculate the average error across analysts as:

$$\overline{\Delta}_{t,s} = \frac{1}{k} \sum_{k=1}^{k} \Delta_{t,k,s} \tag{14}$$

Next, PMAFE is given as:

$$\tilde{\Delta}_{t,k,s} = \frac{\Delta_{t,k,s}}{\overline{\Delta}_{t,s}} \tag{15}$$

5.2 Information sets to define the views

To proceed with the trading strategy, we need to establish which information we will be using to build the rankings. These rankings will be the inputs to compute the weighted return estimates ("smart estimates"). Different analysts' ranks are obtained if we select different time horizons. If we use only the most recent information, we will capture the recent performance of the analysts. This, of course, is more sensitive to unique episodes (e.g., a quarter which has been surprisingly good or bad). If, alternatively, we opt to incorporate the entire analyst's performance, the ranking is less affected by such events, yet it may not reflect the current analyst's ability. We use two information sets: the first uses only the information about the analyst's performance in period $t-1$; the second, uses all the available information for that particular analyst. We call the former the *recent* set and the latter the *all-time* set. We use rankings based on these information sets as the baseline rankings.

In addition to these sets, we also create a hypothetical scenario that assumes we anticipate perfectly the future analyst accuracy performance that would only be available at the end of t. This represents the perfect foresight strategy. The perfect foresight refers to analysts' rankings not stock prices. Therefore, it serves a performance reference point to evaluate the other trading strategies. We call this the *true* set.

6 Experimental Results

The results of the trading strategy based on predicted analysts' rankings are presented in (Table 2).

Table 2. Trading strategy performance

Strategy	Annualized cum. return (in %)	Annualized Std. dev (in %)	Sharpe ratio	Average num. stock	Average turnover rate
Panel A					
Market	-3.032	16.654	-0.182	499	0.053
Panel B: TP					
true	1.785	15.312	0.117	240	0.272
recent	0.634	15.444	0.041	240	0.251
all-time	0.587	15.325	0.038	240	0.238
`last`	0.513	15.478	0.033	240	0.262
`diff`	0.779	15.507	0.050	240	0.269
`random`	0.671	15.474	0.043	240	0.258
`roll.sd`	0.634	15.464	0.041	240	0.264

Panel A reports the performance of *market* (passive strategy). This strategy showed annualized cumulative return of -3.03% and annualized Sharpe ratio of -0.18. The average number of stocks used per quarter is 499.98 and the turnover ratio of strategy is 0.05 which demonstrates the ins/outs of the S&P 500 constituents list.

Panel B of Table 2 demonstrates the results of trading with rankings based on price target. Consistent with our assumption, the *true* resulted in the maximum possible annual cumulative return and the Sharpe ratio (1.78% and 0.12 respectively). This implies that in the settings where analysts' expected returns and rankings are based on price targets, an investor can gain a maximum results from trading strategy. Given the hypothetical assumption of *true*, it is not feasible to implement. The next best strategy is `diff` which is based on our algorithm of predicting the rankings. This strategy resulted in annual cumulative return of 0.78% and the Sharpe ratio of 0.05. In addition, the average per quarter turnover ratio of this strategy of 0.27 implies relative low trading costs.

Figure 1 plots the graphical representation of the cumulative returns for all methods of trading strategy. We see that the *true* strategy is always on top of all the others. We observe that the best outcome was achieved for the strategy based on the first difference of the independent variables.

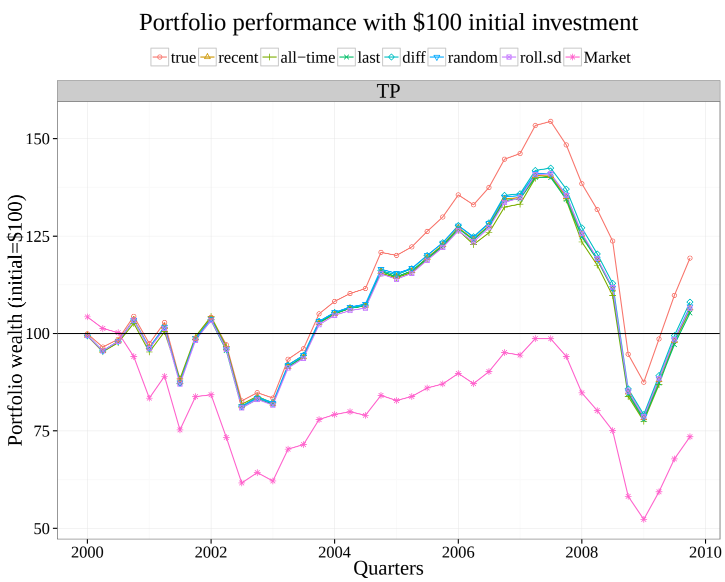

Fig. 1. Performance of the BL model

7 Conclusion

Some institutions, such as StarMine, rank financial analysts based on their accuracy and investment value performance. These rankings are published and are relevant: stocks favored by top-ranked analysts will probably receive more attention from investors. Therefore, there is a growing interest in understanding the relative performance of analysts. In this paper we developed an algorithm that is able to predict the rankings based on state variables that characterize the information environment of the analysts. Further, we designed and operationalized a trading strategy based on the Black-Litterman model with rankings as inputs. We obtained positive successful results from trading that out-performs both the market and the baseline ranking prediction.

References

1. Aiguzhinov, A., Soares, C., Serra, A.: A similarity-based adaptation of naive Bayes for label ranking: Application to the metalearning problem of algorithm recommendation. In: Discovery Science. Lecture Notes in Computer Science, vol. 6332, pp. 16–26 (2010)
2. Black, F., Litterman, R.: Global portfolio optimization. Financial Analysts Journal 48(5), 28–43 (1992)
3. Bouckaert, R.R.: Naive Bayes classifiers that perform well with continuous variables. In: AI 2004: Advances in Artificial Intelligence. Lecture Notes in Computer Science, vol. 3339, pp. 85–116 (2005)
4. Brazdil, P., Soares, C., Costa, J.: Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. Machine Learning 50(3), 251–277 (2003)
5. Cheng, W., Hühn, J., Hüllermeier, E.: Decision tree and instance-based learning for label ranking. In: ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning. pp. 161–168. ACM, New York, NY, USA (2009)
6. Clement, M.: Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics 27(3), 285–303 (1999)
7. Emery, D., Li, X.: Are the Wall Street analyst rankings popularity contests? Journal of Financial and Quantitative Analysis 44(2), 411 (2009)
8. Fama, E.: Efficient capital markets: A review of empirical work. The Journal of Finance 25, 383–417 (1970)
9. Grossman, S., Stiglitz, J.: On the impossibility of informationally efficient prices. American Economic Review 70, 393–408 (1980)
10. Jegadeesh, N., Kim, J., Krische, S., Lee, C.: Analyzing the analysts: When do recommendations add value? The Journal of Finance 59(3), 1083–1124 (2004)
11. Park, C., Stice, E.: Analyst forecasting ability and the stock price reaction to forecast revisions. Review of Accounting Studies 5(3), 259–272 (2000)
12. Sharpe, W.: Mutual fund performance. The Journal of Business 39(1), 119–138 (1966)
13. Vembu, S., Gärtner, T.: Label ranking algorithms: A survey. In: Fürnkranz, J., Hüllermeier, E. (eds.) Preference Learning, pp. 45–64. Springer (2010)
14. Vogt, M., Godden, J., Bajorath, J.: Bayesian interpretation of a distance function for navigating high-dimensional descriptor spaces. Journal of chemical information and modeling 47(1), 39–46 (2007)
15. Womack, K.: Do brokerage analysts' recommendations have investment value? The Journal of Finance 51, 137–168 (1996)
16. Zivot, E., Wang, J.: Modeling financial time series with S-PLUS, vol. 191. Springer Verlag (2003)