Confining value from Neural Networks: A sectoral study prediction of takeover targets in the U.S Technology Sector

SCHOLARONE™
Manuscripts

Confining value from Neural Networks: A sectoral study prediction of takeover targets in the U.S Technology Sector

ABSTRACT

Published studies in the area of predicting M&As have made a relatively limited attempt to use neural network systems (NNs henceforth) in such a decision making process. This paper examines the value of utilising a neural networks approach using M&A data confined in the U.S technology domain. Investors value firms before investing in them to identify their true stock price; yet, technology firms pose a great valuation challenge to investors and analysts alike as the latest IT stock price bubble in Silicon Valley and as the recent stratospheric rise of Financial Technology (FinTechs henceforth) companies have also demonstrated. At the same time, the technology sector in the US commands approximately 8% of GDP and accounts for around 20% of all M&A deals in our sample period. We utilise US technology firms' data from Bloomberg for the period 2000–2016. Our analysis applies and compares a neural network approach to a linear classifier, logistic regression. Our empirical results show that neural networks are a promising method of evaluating M&A takeover targets in terms of their predictive accuracy and adaptability. The higher level of accuracy provided by a neural network approach can provide practitioners with a competitive advantage in pricing merger offers. Trade-offs and limitations of using neural nets as an alternative, general modelling tool are also discussed. Our findings emphasise the value alternative methodologies potentially provide in high-technology industries in order to achieve the screening and explorative performance objectives, given the technological complexity, market uncertainty and the divergent managerial skill sets required for breakthrough innovations in these sectors. This study provides valuable insights to managers aiming to increase the effectiveness of their decision-making for diversification, growth portfolios and investments.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1. **Introduction**

Shareholder theory posits that management has a mandate to maximise shareholder wealth through decisions that add value to investments and stimulate growth. For any large company, growth through Mergers and Acquisitions (M&As henceforth) is often a key part of corporate growth strategy. Growth largely drives value creation and M&As can offer a course to growth when esoteric opportunities are restricted through projected financial, strategic and operational synergies achieved at a fair price. Numerous studies though have shown that M&As more often than not destroy value rather than create it. More than 50% of all M&As lead to a decline in relative total shareholder return after one year. Hence, effective target identification must be built on the foundation of a credible strategy that identifies the most promising market segments for growth, assesses whether organic or acquisitive growth is the best way forward, and defines the commercial and financial hurdles for potential deals. It is thus crucial for companies' upper management to utilise credible and proven methodologies and models to ensure that target identification is based on sound background research. For example, as early as 20 years ago, researchers (Kaastra and Boyd, 1996; Rojas, 1996) argued for the systematic application of neural networks as a method to deal with the problem of non-linearity in financial transactions. Rojas (1996) argues that where there is abundance of data but less theoretical understanding (for example behavioural patterns that are not easily identified through established linear methods or continuously keep changing) neural networks *can discover statistical regularities and keep adjusting parameters even in a changing environment* (p.247).

When valuing a firm three major sources of valuation inputs are considered: (*i*) Current Financial Statements; (*ii*) Firm's Past History; and (*iii*) Peer Group comparisons. While for most firms such crucial information is ready-made, for technology firms such vital sources might be absent. Their financial statements don't include much information about growth prospects either. Most technology firms have limited or no past history. They also possess unique businesses and/or products therefore leading to no directly visible peers or competitors (Damodaran , 2001): *"As more and more technology firms get listed on financial markets, often at very early stages in their life cycles, traditional valuation methods and metrics often seem ill suited to them."* (p. 19). Daniel, et. al (1998) have demonstrated that investors tend to be overconfident when examining unclear information and that mispricing is stronger for stocks whose value is closely tied to their growth. The very fact that technology firms more often than not exhibit unconventional growth patterns makes them difficult to evaluate and can lead to their stocks being massively mis-valued (most of the time over-valued) and therefore increasing M&A activity (Rhodes-Kropf & Viswanathan, 2004; Jovanovic & Rousseau, 2001). While there are idiosyncratic motives for undertaking M&A-led growth strategies, there are also substantial economy-wide factors which cause waves of global M&A activity such as responses to globalization forces and increases in competition, de-regulation and the associated economic reforms and liberalization, block/regional economic integration (i.e. the EU). As such, target firm identification, has become a great research interest area both to the business world and academia alike. The three latest M&A waves (namely, M&As waves 5, 6, and 7) make the case in point:

- **Fifth Wave: 1993 – 2000**

A wave known for its large transactions and overvaluation of firms. Transactions were mostly friendly and financed by equity (Andrade, et al., 2001). It was empowered by cross-border transactions due to the strong economic conditions in the U.S, Europe and Emerging Markets. This wave ended with the burst of the internet bubble causing the market to crash (Dieudonne, et al., 2014).

- **Sixth Wave: 2003 – 2008**

A wave known for producing less overvalued transactions, with the size of both acquirer and target getting smaller. During this wave firms enjoyed more cash, with the excess liquidity causing this wave (Alexandridis, et al., 2012), and having 75 percent of the transactions paid by cash (Gregoriou & Neuhauser, 2007). This wave ended with the 2008 credit crisis.

- **Seventh Wave: 2010 – Present**

The wave which we are currently experience started gradually in 2010, and it coincided with the emergence of FinTechs. Since 2015, it has reached an all-time high of 2.9 trillion dollars in value (Institute for Mergers, Acquisitions and Alliances, 2016). It is aso owed to the system-wide steps taken by central banks after the 2008 credit crisis, such as keeping near-zero interest rate and the quantitative easing procedures which supplied equity and bonds markets with enough liquidity.

With an increasing amount of M&As in the technology sector, it is crucial to identify targets before announcement date, as this can be significantly beneficial for investors, target and acquiring firms. From that comes the motivation of a reliable takeover predication model. *Standard* models have so far been relatively indeterminate in the past, and as a result have not had highly reliable estimations regarding the scale of an outcome or a conclusion on the directional relationships of the variables (Betton et al., 2008; Routledge et al., 2013; Eckbo, 2014;). Hence, we switch our approach to an altered empirical exploration model where this is also tested. We elaborate further on our motivation below where by implication we discuss our reasons for utilising NNs as opposed to the traditional regression techniques.

Having introduced our study topic, we discuss the relevant empirical evidence on the characteristics of M&A deals in the US, valuation challenges and the technology sector in section 2 that follows. In section 3 we discuss methodological issues where determinant variables and neural networks are compared with the traditional statistical techniques of discriminant analysis and logistic regression with regards to the identification of potential takeover targets. Section 4 discusses our methodology and section 5 presents the results of our analysis. The conclusions of the study are presented in section 6.

1.1 Motivation Summary

Financial time series have some characteristics that make them hard to reliably forecast, especially when a traditional statistical method is employed. Such characteristics are as follows (Motiwalla and Wahab 2000; Thawornwong and Enke 2004; Versace et al. 2004):

1. *Non-stationarity* of data, where due to different business and economic cycles, the statistical properties of financial data change randomly over time, which also introduces:
2. *Non-linearity* of data, where the relationship between the financial and economical independent variables and the desired dependent variable may not be linear. Intensified by:
3. *Noisiness* through daily variations in financial time series.

On the other hand, NNs are more flexible and adaptable computing methods that provide the ability to potentially capture the patterns among variables more effectively. Hence, the use of NNs to forecast financial time series as an alternative is justified by some of the in-build qualities they posses. Such characteristics make them reasonably well suited for use in the financial forecasting domain (Hussain et al. 2007; Lin et al. 2006; Lam, 2004; Eakins and Stansell, 2003):

1. Their nonlinearity. NNs can capture nonlinear relations between element (input or independent variables) and response (output or dependent variables).

2. Their data driven nature. No prior explicit relational assumptions on the model are made or modelled between inputs and outputs.

3. Their generalizability. Once trained, NNs can produce relatable results even when the data structure has changed or when they are faced new input patterns.

4. Their assumption neutrality. Dissimilar to traditional statistical techniques, NNs do not employ pre-constructed assumptions on the input data distribution.

Yet, as with any forecasting tool, the robustness of a NN application outcome can equally be questioned. This is addressed in our results and discussion sections at the end of this exposition.

2. Literature Review: Mergers and Acquisitions in the U.S Technology Sector

The U.S is well-known as the most preferred international investment destination measured by Foreign Direct Investment (FDI) flows. In 1989 the FDI position in the U.S (FDIUS) exceeded $400 billion (Harris & Ravenscraft, 1991), while in 2014 for example it totalled $2.4 trillion with an average annual growth of 8.9 percent (Organization for International Investment, 2016; Bureau of Economic Analysis, 2017). It is ranked as the world's top market for 5 years consecutively (Laudicina & Peterson, 2017). Figure 1 illustrates M&As as the most exercised type of investment in the U.S in volume vs. expansions and new establishments.

Figure 1: FDI in the United States by type 1994 – 2016, US Bureau of Economic Analysis
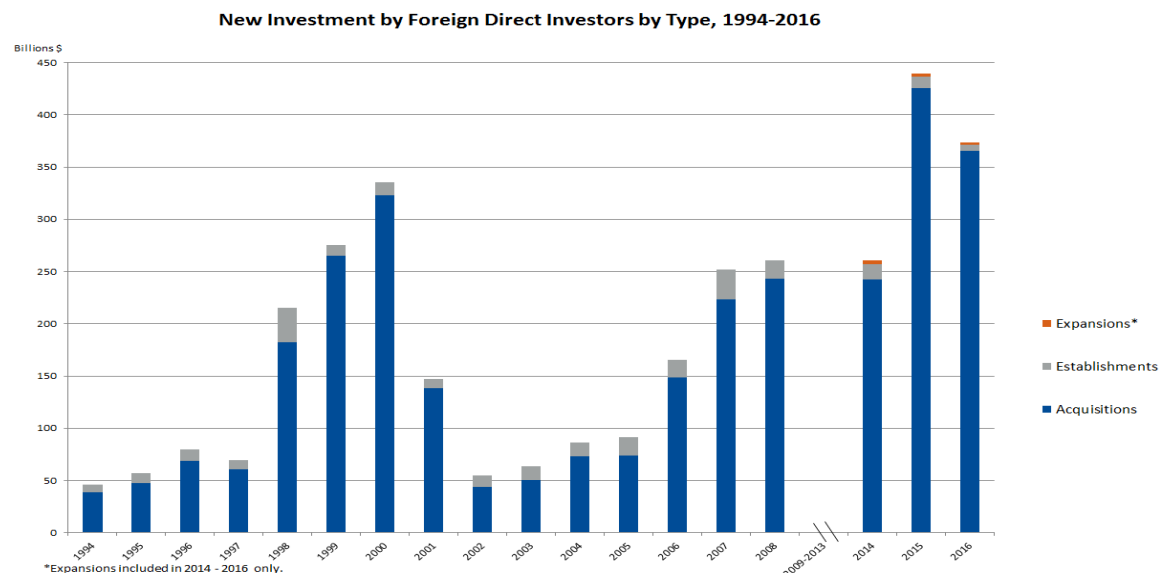
*Due to government budget cuts the Bureau of Economic Analysis was not able to fully conduct a concluding survey from 2009 to 2013.

Rossi & Volpin (2003) suggested the role of the legal system as a factor affecting cross-borders M&A volume. The rationale being that countries with mature legal systems are better able to cope with economic changes, absorb shocks and provide shareholder protection thus improving the liquidity of the market as a whole (Eden & Dobson, 2005; Beck, et al., 2003). Harris & Ravenscraft (1991) claim that FDIUS increases when the dollar is weaker compared to the investor's home currency. Servaes & Zenner (1994) also affirmed that tax regulations have an impact on FDIUS indicating tax benefits for the investors. During the period 2000-2016 the United States occupied the biggest share of worldwide M&A activities. The highest percentage (50 percent) was taken by technology firms in 2000 due to the tech bubble with an average of 37 percent throughout the same 17-year period (WilmerHale , 2017) as shown in figure 2 below; US firms represent on average 20% of global M&A as acquirers and 23% as targets by value (Ernst & Young, 2015).

*Figure 2: M&A Activity; Worldwide vs. US*
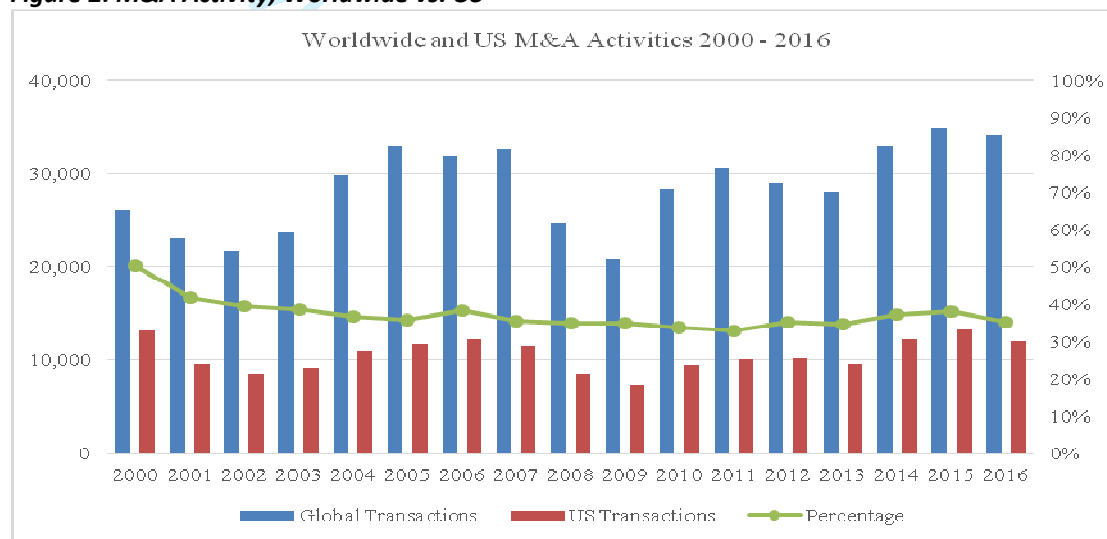
The technology sector experienced the largest M&A activity in the U.S, holding the highest number of transactions in the period between 2000 and 2016 which represented 19.9% of all U.S M&A transactions (Institute for Mergers, Acquisitions and Alliances, 2016). The high volume of M&A transactions in this sector is attributed to three main causes: (*i*) Financial Strength: technology firms enjoying large amounts of cash with high stock prices enabled them to make large acquisitions using cash or stocks. In 2016 technology firms held over $773 billion in cash, accounting for 46% of total cash held by U.S non-financial firms of $1.68 trillion (Moody's, 2016); (*ii*) Industry Trends: i.e. location-based services, digital entertainment, robotics and artificial intelligence, virtual reality, 3-D printing and blockchain. These areas have various applications, they are used by millions of users, and at the same time they are rapidly evolving. Acquisitions are favoured by 41% of technology firms as the path to growth and market share capture on one or more tech areas (Ernst & Young, 2016)[1]. Hagedoorn and Duysters (2002) posit that technology firms in their quest for growth through innovation prefer acquisitions instead of other alternatives such as strategic alliances;

---

[1] Based on a survey conducted by Ernst&Young in October 2016, including 255 respondents from technology firms, of which 51 percent were CEOs, CFOs and other C-level executives.

(*iii*) Talent Acquisition: since 2011 the Big Five[2] technology firms have added more than 418,000 jobs to the market with 76% of technology firms scouting and acquiring other firms in order to secure talent amidst other industries that have lost jobs (Acker, et. al, 2017; Ernst & Young, 2016). The sector accounts for more than 8% of the total US economy. It represents about $1.3 trillion of value (CompTIA, 2017) and it employs more than 4% of all US workforce (U.S. Bureau of Labor Statistics, 2016). The dot com bubble (1999 – 2000) was the peak with 371 and 261 tech IPOs respectively, and in 1996 with 274 IPOs (Ritter, 2017). The growth of IPOs during the 1990s was fuelled by venture capitalists excessively funding start-ups as funding rose from $3 billion in 1990 to $60 billion in 1999 (Lowenstein, 2004); furthermore, 57% of these tech firms going public were less than five years old and in some cases even less than two years old (Westenberg, 2009); institutional investors bought stocks with *thin fundamentals* as they purchased more than 63.6% of technology stocks between 1997 and 2000 (Griffin, et al., 2011). This was coupled with media coverage and narratives from investment bankers, analysts and journalists encouraging individual investors to further invest in the technology sector (Teeter and Sandberg, 2016) making them hold the remaining 36.4% of technology stocks and continue to buy them while institutional investors were rapidly selling them (Griffin, et al., 2011). As a result, the market was extremely overvalued when NASDAQ reached its highest level[3] on March 2000. It lost more than 50% by value in October 2000 (Westenberg, 2000). Its growth has been substantial since the 1990s measured by the number of technology IPOs as indicated by figure 3 below.

Figure 3: US Tech IPOs 1980 – 2016 (Ritter, 2017)
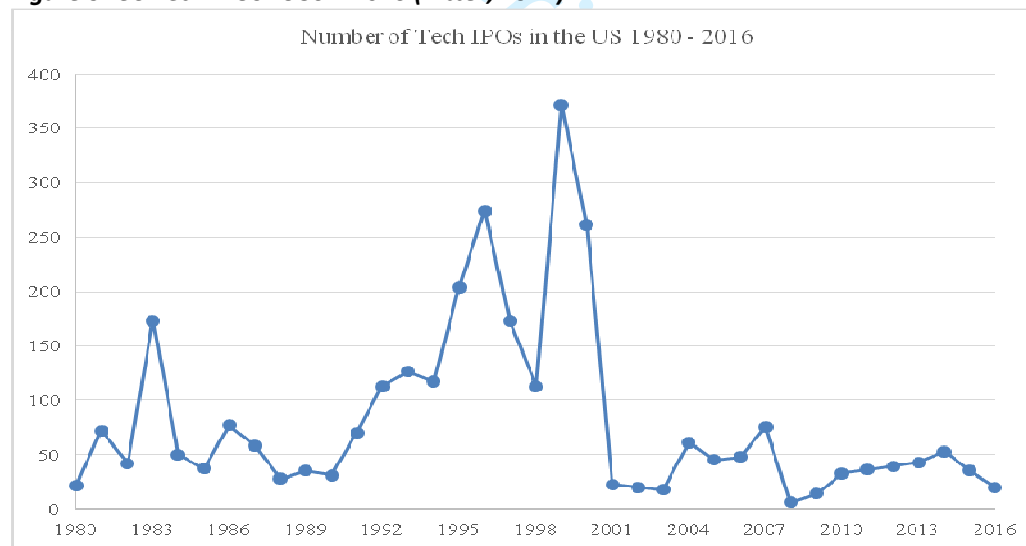
2.1   Valuation of Technology Firms

In an M&A transaction it is as crucial for the acquirer to determine a fair value of synergies for the target, as it is for the target to come to a value for itself. It is also important for the shareholders of both firms to justify the acquisition price (Petitt & Ferris, 2013). The dynamics of the technology sector are characterized by rapidly evolving firms which operate

---

[2] Big Five: Microsoft, Alphabet, Amazon, Apple and Facebook.
[3] NASDAQ Composite Index (IXIC) level of 5,132 was the highest at the time of the tech bubble, it crossed the 5,132 level in 2015 (NASDAQ, 2017).

under high levels of uncertainty and risk (Lev & Zarowin, 1999). This, combined with the lack of positive cash flows (Aydin , 2015) makes their valuation very challenging as also demonstrated by Bakshi & Chenb (2005), where they demonstrate the potential for significant mispricing and departures from fair values. The complexity of valuing technology firms can be attributed to reasons such as:

(i)   Tech firms are often young ones, very dependent on innovation and require huge amounts of upfront investments in intangible assets. Chandra et al. (2011) state: *"...this arises from the uncertain nature of long-run industry prospects as well as competition among firms for market share through first-mover advantages, creation of entry barriers and establishment of property rights in new technology"(p.8)* which leads to the second point;

(ii)   The value of many firms in the technology sector usually comes from intangible assets. These assets however don't always appear on the firm's financial statements due to the lack of accounting standards to accommodate such intangibles, such as innovation, customer satisfaction and human capital, resulting in complexities when it comes to perform an equity valuation (Chan et. al, 2001);

(iii)   Tech firm value is directly dependent on growth; consequently most of the value will originate from future customers or products not from current operations. That makes it challenging for investors to measure firm's beta (risk);

(iv)   The value of a technology is only known after it is commercialized to the market (Park & Park, 2004).

There are various methods to value firms; they are however categorized into three mainstream methods (Hodges, 2007). Table 1 below shows the *pros and cons* as listed by Anadol et al. (2014).

Table 1: Advantages and Disadvantages of Valuation Methods

Method	Advantage	Disadvantage
Discounted Cash Flow **(DCF)**	Has firm theoretical basis; Easy to compare competing opportunities	Estimating future cash flows is difficult at best; Estimating interest rates in the future is uncertain
Comparable Firms **(CF)**	Best when a highly comparable group is available; Units are close in both size and business type	The whole sector may be over/under valued; There are too few comparable examples; Insufficient recent transactions
Asset-Based	Looks at all the underlying values in the firm's assets; Conservative, not likely to be criticised; Traditional method, people are comfortable with it	More relevant if the assets can be liquidated readily; Does not work for initial IPOs; Small firms are disadvantaged; Service firms are difficult to value this way; Growth rates in high-tech firms not included

DCFs as a method, boasts a huge limitation in that firms in the technology sector more often than not either do not pay dividends (even in cases they pay dividends these are often very volatile) or instead choose stock buybacks therefore using this method can undervalue the firm (Palepu, 2003). Any valuation method can be misleading as it does not for example incorporate intangibles, yet typically, acquisition premiums achieve more than 50 percent above market value (DeAngelo, 1990). Also, multiple bidding offers can be significantly different in terms of prices (Bradley, 1980). Even hedge funds in the U.S hire technology consultants to provide expert insights about tech firms as they are hard to value from a financial standalone perspective (Benou and Madura, 2005).

3. Takeover Prediction Techniques

Various researchers have studied the possibility of predicting acquisition targets through statistical aggregation and the associated distress signals (i.e. bankruptcy) using publicly available information of firms and then applying different statistical models on them. It is important to mention that the methodologies used to predict bankruptcy and predict takeover targets are very similar, (discriminant analysis (DA) and logistic regression) therefore we shall consider both broad approaches below.

3.1 Traditional Analytical Techniques

Regression models: Ohlson (1980) utilised logistic regression analysis in order to examine the relationship between binary or ordinal response probability and explanatory variables. He was the first to point out weaknesses in Altman's (1968) model and highlighted the importance of using data from firms' financial statements directly as they will indicate whether the firm filed for bankruptcy before or after releasing them which will help the researcher avoid the "back-casting" issue (i.e. applying the model to firm's data after being bankrupt). This model produced an accuracy prediction rate of 96 percent with a cut-off point of 0.5.The binary logistic regression, a nonlinear model, is one of the predictions' techniques where the dependent variable is a binary or dummy variable. Very few assumptions are required in such model in comparison to other similar dependence techniques such as discriminant analysis. Harris et. al (1982) used a probit model where for example, the dependent variable can take only two values (acquired or not-acquired), in order to produce a probability of a firm to be acquired or not as well as what are the characteristics that affected this probability. Dietrich & Sorensen (1984) used logistic regression model to predict acquisition likelihood. Palepu (1986) used a binomial logit probability model with 9 independent variables; his model suggested a good fit of success in predicting a high number of targets. It however, predicted a high number of non-targets as targets, therefore, it was not sufficient to use this model to gain abnormal returns. Barnes (1990) used multiple discriminant models with 5 chosen industry-related ratios to increase the predictability of his model. While the previous studies, as above, have shown prediction power between 60 to 90 percent Palepu (1986) argued however that these findings are overstated and suffer a biased estimate due to two main flaws in such methodologies: (*i*) state-based sampling for model estimation and prediction testing; (*ii*) using predetermined, arbitrary, optimal cut-off probability. Furthermore, Powell (1997) argued that the characteristics of hostile and friendly takeovers differ therefore using binomial models (treating hostile and friendly takeovers in the same group) will cause misleading results.

Cudd & Duggal (2000) in their study used Palepu's factors (1986) but they added an industry dispersion factor to account for different industries which improved the accuracy of the said model. In addition, they also found that the dummy variable *"industry disturbance"* to be significant therefore indicating that a takeover in the same industry in the past 12 months will increase the probability of takeover for the remaining firms in that industry.

Discriminant analysis (DA): allows the researcher to pair two or more firms (or groups of firms) and compare their differences with respect to several variables simultaneously. Depending on how variables behave (i.e. jointly or independently of one another) DA models can be further applied into two sub-categories namely univariate or multivariate models; multivariate models (MDA) consider simultaneously an entire portfolio of characteristics common to the firms and their interaction; univariate models are limited to only one characteristic at a time. As a technique, DA does very well provided that the variables in every group follow a multivariate normal distribution and the covariance matrices for every group are equal. As early as 1971, Simkowitz and Monroe suggested that target firms tend to be usually smaller, with lower P/Es and dividend payout ratio and lower equity growth. Most importantly, they further observed that non-financial characteristics appeared to be as important as financial. Their multivariate discriminant analysis (MDA) in-sample results predict 83% of the targets and 72% of the non-targets, while the holdout results are slightly worse predicting 64% of the targets and 61% of the non-target.

3.2 Machine Learning Techniques

A differentiated methodological approach used by researchers is the use of Neural Networks (NN), Machine Learning (ML) and Data Mining to predict bankruptcy or takeover targets. Sharda & Odom (1990) compared the use of both neural networks and multivariate discriminant models (MDA) in bankruptcy predictions. In their study, the researchers utilised the same ratios used by Altman (1968) and after executing both models their findings suggest that neural networks seem to outperform MDA based on different holdout samples with an accuracy level ranging from 77.78 to 81.48 percent. In another study, Tsai and Wu (2008) studied the effect of including multiple neural network classifiers in bankruptcy prediction and credit scoring where it was found that single neural network classifiers outperformed multiple neural network classifiers in both credit scoring and bankruptcy prediction. Hongjiu et. al, (2007) used self-organized mapping with Hopfield neural network to cluster data and their model showed accuracy predictions of 80.69 percent for targets and 63.11 percent for non-targets. Their paper suggests also the importance of including non-financial factors to improve the predictability power. Iturriaga and Sanz (2015) used *multilayer perceptron (MLP)* to predict bankruptcy of U.S banks with a 96 percent success rate.

The evidence regarding method and model fit is far from conclusive though. Coats and Fant (1993) for example, confirmed that NN outperformed Multiple Discriminant Analyses (MDA) in their sample 80 percent of the time. Numerous other studies have supported the use of Neural networks (NN) in outperforming logistic regression (LR) in predicting bankruptcy (see for example, Tam and Kiang, 1992; Jo and Han, 1996; Maher and Sen, 1997; Fan & Palaniswami, 2000; Tseng & Hu, 2010). Branch et. al, (2008) utilised both NN and LR to predict whether a takeover attempt will succeed or not with the authors concluding that

'...neural network model outperforms logistic regression in predicting failed takeover attempts and performs as well as logistic regression in predicting successful takeover attempts' (p. 1186). Salchenberger et. al, (1992) compared NN with LR to test healthy and failed thrift institutions and concluded that NN achieved higher accuracy. In all of the above studies a common results' attribution emerges: NNs seem to possess a higher flexibility and ability to address non-linearities. This echoes Zhang's et. al (1999) statement that neural networks can potentially be robust and can provide more reliable estimations when applied on different samples only once the optimal architecture is found.

On the other hand, Altman, et al. (1994) reported that both MDA and NN performed almost the same when trying to predict Italian firms suggesting that contextual and structural considerations as well as firm-characteristics' variables are also important. Equally, Olson et al. (2012) used Logistic Regression, Neural Networks, Support Vector Machines and Decision Trees to predict bankruptcy. They demonstrated that different data with different models present different results. There is trade-off between model accuracy and transparency and transportability. In a sense, in order to increase model transportability (i.e. applying it to new datasets and observations) the accuracy level will decrease. Table 2 below, by Barnes (1998), summarizes other research showing the prediction rates of various methods for North America and the UK.

Table 2. Previous stuies on the characterisitcs of Target firms.

Researcher	Country	Method	Classification (%)	Prediction (%)
Simkovitz and Monroe (1971)	USA	MDA	77	63.2
Tzoannos and Samuels (1972)	UK	GLS	70	–
Stevens (1973)	USA	MDA	70	67
Belkaoui (1978)	Canada	DA	72	70
Wansley and Lane (1983)	USA	MDA	77.3	69.2
Dietrich and Sorensen (1984)	USA	Logit	92.5	–
Rege (1984)	Canada	MDA	–	–
Bartley and Boardman (1996)	USA	MDA	65	–
Palepu (1986)	USA	Logit	–	46
Bartley and Boardman (1990)	USA	MDA	82.5	–
Ambrose (1990)	USA	Logit	65.1	75.6
Barnes (1990)	UK	MDA	68.5	–
Ambrose and Megginson (1992)	USA	Logit	74.3	–
Walter (1994)	USA	Logit	65	66

*DA, discriminant analysis; GLS, generalized least squeres;
MDA, multiple discriminant analysis; -, not reported.
Source: Barnes, Paul (1998), Can takeover targets be identified by statistical techniques?: some UK evidence,
The Statistician, 47, Part4, pp.573-591*

3.3 Takeover Determinant Variables

The main takeover relevant metrics/ratios that have been introduced by the financial literature to identify a takeover target are discussed below.

- **Inefficient Management**

This hypothesis states that managers who fail to maximize their shareholders' wealth and firm's value shall be replaced in accordance with the market for corporate control theory. Therefore, incompetent management increases the probability of their firms to be taken over (Jensen, 1986). Investors will seek to replace the management by purchasing a controlling stake in the firm due to the share prices being below their true value, and target managers will typically get replaced if the bid succeeds (Agrawal and Walkling, 1994). This hypothesis can be measured by EBITDA margin ROE, ROCE, ROA and/or asset turnover.

- **Undervalued Firms**

This hypothesis suggests that firms with low market value compared to book value are targets since they represent a 'cheap buy' (Powell, 1997; Palepu, 1986). It utilises market to book and price to earnings ratios where a bidder will bid for an overvalued firm if it was still less overvalued than the bidder (Dong, et. al, 2006).

- **Firm Size**

Firm size plays a significant role in takeover probability, the bigger the size the lower the probability of it being taken over, (Palepu, 1986), which explains why usually bigger firms acquire smaller ones (Levine and Aaronovitch, 1981). It has been shown that size is a significant factor (Powell, 1997) as measured by market capitalization and total assets.

- **Leverage, Liquidity and Growth**

Powell and Yawson (2007) debated that many takeovers occur as a way to rescue the target firm from a certain bankruptcy due to high debt and poor performance. Therefore firms with low growth and high leverage are more likely to be classified as targets and measured by debt to equity, current ratio and growth in revenues. While low liquidity does not single-handedly affect the takeover likelihood, when coupled with growth and leverage it can have a significant effect. (Palepu, 1986; Cremers et. al, 2008b)

Table 3: Takeover Determinant Variables and Their Ratios

| Takeover Hypothesis	Ratios
Inefficient Management	- EBITDA Margin - ROE - ROCE - ROA - Asset Turnover
Undervalued Firms	- Market to Book - Price to Earnings
Firm Size	- Market Capitalization - Total Assets
Leverage	- Debt to Equity - Equity Multiplier
Liquidity	- Current Ratio - Net Working Capital
Growth	- Growth in Annual Sales

4. Sample, Methodology and Data

Our study required three generic groups of data. M&A transactions records, number of public firms in the technology sector from the year 2000 to 2016, and the relevant financial ratios for the same period. All sample data were gathered from Bloomberg. We define a technology firm as a type of business entity that focuses primarily on the manufacturing and development of technology. This also includes the dissemination of information via high tech companies. It also includes information technology (IT) companies as subsets of technology companies as provided by the NAICS coding system where we placed several restrictions and criteria for selecting our sample. We observe that the number of public technology firms in the United States has been declining over the last 17 years as shown in figure 4. It is important also to mention that the decrease in public firms is not only affecting the technology sector, as it is affecting the whole U.S stock market. Since 1996 the number of public firms in the U.S has decreased by 50 percent, as a result of: (*i*) firms being delisted, acquired or bankrupt; (*ii*) less Initial Public Offering (IPO) activities, where firms remain private due to available capital provided by Venture Capital and Private Equity firms (Mauboussin, et al., 2017).

Figure 4: The Decrease in Number of Public Technology Firms in the U.S from 2000 to 2016
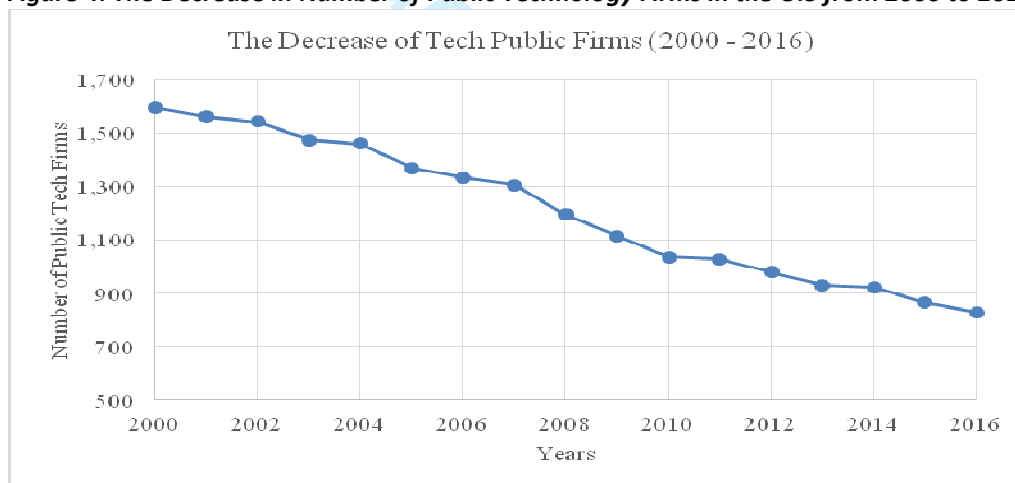
We pose certain sample restriction criteria for the purposes of our study. First, our study period covers the last 17 years where M&A transactions announced between the year 2000 and 2016 are included; second, we eliminate private firms where the target is a publicly traded company and having its domicile in the United States; third, we screen only technology firm targets and exclude firms operating in irrelevant sectors where the target is classified as a technology company by their NAIC code; fourth, we exclude investments, joint ventures, spinoffs and buybacks; fifth, we include only M&As transactions that are *pure* mergers or acquisitions where the acquirer owns more than 50% of the targets' shares. The total number of M&A transactions based on such criteria reached 966 transactions. Figure 5 illustrates the acquirers' industries by number of deals. More than 80% of the M&As were completed. Ninety-three percent were classified as friendly takeovers with 3% representing hostile takeovers. The rest are classified as unsolicited/unsolicited-to-friendly. Technology firms were 53% of the acquirers' transactions. Financial firms came in second at 19%.
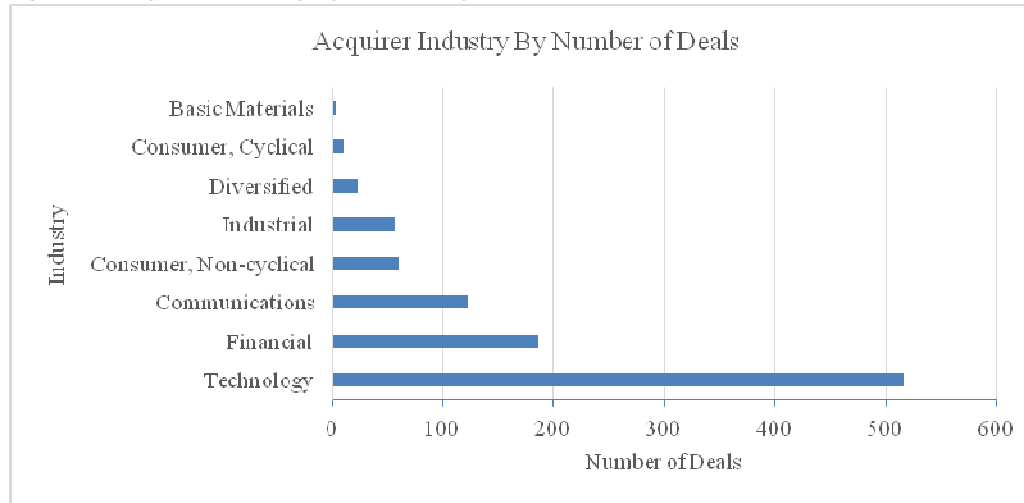
Figure 5: Acquirer Industry by Number of Deals

Figure 6 below, shows the number and value of deals in the technology sector in the U.S. The total dollar value of these transactions for our period of study reached $1.025 trillion. Most acquirers in our sample came from the United States with 88% and the remaining came from Europe with 10%.

Figure 6: M&A Activity in the Technology Sector in United States (2000 – 2016)
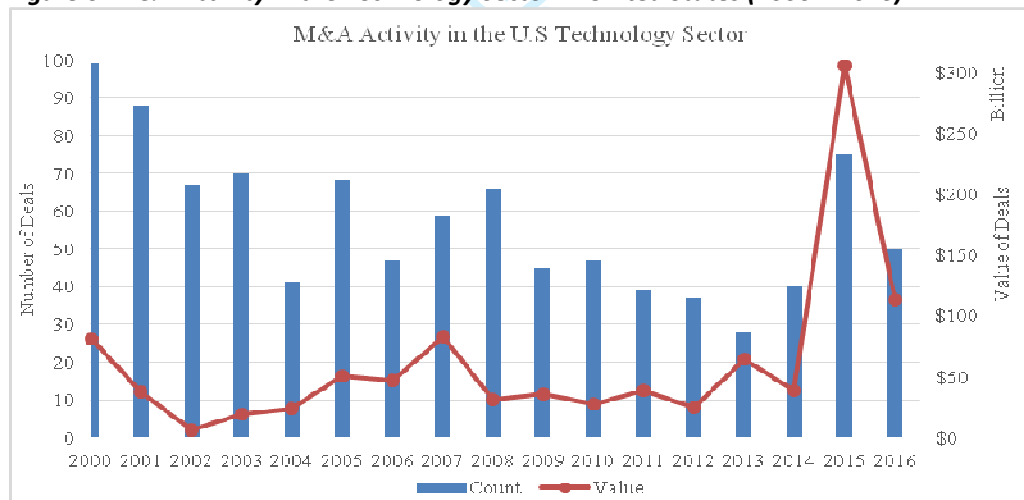
4.1 Datasets

Our study sample consists of two datasets, targets and non-targets. The target-group dataset includes firms which got acquired or received a bid to be acquired within our study period. The non-target group dataset includes firms which did not get acquired or received a bid to be acquired during the same period. The number of firms in our target dataset reached 846. Due to data pre-processing and omitted values this number was brought down to 415. We followed Palepu (1986) in choosing pre-determined ratios for the purposes of consistency and comparability but also in order to avoid the statistical overfitting issue (see also further support in section 4.3.1 below).

From this sample, 102 firms (24.5%) did not provide for a meaningful P/E ratio, and a further 87 firms (21%) did not have information on liquidity ratios. This further resulted in 189 firms been dropped from the sample producing a final 226 usable observations. The non-target dataset reached 2,340 firms.

4.2 Modelling

We apply two distinct methods in order to account for the different predictive accuracy of the two categories (target and non-target). A traditional statistical technique as well as a machine learning, predictive analytics technique, the MLP, has been used to model M&A activity at the developed capital markets and to predict potential targets.

4.2.1   Model 1: Multilayer Perceptron Model (MLP) Analysis Method

Over the last decade, a renewed growing interest in neural networks as a tool for data analysis has been observed. To a certain extent, the attractiveness of artificial neural networks vis-a-vis other statistical methods may have also been partially caused by human issues that merit some mention: often there is a shortcoming of statisticians to clearly communicate their methodologies and algorithms to non-statisticians. A large amount of the extant statistical knowledge raises a hurdle for potential investors of their methods. Neural networks on the other hand, are in a mid-embryonic phase, meaning that the current knowledge is thinner compared to statistical techniques. Artificial Neural Networks (ANN) originate from the biological human brain neurons. It is a network of nodes connected with each other through a weighted connection (Roiger, 2016), which can be greatly beneficial for complex non-linear relationships between variables (Hyndman & Athanasopoulos, 2013). ANN has been used in many industries such as telecommunications, industrials, banking, airlines and healthcare, and has been successfully showcased by (Widrow et. al, 1994). An example/representation of this model is shown below in figure 7.

Figure 7: Multilayer Perceptron Model
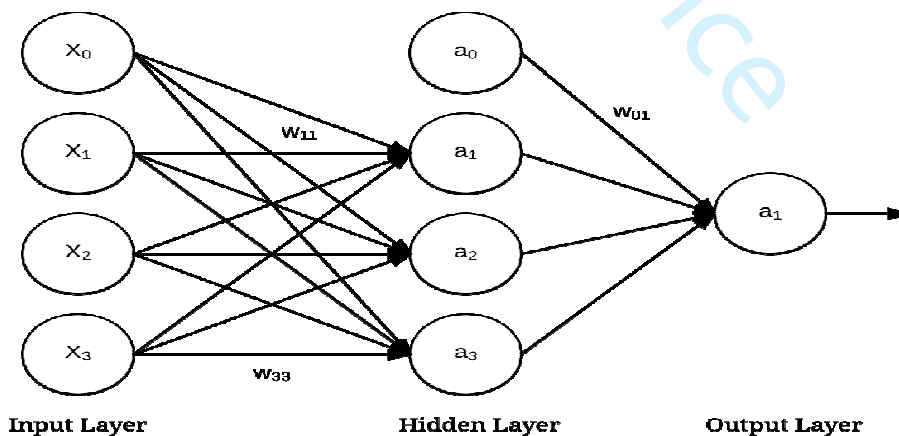
The nodes in the input layers are passive nodes as they only pass the data from the input layer to the hidden layer. In the hidden layer, a weight (W_n) will be generated for each input node.

For the first iteration it is a randomly generated number based on Gaussian distribution. Then each input (X_n) will be multiplied by its weight (W_n) to produce a weighted input (XW_n). The summation of these weighted inputs goes into the activation function to produce an output between (0, 1). The output number gets transferred to the output layer, there they get multiplied again with another set of randomly generated weights to produce the final output number between (0, 1). The model then compares the output number with the target number and calculates the difference. It then adjusts the weights in order to decrease the sum of margin errors (i.e. the cost function).

Input Layer Variables

This layer consists of pre-determined, industry related financial ratios derived and in line with established financial literature. Table 4 shows the financial ratios used in this study which have been used by a number of influential research papers (Ohlson, 1980; Palepu, 1986; Powell & Yawson, 2007). EBITDA and ROA for most technology firms in our sample were non-existent hence had to be dropped as candidate variables as they would limit the sample to less than 100 firms. The number of nodes in this layer simply equals the number of independent variables, in our case 6 for each instance.

Table 4: Financial Ratios Used in our Study

Takeover Hypothesis	Ratio
Inefficient Management	Return on Equity
Undervalued Firms	Price to Earnings Ratio
Firm Size	Market Capitalization
Leverage	Debt to Equity Ratio
Liquidity	Current Ratio
Growth	Rate of change of Annual Revenues

Hidden layer

This layer will receive the nodes sent from the input layer. It will generate a weight for each connection between any node in input layer and any node in the hidden layer. Then it will multiply each node with its weight as shown in equation 1 below:

Equation 1: Summation of Weighted Inputs

$$Net\ Input\ Function(x) = z = b + X_1W_1 + X_2W_2 + \cdots X_nW_n$$

b   :   Bias node

X_1 :   Financial ratio (Ex: ROE for the first instance)

W_1 :   Weight associated with X_1 (Randomly generated number between 0 and 1)

The net input function z, will go into a non-linear activation function (sigmoid function). It will act as a smooth thresholding function to determine the relationship between inputs and outputs. Our sigmoid function performs better for negative variables and classifiers (Zhang, et al., 1998) based on equation 2 below.

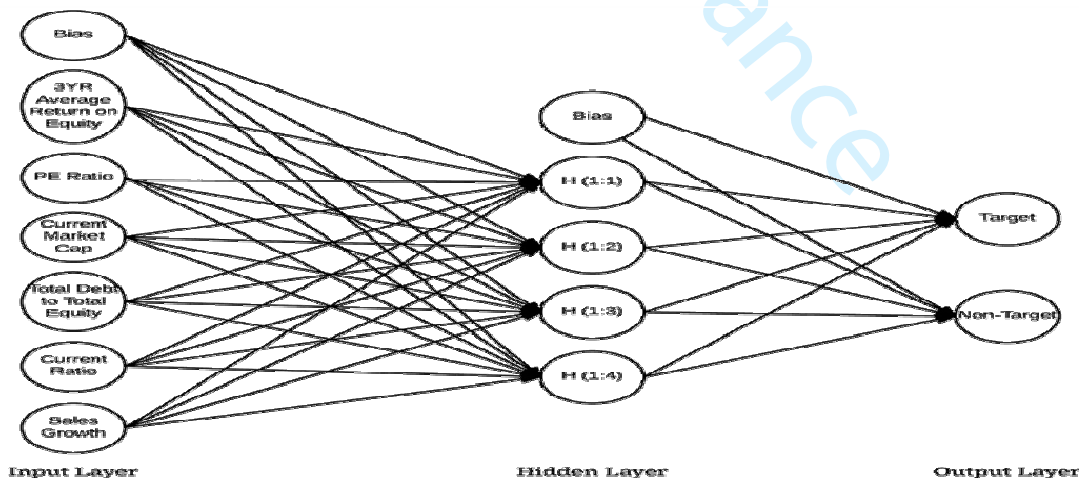***Equation 2: Sigmoid Activation Function***

$$\phi(z) = \frac{1}{1 + e^{-z}}$$

With differentiation $\varphi'(z) = \varphi(z)(1 - \varphi(z))$, in updating the curve. The cost function used in the study was sum of squared errors using an optimisation Gradient Decent method with the following parameters: Initial Learning Rate = 0.4 and Momentum = 0.9. The Cost Function C, given as:

$$C = 1/n \sum_{i=1}^{n}(z_i - \phi(x_i))^2, \qquad \text{with } \phi(x_i) \text{ the output for } xi.$$

Output layer

The output value will then be multiplied with its connection weights again and the final value will go into the Output layer. Hidden layers adjust the weightings on those inputs until they reach the optimization stage that is, the error of the neural network is minimized. An interpretation of this is that the hidden layers extract salient features in the input data which have predictive power with respect to the outputs. This is the discussed feature extraction function and it is parallel to the function of statistical techniques such as principal component analysis. This layer consists of a binary node[4]; it will receive the value from the hidden layer, indicating the dataset which the firm is predicted to be in. One indicates a target, zero indicates a non-target. The final output value will be compared with the desirable target value. This whole process is called *Standard Forward Propagation.* The final model architecture is shown in table 8 below:

Figure 8: Model I Final Architecture

---

[4] The end-outcome will be either output 'target' or 'non-target'; i.e. 1 or 0

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Validation

We used the cross-validation method that involves dividing the data records into three sets: (*i*) Training data set: data records that are used to train the model; (*ii*) Testing data set: records that are used to observe the error rate while training in order to further tweak the model; (*iii*) Holdout data set: this set of records is used to assess the model's final error rate and performance. Validation is used to measure the performance and the generalization ability of this model (Kaastra & Boyd, 1996). While there is no standardized rate of division in the literature some researchers (Hammerstrom, 1993) recommend using the 70/30 ratio. In our study, our data is randomly divided into three groups as follows: 70% for training; 20% for testing; and 10% as a holdout. We first clustered the data into years (Sample 1). Then it was clustered into target and non-target firms. Next, the records were randomly sorted, and the analysis was performed on 3 sets:

- Sample 1: All records of all years, randomly sorted.

  However, once this approach was finalised we discovered that this would potentially create a considerable over-training issue because then the requirement would be to repeat the steps above 17 times (17 years) with the same companies appearing on all data sets. We then took a 2nd sample approach.

- Sample 2:
  - 50/50: 50% of the training data consisted of non-target firms, and 50% target.

Data itemization in our study could potentially suffer from unreliability owed to sample limitations where the data available were not enough to train different networks on different subsets of the data. Consistent with Srivastava et al. (2014), Dekel et al. (2010) and Hinton & Salakhutdinov (2006), at this stage we only performed the training once in order not to fall into the over fitting and overtraining where the network would just *memorise* the outcome and not learn thus making it only usable in our specific data set. The data was fed to the network in at once but it used the data 10 times (learning epochs = 10) to update the weights. Following the above authors' prior work on data size and data diversity considerations we performed the experiment based on 3 trials and then took the average of these trials as shown in the analysis.

4.2.2 Model 2: Logistic Regression Model (LR)

Logistic regression method was used as the nature of our study is to forecast takeover targets. Therefore, the output is always binary (i.e. target, non-target) so it is important to use a technique that can classify a data instance into two classes by predicting the probability of an input being in a certain class. We convert the values of our independent variable from a string format to a numerical format assigning the following codes: 0 = Non-Target, 1 = Target. The logistic regression model starts with no predictive variables and only includes the intercept (constant) and measures the prediction power of this model using -2 Log Likelihood.

It then adds one predictive variable per step and calculates -2 log likelihoods again to measure if the new variable improved the prediction accuracy for all predictive variables. The model allows us to calculate the odds of an input (firm) to be acquired or not using ($Odds = e^{a+b_i X}$) where, a is the intercept (constant), b is the predictive variable added in step i, and X is the independent variable (Acquisition Status, 0 or 1).

Next, we convert the odds to probabilities using ($Probabilities = \frac{Odds}{1+Odds}$). Based on the probabilities result for each input the model classifies them into target or non-target based on a threshold (0.5), any input with a probability equals 0.5 or more will be classified as target, anything less than that will be classified as non-target. Based on this classification, the model produces a classification showing the number of cases correctly classified versus incorrect classifications in order to produce an overall prediction accuracy.

Predictive Variables

The same inputs from Model 1 are utilised, therefore maintaining consistency in our predictive variables (independent variables) namely, Return on Equity, Price to Earnings, Market Capitalization, Debt to Equity, Current Ratio, Rate of change of Annual Revenues. Our model is based on 3 traditional empirical formulae as proposed by Swaminathan and Rogers (1990) formulation of the logistic regression procedures where:

Equation 3: Odds Function

$$Odds = \frac{P_i}{1 - P_i}$$

Equation 4: Probability Using Odds

$$P_i = E\left(Y = 1 \mid X_i\right) = \frac{1}{1 + e^{-Z_i}}$$

Equation 5: Logistic Regression Equation

$$\ln\left(odds\right) = \ln\left(\frac{P_i}{1 - P_i}\right) = Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Odds   :   Ratio of probability occurring divided by the probability of it not occurring

P_i :   The probability of firm i being taken over

β_0 :   The intercept

Z_i :   The weighted sum of the predictive variables

β_n :   The coefficients for the financial ratio X_n

5. Results and Findings

Model I: MLP

The results of our sample include a total number 226 technology firms, 50 percent target firms and 50 percent non-target firms. Table 5 shows the prediction percentages for training, testing and holdout datasets based on three trials. As the table shows 70% of the data is reserved for training, 20% for testing and the final 10% for our final holdout sample.

Table 5: 50/50 Sample Cases Summary – Model I

Case Processing Summary

		N	Percent
Sample	Training	157	69.45%
	Testing	46	20.35%
	Holdout	23	10.20%
Valid		226	100.0%
Excluded		0	
Total		226	

We apply a standard feedforward propagation neural network with a single hidden layer in our sample in order to identify potential takeover targets (and hence for example, the possibility to yield positive abnormal returns from investing in these targets stocks).

Below, table 6 summarizes the results of our analysis where the predictive ability of the model is tested. The variables applied to the neural network models are the return on equity, price to earnings ratio, market capitalization, debt-to-equity, current ratio, and rate of change of annual revenues and by default, industry. Overall the results are promising compared to the standard binary regression technique. Our sample had an out-of-sample overall average prediction accuracy of 71.4 percent, with an average of 28.6 percent of incorrect predictions.

Table 6: Model 1, 50/50 Sample Cases Results

		Trial 1	Trial 2	Trial 3
Training				
	Non-Target	86.1	71.6	58.3
	Target	40.5	49.4	61.9
	Overall	61.5	61.1	60.1
Testing				
	Non-Target	72.0	52.9	60.9
	Target	28.6	60.9	52.4
	Overall	52.2	57.5	56.8
Holdout				
	Non-Target	87.5	87.5	50.0
	Target	25.0	54.5	87.5
	Overall	66.7	68.4	71.4

The attempt rate (i.e. trials) at 3 trials showed improvement in our holdout sample for correctly identifying the target companies with a 87.5% accuracy prediction rate. Yet it has to be recognised that it also correctly identified non-targets only 50% of the time giving an overall prediction accuracy rate of 71.4% at trial three. Our neural network model attempts to provide a tool that can adaptively sift through noise and identify patterns in complicated financial relationships where non-linearity might pose problems. Using 6 inputs considered to be the most relevant, and having only 4 hidden nodes our sample gets around the issue of having a relatively small dataset. Adaptability also lies in the recognition of not adding too many nodes which could lead to mode overfitting.

Our results support that such an approach can potentially provide meaningful explanation regarding dependent and independent variables compared to a traditional regression model. We turn to this below in tables 7 and 8.

Table 7: Model 2: Regression: 50/50 Sample Cases Summary

Unweighted Cases		N	Percent
Selected Cases:	Included in Analysis	226	100
	Missing Cases	0	.0
	Total	226	100
	Unselected Cases	0	.0
Total		226	100

Table 8: 3-Step Classification Table[a] - Model II

|	Observed		Predicted		
			Acquisition Status		Percentage Correct
			Non-Target	Target	
Step 1	Acquisition Status	Non-Target	59	54	52.2
		Target	38	75	66.4
	Overall Percentage				59.3
Step 2	Acquisition Status	Non-Target	57	56	50.4
		Target	35	78	69.0
	Overall Percentage				59.7
Step 3	Acquisition Status	Non-Target	65	48	57.5
		Target	38	75	66.4
	Overall Percentage				61.9

a. The cut value is .500

Table 8 above, illustrates the three steps taken by our regression model when adding new predictive variables to the model and the accuracy of correct predictions on each step. The model was able to increase the accuracy with each step, albeit marginally, reaching an overall accuracy of 61.9%. This model correctly identified the target companies with a 66.4% accuracy prediction rate and it also correctly identified non-targets only 57.5% of the time. Comparatively, the first model achieves a higher accuracy overall over model 2 providing some support for the utilisation of NNs.

It has to be said though that, the Training, Testing, and Holdout results differ from each other in each of the 3 trials for the 50/50 samples. We suggest that it is due to the random number generator where the network starts with a random initial numbers to start with and then keeps updating the weights accordingly; this is important in order to create a global optimum solution. In the first instance we actually had an average 6% change from one step to the next but a variability of 16.5% in-between the steps. Compared to the second model the step difference is 1% with a variability of 14.5% in-between the steps. The observations drawn are: (i) the regression model is static throughout the sample and trials whereas the NN model shows evolution and adaptability, (ii) there are large swings in variable values where for example, the RoE, D/E and liquidity swing wildly – deep in negative and high up in positive territory - from year to year, and (iii) the number of observations is relatively limited where the holdout sample is strictly anecdotal data since it covers only a limited number of observations, hence the expressive power of the network is potentially not enough to capture the target function. One alternative would be to add more layers or more hidden units in fully connected layers. So while it's helpful to test different methods, and provide for greater accuracy, it does not by itself, conclusively determine which method is best owed to data limitations. In addition, an examination of the variables also provides some interesting insights.

Table 9 below shows the importance of each variable fed into the model in terms of characterising its output. While the variable importance analysis below shows the input effects on the output it can be also clearly seen that the 3 variables mentioned above account for over 80% of the effects on output.

Table 9. Variable Importance Analysis
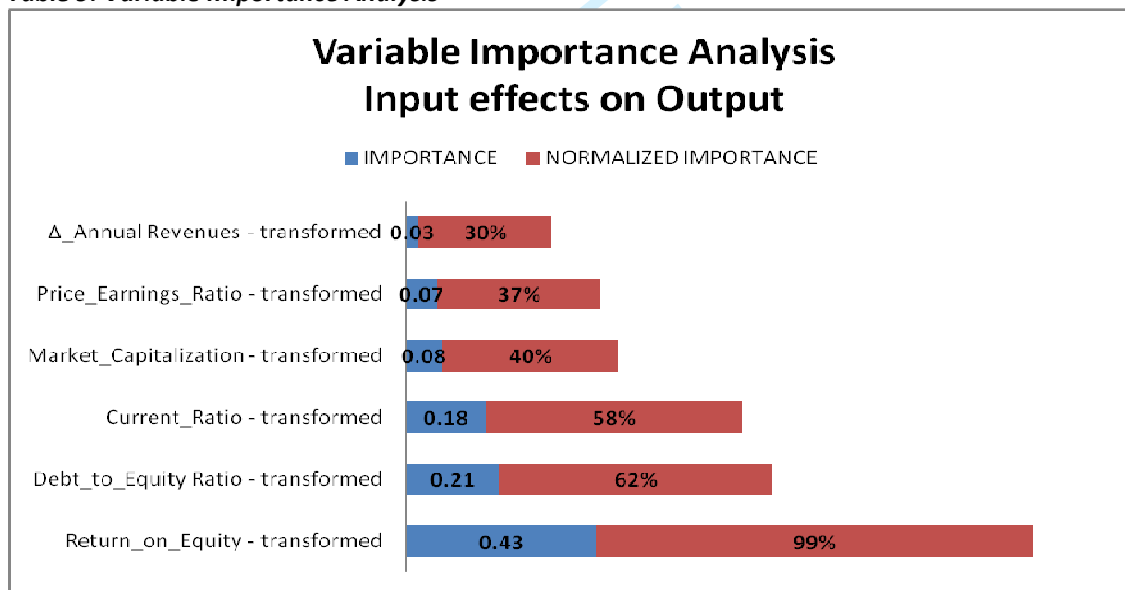
The variable importance analysis showed a great importance for ROE, D/E ratio and liquidity. These are consistent with the inefficient management, leverage and liquidity takeover hypotheses but the direction of the relationship between the independent and dependent variables is not clear. It is also these 3 variables that showed the greatest volatility throughout our sample period.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

6. Conclusions and limitations

This paper examines the use of a neural network method for pricing mergers. With over 50% of mergers failing, it is critical for acquiring firms to identify the characteristics of a target prior to a merger that will provide synergies once the merger is complete. The neural network model presented in this paper simplistic as it may be at this stage, shows overall improvements on the accuracy of predicting merger targets over linear regression results. ANN has outperformed logistic models in both senses of discrimination and calibration, although from the arbitrary standpoint of accuracy (cutoff point 0.5), logistic models can be superior to ANN models. The fact is that in some applications neural networks fit better than other models such as linear regression and this usually occurs when there are nonlinearities involved though it is important to evaluate other aspects. For example: a linear regression model will have less parameters to estimate compared to a NN for the same set of input variables. Hence, a NN will require a larger dataset for its calibration and subsequent optimization in order to get the required benefit of generalization, applicability and nonlinear mapping. In the absence of critically enough data, despite existing nonlinearities involved, a linear regression model may indeed be better calibrated.

Improvements as per the accuracy of target prediction can translate into significant savings in offering prices for target companies. Reliable predictions can improve the quality of decisions and business strategy in target determination and fair price decisions. Neural network methods permit the use of an expanding number of prospective venture opportunities with the added benefit that as market changes are introduced and more dynamic analysis is eventually involved new and more inputs can be loaded onto the model with less resource devotion. Having said that, it is also important to identify that neural network methods do not provide for a fuller analysis of significance for each of the autonomous variables in the model as traditional regression methods do.

Using a different activation function and a 'deeper' network with more hidden layers could potentially account for how each successive layer uses the output from the previous layer as input. It could also further show how the algorithm self learns from multiple levels of representations that correspond to different levels of abstraction (i.e. the levels form the hierarchy of concepts above). The quantity of data at our disposal though is relatively limited for more hidden layers to be involved in terms of describing potential causal connections between input and output. The transfer function is the calculated derivative sigmoid function utilized; we see this as important when calculating the weight updates in the network based on the amount of data and the computational load of our simulation. Lastly, while the extra layers could potentially help in learning features indeed the authors felt that with such a sample introducing LeRu we may also run the risk of naively training a 'deeper' neural network. As argued above, the possibility of added layers of abstraction could also show rare dependencies modelling in the training data.

It could arguably have been also interesting to investigate how model performance is influenced by using different activation functions (for example utilising the so called ReLU method – the Rectified Linear Unit - instead of Sigmoid,) or also involve in the analysis a higher number of hidden nodes. This is another area for research where traditionally, machine learning evaluation works best in producing an extrapolative model. The trade-off

though, of creating a flexible, nonparametric predictive model on the other hand, is that causal interpretations can potentially be lost. Equally, linear regression is a relatively inflexible approach yet it is less complicated in its interpretation. Flexible constructs avoid the assumptions of a particular functional form for a model, but they also require a larger number of observations and are more complicated and challenging to interpret. In addition, it can also be supported that NNs with different initializations produce different signals for a certain feature. As seen above, our NN with a certain initialization produced better signals in some cases and incorrect signals in some other.

Our results are also consistent with 20 years of research and some seminal papers that date as back as the 90s until today (see for example, Sen and Gibbs, 1994; Sinha and Richardson, 1998; Fescioglu-Unver and Tanyeri, 2013; Spangler et al., 2015; Tkáč' and Verner, 2016). Such studies indicate that although neural networks map the data satisfactorily, it is still questionable whether they predict merger targets significantly better than logistic regression. This strongly suggests that the financial models used to predict mergers are relatively inadequate. Firms should approach the development of merger prediction models cautiously and identify other factors that are more likely to predict mergers. Neural networks give the best overall results for the largest multiple classification cases. There is substantial room for improvement in overall performance for all techniques. The results indicate that data mining methods and data proportions and characteristics have a significant impact on classification accuracy. Zhu et al. (2001) for example, state that within data mining methods, rough sets provide better accuracy, followed by neural networks and inductive learning.

The generalization breadth of this study is limited within a specific sector (technology) in a specific country (United States) covering a specific period (2000–2016). One of the most important limitations was data collection, as we had to omit approximately 50 percent of the initial sample due to unavailable data on firms and their financial ratios. The takeover determinants were chosen from previous studies done by other researchers that showed statistical significance; this may affect the results of this analysis as the sample size, sector and study period are different. Further research can be done to extend this model and maybe improve the accuracy of it by including for example: (*i*) Technology Firms' Specific Ratios, this will allow the model to study technology firms not just from a financial but from an operational perspective too; (*ii*) Social Profiling, Social Media and softer Social variables not captured or modelled by standardised techniques, where these can be leveraged in order to discover opportunities or create maps for those interested audiences (Beese, 2015). Monitoring social media impression of the firm or its management might give an indication of its takeover probability.[5]

Old may be, but this echoes also Kuo and Reitsch's (1995) early research in the managerial forecasting problem; many managers value the 'softer' features of neural nets, particularly when standard regression models tend to emphasize the causal interpretations (more the *why*) of the problem and not the solution.

---

[5] Similar research has been done in this field by Xiang et al. (2012). *A Supervised Approach to Predict Company Acquisition with Factual and Topic Features Using Profiles and News Articles on TechCrunch*. in ICWSM.

1
2
3   **References**
4
5   Acker, O., Hagen, H. and Hajj, J., (2017) *2017 Technology Trends.* [Online] available at:
6   https://www.strategyand.pwc.com/trend/2017-technology-trends, [Accessed 16 May 2017].
7
8   Agrawal, A. and Walkling, R. A., (1994) Executive Careers and Compensation Surrounding Takeover
9   Bids, *The Journal of Finance,* 49(3): 985 - 1014.
10
11 Alexandridis, G., Mavrovitis, C. F. and Travlos, N. G., (2012) How Have M&As Changed? Evidence
12 from the Sixth Merger Wave*, The European Journal of Finance,* 18(8): 663-688
13
14 Altman, E. L., Marco, G. and Varetto, F., (1994) Corporate distress diagnosis: comparisons using
15 linear discriminant analysis and neural networks (the Italian experience), *Journal of Banking and*
16 *Finance,* 18(3): 505 - 529.
17
18 Altman, E. L., (1968) Financial Ratios, Discriminant Analysis And The Prediction Of Corporate
19 Bankruptcy, *The Journal of Finance,* 23(4): 589 – 609.
20
21 Anadol, B., Paradi, J. C., Simak, P. and Yang, X., (2014) Valuing Private Companies: A DEA Approach,
22 *International Journal of Business and Management,* 9(2): 16 - 28.
23
24 Andrade, G., Mitchell, M. & Stafford, E., (2001) New Evidence and Perspectives on Mergers, *Journal*
25 *of Economic Perspectives,* 15(2): 103 – 120.
26
27 Aydin , N., (2015) A Review of Models for Valuing Young and Innovative Firms, *International Journal*
28 *of Liberal Arts and Social Science ,* 3(9), p. 1 - 8.
29
30 Bakshi, G. and Chenb, Z., (2005) Stock Valuation in Dynamic Economies, *Journal of Financial*
31 *Markets,* May, 8(2): 111 - 151.
32
33 Barnes, P., (1998) Can Takeover Targets be identified by statistical techniques? some UK evidence,
34 The Statistician, 47(4): 573-591
35
36 Barnes, P., (1990) The Prediction of Takeover Targets in the U.K. by means of Multiple Discriminant
37 Analysi, *Journal of Business Finance & Accounting,* 17(1): 73–84.
38
39 Beck, T., Demirgüç-Kunt, A. and Levine, R., (2003) Law and finance: why does legal origin matter?,
40 *Journal of Comparative Economics,* 41(4): 653 - 675.
41
42 Beese, J., (2015) *What Is Social Listening & Why Is It Important?.* [Online]
43 Available at: https://sproutsocial.com/insights/social-listening/, [Accessed 21 July 2017].
44
45 Benou, G. and Madura, J. (2005) High-tech acquisitions, firm specific characteristics and the role of
46 investment bank advisors, *The Journal of High Technology Management Research*, 16(1):101-120
47
48 Betton, S. Eckbo, B. E. and Thorburn, K. S. (2008) Corporate takeovers, *Handbook of Corporate*
49 *Finance: Empirical Corporate Finance*, edited by B. E. Eckbo, Elsevier, Oxford, 2: 291-416
50
51 Bradley, M., (1980) Interfirm Tender Offers and the Market for Corporate Control, *The Journal of*
52 *Business,* 53(4): 345 - 376.

Branch, B., Wang, J. and Yangc, T., (2008) A note on takeover success prediction, *International Review of Financial Analysis,* 17(5): 186–1193.

Bureau of Economic Analysis (2017), Available: https://www.bea.gov, [3 September 2017].

Chandra, U., Wasley, C. E. and Waymire, G. B., (2011) Income Conservatism in the U.S. Technology Sector, *Accounting Horizons,* 25(2): 285 - 314.

Chan, L. K. C., Lakonishok, J. and Sougiannis, T., (2001) The Stock Market Valuation of Research and Development Expenditures, *The Journal of Finance,* 56(6): 2431 – 2456.

Coats, P. K. and Fant, L. F., (1993) Recognizing Financial Distress Patterns Using a Neural Network Tool, *Financial Management,* 22(3): 142 - 155.

CompTIA (2017), Cyberstates 2017, Research Report, accessed September 2017:
http://www.cyberstates.org/pdf/CompTIA%20Cyberstates%202017.pdf

Cremers, M., Driessen, J., Maenhout, P., and Weinbaum, D. (2008b) Individual stock-option prices and credit spreads, *Journal of Banking and Finance*, 32(12): 2706–2715.

Cudd, M. and Duggal, R., (2000) Industry Distributional Characteristics of Financial Ratios: An Acquisition Theory Application, *The Financial Review,* 35(1): 105 – 120.

Damodaran , A., (2001) *The Dark Side of Valuation : Valuing Old Tech, New Tech and New Economy Companies.* s.l.:Financial Times.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A., (1998) Investor Psychology and Security Market Under- and Overreactions, *The Journal of Finance,* 53(6): 1839 – 1885.

DeAngelo, L. E., (1990) Equity Valuation and Corporate Control, *The Accounting Review,* 65(1): 93 - 112.

Dekel, O., Shamir, O. and Xiao, L. (2010) Learning to classify with missing and corrupted features, *Machine Learning*, 81(2):149–178

Dietrich, K. J. and Sorensen, E., (1984) An application of logit analysis to prediction of merger targets, *Journal of Business Research,* 12(3): 393-402.

Dieudonne, S., Fabienne, C. and Slimane, B. (2014), M&A Activity: Where are we in the Cycle? September 17
Available at SSRN: https://ssrn.com/abstract=2504412 or http://dx.doi.org/10.2139/ssrn.2504412

Dong, M., Hirshleifer, D., Richardson, S. and Hong Teoh, S., (2006) Does Investor Misvaluation Drive the Takeover Market?, *The Journal of Finance,* 61(2): 725 - 762.

Eakins, S.G. and Stansell, S. R. (2003) Can value-based stock selection criteria yield superior risk-adjusted returns: an application of neural networks, *International Review of Financial Analysis*, 12:83–97

Eckbo, B. E. (2014) Corporate takeovers and economic efficiency, Annual Review of Financial Economics 6(1): 51-74.

Eden, L. and Dobson, W., (2005) *Governance, Multinationals, and Growth.* s.l.:Edward Elgar Publishing.

Ernst & Young, (2015) *Buying and Selling: Cross-border mergers and acquisitions and the US corporate income tax ,* s.l.: Business Roundtable.

Ernst & Young, 2016. *Technology: Capital Confidence Barometer.* Available at: http://www.ey.com/Publication/vwLUAssets/ey-ccb-15-technology/$FILE/ey-ccb-15-technology.pdf, [Accessed 16 May 2017].

Fan , A. and Palaniswami , M., (2000) *Selecting Bankruptcy Predictors Ssing a Support Vector Machine Approach.* Como, IEEE.

Fescioglu-Unver, N. and Tanyeri, B. (2013) A comparison of artificial neural network and multinomial logit models in predicting mergers, *Journal of Applied Statistics,* 40(4): 712-720

Gregoriou, G. N. and Neuhauser, K., (2007) *Mergers and Acquisitions: Current Issues.* United Kingdom: Palgrave Scholarly.

Griffin, J. M., Harris , J. H., Shu, T. and Topaloglu, S., (2011) Who Drove and Burst the Tech Bubble?, *The Journal of Finance,* 66(4): 1251 - 1290.

Hagedoorn, J. and Duysters, G., (2002) External Sources of Innovative Capabilities: The Preferences for Strategic Alliances or Mergers and Acquisitions, *Journal of Management Studies,* 39(2): 167 – 188.

Hammerstrom, D., (1993) Working With Neural Networks. *IEEE Spectrum,* 30(7): 46 - 53.

Harris, R. S. and Ravenscraft, D., (1991) The Role of Acquisitions in Foreign Direct Investment: Evidence from the U.S. Stock Market, *The Journal of Finance,* 46(3): 825 - 844.

Harris, R. S., Stewart, J. F. & Guilke, D. K., (1982) Characteristics of Acquired Firms: Fixed and Random Coefficients Probit Analyses, *Southern Economic Journal,* 49(1): 164 - 184.

Hinton G. and Salakhutdinov R. (2006) Reducing the dimensionality of data with neural networks, *Science*, 313(5786):504 – 507

Hodges, M. B., (2007) Three Approaches?, *The Appraisal Journal ,* 75(1): 34 - 44.

Hongjiu, L., Huimin, C. and Yanrong, H., (2007) 'Financial characteristics and prediction on targets of M&A based on SOM- Hopfield neural network,' *2007 IEEE International Conference on Industrial Engineering and Engineering Management*, Singapore, 2007, p. 80-84.

Hussain, A.J., Knowles, A., Lisboa, P.J.G. and El-Deredy, W. (2007) Financial time series prediction using polynomial pipelined neural networks, *Expert Systems with Applications*, 35:1186–1199

Hyndman, R. J. and Athanasopoulos, G., (2013) *Forecasting: Principles and Practice.* s.l.:OTexts.
Institute of Mergers, Acquisitions and Alliances, (2016) M&A Statistics.
Retrieved October 8, 2017, from
http://www.imaa-institute.org/recources/statistics-mergers-acquisitions/

Iturriaga, F. J. L. and Sanz, I. P. (2015), Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks, Expert Systems with Applications, 42(6): 2857-2869

Jensen, M. C., (1986) Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, *The American Economic Review,* 76(2): 323 - 329.

Jo, H. and Han, I., (1996) Integration of Case-Based Forecasting, Neural Network, and Discriminant Analysis for Bankruptcy Prediction, *Expert Systems with Applications,* 11(4): 415 - 422.

Jovanovic, B. and Rousseau, P. L., (2001) *Mergers and Technological Change: 1885-1998,* Vanderbilt University, Working Paper No. 01-W16, August

Kaastra, I. and Boyd, M., (1996) Designing A Neural Network For Forecasting Financial And Economic Time Series, *Neurocomputing,* 10(3): 215 – 236.

Kuo, C. and Reitsch, A. (1995) Neural networks vs. conventional methods of forecasting, *Journal of Business Forecasting Methods and Systems*, 14(4): 17-22.

Lam, M. (2004) Neural network techniques for financial performance prediction: integrating fundamental and technical analysis, *Decision Support Systems*, 37:567–581

Laudicina, P. A. & Peterson, E. R., (2017) *The 2017 A.T. Kearney Foreign Direct Investment Confidence Index.*[Online] Available at: https://www.atkearney.com/foreign-direct-investment-confidence-index [Accessed 3 September 2017].

Levine, P. and Aaronovitch, S., (1981) The Financial Characteristics of Firms and Theories of Merger Activity, *The Journal of Industrial Economics,* 30(2): 149 - 172.

Lin, C., Huang, J., Gen, M. and Tzeng, G. (2006) Recurrent neural network for dynamic portfolio selection, *Applied Mathematics and Computation*, 175:1139–1146

Lowenstein, R., (2004) *Origins of the Crash: The Great Bubble and Its Undoing.* s.l.:Penguin Books.

Maher, J. J. and Sen, T. K., (1997) Predicting Bond Ratings Using Neural Networks: A Comparison with Logistic Regression, *Intelligent Systems in Accounting, Finance and Management,* 6(1): 59–72.

Mauboussin, M. J., Callahan, D. and Majd, D., (2017) The Incredible Shrinking Universe of Stocks: The Causes and Consequences of Fewer U.S. Equities. *Global Financial Strategies*, 22 March.

Moody's Corporation, (2016) *US non-financial corporates' cash pile increases to $1.68 trillion, tech holding the lead.*
[Online] Available at: https://www.moodys.com/research/Moodys-US-non-financial-corporates-cash-pile-increases-to-168--PR_349330 [Accessed 16 May 2017].

Motiwalla, L. and Wahab, M. (2000) Predictable variation and profitable trading of US equities: a trading simulation using neural networks, *Computers and Operations Research*, 27:1111–1129

NASDAQ, (2017) *NASDAQ Composite Index (COMP).*
[Online] Available at: http://www.nasdaq.com/markets/nasdaq-composite [Accessed 6 September 2017].

Ohlson, J. A., (1980) Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research,* 18(1): 109 - 131.

Olson, D. L., Delen, D. and Meng, Y., (2012) Comparative Analysis of Data Mining Methods for Bankruptcy Prediction, *Decision Support Systems,* 52(2): 464 – 473.

Organization for International Investment, (2016) *Foreign Direct Investment in the United States,* Washington DC: www.ofii.org.

Palepu, K. G., (1986) Predicting Takeover Targets: A Methodological and Empirical Analysis, *Journal of Accounting and Economics,* 8(1): 3-35.

Palepu, K. G., (2003) *Business Analysis and Valuation Using Financial Statements.* Third Edition ed. s.l.:South-Western.

Park, Y. and Park, G., (2004) A New Method for Technology Valuation in Monetary Value: Procedure and Application, *Technovation,* 24(5): 387 - 394.

Petitt, B. S. and Ferris , K. R., (2013) *Valuation for Mergers and Acquisitions.* s.l.:Pearson Education, Inc.

Powell, R. and Yawson, A., (2007) Are Corporate Restructuring Events Driven by Common Factors? Implications for Takeover Prediction, *Journal of Business Finance and Accounting ,* 34(7-8): 1169 – 1192.

Powell, R. G., (1997) Modelling Takeover Likelihood, *Journal of Business Finance and Accounting,* 24(7 - 8): 1009 – 1030.

Rhodes-Kropf, M. and Viswanathan, S., (2004) Market Valuation and Merger Waves, *The Journal of Finance,* 59(6): 2685 - 2718.

Ritter, J. R. (2017) Initial Public Offering: updated statistics. *Working paper*, University of Florida

Roiger, R. J., (2016) *Data Mining: A Tutorial-Based Primer,* Second Edition ed. s.l.:Chapman and Hall/CRC.

Rojas, R (1996), Neural Networks: A Systematic Introduction, Springer-Verlag, Berlin

Rossi, S. and Volpin, P. F., (2003) Cross-Country Determinants of Mergers and Acquisitions, *Journal of Financial Economics,* 74(1): 277 – 304.

Routledge, B. R., Sacchetto, S. and Smith N. A. (2013) Predicting Merger Targets and Acquirers from Text, working paper, Carnegie Mellon University

Salchenberger, L. M., Cinar, E. M. and Lash, N. A., (1992) Neural Networks: A New Tool for Predicting Thrift Failures, *Decision Sciences,* 23(4): 899 – 916.

Sen, T. K. and Gibbs, A. M. (1994) An Evaluation of the Corporate Takeover Model Using Neural Networks, *Intelligent Systems in Accounting, Finance and Management*, 3(4): 279-292

Servaes, H. and Zenner, M., (1994) Taxes and the Returns to Foreign Acquisitions in the United States, *Financial Management,* 23(4): 42 - 56.

Sharda, R. and Odom, M. D., (1990) *A Neural Network Model for Bankruptcy Prediction.* s.l., IEEE Xplore.

Simkowitz, M. and Monroe, R. J. (1971) A Discriminant Analysis Function for Conglomerate Targets, *The Southern Journal of Business*, 38: 1-16

Sinha, A. P. and Richardson, M. A. (1998) A Case-Based Reasoning System for Indirect Bank Lending, *Intelligent Systems in Accounting, Finance and Management*, 5(4): 229-240

Spangler, W. E. May J. H. and Vargas L. G. (2015) Choosing Data-Mining Methods for Multiple Classification: Representational and Performance Measurement Implications for Decision Support, *Journal of Management Information Systems*, 16(1): 37-62,

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, 15: 1929-1958

Swaminathan, H. and Rogers, H. J., (1990) Detecting Differential Item Functioning Using Logistic Regression Procedures, *Journal of Educational Measurement,* 27(4): 361 - 370.

Tam K. Y. and Kiang, M. Y., (1992) Managerial Applications of Neural Networks: The Case of Bank Failure Predictions, *Management Science,* 38(7): 926-947.

Teeter, P. and Sandberg, J., (2016) Cracking the Enigma of Asset Bubbles With Narratives, *Strategic Organization ,* 15(1): 91 - 99.

Thawornwong, S. and Enke, D. (2004) The adaptive selection of financial and economic variables for use with artificial neural networks, *Neurocomputing*, 56:205–232

The Institute for Mergers, Acquisitions and Alliances, (2016) *United States - M&A Statistics.* Available at: https://imaa-institute.org/m-and-a-us-united-states/, [Accessed 21 February 2017].

Tkáč' M. and Verner, R. (2016) Artificial neural networks in business: Two decades of research, *Applied Soft Computing*, 38: 788-804

Tsai, C. F. and Wu, J. W. (2008) Using neural network ensembles for bankruptcy prediction and credit scoring*, Expert Systems with Applications*, 34(4): 2639–2649

Tseng, F.-M. and Hu, Y.-C., (2010) Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks, *Expert Systems with Applications,* 37(3): 1846 - 1853.

U.S. Bureau of Labor Statistics, (2016) *U.S. Bureau of Labor Statistics.* [Online] Available at: https://www.bls.gov/home.htm, [Accessed 7 September 2017].

Versace, M., Bhatt, R., Hinds, O. and Shiffer, M. (2004) Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks, *Expert Systems with Applications*, 27:417–425

Westenberg, D. A. (2009) *Initial Public Offerings: A Practical Guide to Going Public*, Practising Law Institute

Westenberg, D. A., (2000) *The IPO Report,* s.l.: Hale and Dorr LLP.

Widrow, B., Rumelhart, D. E. and Lehr, M. A., 1994. Neural Networks: Applications In Industry, Business and Science, *Communications of the ACM,* 37(3): 93 - 105 .

WilmerHale, (2017), 2017 M&A Report, Accessed November 11, 2017:
https://www.wilmerhale.com/uploadedFiles/Shared_Content/Editorial/Publications/Documents/2017-WilmerHale-MA-Report.pdf

Xiang, G., Zheng, Z., Wen, M., Hong, J., & Rose, C. (2012) A Supervised Approach to Predict Company Acquisition with Factual and Topic Features Using Profiles and News Articles on TechCrunch, *Artificial Intelligence*, 607–610.

Zhang, G., Hu, M. Y., Patuwo, B. E. and Indro, D. C., (1999) Artificial Neural Networks In Bankruptcy Prediction: General Framework And Cross-Validation Analysis, *European Journal of Operational Research,* 116(1): 16 - 32.

Zhang, G., Patuwo, E. B. and Hu, M. Y., (1998) Forecasting With Artificial Neural Networks: The state of The Art, *International Journal of Forecasting,* 14(1): 35 – 62.

Zhu, D., Premkumar, G., Zhang, X. and Chu, C. H. (2001) Data Mining for Network Intrusion Detection: A Comparison of Alternative Methods, *Decision Sciences*, 32(4): 635-660