Confining value from neural networks

A sectoral study prediction of takeover targets in the US technology sector

Ioannis Anagnostopoulos

Kingston Business School, Kingston University, Kingston-Upon-Thames, UK, and

Anas Rizeq

Reach Holding, Dubai, United Arab Emirates

Abstract

Purpose – This study provides valuable insights to managers aiming to increase the effectiveness of their diversification and growth portfolios. The purpose of this paper is to examine the value of utilizing a neural networks (NNs) approach using mergers and acquisition (M&A) data confined in the US technology domain.

Design/methodology/approach – Using data from Bloomberg for the period 2000–2016, the results confirm that an NN approach provides more explanation between financial variables in the model than a traditional regression model where the NN approach of this study is then compared with linear classifier, logistic regression. The empirical results show that NN is a promising method of evaluating M&A takeover targets in terms of their predictive accuracy and adaptability.

Findings – The findings emphasize the value alternative methodologies provide in high-technology industries in order to achieve the screening and explorative performance objectives, given the technological complexity, market uncertainty and the divergent skill sets required for breakthrough innovations in these sectors.

Research limitations/implications – NN methods do not provide for a fuller analysis of significance for each of the autonomous variables in the model as traditional regression methods do. The generalization breadth of this study is limited within a specific sector (technology) in a specific country (USA) covering a specific period (2000–2016).

Practical implications – 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 information technology stock price bubbles, Silicon Valley and as the recent stratospheric rise of financial technology companies have also demonstrated.

Social implications – Numerous studies have shown that M&As are more often than not destroy value rather than create it. More than 50 percent 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.

Originality/value – Technology firm value is directly dependent on growth, consequently most of the value will originate from future customers or products not from current assets that makes it challenging for investors to measure a firm's beta (risk) where the value of a technology is only known after its commercialization to the market. A differentiated methodological approach used is the use of NNs, machine learning and data mining to predict bankruptcy or takeover targets.

Keywords Technology, Neural networks, Prediction, Takeover targets

Paper type Research paper

1. Introduction

Shareholder theory posits that management has a mandate to maximize 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 percent 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 utilize 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 (NNs) 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 (e.g. behavioral patterns that are not easily identified through established linear methods or continuously keep changing) NNs "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: current financial statements; firm's past history; and 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 do not 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) 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 overvalued) and therefore increasing M&A activity (Rhodes-Kropf and Viswanathan, 2004; Jovanovic and 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 USA, 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 and Neuhauser, 2007). This wave ended with the 2008 credit crisis:

• Seventh wave: 2010–present.

The wave which we are currently experiencing started gradually in 2010, and it coincided with the emergence of financial technology. Since 2015, it has reached an all-time high of

$2.9 trillion in value (The Institute for Mergers, Acquisitions and Alliances, 2016). It is also 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 utilizing 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 USA, valuation challenges and the technology sector in Section 2 that follows. In Section 3 we discuss methodological issues where determinant variables and NNs are compared with the traditional statistical techniques of discriminant analysis (DA) and logistic regression (LR) with regard 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; and

(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 non-linearity. 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 modeled 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 an 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 US technology sector

The USA is well known as the most preferred international investment destination measured by foreign direct investment (FDI) flows. In 1989, the FDI position in the USA (FDIUS) exceeded $400bn (Harris and Ravenscraft, 1991), while in 2014 for example it totaled $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 five years consecutively (Laudicina and Peterson, 2017). Figure 1 illustrates M&As as the most exercised type of investment in the USA in volume vs expansions and new establishments.

Rossi and 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 and Dobson, 2005; Beck *et al.*, 2003). Harris and Ravenscraft (1991) claimed that FDIUS increases when the dollar is weaker compared to the investor's home currency. Servaes and Zenner (1994) also affirmed that tax regulations have an impact on FDIUS indicating tax benefits for the investors. During the period 2000–2016 the USA 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; US firms represent on average 20 percent of global M&A as acquirers and 23 percent as targets by value (Ernst & Young, 2015).

The technology sector experienced the largest M&A activity in the USA, holding the highest number of transactions in the period between 2000 and 2016 which represented 19.9 percent of all US M&A transactions (The Institute for Mergers, Acquisitions and Alliances, 2016). The high volume of M&A transactions in this sector is attributed to three main causes: 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 $773bn in cash, accounting for 46 percent of total cash held by US non-financial firms of $1.68 trillion (Moody, 2016); industry trends, i.e. location-based
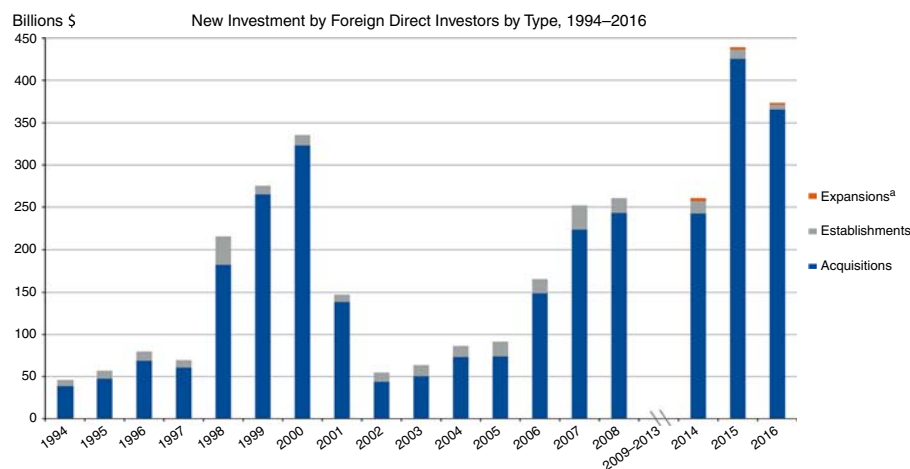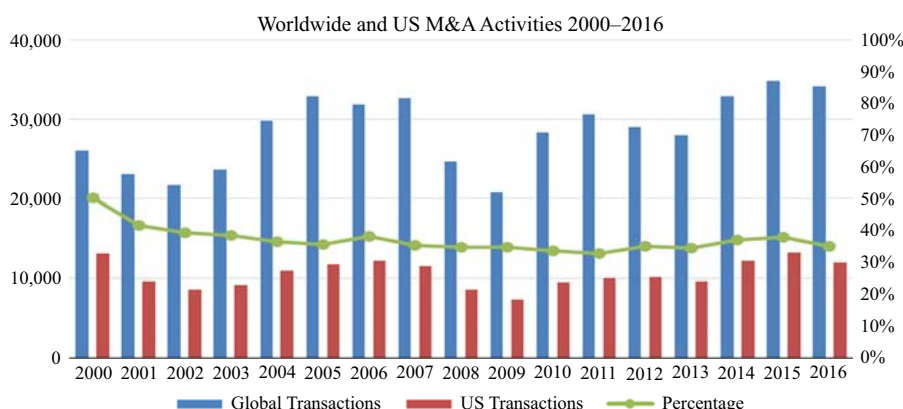

Figure 1.
FDI in the USA by type 1994–2016

Notes: [a]Expansions included in 2014–2016 only. Due to government budget cuts the Bureau of Economic Analysis was not able to fully conduct a concluding survey from 2009 to 2013
Source: US Bureau of Economic Analysis

Worldwide and US M&A Activities 2000–2016

Figure 2.
M&A activity;
worldwide vs USA

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 favored by 41 percent 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; talent acquisition: since 2011 the Big Five[2] technology firms have added more than 418,000 jobs to the market with 76 percent 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 percent of the total US economy. It represents about $1.3 trillion of value (CompTIA, 2017) and it employs more than 4 percent of all US workforce (US Bureau of Labor Statistics, 2016). The dot com bubble (1999–2000) was at the peak with 371 and 261 tech initial public offerings (IPOs), respectively, and in 1996 with 274 IPOs (Ritter, 2017). The growth of IPOs during the 1990s was fueled by venture capitalists excessively funding start-ups as funding rose from $3bn in 1990 to $60bn in 1999 (Lowenstein, 2004); furthermore, 57 percent 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 percent 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 percent 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 percent 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.

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 and Ferris, 2013). The dynamics of the technology sector are characterized by rapidly evolving firms which

Number of Tech IPOs in the USA 1980–2016

Figure 3.
US tech IPOs
1980–2016

Source: Ritter (2017)
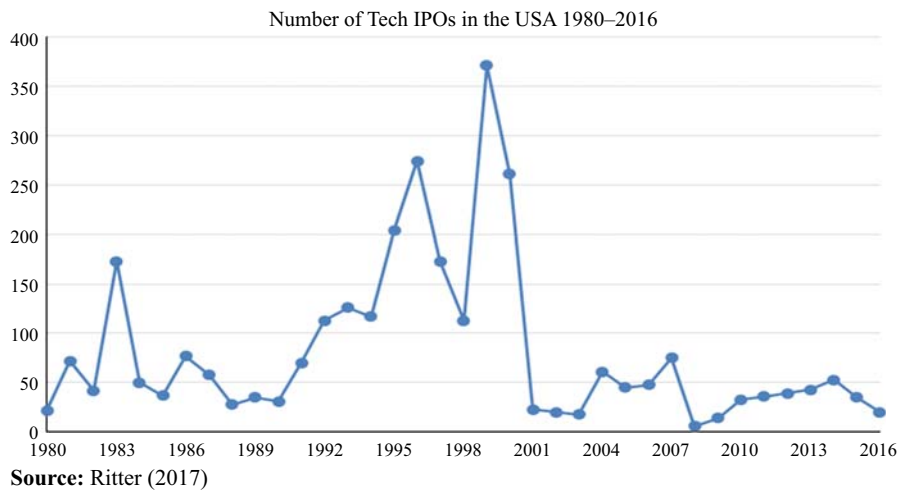
operate under high levels of uncertainty and risk (Lev and Zarowin, 1999). This, combined with the lack of positive cash flows (Aydin, 2015), makes their valuation very challenging as also demonstrated by Bakshi and Chenb (2005), where they demonstrate the potential for significant mispricing and departures from fair values. The complexity of valuing technology firms can be attributed to reasons such as:

(1) 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.

(2) The value of many firms in the technology sector usually comes from intangible assets. These assets however do not 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).

(3) 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).

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

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

DCFs as a method boast 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

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 criticized; 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

different in terms of prices (Bradley, 1980). Even hedge funds in the USA 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 (DA and LR); therefore, we shall consider both broad approaches below.

3.1 Traditional analytical techniques
3.1.1 Regression models. Ohlson (1980) utilized LR 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 LR, 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 DA. 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 and Sorensen (1984) used LR model to predict acquisition likelihood. Palepu (1986) used a binomial logit probability model with nine 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 five chosen industry-related ratios to increase the predictability of his model. While the previous studies, as above, have shown prediction power between 60 and 90 percent, Palepu (1986) argued, however, that these findings are overstated and suffer a biased estimate due to two main flaws in such methodologies: state-based sampling for model estimation and prediction testing; and using pre-determined, 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 and Duggal (2000) in their study used Palepu's (1986) factors 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 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.

   3.1.2 Discriminant analysis (DA). It 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 discriminant analysis (MDA) models consider simultaneously an entire portfolio of characteristics common to the firms and their interaction, whereas univariate models are limited to only one characteristic at a time. As a technique, DA does very well provide 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 MDA in-sample results predict 83 percent of the targets and 72 percent of the non-targets, while the holdout results are slightly worse predicting 64 percent of the targets and 61 percent of the non-target.

3.2 Machine learning (ML) techniques

A differentiated methodological approach used by researchers is the use of NNs, ML and data mining to predict bankruptcy or takeover targets. Sharda and Odom (1990) compared the use of both NNs and MDA models in bankruptcy predictions. In their study, the researchers utilized the same ratios used by Altman (1968) and after executing both models their findings suggest that NNs 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 NN classifiers in bankruptcy prediction and credit scoring where it was found that single NN classifiers outperformed multiple NN classifiers in both credit scoring and bankruptcy prediction. Hongjiu *et al.* (2007) used self-organized mapping with Hopfield NN to cluster data and their model showed accuracy predictions of 80.69 percent for targets and 63.11 percent for non-targets. Their paper also suggests the importance of including non-financial factors to improve the predictability power. Iturriaga and Sanz (2015) used multilayer perceptron (MLP) to predict bankruptcy of US 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 MDA in their sample 80 percent of the time. Numerous other studies have supported the use of NNs in outperforming LR in predicting bankruptcy (see e.g. Tam and Kiang, 1992; Jo and Han, 1996; Maher and Sen, 1997; Fan and Palaniswami, 2000; Tseng and Hu, 2010). Branch *et al.* (2008) utilized 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 nonlinearities. This echoes Zhang's *et al.* (1999) statement that NNs 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 LR, NNs, 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 data sets and observations) the accuracy level will decrease. Table II, by Barnes (1998), summarizes other research showing the prediction rates of various methods for North America and the UK.

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 utilizes 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.

Researcher	Country	Method	Classification (%)	Prediction (%)
Simkowitz 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	–	–
Palepu (1986)	USA	Logit	–	46
Ambrose (1990)	USA	Logit	65.1	75.6
Barnes (1990)	UK	MDA	68.5	–
Bartley and Boardman (1990)	USA	MDA	82.5	–
Ambrose and Megginson (1992)	USA	Logit	74.3	–
Walter (1994)	USA	Logit	65	66
Bartley and Boardman (1996)	USA	MDA	65	–

Notes: DA, discriminant analysis; GLS, generalized least squares; MDA, multiple discriminant analysis; –, not reported
Source: Barnes (1998)

Table II.
Previous studies on
the characteristics
of target firms

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.*, 2008) (Table III).

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 defined 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 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 observed that the number of public technology firms in the USA has been declining over the last 17 years as shown in Figure 4. It is also important to mention that the decrease in public firms is not only affecting the technology sector, as it is affecting the whole US stock market. Since 1996 the number of public firms in the USA has decreased by 50 percent, as a result of: firms being delisted, acquired or bankrupt; and less IPO activities, where firms remain private due to available capital provided by venture capital and private equity firms (Mauboussin *et al.*, 2017).

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 USA; third, we screen only technology firm

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

Table III.
Takeover determinant variables and their ratios

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 percent 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 percent of the M&As were completed. A total of 93 percent were classified as friendly takeovers with 3 percent representing hostile takeovers. The rest are classified as unsolicited/unsolicited-to-friendly. Technology firms were 53 percent of the acquirers' transactions. Financial firms came in second at 19 percent.

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

4.1 Data sets

Our study sample consists of two data sets, targets and non-targets. The target-group data set includes firms which got acquired or received a bid to be acquired within our study period. The non-target-group data set 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 data set 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
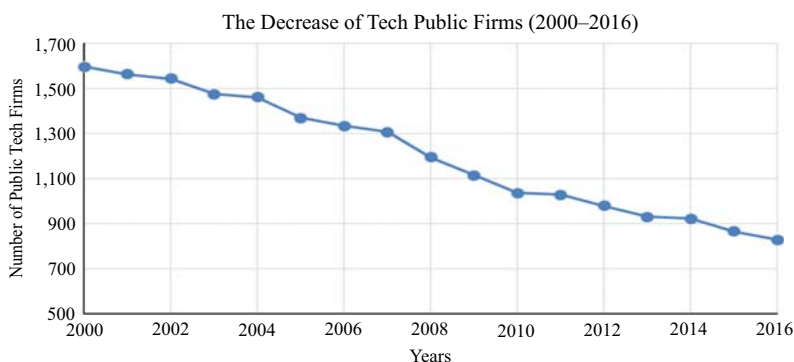

Figure 4.
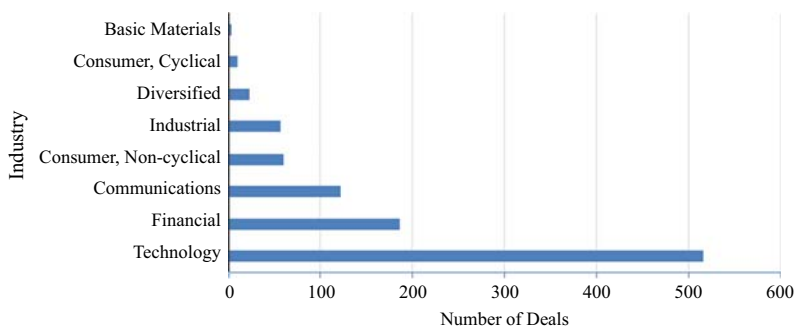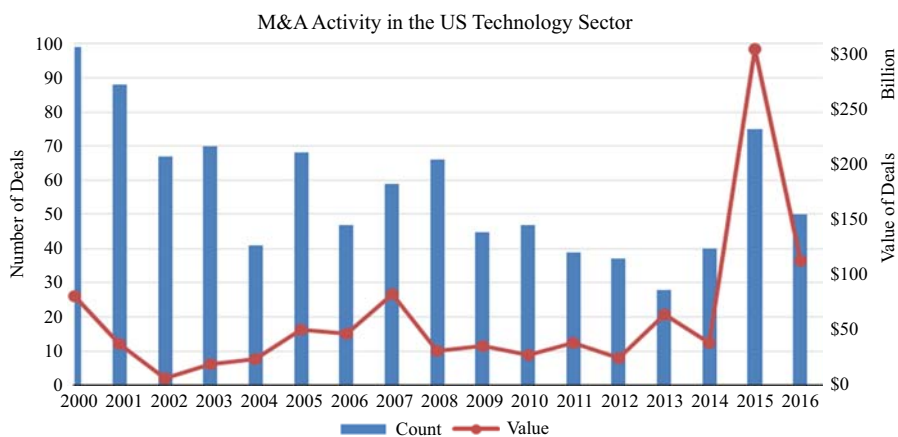The decrease in number of public technology firms in the USA from 2000 to 2016

Figure 5.
Acquirer industry by number of deals

Figure 6.
M&A activity in the technology sector in USA (2000–2016)

consistency and comparability but also in order to avoid the statistical overfitting issue (see also further support in Section 4.3.1).

From this sample, 102 firms (24.5 percent) did not provide for a meaningful P/E ratio, and a further 87 firms (21 percent) 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 data set reached 2,340 firms.

4.2 Modeling
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 an ML, 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 (MLP) model analysis method. Over the last decade, a renewed growing interest in NNs as a tool for data analysis has been observed. To a certain extent, the attractiveness of artificial NNs *vis-à-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. NNs, on the other hand, are in a mid-embryonic phase, meaning that the current knowledge is thinner compared to statistical techniques. Artificial neural networks (ANNs) 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 and 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 in Figure 7.

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 and 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
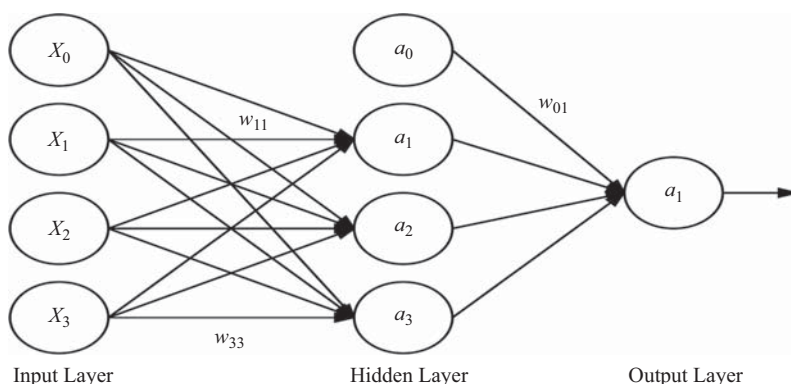
Figure 7.
Multilayer perceptron
model

final output number between 0 and 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 IV shows the financial ratios used in this study which have been used by a number of influential research papers (Ohlson, 1980; Palepu, 1986; Powell and 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 six for each instance.

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 the input layer and any node in the hidden layer. Then it will multiply each node with its weight as shown in the following equation (summation of weighted inputs):

$$\text{Net input function}(x) = z = b + X_1 W_1 + X_2 W_2 + \cdots X_n W_n, \quad (1)$$

where b is the bias node; X_1 the financial ratio (Ex: ROE for the first instance); W_1 the 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 the following equation (sigmoid activation function):

$$\phi(z) = \frac{1}{1 + e^{-z}}. \quad (2)$$

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

Table IV.
Financial ratios used
in our study

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 optimization gradient descent method with the following parameters: initial learning rate $= 0.4$ and momentum $= 0.9$. The cost function C is given as follows:

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

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 NN 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 data set 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 Figure 8.

Validation. We used the cross-validation method that involves dividing the data records into three sets: training data set: data records that are used to train the model; testing data set: records that are used to observe the error rate while training in order to further tweak the model; and 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 and 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 are randomly divided into three groups as follows: 70 percent for training; 20 percent for testing; and 10 percent 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 three sets:

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

     However, once this approach was finalized we discovered that this would potentially create a considerable overtraining issue because then the requirement
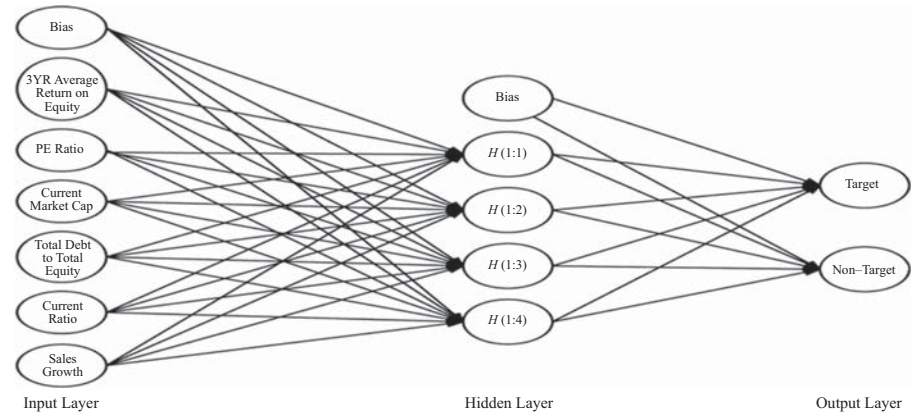

Figure 8.
Model I final
architecture

would be to repeat the steps above 17 times (17 years) with the same companies appearing on all data sets. We then took a second-sample approach.

- Sample 2:
  - 50/50: 50 percent of the training data consisted of non-target firms, and 50 percent 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 and Salakhutdinov (2006), at this stage we only performed the training once in order not to fall into the overfitting and overtraining where the network would just memorize the outcome and not learn thus making it only usable in our specific data set. The data were fed to the network in at once but it used the data ten 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 three trials and then took the average of these trials as shown in the analysis.

4.2.2 Model 2: logistic regression (LR) model. LR 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 LR 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 $(\text{Odds} = e^{a + b_i X})$, 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 = (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 vs incorrect classifications in order to produce an overall prediction accuracy.

Predictive variables. The same inputs from Model 1 are utilized; 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 three traditional empirical formulae as proposed by Swaminathan and Rogers (1990) formulation of the LR procedures where.

The Odds function is as follows:

$$\text{Odds} = \frac{P_i}{1 - P_i}. \tag{3}$$

The probability using Odds is as follows:

$$P_i = E(Y = 1 | X_i) = \frac{1}{1 + e^{-Z_i}}. \tag{4}$$

The LR equation is as follows:

$$\ln(\text{odds}) = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n, \tag{5}$$

where Odds is the 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; and β_n the coefficients for the financial ratio X_n.

5. Results and findings
5.1 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 V shows the prediction percentages for training, testing and holdout data sets based on three trials. As the table shows 70 percent of the data are reserved for training, 20 percent for testing and the final 10 percent for our final holdout sample.

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

Table VI summarizes the results of our analysis where the predictive ability of the model is tested. The variables applied to the NN 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

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	

Table V.
50/50 sample cases summary – model I

	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

Table VI.
Model 1, 50/50 sample cases results

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.

The attempt rate (i.e. trials) at three trials showed improvement in our holdout sample for correctly identifying the target companies with a 87.5 percent accuracy prediction rate. Yet it has to be recognized that it also correctly identified non-targets only 50 percent of the time giving an overall prediction accuracy rate of 71.4 percent at trial three. Our NN 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 six inputs considered to be the most relevant, and having only four hidden nodes, our sample gets around the issue of having a relatively small data set. 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 in Tables VII and VIII.

Table VIII 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 percent. This model correctly identified the target companies with a 66.4 percent accuracy prediction rate and it also correctly identified non-targets only

Unweighted cases	n	Percent
Selected cases		
Included in analysis	226	100
Missing cases	0	0.0
Total	226	100
Unselected cases	0	0.0
Total	226	100

Table VII.
Model 2: regression:
50/50 sample cases
summary

| | | Predicted | |
| | | Acquisition status | |
Observed	Non-target	Target	Percentage correct
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

Note: [a]The cut value is 0.500

Table VIII.
3-step classification
table[a] – model II

57.5 percent of the time. Comparatively, the first model achieves a higher accuracy overall over Model 2 providing some support for the utilization of NNs.

It has to be said though that the training, testing and holdout results differ from each other in each of the three 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 percent change from one step to the next but a variability of 16.5 percent in-between the steps. Compared to the second model the step difference is 1 percent with a variability of 14.5 percent in-between the steps. The observations drawn are: the regression model is static throughout the sample and trials whereas the NN model shows evolution and adaptability; 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 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 is 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.

Figure 9 shows the importance of each variable fed into the model in terms of characterizing its output. While the variable importance analysis shows the input effects on the output, it can also be clearly seen that the three variables mentioned above account for over 80 percent of the effects on output.

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. These three variables also showed the greatest volatility throughout our sample period.

6. Conclusions and limitations
This paper examines the use of an NN method for pricing mergers. With over 50 percent 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 NN model
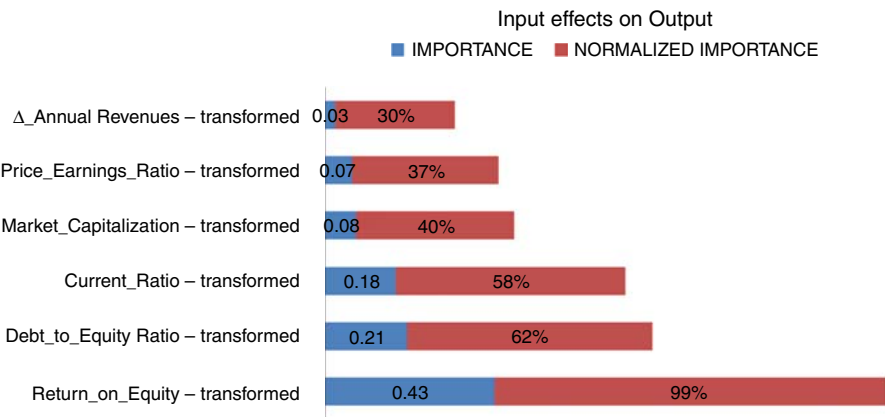

Figure 9.
Variable importance
analysis

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 (cut-off point 0.5), logistic models can be superior to ANN models. The fact is that in some applications NNs 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 an NN for the same set of input variables. Hence, an NN will require a larger data set 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. NN 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 NN 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. Finally, while the extra layers could potentially help in learning features indeed the authors felt that with such a sample introducing Rectified Linear Unit (ReLU) we may also run the risk of naively training a "deeper" NN. As argued above, the possibility of added layers of abstraction could also show rare dependencies modeling in the training data.

It could arguably have also been interesting to investigate how model performance is influenced by using different activation functions (e.g. utilizing the so called ReLU method instead of Sigmoid) or also involve in the analysis a higher number of hidden nodes. This is another area for research where, traditionally, ML 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 1990s until today (see e.g. 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 NNs map the data satisfactorily, it is still questionable whether they predict merger targets significantly better than LR. 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. NNs 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 NNs and inductive learning.

The generalization breadth of this study is limited within a specific sector (technology) in a specific country (USA) 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 improve the accuracy of it by including for example: technology firms' specific ratios, this will allow the model to study technology firms not just from a financial but from an operational perspective too; social profiling, social media and softer social variables not captured or modeled by standardized 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.

Notes

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.

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).

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

5. Similar research has been done in this field by Xiang *et al.* (2012).

References

Acker, O., Hagen, H. and Hajj, J. (2017), "2017 technology trends", available at: www.strategyand.pwc. com/trend/2017-technology-trends (accessed May 16, 2017).

Acquisitions and Alliances (2016), "M&A statistics", available at: www.imaa-institute.org/recources/ statistics-mergers-acquisitions/ (accessed October 8, 2017).

Agrawal, A. and Walkling, R.A. (1994), "Executive careers and compensation surrounding takeover bids", *The Journal of Finance*, Vol. 49 No. 3, pp. 985-1014.

Alexandridis, G., Mavrovitis, C.F. and Travlos, N.G. (2012), "How have M&As changed? Evidence from the sixth merger wave", *The European Journal of Finance*, Vol. 18 No. 8, pp. 663-688.

Altman, E.L. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol. 23 No. 4, pp. 589-609.

Altman, E.L., Marco, G. and Varetto, F. (1994), "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience)", *Journal of Banking and Finance*, Vol. 18 No. 3, pp. 505-529.

Anadol, B., Paradi, J.C., Simak, P. and Yang, X. (2014), "Valuing private companies: a DEA approach", *International Journal of Business and Management*, Vol. 9 No. 2, pp. 16-28.

Andrade, G., Mitchell, M. and Stafford, E. (2001), "New evidence and perspectives on mergers", *Journal of Economic Perspectives*, Vol. 15 No. 2, pp. 103-120.

Aydin, N. (2015), "A review of models for valuing young and innovative firms", *International Journal of Liberal Arts and Social Science*, Vol. 3 No. 9, pp. 1-8.

Bakshi, G. and Chenb, Z. (2005), "Stock valuation in dynamic economies", *Journal of Financial Markets, May*, Vol. 8 No. 2, pp. 111-151.

Barnes, P. (1990), "The prediction of takeover targets in the UK by means of multiple discriminant analysis", *Journal of Business Finance & Accounting*, Vol. 17 No. 1, pp. 73-84.

Barnes, P. (1998), "Can takeover targets be identified by statistical techniques? Some UK evidence", *The Statistician*, Vol. 47 No. 4, pp. 573-591.

Beck, T., Demirgüç-Kunt, A. and Levine, R. (2003), "Law and finance: why does legal origin matter?", *Journal of Comparative Economics*, Vol. 41 No. 4, pp. 653-675.

Beese, J. (2015), "What is social listening & why is it important?", available at: https://sproutsocial.com/insights/social-listening/ (accessed July 21, 2017).

Benou, G. and Madura, J. (2005), "High-tech acquisitions, firm specific characteristics and the role of investment bank advisors", *The Journal of High Technology Management Research*, Vol. 16 No. 1, pp. 101-120.

Betton, S., Eckbo, B.E. and Thorburn, K.S. (2008), "Corporate takeovers", in Eckbo, B.E. (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 2, Elsevier, Oxford, pp. 291-416.

Bradley, M. (1980), "Interfirm tender offers and the market for corporate control", *The Journal of Business*, Vol. 53 No. 4, pp. 345-376.

Branch, B., Wang, J. and Yangc, T. (2008), "A note on takeover success prediction", *International Review of Financial Analysis*, Vol. 17 No. 5, pp. 186-1193.

Bureau of Economic Analysis (2017), *GDP by Industry*, available at: www.bea.gov (accessed September 3, 2017).

Chan, L.K.C., Lakonishok, J. and Sougiannis, T. (2001), "The stock market valuation of research and development expenditures", *The Journal of Finance*, Vol. 56 No. 6, pp. 2431-2456.

Chandra, U., Wasley, C.E. and Waymire, G.B. (2011), "Income conservatism in the US technology sector", *Accounting Horizons*, Vol. 25 No. 2, pp. 285-314.

Coats, P.K. and Fant, L.F. (1993), "Recognizing financial distress patterns using a neural network tool", *Financial Management*, Vol. 22 No. 3, pp. 142-155.

CompTIA (2017), "Cyberstates 2017, research report", available at: www.cyberstates.org/pdf/CompTIA%20Cyberstates%202017.pdf (accessed September 25, 2017).

Cremers, M., Driessen, J., Maenhout, P. and Weinbaum, D. (2008), "Individual stock-option prices and credit spreads", *Journal of Banking and Finance*, Vol. 32 No. 12, pp. 2706-2715.

Cudd, M. and Duggal, R. (2000), "Industry distributional characteristics of financial ratios: an acquisition theory application", *The Financial Review*, Vol. 35 No. 1, pp. 105-120.

Damodaran, A. (2001), "The dark side of valuation: valuing old tech", *New Tech and New Economy Companies*, John Wiley & Sons, New York, NY.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998), "Investor psychology and security market under- and overreactions", *The Journal of Finance*, Vol. 53 No. 6, pp. 1839-1885.

DeAngelo, L.E. (1990), "Equity valuation and corporate control", *The Accounting Review*, Vol. 65 No. 1, pp. 93-112.

Dekel, O., Shamir, O. and Xiao, L. (2010), "Learning to classify with missing and corrupted features", *Machine Learning*, Vol. 81 No. 2, pp. 149-178.

Dietrich, K.J. and Sorensen, E. (1984), "An application of logit analysis to prediction of merger targets", *Journal of Business Research*, Vol. 12 No. 3, pp. 393-402.

Dieudonne, S., Fabienne, C. and Slimane, B. (2014), "M&A activity: where are we in the Cycle?", available at: https://ssrn.com/Abstract=2504412 http://dx.doi.org/10.2139/ssrn.2504412 (accessed September 17, 2017).

Dong, M., Hirshleifer, D., Richardson, S. and Hong Teoh, S. (2006), "Does investor misvaluation drive the takeover market?", *The Journal of Finance*, Vol. 61 No. 2, pp. 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*, Vol. 12 No. 1, pp. 83-97.

Eckbo, B.E. (2014), "Corporate takeovers and economic efficiency", *Annual Review of Financial Economics*, Vol. 6 No. 1, pp. 51-74.

Eden, L. and Dobson, W. (2005), *Governance, Multinationals, and Growth*, Katholieke Universiteit Leuven.

Ernst & Young (2015), *Buying and Selling: Cross-Border Mergers and Acquisitions and the US Corporate Income Tax*, Business Roundtable, New York, NY.

Ernst & Young (2016), "Technology: capital confidence barometer", available at: www.ey.com/Publication/vwLUAssets/ey-ccb-15-technology/$FILE/ey-ccb-15-technology.pdf (accessed May 16, 2017).

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

Fescioglu-Unver, N. and Tanyeri, B. (2013), "A comparison of artificial neural network and multinomial logit models in predicting mergers", *Journal of Applied Statistics*, Vol. 40 No. 4, pp. 712-720.

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

Griffin, J.M., Harris, J.H., Shu, T. and Topaloglu, S. (2011), "Who drove and burst the tech bubble?", *The Journal of Finance*, Vol. 66 No. 4, pp. 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*, Vol. 39 No. 2, pp. 167-188.

Hammerstrom, D. (1993), "Working with neural networks", *IEEE Spectrum*, Vol. 30 No. 7, pp. 46-53.

Harris, R.S. and Ravenscraft, D. (1991), "The role of acquisitions in foreign direct investment: evidence from the US stock market", *The Journal of Finance*, Vol. 46 No. 3, pp. 825-844.

Harris, R.S., Stewart, J.F. and Guilke, D.K. (1982), "Characteristics of acquired firms: fixed and random coefficients probit analyses", *Southern Economic Journal*, Vol. 49 No. 1, pp. 164-184.

Hinton, G. and Salakhutdinov, R. (2006), "Reducing the dimensionality of data with neural networks", *Science*, Vol. 313 No. 5786, pp. 504-507.

Hodges, M.B. (2007), "Three approaches?", *The Appraisal Journal*, Vol. 75 No. 1, pp. 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*, pp. 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*, Vol. 35 No. 3, pp. 1186-1199.

Hyndman, R.J. and Athanasopoulos, G. (2013), *Forecasting: Principles and Practice*, OTexts. Institute of Mergers, London.

Iturriaga, F.J.L. and Sanz, I.P. (2015), "Bankruptcy visualization and prediction using neural networks: a study of US commercial banks", *Expert Systems with Applications*, Vol. 42 No. 6, pp. 2857-2869.

Jensen, M.C. (1986), "Agency costs of free cash flow, corporate finance, and takeovers", *The American Economic Review*, Vol. 76 No. 2, pp. 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*, Vol. 11 No. 4, pp. 415-422.

Jovanovic, B. and Rousseau, P.L. (2001), "Mergers and technological change: 1885-1998", Working Paper No. 01-W16, Vanderbilt University, Chicago, August.

Kaastra, I. and Boyd, M. (1996), "Designing a neural network for forecasting financial and economic time series", *Neurocomputing*, Vol. 10 No. 3, pp. 215-236.

Kuo, C. and Reitsch, A. (1995), "Neural networks vs conventional methods of forecasting", *Journal of Business Forecasting Methods and Systems*, Vol. 14 No. 4, pp. 17-22.

Lam, M. (2004), "Neural network techniques for financial performance prediction: integrating fundamental and technical analysis", *Decision Support Systems*, Vol. 37 No. 4, pp. 567-581.

Laudicina, P.A. and Peterson, E.R. (2017), "The 2017 A.T. Kearney foreign direct investment confidence index", available at: www.atkearney.com/foreign-direct-investment-confidence-index (accessed September 3, 2017).

Lev, B. and Zarowin, P. (1999), "The boundaries of financial reporting and how to extend them", *Journal of Accounting Research*, Vol. 37 No. 2, pp. 353-385.

Levine, P. and Aaronovitch, S. (1981), "The financial characteristics of firms and theories of merger activity", *The Journal of Industrial Economics*, Vol. 30 No. 2, pp. 149-172.

Lin, C., Huang, J., Gen, M. and Tzeng, G. (2006), "Recurrent neural network for dynamic portfolio selection", *Applied Mathematics and Computation*, Vol. 175 No. 2, pp. 1139-1146.

Lowenstein, R. (2004), *Origins of the Crash: The Great Bubble and Its Undoing*, 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*, Vol. 6 No. 1, pp. 59-72.

Mauboussin, M.J., Callahan, D. and Majd, D. (2017), "The incredible shrinking universe of stocks: the causes and consequences of fewer US equities", *Credit Suisse Report, Global Financial Strategies*, available at: https://research-doc.credit-suisse.com/docView?language=ENG&format=PDF&sourceid=em&document_id=1072753661&serialid=h%2B%2FwLdU%2FTIaitAx1rnamfYsPRAuTFRGdTSF4HZIvTkA%3D (accessed October 22, 2017).

Moody's Corporation (2016), "US non-financial corporates' cash pile increases to $1.68 trillion, tech holding the lead", available at: www.moodys.com/research/Moodys-US-non-financial-corporates-cash-pile-increases-to-168--PR_349330 (accessed May 16, 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*, Vol. 27 Nos 11/12, pp. 1111-1129.

NASDAQ (2017), "NASDAQ composite index (COMP)", available at: www.nasdaq.com/markets/nasdaq-composite (accessed September 6, 2017).

Ohlson, J.A. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, Vol. 18 No. 1, pp. 109-131.

Olson, D.L., Delen, D. and Meng, Y. (2012), "Comparative analysis of data mining methods for bankruptcy prediction", *Decision Support Systems*, Vol. 52 No. 2, pp. 464-473.

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

Palepu, K.G. (1986), "Predicting takeover targets: a methodological and empirical analysis", *Journal of Accounting and Economics*, Vol. 8 No. 1, pp. 3-35.

Palepu, K.G. (2003), *Business Analysis and Valuation Using Financial Statements*, 3rd ed., South-Western.

Park, Y. and Park, G. (2004), "A new method for technology valuation in monetary value: procedure and application", *Technovation*, Vol. 24 No. 5, pp. 387-394.

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

Powell, R. and Yawson, A. (2007), "Are corporate restructuring events driven by common factors? Implications for takeover prediction", *Journal of Business Finance and Accounting*, Vol. 34 Nos 7-8, pp. 1169-1192.

Powell, R.G. (1997), "Modelling takeover likelihood", *Journal of Business Finance and Accounting*, Vol. 24 Nos 7-8, pp. 1009-1030.

Rhodes-Kropf, M. and Viswanathan, S. (2004), "Market valuation and merger waves", *The Journal of Finance*, Vol. 59 No. 6, pp. 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*, 2nd ed., Chapman and Hall/CRC, New York, NY.

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*, Vol. 74 No. 1, pp. 277-304.

Routledge, B.R., Sacchetto, S. and Smith, N.A. (2013), "Predicting merger targets and acquirers from text", working paper, Carnegie Mellon University, Pittsburgh.

Salchenberger, L.M., Cinar, E.M. and Lash, N.A. (1992), "Neural networks: a new tool for predicting thrift failures", *Decision Sciences*, Vol. 23 No. 4, pp. 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*, Vol. 3 No. 4, pp. 279-292.

Servaes, H. and Zenner, M. (1994), "Taxes and the returns to foreign acquisitions in the United States", *Financial Management*, Vol. 23 No. 4, pp. 42-56.

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

Simkowitz, M. and Monroe, R.J. (1971), "A discriminant analysis function for conglomerate targets", *The Southern Journal of Business*, Vol. 38 No. 2, pp. 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*, Vol. 5 No. 4, pp. 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*, Vol. 16 No. 1, pp. 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*, Vol. 15 No. 1, pp. 1929-1958.

Swaminathan, H. and Rogers, H.J. (1990), "Detecting differential item functioning using logistic regression procedures", *Journal of Educational Measurement*, Vol. 27 No. 4, pp. 361-370.

Tam, K.Y. and Kiang, M.Y. (1992), "Managerial applications of neural networks: the case of bank failure predictions", *Management Science*, Vol. 38 No. 7, pp. 926-947.

Teeter, P. and Sandberg, J. (2016), "Cracking the enigma of asset bubbles with narratives", *Strategic Organization*, Vol. 15 No. 1, pp. 91-99.

Thawornwong, S. and Enke, D. (2004), "The adaptive selection of financial and economic variables for use with artificial neural networks", *Neurocomputing*, Vol. 56 No. 1, pp. 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 February 21, 2017).

Tkáč, M. and Verner, R. (2016), "Artificial neural networks in business: two decades of research", *Applied Soft Computing*, Vol. 38 No. 3, pp. 788-804.

Tsai, C.F. and Wu, J.W. (2008), "Using neural network ensembles for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, Vol. 34 No. 4, pp. 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*, Vol. 37 No. 3, pp. 1846-1853.

US Bureau of Labor Statistics (2016), "US bureau of labor statistics", available at: www.bls.gov/home.htm (accessed September 7, 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*, Vol. 27 No. 3, pp. 417-425.

Westenberg, D.A. (2000), "The IPO report", Hale and Dorr LLP, New York, NY.

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

Widrow, B., Rumelhart, D.E. and Lehr, M.A. (1994), "Neural networks: applications in industry, business and science", *Communications of the ACM*, Vol. 37 No. 3, pp. 93-105.

WilmerHale (2017), "2017 M&A report", April 28, p. 36, available at: www.wilmerhale.com/uploadedFiles/Shared_Content/Editorial/Publications/Documents/2017-WilmerHale-MA-Report.pdf (accessed November 11, 2017).

Xiang, G., Zheng, Z., Wen, M., Hong, J. and Rose, C. (2012), "A supervised approach to predict company acquisition with factual and topic features using profiles and news articles on TechCrunch", *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media, Dublin, Ireland, June 4-7*, pp. 607-610.

Zhang, G., Patuwo, E.B. and Hu, M.Y. (1998), "Forecasting with artificial neural networks: the state of the art", *International Journal of Forecasting*, Vol. 14 No. 1, pp. 35-62.

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*, Vol. 116 No. 1, pp. 16-32.

Zhu, D., Premkumar, G., Zhang, X. and Chu, C.H. (2001), "Data mining for network intrusion detection: a comparison of alternative methods", *Decision Sciences*, Vol. 32 No. 4, pp. 635-660.

Corresponding author
Ioannis Anagnostopoulos can be contacted at: Y.ANAGNOSTOPOULOS@KINGSTON.AC.UK