

Investigating the use of the method of "nearest neighbors" to predict the behavior of stock prices in the implementation of arbitrage

Perminov G. (National Research University-
Higher School of Economics, Moscow, Russia)

Abstract

In this paper, it was considered one of the types of trading in the stock market - Implementation of arbitrage. The aim of the study was to examine the possibility of using the method of "nearest neighbors" with heuristic rules to predict short-term stock price behavior during arbitrage transactions. In a paper tests the hypothesis that the two parameters - TotalRise (percentage price change over the entire period of growth) and LastChange (percentage change in price over the last day) - are crucial for predicting the behavior of stocks after a sharp rise on positive news. Consequently, an investor might assume, how to behave in the price of the shares, if the result will analyze arbitrage other stocks have close to the value of the shares and TotalRise LastChange.

For this action, the risk of loss was defined as the ratio of the neighbors with a loss to all of its neighbors (if 20 neighbors of the action 5 neighbors were unprofitable (i.e. the respective shares rose in price), the riskiness of the operation can be set equal to 25%). Win value was defined as the average of the gains (and losses) of all the neighbors.

As a result, developed a model is plausible determines the behavior of the shares.

Key words: Technical trading rules, Nearest neighbour predictors, Arbitrage trade.

1. Description of the subject area

The stock market is called the totality of institutions, organizations and participants interacting with each other in the process of trading securities - stocks, bonds, options, futures . In addition to the actual functions of the organization of trade in the stock market is the problem and identifying the most "fair" price of the asset at any given time. At the price of assets is influenced by many factors - macro and micro-economic, political, climatic, etc., and therefore the value of securities fluctuates constantly, is in constant motion. One of the current hypotheses regarding the nature of the behavior of stock prices is a hypothesis about the desire of the market to equilibrium. Speculation assets, i.e. making purchases and sales in order to profit from price differences of transactions on an ideal market equilibrium can not bring any money, since the price movement is subject to the laws of Brownian motion with zero expectation of price change. However, in the real market, as opposed to the ideal, or speculative, as they say, there are arbitrage opportunities. Moreover, it is speculation and are the mechanism that allows the participants of the stock market empirically find the current "fair" price of traded assets. Arbitration is a mechanism to ensure continuous movement of prices to equilibrium.

Among the many types of arbitrage is probably the most common is a temporary arbitration when a trader buys the asset (as they say, is a "long" position) in anticipation of the growth of its price, based on its further resale at a profit. There is also the opposite direction arbitrage transaction where a trader , waiting for the fall of the asset price, it takes a loan from a broker, sells at the same price, and then, after some time, buys the asset at a lower price and returns it to the broker. In this case, the winning trader will be the difference between the original sale price and the purchase price in the future. At the same risk for the trader is that his forecast fall in the price of the asset would be wrong, and he will be forced to buy it back at a lower price.

The aim of the study was to examine the possibility of using the method of "nearest neighbors" (NNP) with heuristic rules to predict short-term stock price behavior during arbitrage transactions.

Arbitration - is making a profit through the use of different prices for the same products or securities. Arbitration is a widespread speculative tactics usually consists of the sale of securities at a relatively high price and the simultaneous purchase of the same securities at a relatively low price.

Let us consider a type of arbitrage associated with the output of positive news about the company. Change the presentation of the news of the fair market price of the shares, and very often the company's shares begin to grow rapidly, as many traders rush to buy them as long as they have not yet exhausted his supply growth. Purchase and subsequent resale of such shares - an obvious arbitrage transaction, the success of which depends on access to news.

But , as often happens , the rapid growth of quotations of the shares may cause them to "re-evaluation" and follow the "adjustment" or "rollback" - fall in the share price due to the fact that short-term traders to close their "long" positions . From this particular share's behavior when leaving positive news follows the possibility of arbitrage using a "short" sale. Short selling (short sell) means a sale of a security borrowed. This operation is often used for speculation, when the investor is bearish on hopes that when he would have to return a security, it will be able to buy it at a lower rate than that at which he sold (Fig. 1).

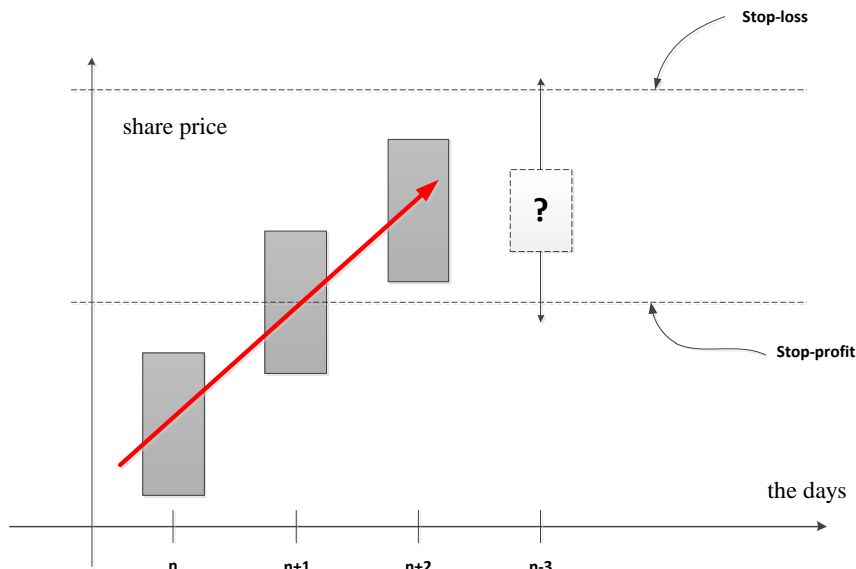


Figure 1 Temporary arbitrage transaction in a short sale

To carry out arbitrage trader must go through several stages.

First, it must select only the volatile stock. Volatility - a statistical indicator of the ability of the market price of hard and fast change over time. If the shares are not sufficiently volatile, the arbitrageurs simply can not earn enough money to weak movements, and all his money will go to the commission to brokers.

Shares for short sales can be selected according to the following criteria:

1. The shares are to be liquid - the amount of the purchase and sale of shares in a day should be significant (more than 300,000 per day). There must be times when you can not quickly buy or sell a large block of shares.
2. The price per share should be $\geq \$5$ - many exchanges do not allow you to sell "short" stocks under \$5.
3. Growth stocks should last more than a day - to increase the likelihood of the situation, "enumeration" of shares. If the stock rises sharply only one day, then it is likely that not all potential buyers have already learned about the positive news, and it is possible to expect continuation of growth and the next day.

For all selected quotations to potential arbitrage opportunity, the investor must determine the values of stop-loss levels and stop-profit. These levels can reduce the risk of the investor.

Thus, the main challenge that the trader is to determine what will be the result of arbitrage - losing or winning.

2. Using nearest neighbor models

In technical analysis when working with stock markets dominated by models that use different types of moving averages. In 2003, Fernandez Rodriguez , Sosvilla - Rivero and Andrada - Felix (2003a, 2003b) found that, given the interest rate and transaction costs, trading rules based on models of the nearest neighbor (NNP) superior results, obtained using the moving average. The idea of applying the methods of NNP is based on the fact that segments of the time series, taken from the past, may have similarities with the future of the segment, this approach falls into the general class of models known as non-parametric regression. Currently used methods of NNT in various versions with trade (empirical) rules (Qian and Rasheed, 2007), Arya S., Mount D. (1993, 1998). Temporary Arbitration is the work of Cao L. J., & Tay F. E. H. (2003) and Nagarajan V., Wu Y., Liu, M., & Wang Q. (2005). It should be noted that recently successfully and effectively develop various hybrid methods (Afolabi MO, & Olude O. (2007); Kim H., & Shin K. (2007); Mandziuk J., & Jaruszewicz M. (2007)).

3. Practical verification of the proposed methodology

In a paper tests the hypothesis that the two parameters - TotalRise (percentage price change over the entire period of growth) and LastChange (percentage change in price over the last day) - are crucial for predicting the behavior of stocks after a sharp rise on positive news. Consequently, an investor might assume, how to behave in the price of the shares, if the result will analyze arbitrage other stocks have close to the value of the shares and TotalRise LastChange.

To this were the following tasks:

1. Unloading quotes.
2. Carrying on the volatility of the stock filter.
3. Application of the method of nearest neighbors.

3.1. Unloading quotes

The source of data (quotes) used a historical database of quotations Quotes Plus and a special DLL to work with this database.

3.2. Carrying on the volatility of the stock filter

As mentioned earlier, the arbitrage opportunities appear only in volatile action. Volatile stocks meet the following criteria:

1. The company's stock is trading more than 50 days (the company to "calm down" after the IPO - Initial Public Offering)
2. Not less than 300,000 of sales and sales per day for the stock.
3. The price per share should be more than \$5.

After filtering the volatility of the resulting sample should be taken of the situation is quite sharp and prolonged growth (over 2 days). The sharp rise in the share price over a few days can be associated with the

speculative boom, and not with the success of the company, which means that at some point may experience a recession prices. In such situations, there is an opportunity for arbitrage trader.

3.3. Application of the method of nearest neighbors

After filtration will get a list of options arbitrage opportunities (company name, date of occurrence of arbitrage opportunities, the price of shares on the first, second, third, ..., n-day percentage price change over the entire period of growth (TotalRise), the percentage price change over the last day (LastChange), and so on).

To conduct the study were created by the following table:

1). Table NASDAQ - contains a list of symbols (tickers) of companies whose shares are traded on the NASDAQ and to be discharged from the system Quotes Plus.

Column Name	Description	Example
Company Name	Ticker Short	MSFT - Microsoft has

2). Table Trade_Data - contains data on quotations unloaded from the system Quotos Plus.

Column Name	Description	Example
Ticker	Short name	ZEUS
Date	The date of formation of the desired pattern of price action (pattern)	23.01.2006
Last Day Change%	Changes in the share price over the last day	13.8289 %
Gap%	What percentage of the share price rose at the opening of the market. The difference between the Open (0) - Close (-1), where Open (0) at the opening price of the day, and Close (-1) on the closing price of the previous day.	2.149 %
MaxLoss,% (0)	The maximum loss during the day Maxloss = (1 - High (0) / Close (0))	-4.5752 %
MaxGain,% (0)	The maximum possible gain for the current day. MaxGain (0) = (1 -Low (0) / Close (0))	1.2465 %
Gain	Winning at the close of the 0 day Gain = (1 -Open (0) / Close (0)).	-0.11 %
Volume	The volume of day trading in units on the last day	988733
Days of Rise	No. days of growth	7
Total Rise,%	Percentage increase the share price over a few days - the closing price of the last day of the fall to the closing price of the last day. TR = (Close (-1) / Close (-8) - 1)	22.30444 %
\$, Million	dollar trading volume = Volume * Close (-1)	4,74
Stoploss	The optimal stoploss, calculated by the method of nearest neighbors	18%
Rew	The average gain of 20 neighbors (not including this piece)	1.2367 %
Risk	Medium Risk by 20 neighbors (not including this piece)	0.3611 %

3). Table Trade_Data(2) - is identical in structure to the table Trade_Data. It will store data for testing.

4). Table Extract - is identical in structure to the table Trade_date. It Trade_data from the table or Trade_data (2) The data will be unloaded c filtered through the fields REW and RISK.

5). Table of Trade - in this table are unloaded Extract data from a table and given a form convenient for plotting trade.

Column Name	Description	Example
%	The percentages of actual win	-2.37%
Date	Date patterning	22.07.2009
20000	The value of gain / loss if the initial amount of investment was \$ 20,000	19642.8401
Ticker	Company Name	ONXX
Down	1 – loss 0 - win	-1
Risk	Probability lose	0.286%
Rew	Potential value gain	3.32%

To handle the tables were written by the following methods:

1). Load_data method performs the following operations:

1. Takes the name of the table NASDAQ companies.
2. With the help of a specialized library terminates processes Quotes Plus volatile stocks, that is, satisfying the following criteria:
 - a. the stock was trading for more than 50 days;
 - b. the amount of the purchase and sales of shares of more than 300,000 pieces a day;
 - c. share price more than \$5;
 - d. growing share of two days or more;
 - e. the last day of the share price rose to not less than 3%.
3. The selected data is loaded into the table Trade_data.

2). The method of K-nn - finds the optimal value of the index stoploss for all the shares, as well as the expected value of risk (Risk) and gain (Rew). To search for all values of the nearest neighbor method is used (Fig. 2). As the coordinate system used by field TotalRise (total height of the share price over the last few days of continuous growth) and LastDayRise (growth of the share price over the last day). The method according to the values and TotalRise LastDayRise is 20 nearest neighbors, and it is believed that in the calculation of the Euclidean distance to the nearest neighbor and TotalRise LaseDayRise considered with the same weight.

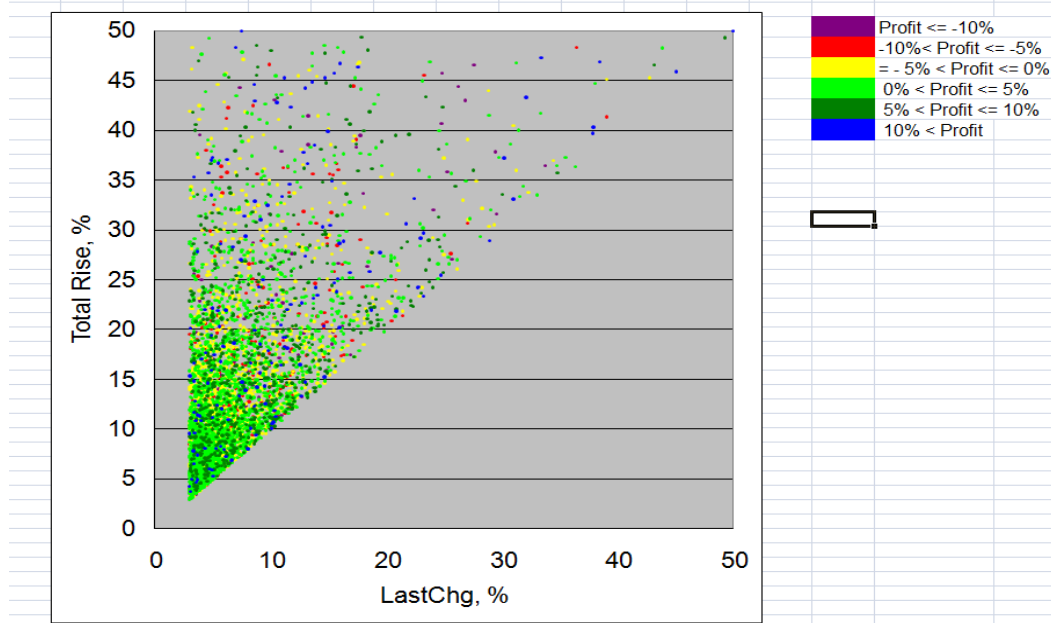


Fig. 2 Applications of the method of nearest neighbors, to estimate the values of gain and risk

3). Method Extract - Trade_data selects from the table and writes the data table Extract stock in which risks , in accordance with the method of the nearest neighbors does not exceed 30% and the expected payoff is greater than 2%.

4). Method Trade - unloads data from the table in the table Extract Trading. Put down the initial amount. In our case, the amount is \$20,000. Then, for each row is a comparison between the actual and the expected gain. If expectations are not met, the quote was closed at a high price , the investment sum will be reduced by the percentage of loss. If the expectation proved to be correct, and managed to get a win in arbitrage, the amount is added to the invested amount of the winnings. After processing all the lines drawn chart trading.

4. The practical implementation of the proposed methodology

The implementation of the proposed method is shown on a real example.

4.1. After opening the table calls the NASDAQ Load_data. As a result of the method Trade_date table will be filled with 7000 lines (Fig. 3). Each row represents a different specification of arbitrage opportunities. It filled all the columns except for columns Stoploss, Rew and Risk. The values in these fields will be added in the performance of the method K-nn. To test were selected last 2000 lines and copied to the table Trade_data (2).

1	Ticker	Date	Last Day Change%	Open(0)	Gap %	MaxLoss,%(0)	MaxGain,%(0)	Close(0)	Gain	Volume	Days of Rise	Total Rise,%	\$, Million	StopLoss	Rew	Risk
2	ZEUS	1/23/2006	13.0309	28	2.15	-4.575	1.246	28.03	-0.11	988733	2	14.2083	27.10			
3	SURG	1/24/2006	11.1111	6.14	2.33	-1.629	8.306	5.92	3.55	428350	2	14.2857	2.57			
4	CRAY	1/24/2006	12.2642	9.76	2.52	-2.459	1.639	9.64	1.23	950773	2	12.2642	9.05			
5	HILL	1/24/2006	6.91058	8.14	3.17	-0.123	7.862	7.60	6.63	847203	2	7.78888	6.68			
6	SBIB	1/24/2006	5.05691	11.63	4.99	-0.287	3.61	11.35	2.46	1637430	2	5.85989	18.14			
7	ISRG	1/24/2006	3.50362	138.3	3.1	-0.174	5.337	132.29	4.33	770449	2	4.7158	103.33			
8	SLAB	1/24/2006	3.72038	48.27	16.2	-5.718	0.269	49.04	-1.60	2191948	2	4.06307	91.05			
9	PNRA	1/24/2006	3.09539	67.33	3.14	-1.485	0.401	67.67	-0.50	567658	2	3.86635	37.06			
10	PWER	1/25/2006	9.22509	8.11	3.21	-1.473	2.782	6.14	-0.49	1086169	3	12.9771	6.43			
11	BOLT	1/25/2006	3.3275	12.23	3.67	0	5.45	11.96	2.23	588450	2	8.58897	6.94			
12	NETL	1/25/2006	4.09707	17.17	3.94	-2.68	8.04	17.63	-2.68	1277680	2	8.50854	21.10			
13	JACK	1/25/2006	4.67025	19.27	4.25	-10.84	0.467	20.77	-7.78	1029800	2	5.95988	19.04			
14	ISIL	1/25/2006	3.50809	28.25	5.21	-4.673	1.735	28.89	-2.27	4246019	2	5.7503	114.01			
15	COGO	1/26/2006	3.18022	9.1	3.88	-0.769	5.385	8.73	4.07	1213436	4	21.3296	10.63			
16	TASR	1/26/2006	8.88517	9.97	2.08	0	4.413	9.66	3.11	6592070	3	21.3292	64.38			
17	VDSI	1/26/2006	9.40695	11.23	4.95	-0.801	2.493	11.01	1.96	963558	3	15.6757	10.31			
18	NDQA	1/26/2006	11.4264	44.24	2.41	-0.588	5.289	43.55	1.56	3502227	3	15.3538	151.30			
19	ZOLT	1/26/2006	7.88177	13.53	2.97	-0.148	2.661	13.42	0.81	812612	2	11.6398	10.68			
20	SPIL	1/26/2006	4.65444	5.993	2.29	-4.743	1.186	6.09	-1.58	2777038	4	10.9118	16.27			
21	EGOV	1/26/2006	5.10204	6.39	3.4	-3.13	3.599	6.29	1.56	569392	4	9.57447	3.52			
22	CTRP	1/26/2006	6.6179	7.986	2.42	-0.078	2.176	7.91	1.00	3450552	3	9.22781	26.91			
23	BEBE	1/26/2006	4.92578	17.44	12.2	-7.798	1.95	18.53	-6.25	1498016	4	5.7104	23.29			
24	GROW	1/27/2006	4.15512	9.625	2.39	-1.299	10.91	8.81	8.42	661522	4	24.2564	6.22			
25	NETL	1/27/2006	4.11347	18.95	3.3	-0.106	2.796	18.81	0.79	1039640	4	20.565	19.08			
26	SMSI	1/27/2006	10.6299	8.64	2.49	-1.157	5.566	8.25	4.51	456685	2	14.3826	3.85			
27	CTRN	1/27/2006	6.18322	48	2.76	-1.354	3	46.96	2.17	477397	2	9.49367	22.30			
28	BTUI	1/27/2006	6.66199	15.78	3.75	-2.978	3.042	15.36	2.66	419110	2	8.25622	6.37			
29	ANAD	1/27/2006	4.80621	6.91	2.22	0	5.933	6.55	5.21	911115	3	7.30159	6.16			
30	FCFS	1/27/2006	5.03553	17.61	3.59	0	2.953	17.43	1.02	2596568	2	6.75039	44.14			
31	XRTX	1/27/2006	5.99656	22.5	3.5	-3.867	3.111	21.92	2.58	529753	2	6.25612	11.52			
32	RIGL	1/27/2006	3.12926	7.8	2.9	-1.026	3.205	7.85	-0.64	457348	2	4.69813	3.47			
33	SSRI	1/30/2006	3.41544	18.96	2.65	-1.16	4.536	18.35	3.22	1295217	5	21.1148	23.92			
34	REDF	1/30/2006	8.44287	21.9	2.1	-2.055	4.064	21.76	0.64	2428186	2	14.2173	52.08			
35	PPHM	1/30/2006	3.44828	6.175	2.92	-13.36	0.405	6.95	-12.55	386293	2	8.1081	2.32			
36	WIRE	1/31/2006	6.34298	27.99	3.06	-3.323	1.858	28.92	-3.32	588809	3	13.2138	15.99			
37	CHNR	2/1/2006	22.6326	19.75	8.16	-9.873	7.089	20.90	-5.82	2720114	2	59.0592	49.67			
38	ACAD	2/1/2006	10.728	15.17	4.98	-7.119	1.187	15.40	-1.52	2706370	3	41.1133	39.11			
39	EXFO	2/1/2006	10.6178	5.95	3.84	-0.84	7.227	5.67	4.71	511089	7	20.6316	2.93			
40	RIMM	2/1/2006	9.0034	25.07	2.16	0	5.253	24.00	4.26	5.5E+07	3	11.9373	1353.07			
41	CTDC	2/2/2006	104.202	12.47	2.63	-16.28	23.5	9.64	22.69	1E+07	2	247.133	124.51			
42	HIHO	2/2/2006	37.6558	5.6	2.21	-19.64	11.79	5.00	10.71	637144	8	63.5433	3.49			
43	CAAS	2/2/2006	46.0298	11.5	5.99	-22.17	16.78	9.98	13.22	1695177	3	53.4653	18.39			
44	CHDX	2/2/2006	13.0081	5.8	4.32	0	21.84	4.75	18.05	1538839	2	44.2906	8.56			
45	JRJC	2/2/2006	18.7126	8.38	5.67	-2.864	14.32	7.18	14.31	1165676	2	22	9.24			
46	FTEK	2/3/2006	11.043	10.81	4.68	-6.383	1.48	11.30	-4.53	600064	2	12.25	6.20			
47	COGO	2/3/2006	4.28101	9.76	2.74	-2.459	3.074	9.94	-1.84	1185525	2	12.1606	11.26			
48	REDF	2/3/2006	4.31267	24.4	5.08	-18.03	2.664	28.35	-16.19	1440089	2	11.9036	33.44			
49	GERN	2/6/2006	6.20155	8.425	2.49	-0.89	5.045	8.00	5.04	2527024	2	7.73264	20.77			
50	TXRH	2/6/2006	4.72653	16.55	6.71	-0.544	3.263	16.23	1.93	739578	2	6.96552	11.47			
51	ISIS	2/6/2006	5.36672	6.13	4.07	-3.426	8.972	5.59	8.81	912159	2	5.36672	5.37			
52	SIRI	2/7/2006	4.91804	5.97	3.65	-2.848	1.174	5.98	-0.17	7.9E+07	2	8.06755	454.08			
53	ZIGO	2/7/2006	5.37765	17.89	2.58	-2.739	1.789	18.19	-1.68	474344	2	6.40635	8.27			
54	ENDP	2/7/2006	6.40868	27.84	3.11	-0.539	1.473	27.75	0.32	1620905	3	6.34108	43.76			

Figure 3. Fragment of Trade_data

4.2. Next, press the button k-nn (Figure 3). You will see two dialog boxes for entering parameters (Fig. 4). The first dialog box prompts you to enter the value of stoploss. Stoploss value can be set firmly on all transactions , or set to "-1" . In this case , the method calculates the optimum value for each operation stoploss. In the second dialog box, you enter values stopprofit. As a result, the method will fill the columns stoploss, risk and rew. To calculate the value of a column used rew gain the 20 nearest neighbors and possible losses when exiting some of those neighbors stoploss (Fig. 5).

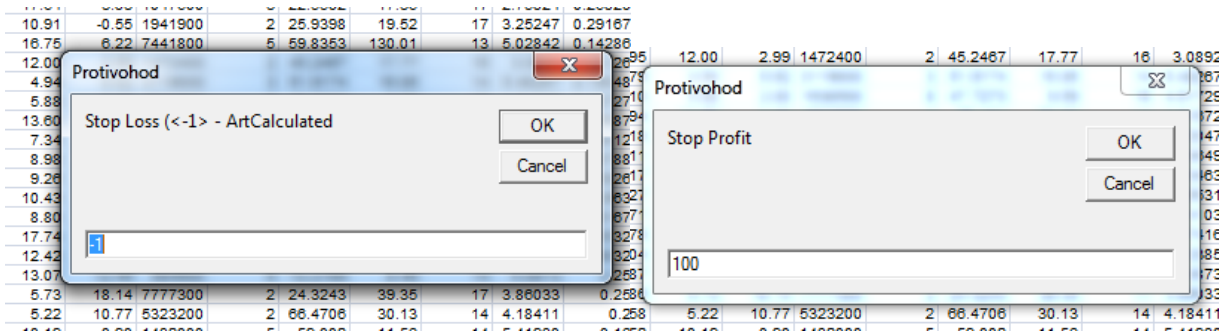


Figure 4. Parameter settings and stoploss stopprofit

	A	B	D	I	J	K	L	M	O	Z	AA	AB	AC	AT	AU	AV	Z
1	Ticker	Date	Last Day Change%	Open(0)	Gap %	MaxLoss,%(0)	MaxGain,%(0)	Close(0)	Gain	Volume	Days of Rise	Total Rise,%	\$, Million	StopLoss	Rew	Risk	
2	KONG	7/17/2009	6.34191	11.84	2.33	-1.351	2.449	11.89	-0.42	409300	7	22.3044	4.74	18	1.237	0.361	
3	YHOO	7/17/2009	4.01483	17.18	2.02	-1.455	3.085	17.01	0.99	3.3E+07	7	17.1071	547.55	17	0.841	0.399	
4	WRPT	7/17/2009	5.44218	9.5	2.15	-4.211	1.263	9.48	0.21	315900	6	15.3846	2.94	13	1.04	0.397	
5	AMED	7/17/2009	3.46988	37.16	2.14	-0.188	2.395	36.97	0.51	797200	3	10.7458	29.00	15	0.412	0.438	
6	SYNA	7/17/2009	3.10786	34.75	2.69	-1.353	1.151	34.93	-0.52	1931600	3	4.09106	65.37	13	-0.15	0.511	
7	CLWR	7/20/2009	8.56698	7.19	3.16	-1.53	2.503	7.24	-0.70	4022100	8	32.2581	28.03	17	1.027	0.461	
8	PWRD	7/20/2009	10.3057	35.75	2.14	-0.14	8.224	33.50	6.29	3473500	5	21.6122	121.57	17	1.828	0.367	
9	KIRK	7/20/2009	4.75423	13.35	2.69	-1.873	4.419	13.50	-1.12	516400	6	20.3704	6.71	14	1.207	0.398	
10	CRNT	7/20/2009	4.60705	7.88	2.07	-0.508	7.995	7.40	6.09	552600	6	17.6829	4.27	14	1.196	0.398	
11	SPP1	7/20/2009	5.76576	6.25	6.47	-5.92	4	6.59	-5.44	2055600	2	12.2371	12.07	13	0.367	0.455	
12	HGSI	7/21/2009	10.6315	14.5	4.77	-2.276	7.931	14.05	3.10	8.2E+07	2	316.867	1140.06	13	1.538	0.426	
13	OREX	7/21/2009	4.44444	7.68	2.13	-23.31	0.391	8.27	-7.68	2538900	3	37.9817	19.09	13	2.668	0.242	
14	SPAR	7/21/2009	3.36957	11.03	16	-0.272	10.52	9.88	10.43	615400	6	14.8551	5.85	17	0.933	0.389	
15	ONXX	7/22/2009	21.0178	35.49	2.22	-3.409	2.001	36.33	-2.37	1E+07	6	26.6229	348.19	19	3.317	0.286	
16	MKSI	7/22/2009	11.9213	19.75	2.12	-1.873	1.924	19.87	-0.61	924900	4	15.5317	17.89	20	1.803	0.366	
17	AATI	7/22/2009	5.02092	5.35	6.57	-4.488	9.346	5.00	6.54	498800	2	14.6119	2.50	13	0.919	0.405	
18	MMLP	7/22/2009	8.24786	26	2.65	-3.423	0	26.50	-1.92	323500	3	10.708	8.19	17	0.548	0.41	
19	TESO	7/22/2009	5.61942	9	8.83	-13.56	0.222	9.81	-9.00	986000	3	8.8158	8.15	19	0.402	0.424	
20	CWTR	7/23/2009	11.3677	6.5	3.67	-5.538	3.692	6.85	-5.38	3448400	2	14.4161	21.62	20	1.479	0.377	
21	MEOH	7/23/2009	6.20314	16.16	3.72	-2.97	3.166	16.53	-2.29	816300	3	11.3653	12.72	14	0.148	0.442	
22	SOEN	7/23/2009	6.8596	10.51	5.42	-9.42	5.138	11.34	-7.90	1201100	4	10.7778	11.97	14	0.196	0.422	
23	SIVB	7/23/2009	5.68097	29.52	5.09	-9.722	0.237	32.09	-8.71	570000	2	6.92805	16.01	19	0.528	0.424	
24	ENER	7/23/2009	3.85208	14.13	4.82	-7.928	1.982	15.22	-7.71	3386000	3	5.64263	45.64	20	-0.05	0.472	
25	BCRZ	7/24/2009	10.6061	9.99	5.27	-12.11	0.601	10.95	-9.61	4811500	11	159.29	45.66	17	3.01	0.313	
26	INCY	7/24/2009	6.39098	5.9	4.24	-0.847	3.729	5.84	1.02	3565500	2	27.7652	20.18	20	1.814	0.378	
27	CSUN	7/24/2009	11.8952	5.76	3.78	-2.431	6.597	5.44	5.56	3368100	2	20.3905	18.69	15	1.605	0.372	
28	SQNM	7/24/2009	8.98438	5.87	5.2	-14.48	0.852	6.35	-8.18	1.5E+07	2	19.7425	82.09	17	1.191	0.38	
29	ENER	7/24/2009	12.908	15.54	2.1	-0.965	1.866	15.55	-0.06	5872800	4	19.279	89.38	15	1.389	0.383	
30	JASO	7/24/2009	11.2971	5.55	4.32	0	8.108	5.17	6.85	1.1E+07	2	11.7647	60.55	15	1.499	0.338	
31	GROW	7/24/2009	3.11186	12.6	2.77	0	4.762	12.50	0.79	312700	3	9.6601	3.83	13	0.059	0.466	
32	OXPS	7/24/2009	7.44337	17.26	3.98	-0.29	3.65	16.72	3.13	1210600	2	8.85246	20.10	14	0.572	0.391	
33	JAZZ	7/27/2009	43.8356	6.61	4.92	-15.73	3.631	6.58	0.45	8953600	3	48.2353	56.41	19	1.043	0.364	
34	HITK	7/27/2009	7.31071	17.35	5.54	-4.15	2.997	17.73	-2.19	490800	4	20	8.07	15	1.085	0.363	
35	CISG	7/27/2009	5.69883	19.78	2.54	0	4.196	19.28	2.53	556400	2	18.6347	10.73	14	1.176	0.383	
36	MPEL	7/27/2009	4.14313	5.7	3.07	0	6.491	5.60	1.75	3216500	4	5.73614	17.79	20	-0.1	0.471	
37	HITK	7/28/2009	7.84671	18.26	2.99	-0.055	16.21	15.86	13.14	911200	5	29.4161	16.16	18	1.301	0.425	
38	WTFC	7/28/2009	12.313	20.28	3.89	-13.36	0.049	22.61	-11.49	700600	4	28.5056	13.68	15	1.646	0.356	
39	KLIC	7/29/2009	15.6313	5.89	2.08	-1.188	2.547	5.85	0.68	3413600	7	36.7299	19.70	13	1.262	0.4	
40	UEPS	7/29/2009	10.0335	16.84	2.37	-0.891	1.425	16.82	0.12	933000	8	27.027	15.35	17	0.59	0.42	
41	SYNT	7/29/2009	10.1081	42.22	3.63	-3.768	5.59	40.34	4.45	634500	2	11.3115	25.85	17	1.349	0.358	
42	CY	7/29/2009	3.17757	11.27	2.08	-3.461	3.283	10.91	3.19	1.1E+07	2	3.56473	122.74	14	-0.3	0.509	
43	FIRE	7/30/2009	18.9519	16.99	2.53	-7.357	5.65	17.64	-3.83	1047600	3	22.5592	17.36	17	2.758	0.268	
44	ENDP	7/30/2009	3.05569	21.56	3.11	-1.902	2.69	21.01	2.55	2524000	2	4.08163	52.78	13	-0.05	0.503	
45	CENX	7/31/2009	8.97289	9.04	7.88	-10.07	1.659	9.66	-6.86	8622800	2	15.5862	72.26	14	1.22	0.394	
46	DPTR	7/31/2009	10.3448	20.6	7.29	-0.971	4.854	19.90	3.40	760440	2	14.9701	14.60	17	1.177	0.422	
47	MPEL	7/31/2009	6.71785	5.71	2.7	-5.954	1.926	6.03	-5.60	3584900	2	11.4228	19.93	14	0.273	0.435	
48	SIRO	7/31/2009	8.29166	28.66	10.3	-0.593	7.223	28.00	2.30	922600	2	10.6899	23.98	17	0.576	0.409	
49	CAR	7/31/2009	3.38572	8.81	3.04	-2.27	2.157	8.84	-0.34	2312200	2	8.50254	19.77	18	0.162	0.448	
50	AFFX	7/31/2009	5.48927	9.04	2.26	-6.748	1.991	9.50	-5.09	1905300	2	6.76329	16.84	19	0.386	0.444	
51	PAAS	7/31/2009	5.23225	20.35	3.25	-3.047	0	20.73	-1.87	1173900	2	6.59816	23.14	20	0.257	0.451	
52	SSRI	7/31/2009	3.47185	20.47	4.07	-2.345	2.052	20.86	-1.91	543900	2	6.55472	10.70	20	-0.16	0.481	
53	FWLT	7/31/2009	3.03301	23.79	2.99	-0.757	1.892	23.49	1.26	1729800	2	5.33515	39.96	14	0.332	0.431	
54	GOLD	7/31/2009	3.87959	64.33	3.56	-1.663	0.466	64.99	-1.03	1255900	2	4.10591	78.02	18	-0.35	0.507	

	A	B	C	D	E	F	G	H	I	J
1	%	Date	20000	Ticker	Chg0	\$/<\$>	Down	Risk	Rew	
2	-5.82	2/1/2006	18795.444	CHNR	22.6	59.1	1	0.23	2.37	
3	-19	2/2/2006	15184.31	CAAS	46.0	53.5	1	0.29	2.83	
4	-3.38	2/22/2006	14631.196	GKNT	17.2	101.0	1	0.29	2.32	
5	-2.99	3/2/2006	14154.444	INAP	20.4	22.6	1	0.26	3.42	
6	9.55	3/6/2006	15465.873	AFOP	14.7	47.7	0	0.24	4.30	
7	3.33	5/17/2006	15941.403	LTXC	22.9	28.4	0	0.26	3.48	
8	-10.9	5/30/2006	14155.993	SGMO	3.9	39.5	1	0.2	2.84	
9	10.5	5/31/2006	15604.188	SGMO	15.2	60.7	0	0.17	5.11	
10	-6.3	8/15/2006	14580.542	KONG	21.4	21.4	1	0.29	3.54	
11	6.4	8/23/2006	15473.018	CTIC	22.3	22.3	0	0.27	3.60	
12	16.3	9/11/2006	17947.384	ANGN	109.4	131.6	0	0.27	4.96	
13	4.66	11/13/2006	18743.874	CHDX	21.4	23.2	0	0.28	3.46	
14	-0.91	11/17/2006	18532.374	JST	34.4	37.0	1	0.28	2.31	
15	5.19	1/8/2007	19453.311	SGEN	24.0	24.2	0	0.29	3.27	
16	6.02	3/22/2007	20584.112	MSON	20.5	23.4	0	0.26	3.40	
17	5.33	3/27/2007	21641.236	ABIO	24.9	25.7	0	0.26	3.74	
18	-7.92	3/30/2007	19886.269	DNDN	147.7	249.5	1	0.27	6.83	
19	38	4/18/2007	27402.419	AVNR	308.7	332.5	0	0.29	4.69	
20	1.53	5/4/2007	27782.023	CROX	19.9	25.0	0	0.28	3.43	
21	12.6	5/18/2007	31248.492	OMEX	80.9	106.5	0	0.28	4.49	
22	0.69	5/31/2007	31423.575	ARUN	4.6	39.0	0	0.28	2.41	
23	4.03	6/19/2007	32649.724	CMED	3.0	28.8	0	0.21	2.77	
24	-5.71	7/5/2007	30744.026	BCON	11.6	59.8	1	0.17	4.42	
25	-15	7/6/2007	26092.422	BCON	13.5	81.4	1	0.29	4.10	
26	9.19	7/9/2007	28450.953	CHNR	14.5	55.0	0	0.16	5.38	
27	-4	8/3/2007	27272.916	HMSY	19.7	22.1	1	0.27	2.76	
28	-19	9/5/2007	22051.062	CRDC	49.1	54.8	1	0.23	3.68	
29	13.2	9/20/2007	24915.457	EFUT	49.9	58.1	0	0.27	2.18	
30	3.46	9/21/2007	25738.14	ASTI	19.0	70.1	0	0.23	5.46	
31	-0.25	9/25/2007	25633.688	FMCN	19.2	24.5	1	0.27	3.06	
32	15	9/26/2007	29449.042	CDII	18.7	66.7	0	0.19	5.29	
33	16.4	9/27/2007	34231.323	CNTF	12.5	56.4	0	0.29	2.46	
34	9.7	10/3/2007	37512.006	CDII	20.4	26.9	0	0.29	3.31	
35	0.32	10/9/2007	37593.797	LLNW	3.2	28.4	0	0.19	2.75	
36	-1.71	10/24/2007	36910.81	ISRG	4.3	34.1	1	0.26	2.33	
37	12.5	11/16/2007	41487.138	CSIQ	14.7	77.4	0	0.27	3.56	
38	-1.11	11/29/2007	40984.888	REGN	23.1	24.2	1	0.25	3.68	
39	-1.43	11/30/2007	40356.762	REGN	3.2	28.2	1	0.17	2.47	
40	-1.71	12/6/2007	39626.793	PWRD	3.0	26.7	1	0.17	2.53	
41	3.58	12/7/2007	41006.865	ZIXI	15.8	48.3	0	0.22	4.68	
42	-0.37	12/13/2007	40816.82	RIGL	224.4	236.6	1	0.28	5.12	
43	2.35	12/19/2007	41736.187	PEIXD	18.4	51.5	0	0.19	4.74	
44	-9.36	12/26/2007	37790.604	CSII	39.0	41.3	1	0.24	4.14	

Figure. 7. Table View Trading (on training data)

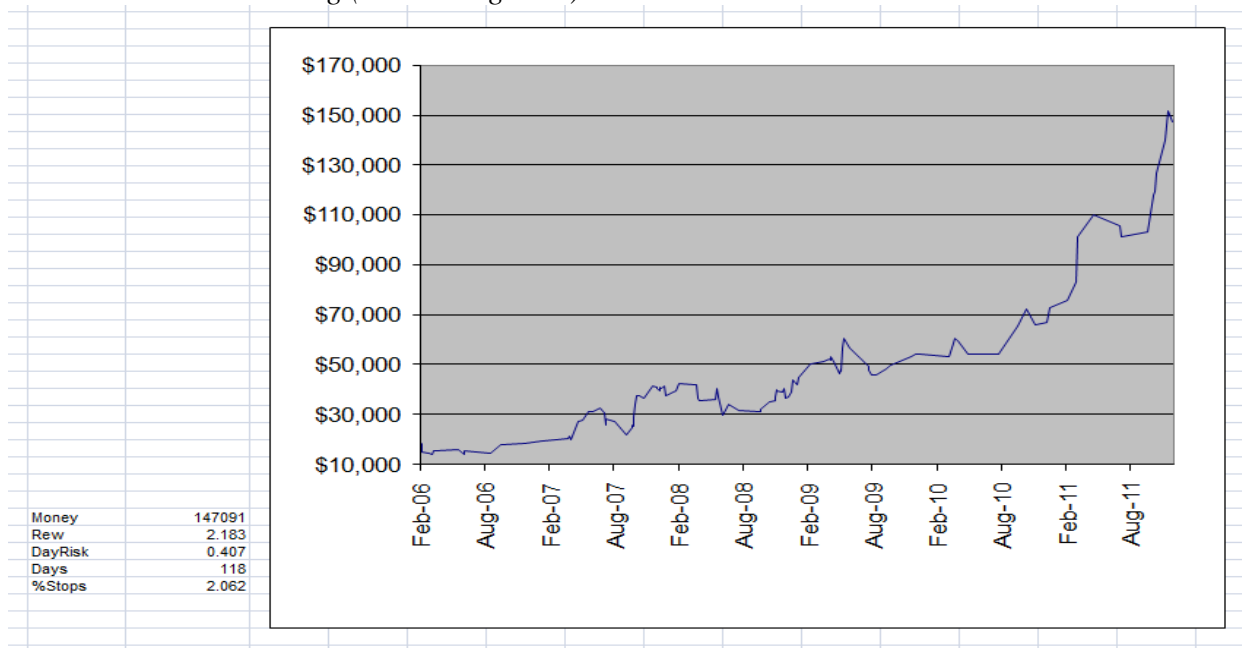


Figure 8. Schedule of trade (on training data)

4.6. Now, to test models of repeat trade, but for the data from the table Trade_data (2) (Fig. 9).

	A	B	C	D	E	F	G	H	I	J
1	%	Date	20000	Ticker	Chg0	\$/<\$>	Down	Risk	Rew	
2	-13	7/21/2009	17360	OREX	4.4	38.0	1	0.24	2.67	
3	-2.37	7/22/2009	16909.112	ONXX	21.0	26.6	1	0.29	3.32	
4	-3.83	7/30/2009	16222.208	FIRE	19.0	22.6	1	0.27	2.76	
5	-0.55	8/10/2009	16092.501	ELON	21.1	25.9	1	0.29	3.25	
6	6.22	9/10/2009	17052.651	ATPG	16.1	59.8	0	0.14	5.03	
7	2.99	9/22/2009	17522.714	FFBC	17.3	45.2	0	0.26	3.09	
8	6.62	11/17/2009	18642.061	NABI	14.4	51.9	0	0.19	5.48	
9	2	11/30/2009	18974.902	CYTX	13.6	47.7	0	0.23	4.52	
10	0	12/16/2009	18934.902	DPTR	20.8	68.4	0	0.26	3.62	
11	-1.8	3/4/2010	18553.496	CWTR	24.0	29.3	1	0.29	2.72	
12	10.2	3/17/2010	20405.953	KONG	23.2	29.7	0	0.29	3.11	
13	3.54	3/24/2010	21088.664	POZN	3.1	34.0	0	0.26	2.53	
14	-1.36	4/1/2010	20761.743	PEXD	26.1	26.1	1	0.3	3.26	
15	-8.78	4/30/2010	18899.636	PWER	36.5	48.3	1	0.29	2.14	
16	-0.51	7/26/2010	18763.263	IDSA	13.4	46.1	1	0.26	4.27	
17	2.44	7/30/2010	19180.184	PWER	24.8	27.9	0	0.29	3.12	
18	12.5	8/11/2010	21529.678	CALL	18.2	72.2	0	0.25	5.09	
19	18.1	9/17/2010	25395.776	CHTP	22.8	24.3	0	0.25	3.86	
20	10.8	10/12/2010	28090.706	ASTI	23.9	66.5	0	0.2	4.18	
21	-8.98	11/4/2010	25527.052	AERL	17.5	59.8	1	0.16	5.42	
22	1.89	12/10/2010	25968.696	LOCM	20.9	26.5	0	0.28	3.36	
23	8.57	12/17/2010	28154.584	NYMX	50.0	51.4	0	0.27	3.38	
24	4.31	2/2/2011	29329.016	APKT	20.4	25.4	0	0.26	3.58	
25	6.33	2/23/2011	31146.785	BDCO	80.7	181.7	0	0.29	4.43	
26	9.3	2/28/2011	34004.16	DPTR	18.2	60.3	0	0.18	5.48	
27	22.2	3/7/2011	41520.641	ATRN	90.9	126.7	0	0.26	5.05	
28	8.35	4/19/2011	44946.712	REDF	18.4	48.0	0	0.22	4.23	
29	-3.77	7/5/2011	43213.601	SODA	3.3	33.8	1	0.21	2.78	
30	-3.97	7/8/2011	41459.138	JVA	19.7	72.6	1	0.29	4.65	
31	12.5	7/18/2011	46594.262	REDF	3.7	27.9	0	0.21	2.10	
32	1.79	9/20/2011	47388.785	KONA	22.8	36.0	0	0.28	2.23	
33	14.3	10/6/2011	54118.611	YRCW	50.0	50.0	0	0.29	2.34	
34	1.43	10/11/2011	54853.72	GPOR	3.3	26.8	0	0.24	2.10	
35	6.01	10/13/2011	58112.113	TQNT	24.6	28.3	0	0.29	3.06	
36	-0.73	11/4/2011	57650.567	INHX	115.7	125.3	1	0.27	5.17	
37	10.5	11/7/2011	63656.088	TSTC	14.8	50.3	0	0.23	4.36	
38	8.39	11/16/2011	68959.146	MU	23.4	26.0	0	0.26	3.43	
39	-3.03	11/29/2011	66826.526	INHX	20.2	22.9	1	0.28	3.14	
40			66786.526							
41			66746.526							

Figure 9. Trading testing a table with data from the table Trade_data (2)

As can be seen from a comparison of Figures 7 and 9, the trade was successful. Moreover, the results of the test were trading even more successful than the results of the training of trade, which shows a certain margin of safety of the method. Thus, the hypothesis that the nearest neighbor method can be used to search for short-term arbitrage opportunities seems plausible.

References:

Fernández-Rodríguez F., Sosvilla-Rivero S. and J. Andrada-Félix (2003a): "Nearest-neighbour predictions in foreign exchange markets", in Shu-Heng Chen and Paul Wang, eds., Computational Intelligence in Economics and Finance, Berlin: Physica Verlag.

Fernández-Rodríguez F., Sosvilla-Rivero S. and J. Andrada-Félix (2003b): "Technical Analysis in Foreign Exchange Markets: Evidence from the EMS", Applied Financial Economics, 13, 113-122.

Qian B., Rasheed, K. (2007). Hurst exponent and financial market predictability. Proceedings of the 2nd IASTED international conference on financial engineering and applications. Cambridge, MA, USA, 203–209.

Arya S., Mount D. (1993): Approximate nearest neighbor queries in fixed dimensions. In: Fifth Annual ACM-SIAM Symposium on Discrete Algorithms. (1993) 271–280.

Arya S., Mount D., Netanyahu N., Silverman R., Wu A. (1998): An optimal algorithm for approximate nearest neighbor searching. *Journal of the ACM* 45 (1998) 891–923.

Cao L. J., & Tay F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, 14(6), 1506–1518.

Nagarajan V., Wu Y., Liu, M., & Wang Q. (2005). Forecast studies for financial markets using technical analysis. In *International Conference on Control and Automation (ICCA)* (pp. 259–264).

Afolabi M. O., & Olude O. (2007). Predicting stock prices using a hybrid Kohonen self organizing map (SOM). In *Proceedings of the 40th Hawaii international conference on system sciences* (p. 48).

Kim H., & Shin K. (2007). A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. *Applied Soft Computing*, 7, 569–576.

Mandziuk J., & Jaruszewicz M. (2007). Neuro-evolutionary approach to stock market prediction. In *Proceedings of international joint conference on neural networks*, Orlando, Florida, USA (pp. 2515–2520).