International Journal of Computer and Information Technology (ISSN: 2279 – 0764)
Volume 05 – Issue 06, November 2016

Using Artificial Neural Networks to Predict the E-Business Startup Success

Ayebazibwe Abel
Department of Computing and Information Technology
Jomo Kenyatta University of Agriculture and Technology
Email: Abel82012 [AT] gmail.com

Abstract—**The purpose of this research is to develop a model to predict e-business startup success using artificial neural networks. This model uses both quantitative and qualitative variables. The neural networks used are feed forward networks with one hidden layer and the resilient back propagation training method is used to train the neural networks. A sample set of 110 startups was got from startup's crunchbase profile website. The data records were divided in to 3 groups: learning, validating and testing groups. The network was created, trained and tested using neural network software, easyNN. After training the network, the training average error was 0.009073 which was below the target 0.01000000 and the validating error was 0.16974157 which was also below the target. ANN was able to perfectly predict the score of 4 (25%) of its test cases, its scores were within +/- 1 of the real score. Variable which mostly influences the success of an e-business startup are Marmer stage, Number of founders, keywords and driven by users. This research successfully developed a model that predicts an e-business startup's success rate, and learning from these findings can guide entrepreneurs and investors in making decisions about future e-business startups.**

***Keywords*--- artificial neural network; Prediction; e-business; startup; success**

I. INTRODUCTION

Artificial Neural Networks (ANNs) are artificial intelligence-based computational procedures for mapping input patterns to outputs consisting of real-valued or discrete valued functions. Traditional statistical learning techniques can, in general, only learn combinations of linear functions, whereas neural networks can learn non-linear functions of arbitrary complexity. For problems where the mapping of inputs into outputs is complex or obscure, ANNs are among the most efficient of learning techniques known [1].

Prediction is a forecast, but not only about the weather. Pre means "before" and "diction" has to do with talking. So a prediction is a statement about the future. It's a guess, sometimes based on facts or evidence, but not always. There are different modeling techniques to predict an outcome through using historical data to estimate the parameters of a model; it then delivers the resulting exactly specified model to the new case for use in prediction and classification [2]. Having considered the strengths and limitations of each method, the one which outperforms other under the given conditions is chosen leading to a significant solution for a given application problem [3]. There are multiple ways to define *success* in the startup world. Success can be a reflection of profit, the number of users or customers, an acquisition, etc. With the rising importance of the social aspects of business, this project will define startup success by its *social media capital*. Social media capital is defined as the total number of social media connections that a company has among its social media accounts. Although social media followers and "likes" do not directly correlate with customer or user numbers, it instead represents public endorsement. E-business is the application of information and communication technologies (ICT) in support of all the activities of business. E-business stands for Electronic Business, which also means the administration of conducting business via the Internet.

This would include the buying and selling of goods and services, along with providing technical or customer support through the Internet. The increasing use of the Internet and web technology has introduced various web-based applications and strategies for businesses such as the e-Business development. Today, the e-Business concept is widely adopted by large organizations and gradually followed by small-to medium sized enterprises (SMEs) [4]. E-Business can be seen as a new way of conducting a business that has changed the traditional buying and selling process into an online-based process. It has also changed the way people perceive their internet technology investment by focusing on new business models and concepts [5].

Digital entrepreneurship has become a defining factor of the world economy, and startups are the gateway into the business world for many entrepreneurs. In this research, a startup is defined as a new company that has plans for significant growth. Despite the fact that startups have nearly insurmountable odds for success, thousands of new Internet entrepreneurs dedicate their time and money in hopes of becoming the next *Facebook* or or *Angry Birds*. With an average of 15,000 applications added to the mobile and web application market every week, the one-millionth mobile application went live in early December of 2011 [6]. This influx of new applications and website startups on the market has caused investors and new entrepreneurs to ask what makes a startup fail, and more importantly, what makes a startup succeed? Little research has been conducted regarding these questions, and even less research has been published by firms who have the resources to compile such data.

Over the last decade, we have seen a rapid acceptance of new technologies like neural networks and data mining methodologies for solving a wide range of business problems such as forecasting, modeling, clustering, and classification. As world becomes more excited about e-commerce, however, it is important for the investor to predict the success of the e-business startup.

Investors use different tools to evaluate the survival capabilities of e-businesses startup, most of the tools are based on traditional statistic techniques with low performance. The inaccuracy of these tools is caused by the nonlinearities and non stationary of economic and financial data [7]. The primary goal of this research project was to develop a success predicting model of an e-business startup using artificial neural network (ANN), both qualitative and quantitative variables were identified; startup success was defined using by its *social media capital* which is the total number of social media connections that a company has among its social media accounts. Although social media followers and "likes" do not directly correlate with customer or user numbers, it instead represents public endorsement. This research focused on the backpropagation algorithm learning method [8].

II. Related work

Many different techniques have been applied to business success prediction since its beginnings in the 1960's. The field arguably started earlier, but the first statistical and mathematical models were published. The various techniques used since then are reviewed and referenced in this research. Where the terms Type I and Type II Error are used, Type I Error refers to misclassifying a failing business as successful and conversely Type II Error refers to misclassifying a successful business as a failure [9].

A. *Univariate Analysis*

Early attempts to use financial ratios to predict business failure stem from the work of Patrick (1932). This work was later extended by Beaver (1966), who presented the first modern statistical model for business success prediction. Beaver's univariate approach did not contain an overall measure of financial distress, which led to the problem that different ratios made conflicting predictions about a given business. In addition, it was noted that one ratio alone could not encompass the complexity of business failure. This model's error was estimated at 22% Type I Error and 5% Type II Error for one year prediction intervals.

B. *Discriminant Analysis (DA)*

This was the first multivariate approach applied to business success prediction. This extended the work of Beaver (1966) by addressing the problem that various ratios made conflicting predictions, and incorporates the concept of a composite measure of business distress. Altman (1968) presented a DA model to predict business failure, in which the information from several variables (ratios) was combined into a single weighted score for each business. However, the practical significance of the difference may be minor. Moreover, to overcome the second requirement quadratic discriminant analysis can be used, which involves squared independent variable equal misclassification costs of Type I and II Error, which is usually violated.

C. *Logit and Probit Analysis*

Logit Analysis (LA) generates a score for each business similar to DA. However, it is free from the normality and equal covariance assumptions of DA. LA is based upon the cumulative logistic function (CLF), and due to its non-linear nature the coefficients are usually estimated using the maximum likelihood method. Furthermore, unlike the difficult interpretation of the Z-score in the DA model, the score in LA can be directly interpreted as the probability of failure (where the CLF ranges from 0 to 1). The empirical results for his model were disappointing, but he showed that LA is more statistically valid and easier to interpret than DA. In addition, subsequent studies on LA have shown that it is usually slightly empirically superior to DA in both classification and prediction accuracy.

D. *Human Information Processing*

Human Information Processing (HIP) involves utilizing the existing ability of human decision makers to use information (usually accounting information) to group companies according to their probability of future bankruptcy. Hence, HIP attempts to model the relationship between cues and decisions, rather than the information processing used to form the decisions. Consequently, interviewing and questionnaires are frequently used research techniques. The pioneering studies in this area are by Libby (1975) and Casey (1980). Abdel-Khalik and ElSheshai (1980) presented results of a quantitative comparison that found mathematical and statistical (quantitative) models to be superior decision makers compared with experienced professionals. Moreover, the HIP model produces decisions without a level of certainty

E. *Sequential Procedures*

Healy (1987) presented cumulative sum (CUSUM) procedures that detect a shift in a series of variables' values from a 'good' distribution to a 'bad' distribution. CUSUM procedures, which date back to 1954, are a set of sequential procedures based on likelihood ratios. These procedures attracted the CUSUM name as they reduce to calculating cumulative sums for many common distributions. CUSUM procedures detect the optimal starting point of the shift and then provide a signal of the shift as soon as possible after the shift occurs. Healy demonstrated the application of CUSUM procedures for detecting shifts in the mean and covariance matrix of a multivariate normal distribution. He noted that this method is very efficient at detecting a shift in the mean, when the mean of the 'good' and 'bad' distributions are known. However, it was also noted that the CUSUM procedures were still applicable in situations when at least the direction of slide towards the 'bad' distribution was known, as is the case in business success prediction. Kahya and Theodossiou (1999) further extended the work of Theodossiou (1993) by overcoming the problems of non-stationary variables and improving their definition of financial success. In addition, several minor statistical refinements were made to their CUSUM procedures. The best CUSUM model was then chosen by a neural network search procedure from an initial set of 54 explanatory variables (comprising 27 popular variables and their first differences). Interestingly, models using the most popular explanatory variables were found to be non-stationary with deteriorating forecasting performance over time. The final CUSUM model chosen had superior predictive power compared to both DA and LA. Despite these positive empirical results, CUSUM procedures have not been widely used in business success prediction. The exact reason for this is unknown, but the greater complexity of the CUSUM procedures appears to be the most likely reason for its low popularity.

III. RESEARCH METHODOLOGY

This chapter elaborately discusses the methodology used in this study. The research question of what kind of model to develop proposed in chapter one is presented here. All phases of the research design, the identification of indicators, the identification of the appropriate neural network model to use, the analysis of a training algorithm and examination neural network whether accurately forecasts the success of an e-business. The research can be categorized as a combination of exploratory and descriptive study seeking insights into using Artificial Neural Networks to predict the Startup e-Business Success.

A. *Choosing variables*

Due to the lack of access to data regarding the entrepreneurs and their funds, the variables were ones that could be gathered by simply looking at a mobile applications, website, or social media page. These variables included factors such as type of startup (or its "category", such as social networking, games, finance, etc.), whether the startup had keywords in its name, target gender, target age group, and whether the startup was social media relevant or not. A list of all variables collected in a dataset is:

- **Marmer Stage** (Stage 3 or Stage 4)
- **Location** (country)
- **Number of Founders** (Numeric, 0 if unknown)
- **Type of e-business** (Online Shopping, Marketing, Information Service Providers, Consultant, Subscription, Manufacturer, utility, Broker, Community)
  Subtype of Startup (Software, Tangible products, Education, Game, Travel, Social media network, Webhosting, Fashion, Aggregator, food, Messaging, Dating, Music)
- **Social Media Relevancy** (Yes, No)
- **Driven by** (Users, Administrative)
- **Keywords (**Yes, No)
- **Target Age (**Kids and Teens, Teens and Adults, Adults, All)
- **Target Gender** (Men, Women, Unisex, Both)

B. *Gathering Data*

Data collection involved researching a startup's Crunchbase profile, and if Crunchbase did not report needed information, it was pulled directly from the e-business's website, or from articles written about the e-business or its founders. Crunchbase is a database website that specifically lists factual variables, such as the number of founders and year of founding of active and inactive e-business and companies. All other conceptual data was pulled directly from the website itself. If enough information about an e-business was not found, it was removed from the dataset. Less than ten startups were removed the dataset due to issues with missing data, yielding 110 data records. Neural networks cannot function if some data is missing from training examples, and so training examples were removed that had two or more missing data pieces.

C. *Scoring Social Media Capital*

After identifying the training examples for the dataset, a social media capital score, or success score, was assigned to each e-business startup. This was calculated using Facebook likes and Twitter numbers for all training examples. Social media numbers was collected between September 2015 and November 2015 and e-business social media scores are representative of this specific time period. Social media numbers was rounded to the nearest thousand if over 10,000, rounded to the nearest hundred if between 1000 and 10,000, rounded to the nearest multiple of 50 if between 50 and 50 and 1000, and rounded to the nearest ten if below 50.

The equations that was used to calculate social media capital *totals* assert that one Facebook like is worth two raw Twitter followers. A *raw* Twitter follower is the number of Twitter accounts the e-business *is* following subtracted from the number of accounts *following* the e-business. Calculating raw Twitter followers in this manner versus recording the number of accounts following an e-business on Twitter was necessary to control for e-business that followed hundreds or thousands of Twitter users in hopes of getting blind "follow-backs".

In the Twitter-sphere, an account with a high ratio of follow*ers* to follow*ing* is indicative of Twitter popularity.

To calculate a rating for each e-business between 1 and 10, it was necessary to find the log10 value of the total social media score in order to scale scores down to reasonable numbers. After this was done, the largest value e-business in each was used to factor all other e-business numbers into this scale. Once a social media score was calculated for each e-business, it was rounded to the nearest whole number. Startups that scored below the integer value of 1 was rounded to 1. The base formula for finding the social media capital score for e-business startups is defined as follows:

Social Media Capital Total=
Log [(Facebook likes) x (2 (Twitter Followers - Twitter Following))]

Fab had the highest social media capital total with 1,604,426 Facebook likes, 73, 224 Twitter followers, and the Fab Twitter account was following 41, 738 users on Twitter at the time of data collection. This startup was manually set to the rating of 10.

Log [(1,604,426 Facebook likes) x (2(73, 224 Twitter followers - 41, 738 Twitter accounts following))] = 11.004467
X (11.004467) - 1 = 10
X = 0.999594= Constant
0.999594 (Social Media Capital Total) - 1 = Social Media Capital Rating.

The remaining startup social media scores were calculated with this constant. For example, the web-based startup Warby Parker had 449, 632 Facebook likes, 75, 627 Twitter followers, and was following 254 Twitter accounts at the time of data collection.

Log [(449, 632 Facebook likes) x (2 (75, 627 Twitter followers – 254 Twitter accounts following)] = 10.831103
0.999594 (10.831103) - 1 = 9.82671
(This score was then rounded up to the integer value of 10 in the neural network).

D. *Using Neural Networks*

After a dataset was coded, it was loaded into an ANN program, *EasyNN*. Neural networks create more accurate models with the least possible amount of independent variables, and so a handful of variables were removed that were deemed redundant or inconclusive for neural network training purposes. For example, “Subtype of an e-business startup” was removed from the neural networks, as it was determined that the data collection on this variable had too much potential for sampling error in determining the subtype of a startup. (Startups can also have multiple subtypes such as it provides information, allows adverts and does online shopping as well which added to the complexity of this variable.) Also removed were “acquisitions”, the number of times a startup had been purchased, and “investment amount”, the total amount of money given to a startup by investors, as these numbers could be found for only for some of the startups in the datasets.

The variables “social media relevancy”, (if a startup was relevant to social networking) and whether a startup was “user driven or administrative driven” (if a startup relied on users for functionality, it is user driven, but if a startup’s functionality is unaffected by its user count and its experience is the same if there are few users or thousands of users, then it is administrative driven) were deemed to be too similar, and the variable “social media relevancy” was removed.

E. *Training Neural networks*

ANNs are fundamentally nonlinear models that distinguish patterns and classify variables. To do so, a researcher developed unsupervised training methods. Dasgupta et al. found that the Back Propagation (BP) model and Levenberg–Marquardt algorithm of ANNs training outperform other models in classification. Therefore, a feed-forward BP model and Levenberg–Marquardt (trainlm) was employed in the current research [9]. The whole dataset was randomly divided into three groups with no repetition. The training set contained 60 data records, validation set contained 40 data records and testing set contained 10 data records.

The resilient back propagation training method was used for training.

This method is very adequate when sigmoid transfer functions are used [10]. To prevent over fitting, a very common problem during training, the training set error and the validation set error was compared every 100 epochs. Training was set to be finished when training set error began to decrease while validation set error increase. Considering the research variables, the architecture of the ANN was made up of two main layers, an input layer for e-business success indicators and an output layer for e-business post-implementation success as shown Figure 1.

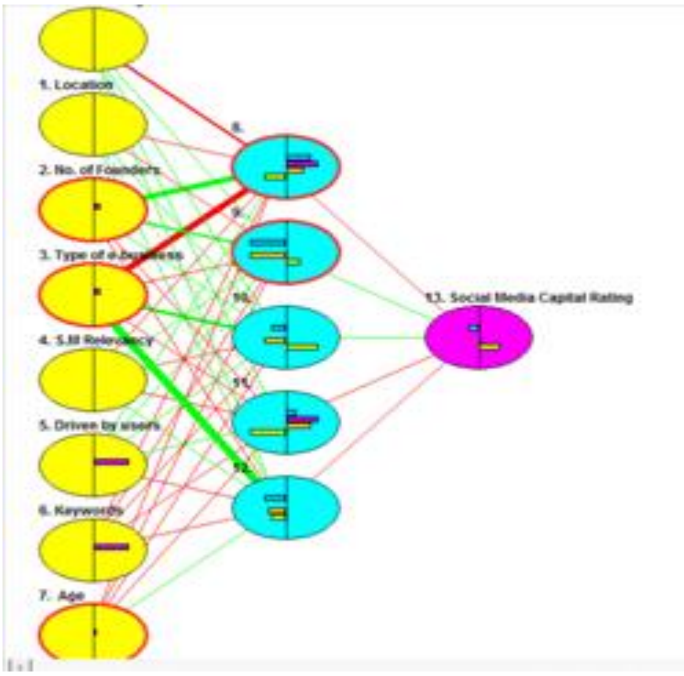

Figure 1. Artificial neuron network architecture.

The model was tested as shown in Figure 3.

IV. RESULTS

After training the network, the training average error was 0.009073 which was below the target (0.01000000) and the validating error was

0.16974157 which was also below the target as shown in Figure 2.

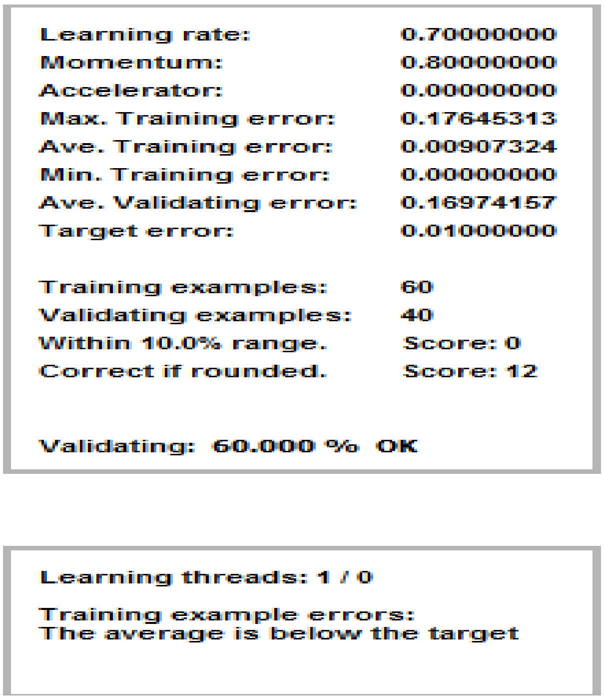

Figure 2. Training and validating error Summary

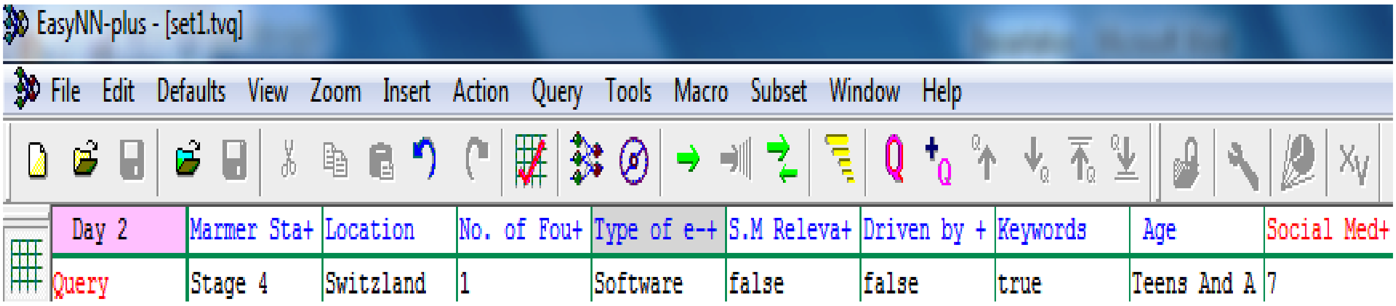

Figure 3. Testing the Model

It is observed that the variables, Number of founders and type of business are of big importance in the success of the e-business startup as shown in Figure 4.

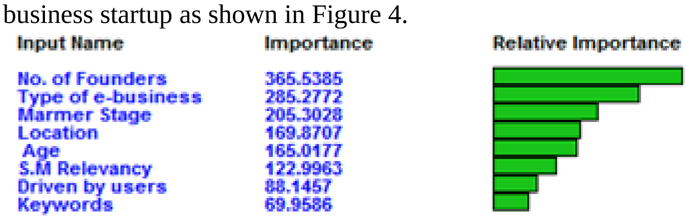

Figure 4. Variables importance

The novelty of this project was to create model that could approximate a startup's success score. To test the accuracy of the neural networks, training examples were removed representatively across ratings from the dataset to use as test cases for testing the accuracy of the neural network. The ANN was able to perfectly predict the score of 4 (25%) of its test cases, its scores were within +/- 1 of the real score as shown in table 1 and Figure 5.

Table 1: Predicted social media compared to actual Score

Startup Name	Real Score	ANN Score	Score Difference
Plowz and Mowz	6	5	1
PasswordBox	6	5	1
CreativeLive	9	9	0
Liki	5	5	0
Koru	5	5	0
SupplyHog	5	5	0
Codecademy	9	9	0
Tuskhub	5	5	0
Duolingo	9	10	-1
MediaCore	5	6	-1

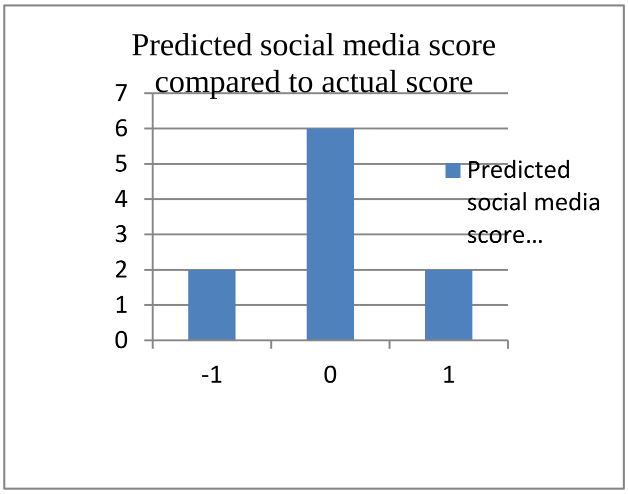

Figure 5. Predicted social media compared to actual score

Variable Influence

Not surprisingly, startups that had reached Stage 4 of the Startup Genome's Marmer Scale had a leg up on startups that were recorded as being in Stage 3 development.Training examples in the 7-10 range were all in Stage 4 of the Marmer scale, startups in the 5-7 were mixed in stage, and all but one startups rated a 4 or less were in Stage 3. Because the Marmer Stages have very specific definitions, including them in the final neural network analysis was a simple way of ensuring that training example startups fell into an accurate score range. E-business Startups having multiple founders had a slight positive impact on a startup's score. This is consistent with Startup Genome's research that startups with solo founders take longer to reach Stage 4 of the Marmer Scale than startups with multiple founders [11]. Paul Graham, founder of startup incubator Y Combinator, states that most successful startups have more than one founder, and this is because the hard work that is involved in launching a startup requires multiple people to implement [12]. This study supports these prior testimonies that multiple founders have a positive effect on a startup's overall outcome as shown in Table 2.

Table 2: Tabulation of the number of founders

Rating	No of Founders		Total
	1	2+	
5	1	10	10
6	2	14	16
7	5	22	27
8	5	17	22
9	2	9	11
10	2	8	10
Total		96	

Keywords existing in a startup's name had a slightly positive effect on rating in the neural network. Search engine optimization is one way to optimize startup discovery, as including keywords in a startup's name is a simple and effective way of increasing the chances that a potential user will be led to a website. 66.6 % of E-business startups rated between 5 and 6 did not have keywords in their names, 67.4% 0f startups rated between a 8 and 10 had keywords in their names, which suggests that keywords are more crucial for e-business startups as shown in Table 3.

Table 3. Tabulation of keywords

Rating	Keyword?		Total
	Yes	No	
5	4	8	12
6	2	12	18
7	20	7	27
8	13	6	22
9	7	4	11
10	9	1	10
Total		100	

E-business startups that were user-driven had a slight negative effect on score, while startups that were administrative driven had a slight positive effect on score. More startups were recorded as user driven than administrative driven, which could account for this difference in datasets: 71% of the startups were listed as user driven, while 29% were listed as administrative driven. A cross tabulation showed that of the 43 startups rated an 8 or more, 65% were user driven and 35% were administrative driven.

Social media relevancy was also examined in comparison to rating. In a new neural network, "social media relevancy" was analyzed to see how this variable affected a startup's rating, and it was found that social media relevancy had a positive effect on web-based startups' rating. Despite that the "user or administrative" variable and social media relevancy variables are very similar yet have opposite affects on score, these relationships can be explained from the high number of social media *marketing* startups included in the dataset. These marketing startups were social media relevant yet were also technically administrative driven, and scored in the mid to upper range of the dataset.

V. DISSCUSSION

Predicting an e-business startup success is one of the most important problems faced by e-commerce investors, many different tools have been used to evaluate the survival capabilities of middle-aged e-businesses but there is no known tool for startup ones. Most of the tools are based on regression models and in quantitative variables. Nevertheless, qualitative variables which measure the company way of work and the manager skills can be considered as important as quantitative ones. Thus, we used ANNs to efficiently discriminate high influential variables and low influential variables. Neural network's ability of detecting underlying interactions between independent and dependent variables, as well as the nonlinear relationship among applied attributions has made an opportunity to develop a predictive model which predicts the success of e-business startup.

The valid ANN predictive model may play a crucial role in e-business as it can be embedded in computerized e-business decision support systems in which, through different social media likes as system inputs can provide a feasible approach to predict the success of an e-business startup.

By defining startup success by social media capital, this project ultimately identified variables that increased or decreased a startup's social media capital. Therefore, it is consistent that variables that added social components to startups, such as a startup having a complimentary website, had positive relationships with rating. These findings suggest that to increase a startup's social media capital, it is necessary to make the startup as user friendly as possible in order to attract customers (and potential customers). "User friendly" variables include making a startup easy to find by including keywords in its name, and by not having advertisements on a startup's interface.

There was a concern that social media relevant startups would score better than non-social media relevant startups as startup success was measured by social media capital, but in fact this was not an overwhelming issue in the dataset. This suggests that accumulating social media capital is not just important for social media relevant startups, but for startups in general. This also suggests that social media capital is a valid way to measure success across e-business startups.

In terms of founders, this project supported previous notions that multiple founders correlated with a startup's success. A startup having multiple founders in either dataset had a positive impact on score in the neural networks, and this was supported by the cross tabulation. This finding reiterates the idea that achieving startup success requires work and commitment from multiple individuals in order to succeed.

In the analysis stage of this project, it was difficult to compare startups that were completely different *types* of businesses, and so the next step to finessing the methods for this project would be to focus datasets on one specific category of startup (such as games, or productivity startups). Focusing on one or two startup categories would have presented an opportunity to have more specific variables relevant to a startup's category, which would offer more insight to success and failure patterns for those kinds of startups.

Selection bias in selecting the training examples was another issue that could have been resolved with a different selection method. Although training examples from "watch lists" that catered to a particular category of startups were deliberately not collected, it was still found that the startups within some "watch lists" were not representative of all startup types. It would be necessary to increase the same size for the dataset from 100 to a larger number to better represent all categories of startups in the datasets. The alternative, as just discussed, would be to simply focus on one or two categories of startups so that it would not be necessary to have a categorically representative startup sample. This work came across with a midway sensitivity, which might be because of limited data records about e-business startups and not all factors that influence the success of an e-business startup due to lack of access to data regarding the entrepreneurs and their funds.

Regardless, the project's categorical scope was general and was an adequate starting point to determine if there were general conceptual patterns to find within the datasets. While some of the project's findings may seem self-evident, this project is one of the few pieces of research available regarding data of this nature. There is much misguided advice within the e-business startup world due to entrepreneurs being told that there is no rhyme or reason to determining which startups make it and which do not, and advice backed from methodological research rather than from blind speculation can direct entrepreneurs in a more reliable direction. This project successfully develops a model to predict the e-business success and using from this model can guide e-business entrepreneurs and investors in making decisions about future e-business startups.

REFERENCES

[1] Smith, M. Neural networks for statistical modeling, Van Nostrand Reinhold. 1993; 115 Fift Avenue, New York, NY 10003.

[2] Han, J., &Kamber, M. Data Mining: Concepts and Techniques. 2006 (2nd Ed.). USA: Morgan Kaufmann Publishers.

[3] Kotsiantis, S. B. Supervised machine learning: A review of classification techniques. 2007; 249 (31) pp. 249-268.

[4] Al-Qirim, N. Global Electronic Business Research: Opportunities and Directions. IGP: London. 2006; 2(2) pp. 2289-2818.

[5] Damanpour, F. E-Business E-commerce Evolution: Perspectives and Strategy. Journal of Managerial Finance. 2001.

[6] Freierman, S. (2011, December 11). One Million Mobile Apps, and Counting at a Fast Pace. The New York Times. Retrieved March 13, 2014, from http://www.nytimes.com/2011/12/12/technology/one-million-apps-and-counting.html?_r=0

[7] Saiful Anwar, Yoshiki Mikami (2011). Comparing Accuracy Performance of ANN, MLR, and GARCH Model in Predicting Time Deposit Return of Islamic Bank.2011; 2(1) pp 44-51.

[8] Kartalopoulos, Stamatios V., Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications, IEEE Press, New York, New York, 2012.

[9] Mikolov, Tom. Statistical language models based on neural networks. PhD thesis, PhD. thesis, Brno University of Technology. 2012; retrieved from http://www.fit.vutbr.cz/~imikolov/rnnlm/thesis.pdf

[10] Bikrampal Kaur: Dr.Himanshu Aggrawal (2010). Artificial Neural Network based Diagnostic Model for Causes of Success and Failures. 2010; 8(1) pp. 1947-5500.

[11] Demuth, H., Beale, M., & Hagan, M. Neural Network Toolbox User's guide, version 4. Natick The Mathworks Inc. MA 01760, USA. 2002.

[12] Startup Genome Report. Compass. Retrieved March 12, 2014, from https://s3.amazonaws.com/startupcompass-public/StartupGenomeReport1_Why_Startups_Succeed_v2.pdf.

[13] Graham, P. (n.d.). The 18 Mistakes That Kill Startups. Paul Graham. Retrieved March 13, 2014, from http://www.paulgraham.com/startupmistakes.html

[14] Gershenson, C. (n.d.). Artificial Neural Networks for Beginners. arXiv. Retrieved March 14, 2014, from http://arxiv.org/pdf/cs/0308031.pdf

[15] Power, D. (2012, May 4). Social Media Capital: Earn It First, Then Spend It. Sprout Social. Retrieved March 15, 2014, from http://sproutsocial.com/insights/social-media-capital/

[16] Discover The Patterns Of Successful Internet Startups In The Startup Genome Report. (2010, May 28). Compass. Retrieved March 13, 2014, from http://blog.startupcompass.co/discover-the-patterns-of-successful-internet

[17] Siganos, D. & Stergiou, C. (n.d.). Introduction to neural networks. Neural Networks. Retrieved March 14, 2014, from http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Introduction to neural networks