Pair Trading Strategies using Machine Learning: A Case of PSX Firms

Muhammad Khalid Sohail*, Abdul Raheman**, Iftikhar Hussain Adil***, M. Faisal Rizwan****, Safi Ullah Khan*****

Abstract

Pair trading is generally known as a profitable strategy of an investment. Here, in this study DBSCAN clustering algorithm (machine learning) in addition to the traditional pair trading technique. is used. By using this algorithm 3 cluster are identified from EPS, Market-Cap, sector classification and BVS (fundamental variables) with other factors formed from PCA on the returns of daily data of two years of the sample firms. Pairs are also formed based on traditional distance approach of Gatev et al. (2006). *Sample consists of 80 stocks from five different sectors; banking, chemicals, cement, textile, food and care products from year 2011 to 2019.Under the machine learning a remarkable 1.16% average excess monthly return with Sharpe ratio of 2.48 is to be observed. For risk adjusted returns, Jensen's alpha under CAPM is also to be observed positive and significant. The results also authenticate mean revision and market neutrality at PSX. Investors can get positive returns through pair trading at PSX. Portfolio and fund managers can form the pair trading strategy to reap the profitability for their clients and specially they can get higher returns by using machine learning approach.*

Keywords: Machine learning, Pair trading, Jensen's alpha, Distance, DBSCAN, PSX, PCA.

JEL Classification: F120

1. Introduction

Pairs trading means finding the pair of investments (stocks, indices, currencies etc.) that were closely related in past (historical data) and taking long and short positions in pairs when their prices diverge and close the positions when prices converge. Pair trading is contrarian strategy as when stock diverges, winner stock is sold while loser is bought.

*Faculty Member, Department of Management Studies, Bahria University, Islamabad, Pakistan.
Email: ksohail.buic@bahria.edu.pkand mkhalids33@live.com

**Faculty member, Faculty of Management Sciences, International Islamic University, Islamabad.
Email: abdul.raheman@iiu.edu.pk

***Faculty member, School of Social Sciences and Humanities (S3H), NUST, Islamabad, Pakistan.
Email: iftikhar.adil@s3h.nust.edu.pk

****Faculty member, Department of Management Sciences, Islamic International University, Islamabad,Pakistan.
Email: faisal.rizwan@iiu.edu.pk

*****Faculty member, Universiti Teknologi Brunei, Brunei Darussalam. Email : safiullah.khan@utb.edu.bn

Pair trading methodology was initially used by Morgan Stanley in 1983 to make profit from market inefficiencies. The theme is based on mean reversion theory. Different methods: stochastic spread, distance cointegration and recently clustering approaches are used for pair trading. However, the literature is deficient regarding the use of machine learning in pair trading.

Machine learning, nowadays, is the area that was used by the researchers in social sciences like finance, accounting, etc. (Lin, Hu, & Tsai, 2012; Mitchell, 2006; Huang et al., 2004; Ecer, 2013). According to Mitchell (2006), machine learning models provide high accuracy and more predictive power as compared to the conventional techniques. Machine learning algorithms are usually separated into two classes based upon the desired final results of the algorithms. In supervised machine learning, outcomes are already known, which is predicted based on some features. Nevertheless, in unsupervised machine learning, features are used to generate unknown outcomes. An unsupervised mean that all the variables take part to generate in an output. In unsupervised machine leanings, an output is unknown, which is generated by the algorithm itself. The most widely used unsupervised machine learning is the clustering.

According to Myatt and Johnson (2014) clustering of data may help to identify and summarize the classes of individuals from which it belongs. Clustering helps to make groups into related sets of observations or clusters, so that all the observations move to those observations which are similar to one another. Clustering is formed in such a way that the data is subdivided into groups through a structured pattern in the data.

Mainly, three types of clustering are used in literature: (a) hierarchal clustering also called agglomerative hierarchal clustering that organize the data into hierarchically which provide insight into the problem under investigation. The Second type of clustering is k-means which partition the data into lists of clusters which are already predefined number of groups. K-mean clustering is often used where dataset is large, and no hierarchy existed in the data. The Third, is Density-based spatial clustering of applications with noise, called DBSCAN.

The DBSCAN is powerful algorithm to find patterns, associations and to predict. To calculate the similarities between observations, the distance between these observations are computed. To observe the similarities between observation, it is required that all the variables are normalized on the same scale to prevent disproportionate weight and bias in the data for analysis. Therefore, it was required to standardize the data on a similar measuring scale so that the distance between each observation could be calculated.

While going through detailed literature review, there is no such study in Pakistan Stock Exchange (PSX) to explore this issue except the Sohail, Sindhu, and Imran (2020) where simple co-integration approach is used, and authors documented profitability in pair trading. However, in this study the more advanced techniques of machine learning were ignored. Therefore, the main objective of this study focuses to look for evidence on pairs trading's profitability and to observe the edge of machine learning approach over traditional approach at PSX.

The clustering process is tested under unsupervised machine learning approach using the Principal Component Analysis and combined with DBSCAN Clustering to group the stocks with higher similarities.

Therefore, research questions in this study are as follow:

Does pair trading under distance approach provide profitability?

Does pair trading under machine learning approach provide profitability?

Does pair trading under machine learning approach and distance approach provide positive risk adjusted performance?

In Pakistan, there is no single study regarding pair trading except the study of Qazi, Rahman, and Gul (2015) and Sohail, Sindhu, and Imran (2020). In the study of Qazi, Rahman, and Gul (2015) only pairs were formed and no trading algorithm was used. However, in the study of Sohail, Sindhu and Imran (2020) only co-integration approach was used. So, this study is the first one to take this aspect of machine learning and will contribute in the context of use of unsupervised machine learning for pair formation.

2. Literature Review

One of the main themes in pair trading is the concept of the market neutrality. Haque and Haque (2014) that pair trading involves risk also. However, according to Elliott, Hoek, and Malcolm (2005), the risk in pair trading can be neutralized by taking long and short position in financial assets. The strategy is market neutral only if the stocks of each pair have identical exposure to the market.

In the study of Gatev, Goetzmann, and Rouwenhorst (2006), 11% annualized return was observed. Alexander and Carol (2001) used the cointegration approach for pair trading in DJIA and found positive returns. The same approach was also witnessed in the studies of Caldeira and Moura (2013) and Rudy, Dunis, and Laws (2010). They also showed positive returns in pair trading.

In Australia, Bogomolov (2011) used three different methods; stochastic spread, distance and cointegration approaches for pair trading. Mudchanatongsuk, Primbs, and Wong (2008) also provided a different model and validating the mean reversion theory. The most common approaches to pair trading are: Distance Approach, Stochastic Approach, Stochastic Residual Spread and Co-integration Approach (Elliott, Van Der Hoek, & Malcolm, 2005; Gatev et al., 2006; Do & Faff, 2010;Liew & Wu, 2013).

Five different pair trading strategies are reported by Krauss (2017), however, Blázquez and Román (2018) favored co-integration approach after analyzing different methodologies under the umbrella of pair trading. The profitability in pair trading strategies are witnessed by different studies like the Namwong, Yamaka,and Tansuchat (2019) study. They used the data of Thailand and reported profitability. Recently, Ramos-Requena, Trinidad-Segovia, and Sánchez-Granero (2020) and Sohail, Sindhu and Imran (2020) also reported profitability in pair trading.

3. Methodology

The Sample consists of Eighty firms from five sectors including banking, cement, textile composite, chemical, and food & care products where 16 firms are included from each sector. The formation period of pair trading consists of one year for distance approach while the two years for machine learning. The trading period is same that consists of six months for the two approaches; the distance and unsupervised machine learning.

To follow the methodology of distance approach, the prices of stock should be normalized (i.e. begins at one (100%, and its value will change according to the subsequent returns) as discussed below.

$$P_t^i = \pi^T{}_{t=1}\left(1 + r^i{}_t\right) \text{ --- (1)}$$

Where $p\ _t\ ^i$ represents prices of stock i that are now normalized. $r\ _t\ ^i$ represents the stock returns on daily basis .For formation period, total number of trading days are represented by T. At the end of day each stock normalized price is t = 1, 2, . …. . T. The distance, $D_{(i,j)}$is measured as:

$$D_{(i,\,j)} = \frac{\sum_{t=1}^{T}\left(P\ _t^i - P\ _t^j\right)^2}{T} \text{ --- (2)}$$

These distances are sorted in ascending order and then ranking of pairs take place. For example, the top 5 pairs are selected to match with machine learning (clustering) approach where three clusters were identified containing two 3’s and one 4’s stocks.

In machine learning approach, first, PCA analysis are conducted on the daily returns of two years’ data (495 days) and 30 factors are formed. PCA creates new variables, that are linear patterns from the original variables in this case daily returns and other variables. In summary 30 factors were formed. These 30 factors are combined with the fundamental factors; EPS, Sector classification, BVP and Market capitalization to apply DBSCAN clustering algorithm in Python.

In machine learning, DBSCAN is a clustering technique method that is used to break up clusters of high density to low density of clusters. It also assists in handling outliers within the dataset. Two criteria are required for DBSCAN:"eps: the minimum distance between two points. It implies that if the distance between two points is lower or equal to this value (eps), these tips are considered neighbors. Midpoint's: the minimum number of points to form a dense region. For example, if we set the min Points parameter as 5, then we need at least 5 points to form a dense region".

DBSCAN starts by splitting the data into n dimensions. A random point is selected and then nearby points are selected to this random point and so on, and finally clusters are identified with the help of this algorithm. Minimum four points are required to from cluster. The point outside is not contained within any cluster and treated as outliers. Based on DBSCAN clustering, 3 clusters are identified, and each cluster carry 3, 3, and 4 firms. The pairs are formed from three clusters to start the trading algorithm.

After identification of pairs under the two approaches trading algorithm is applied to calculate the portfolio returns. In trading algorithm, under the two approaches, the spread of each pair is calculated as: $|P_t^i - P_t^j|$

While trigger gets the following forms:

Trigger (i, j) = n x std (i, j)

$$\text{std } (i,j) = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left[\left(P^i_t - P^j_t\right)^2 - D(i,j)\right]^2} \quad \text{------------------------------ (3)}$$

A trade remains opened when $|P_t^i - P_t^j|$ is greater than trigger (i, j). It involves the long position and short position for the same amount by committed capital (i.e. no investment).The values for n are 1.5, 2 and 2.5 SD's.A pair trade remains opened and closed in the six months period until above conditions met,i.e. divergence and convergence of stocks takes place, however, trading set to close at the end of six months where stocks do not converge.

The calculation of returns is based on the cash flows by investing capital, however, since in pair trading no initial capital is required, the returns are calculated by dividing the payoff of each period by the total amount of pairs. $R_t(l^k)$ and $R_t(s^k)$, represent one-day return on long and short positions. The daily returns of pair are calculated as:

$$R_t(p^k) = R_t(l^k) - R_t(s^k) \quad \text{--(4)}$$

So the daily return of N_t^* pairs is

$$R_t^{port} = \sum_{k=1}^{Nt*} W_t^k \, \mathrm{Rt}(p^k) \quad \text{---(5)}$$

Where W_t^k represents the equally weighted.

These daily returns are converted to monthly returns. These monthly returns are calculated for the period of six years from Jan. 2011 to Dec 2019. Further due to the market-neutrality of the strategy, the seportfolio returns can be considered as excess returns.

Average AR is calculated as:

$$\overline{AR} = \frac{1}{n}\sum_{i=1}^{n} AR \text{ --(6)}$$

The significant of these AR's are checked by t statistics.

$$t = [\ \overline{AR}\ ] / [\ \frac{s}{\sqrt{n}}\ ] \text{ --(7)}$$

Further, the performance of pair trading portfolio is checked by Sharpe, Treynor and Jensen's alpha (using CAPM

SR= Excess pair - trading portfolio return / SD of portfolio ---(8)

T= Excess pair - trading portfolio return / systematic risk portfolio -------------------------------------(9)

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \varepsilon_{p,t} \text{ --(10)}$$

The two hypothesis "abnormal returns of pairs trading is different from zero", "Jensen's alpha of pairs trading is different from zero" under the two approaches is tested.

4. Results

The sample of this study comprises of 5 sectors of PSX named as Textile, Cement, Chemical, Banking and food & care products. 16 companies in each sector are selected based on their market capitalization for distance approach and 80 companies are selected for machine learning. The formation period of pair trading consists of one year for distance approach while two years for machine learning. The trading period is the same that consists of six months for the two approaches. The data which is used in this study is (daily data) over the period 2011-2019. For machine learning fundamental characteristics like BVP, EPS, Mkt Cap are also used.

Three cluster were identified through DBSCAN clustering. The graphical representation obtained through python language is represented as figure-1 and figure-2. The results of these three clusters are compared with three sectors; banking, cement and textile.

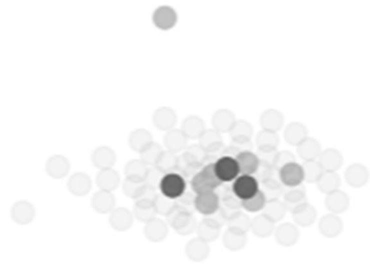

Figure 1: T-SNE of all Stocks with DBSCAN Clusters Noted

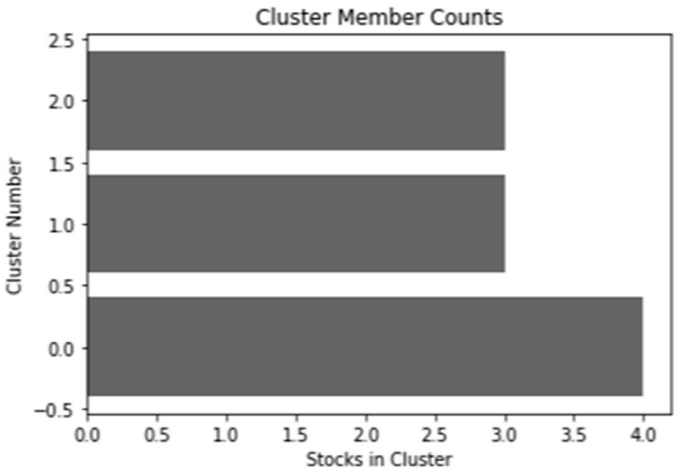

Figure 2: Stocks in cluster

The results of descriptive analysis under various parameterization of distance and machine learning approach are presented in Table 1 under Panel-A and Panel-B respectively. In Table 1, Panel-A represents the analysis of distance approach, while Panel-B represents the analysis of Machine learning approach In traditional and machine learning approach, the average values for various parametrizations are positive to show that pair trading strategy is profitable at PSX. In distance approach, under the textile sector, the highest average returns of 0.0207 are to be observed under the parameter of 2.5 standard deviations, however, the volatility is also maximum. In the same way, in machine learning approach under cluster 2, the highest average returns of 0.0255 are to be observed under the parameter of 1.5 standard deviations. These results are with accordance to prior studies of pair trading like Namwong, Yamaka, and Tansuchat (2019) and Zhang and Urquhart (2019) etc.

Table 1
Descriptive Analysis under Different Parametrization

Panel-A	**b_1.5_SD**	**b_2_SD**	**b_2.5_SD**	**c_1.5_SD**	**c_2_SD**	**c_2.5_SD**	**t_1.5_SD**	**t_2_SD**	**t_2.5_SD**
Average	0.0082	0.0089	0.0093	0.0095	0.0099	0.0126	0.018	0.019	0.0207
Median	0.0063	0.007	0.0077	0.0073	0.0081	0.0094	0.0174	0.017	0.016
Minimum	-0.0016	-0.0027	-0.0033	-0.0058	-0.0117	-0.0125	0.0028	0.0034	0.0021
Maximum	0.0447	0.0544	0.0508	0.0279	0.0317	0.0559	0.0641	0.0877	0.1654
SD	0.0067	0.0081	0.0079	0.0079	0.0091	0.0127	0.0112	0.0134	0.0225
Panel-B	**cl1_1.5_SD**	**cl1_2_SD**	**cl1_2.5_SD**	**cl2_1.5_SD**	**cl2_2_SD**	**cl2_2.5_SD**	**cl3_1.5_SD**	**cl3_2_SD**	**cl3_2.5_SD**
Average	0.0102	0.0129	0.0159	0.0121	0.0154	0.0195	0.0255	0.0111	0.0154
Median	0.0088	0.0093	0.0102	0.0096	0.0108	0.0116	0.0118	0.0014	0.0014
Minimum	0.0008	0.0002	0.0004	0.0002	0.0003	0	0	0	0
Maximum	0.0272	0.0398	0.0582	0.0401	0.0638	0.1104	0.2703	0.0837	0.2124
SD	0.0074	0.0104	0.0162	0.0106	0.0145	0.0225	0.0442	0.0194	0.0358

To test the significance of portability under the two approaches of distance and machine learning, traditional t-statistic is applied. Similarly, Jensen's alpha under CAPM is used for risk adjusted returns of pair trading portfolio. Further, some other well-known performance measure ratios Treynor and Sharpe ratio is also used.

After calculating the average returns of these trading portfolios under distance and machine learning approaches average profitability (abnormal returns) of these trading portfolios are presented in Table 2. Overall, significant, and positive results are to be observed under these two approaches.

In distance approach, these results are highly significant with an average return of 1.19% and t statistic of 19. 2035. These results are also showing that pair trading strategy is also profitable at PSX and investors can earn positive returns as in case of other developing and emerging markets of the world. These results are also supportive of market neutrality and mean reversion theory. For different portfolio and fund managers, it is suggested that they can form the pair trading strategy to reap the profitability for their clients. The Sharpe ratio and Treynor ratios are also positive with values of 2.48 and 1.25 respectively.

Table 2
Comparison of Distance v/s Machine Learning

	AR's	SD	t	p value	Sharpe Ratio	Treynor Ratio
Avg_distance	0.0119	0.0048	19.2035	0.0000	2.4800	1.4500
Avg_machine	0.0160	0.0162	7.6503	0.0000	0.9900	0.7500

On the other hand, in case of machine learning approach, these results are also highly significant with an average positive return of 1.6% and t statistic of 7.6503. These results are higher than distance approach. These results are also showing that pair trading strategy is also profitable at PSX and investors can earn some additional positive returns by using machine learning approach. These results are also supportive of market neutrality and mean reversion theory. For different portfolio and fund managers, it is suggested that they can form the pair trading strategy to reap the higher profitability for their clients by using machine learning approach. The Sharpe ratio and Treynor ratios are also positive with values of 0.99 and 0.75, respectively, lower than as compared with distance approach.

Therefore, under the two approaches of distance approach and machine learning, the null hypothesis is rejected, and we are now in position to answer the research question number 1 that pair trading is profitable at PSX. For risk adjusted return of pair trading portfolio, the results of all the three sectors in detail (separately) and as a whole are presented in table 3. The Jensen's alpha for CAPM under various parameterization shows that risk adjusted returns of pair trading portfolio for all the three sectors and as a whole, is significantly positive. These results are also with the accordance of prior studies. Therefore, null hypothesis is rejected for Jensen's alpha. Therefore, investors can get positive risk adjusted returns at PSX using distance approach. Textile sector under 2.5 standardization parameter rewarding higher risk adjusted returns as compared with other two sectors: banking and cement.

Table 3
Results of Jensen's Alpha of Distance Approach

	Variable	Coefficient	Std. Error	t-Statistic	Prob.
b_1.5_SD_5_CAPM	Intercept (alpha)	0.0078	0.0009	8.2678	0.0000
b_2_SD_5_CAPM	Intercept (alpha)	0.0084	0.0011	7.4300	0.0000
b_2.5_SD_5_CAPM	Intercept (alpha)	0.0088	0.0011	7.9749	0.0000
c_1.5_SD_5_CAPM	Intercept (alpha)	0.0102	0.0011	9.4083	0.0000
c_2_SD_5_CAPM	Intercept (alpha)	0.0107	0.0013	8.5635	0.0000
c_2.5_SD_5_CAPM	Intercept (alpha)	0.0133	0.0018	7.4473	0.0000
t_1.5_SD_5_CAPM	Intercept (alpha)	0.0172	0.0016	10.9779	0.0000
t_2_SD_5_CAPM	Intercept (alpha)	0.0184	0.0019	9.7770	0.0000
t_2.5_SD_5_CAPM	Intercept (alpha)	0.0203	0.0032	6.3710	0.0000
dist_avg_CAPM	Intercept (alpha)	0.0117	0.0007	17.5458	0.0000

For risk adjusted return of pair trading portfolio, the results of all the three clusters formed through machine learning in detail (separately) and as a whole are presented in table 4. The Jensen's alpha for CAPM under various parameterization shows that risk adjusted returns of pair trading portfolio for all the three clusters and as a whole, is significantly positive. These results are also with the accordance of prior studies. Therefore, null hypothesis is rejected for Jensen's alpha under machine learning approach. Therefore, investors can get also positive risk adjusted returns at PSX by using machine learning approach. Cluster 3 under 1.5 standardization parameter rewarding higher risk adjusted returns as compared with other two clusters.

The higher risk adjusted performance of pair trading portfolio is to be observed under machine learning approach as compared with the distance approach.

Table 4
Results of Jensen's Alpha of Machine Learning Approach

	Variable	Coefficient	Std. Error	t-Statistic	Prob.
cl1_1.5_SD_5_CAPM	Intercept (alpha)	0.0093	0.0010	9.3485	0.0000
cl1_2_SD_5_CAPM	Intercept (alpha)	0.0114	0.0014	8.2249	0.0000
cl1_2.5_SD_5_CAPM	Intercept (alpha)	0.0140	0.0022	6.3927	0.0000
cl2_1.5_SD_5_CAPM	Intercept (alpha)	0.0117	0.0015	7.7970	0.0000
cl2_2_SD_5_CAPM	Intercept (alpha)	0.0150	0.0021	7.2922	0.0000
cl2_2.5_SD_5_CAPM	Intercept (alpha)	0.0186	0.0032	5.8488	0.0000
cl3_1.5_SD_5_CAPM	Intercept (alpha)	0.0257	0.0063	4.1029	0.0001
cl3_2_SD_5_CAPM	Intercept (alpha)	0.0120	0.0027	4.3802	0.0001
cl3_2.5_SD_5_CAPM	Intercept (alpha)	0.0140	0.0051	2.7748	0.0074
mach_avg_CAPM	Intercept (alpha)	0.0156	0.0023	6.8442	0.0000

5. Conclusion

Overall, the results of pair trading are profitable at PSX. The results show that for various parameterization, pair trading under machine learning approach have superior returns as compared with distance approach. In the same way, the risk adjusted returns of machine learning were observed to be higher as compared with distance approach. The sector-wise results under distance approach and cluster-wise results under machine learning approach separately and as a whole are statistically significant and positive at PSX. The Sharpe ratio and Treynor ratios are also positive under the two approaches.

The results also authenticate mean revision and market neutrality at PSX. Investors can get positive returns through pair trading at PSX. As a policy implication of this study, for different portfolio and fund managers, it is suggested that they can form the pair trading strategy to reap the profitability for their clients and specially they can get higher returns by using machine learning approach. various markets with different market conditions can yield best opportunities for pairs trading.

The study has some limitations like small sample, only one market, short sale restrictions and transaction cost was ignored. Recommendation for future research: More powerful algorithms may be developed to grasp the significant profits through pair trading amongst different financial assets.

References

Alexander, Carol, 2001. Market Models: *A Guide to Financial Data Analysis*. West Sussex: John Wiley & Sons Ltd.

Blázquez, M. C., & Román, C. P. (2018). Pairs trading techniques: An empirical contrast. *European Research on Management and Business Economics*, *24*(3), 160-167.

Bogomolov, T., Pairs trading in the land down under. In Finance and Corporate Governance Conference, Melbourne, Australia, 2011.

Caldeira, J.F. and Moura, G.V. (2013) Selection of a portfolio of pairs based on cointegration: a statistical arbitrage strategy. *Brazilian Review of Finance, 11*(1), 49–80.

Do, B., & Faff, R. (2010). Does simple pair trading still work?. *Financial Analysts Journal*, *66*(4), 83-95.

Ecer, F. (2013). Comparing the bank failure prediction performance of neural networks and support vector machines: The Turkish case. *Economic Research-Ekonomska istraživanja*, *26*(3), 81-98.

Elliott, R. J., Van Der Hoek, J., & Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, *5*(3), 271-276.

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, *19*(3), 797-827.

Haque, S. M., & Haque, A. E. (2014). Pairs trading strategy in dhaka stock exchange: Implementation and profitability analysis. *Asian Economic and Financial Review*, *4*(8), 1091.

Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision Support Systems*, *37*(4), 543-558.

Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, *31*(2), 513-545.

Liew, R. Q., & Wu, Y. (2013). Pairs trading: A copula approach. *Journal of Derivatives & Hedge Funds*, *19*(1), 12-30.

Lin, W. Y., Hu, Y. H., & Tsai, C. F. (2011). Machine learning in financial crisis prediction: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(4), 421-436.

Mitchell, T. M. (2006). The discipline of machine learning (Vol. 9). Pittsburgh: Carnegie Mellon University, School of Computer Science, Machine Learning Department.

Mudchanatongsuk, S., Primbs, J.A. and Wong, W. (2008,June) Optimal pairs trading: a stochastic control approach. In *American Control Conference*, 1035–1039.

Myatt, G. J., & Johnson, W. P. (2014). Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining (2nd ed.). New Jersey: John Wiley & Sons, Inc.

Namwong, N., Yamaka, W., & Tansuchat, R. (2019, January). Trading Signal Analysis with Pairs Trading Strategy in the Stock Exchange of Thailand. In *International Conference of the Thailand Econometrics Society* (378-388). Springer, Cham.

Qazi, L. T., Rahman, A. U., & Gul, S. (2015). Which pairs of stocks should we trade? Selection of pairs for statistical arbitrage and pairs trading in Karachi Stock Exchange. *The Pakistan Development Review*,*54*(3), 215-244.

Ramos-Requena, J. P., Trinidad-Segovia, J. E., & Sánchez-Granero, M. Á. (2020). Some Notes on the Formation of a Pair in Pairs Trading. *Mathematics*, *8*(3), 348-361.

Rudy, J., Dunis, C., & Laws, J. (2010). Profitable Pair Trading: A Comparison Using the S&P 100 Constituent Stocks and the 100 Most Liquid ETFs. Available at SSRN 2272791.

Sohail, M. K., Sindhu, M. I., & Imran, M. (2020). The Secret to Cash the Profits in Pair Trading Strategies: A Case of KSE Firms, *City University Research Journal*, *10*(2),189-204.

Zhang, H., & Urquhart, A. (2019). Pairs trading across Mainland China and Hong Kong stock markets. *International Journal of Finance & Economics*, *24*(2), 698-726.