Application of Artificial Intelligence Methods for the Semantic Web

Alete-Omoni Chimele O., & E. O. Bennett
Department of Computer Science,
Rivers State University,
Port Harcourt,
Nigeria
fabolous4real03@yahoo.com, bennett.okoni@ust.edu.ng

Abstract
A web based Artificial Intelligent (AI) system is proposed for building semantic representations in a Question and Answer (Q&A) framework. The system is built into a software tool for semantic Q&A using the PHP/MySQL language which is the proven tool for the web. The tool supports semantic Q&A data gathering from a front-end interface. The data is then used in an AI system trained by the Long Short-Term Memory (LSTM) algorithm. Simulation results show that the system can indeed perform as expected. It further shows that increasing the hidden layer size leads to improved perplexity or negative-log likelihood estimates with values as low as 1.1 and 3.5 respectively.

Keywords: *Artificial Intelligence, LSTM, Q&A, Semantic Web*

1. Introduction

With the flow of vast amount of information over the web/internet, and as the web becomes more readily accessible, the use of meaningful representations of data for timely and reliable extraction of knowledge is not necessarily apparent. Nevertheless, research studies have shown the feasibility of extracting meanings of data and about data leading to the evolution of the semantic web. However one major challenge is how to efficiently develop a reasoner for structured information derived by an OWL (Da Silva et al., 2002). In particular, the OWL tools have been shown in the research and industry field as a standard norm for understanding the knowledge generation and mining process in the semantic web. For instance the Resource Description Framework (RDF), a system for building abstract and hierarchical knowledge schemas for the semantic web has been widely applied since the late 1990s (Powers & Practical, 2003). AI is a rapidly evolving computational intelligence field that shows a lot of promises in solving a whole lot of real world and synthetic world problems. The field had initially had a slow start with non-beliefs from a subset of the computing field (Minsky & Papert, 1969) Paper on the limitations of the Perception; (Schmidhuber, 2015).

2. Related Literature

The semantic web is a major feature of modern day web architectures as developers and industry stakeholders strive to make the web more useful and efficient. In this field of study, key areas include semantic indexing and tagging for fast web document retrieval (Najafabadi et al., 2006). In (Bernstein et al., 2016) the value of a semantic web for community wide internet applications had been described with notable contributions coming from social web communities such as facebook.com and google.com. A cloud computing analytic study for a sporting event is currently been investigated in (Mahmood & Takahashi, 2015). With the help of semantic web tools and deep learning they were able to perform analytics for a football match. This analytics involves semantically identifying and discovering hidden causes of match performances from real time video and sensory data and using the information gained for predictions. (Tang et al., 2015) presented an overview of deep learning systems including those with semantic capabilities for sentiment analysis. Sentiment analysis involves analysing sequences of sentiments online. Semantic information embedded in text and the needs to use deep learning algorithms to discover knowledge in sentiments have been stressed. They further used DNNs – a cascade of Restriction Boltzmann Machines (RBMs) and a classification RBM for learning sentiments from a twitter dataset and the Sina microblog with better performances over the existing state-of-the art model such as Support Vector Machine (SVM) and Naïve Bayes (NB).Our primary purpose here is to investigate if a semantic compliant system trained with a temporal DNN can be useful in sentiment relevance modelling tasks i.e. in the ordering of sentiments based on the degree of importance. This is especially useful in product adverts and political campaigns. A novel system for learning semantic representations of community question answering has been attempted in (Zhou et al., 2016). The proposed approach was based on neural networks to learn the semantic representation of queries. The neural network uses a de-noising auto-encoder with continuous word embedding's based on skip-gram models (Mikolov et al., 2013). They used the Yahoo Question and Answer (Q&A) dataset for their experimental studies. From their analysis, they were able to obtain important improvements as compared to other approaches.

3. AI Methods for the Semantic Web

Artificial Intelligence (AI) is concepts that have evolved over time and cuts across many disciplines – neuroscience, psychology, computer science, engineering, sociology etc. As the name implies AI mimics intelligent behaviour by learning or borrowing from real world phenomena and actors – mammals, insects, environmental effects etc. AI has grown into a very large discipline studied worldwide by many universities, laboratories and even Government agencies.

LSTM network introduces a new structure called a memory cell. Each memory cell contains four main elements: the input gate, forget gate, output gate and a neuron with a self-recurrent. These gates allow the cells to keep and access information over long periods of time. LSTM calculates the hidden states by a set of equation as follows:

$$i = \sigma\left(x_t U^i + s_{t-1} W^i\right) \quad (1)$$
$$f = \sigma(x_t U^f + s_{t-1} W^f) \quad (2)$$
$$o = \sigma(x_t U^o + s_{t-1} W^o) \quad (3)$$
$$g = \tanh(x_t U^g + s_{t-1} W^g) \quad (4)$$
$$c_t = c_{t-1} \circ f + g \circ i \quad (5)$$
$$s_t = \tanh(c_t) \circ o \quad (6)$$
$$y = softmax(V s_t) \quad (7)$$

In these equations $i, f, oandg$are related to the input gate, forget gate, output gate and self-recurrent respectively. As they look complicated, the LSTM equations will be explained more in detail. i Indicates how much of the new information will be let through the memory cell. f is responsible for information should be thrown away from memory cell. Everyone in f means keeping information, while every zero means get rid of this information. o decides how much of the information will be passed to expose to the next time step and also to output. c_t can be called as the internal memory of the memory cell which is the sum of element wise multiplication of previous internal memory state by the forget gate, and element wise multiplication of self-recurrent state with input gate. Finally, $s_t$is related to the hidden state which can be calculated by element wise multiplication of the internal memory with the output gate. Additionally the final output can be calculated by Eq.7 which is equal with Eq.2.

4. Methodology

4.1 Constructive Research

In this research work, the researcher used the constructive research approach. Constructive research is one of the very common computer science research procedures. The constructive approach means problem solving through the construction or use of models, diagrams, plans, organizations, etc. This style of research is generally used in technical sciences, operations analysis, mathematics, clinical medicine and in operations research (Hochreiter & Schmidhuber, 1998). The word "construct" is frequently used in this context to indicate a new contribution being developed and the construct can be a new theory, model, software, algorithm, or a framework. The conclusions have to be objectively defined and this may include assessing the "construct" being developed analytically against some predefined criteria or performing some benchmark tests with the prototype. Constructive method solves practical problems while producing an academically valued theoretical contribution (Liisa et al., 2017). The use of the constructive research approach has made the researcher gain made understanding of the domain being researched.

5. Confronting the Semantic Web System

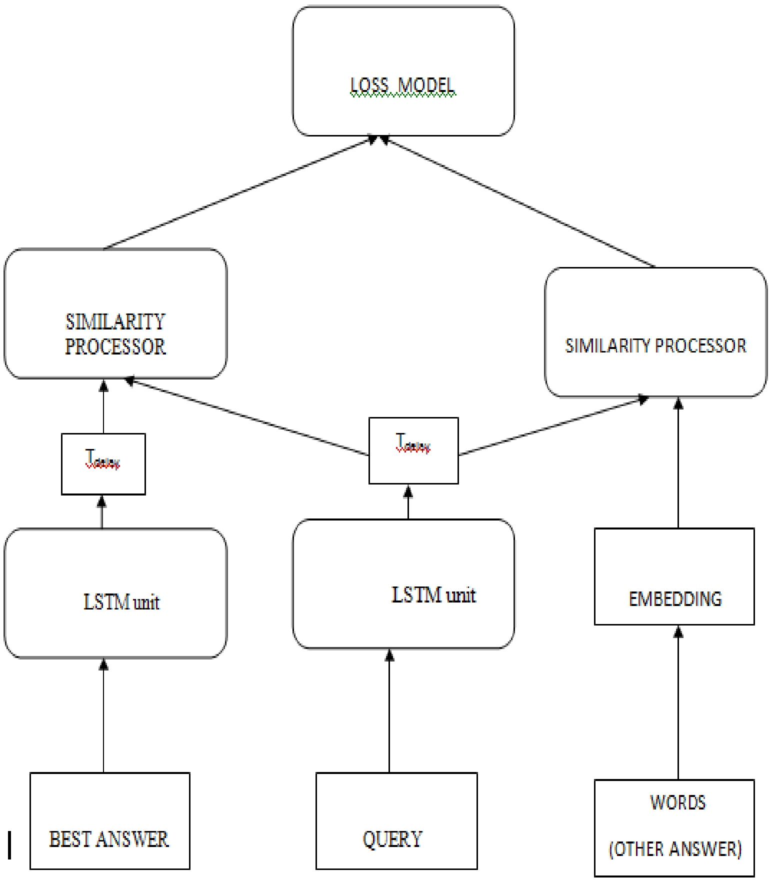

Figure 1: Architecture of Proposed System

The proposed system uses a deep Long-short term memory (LSTM) Artificial Neural Network (ANN) as machine learning approach for the learning of semantic feature representation. Machine learning approaches has been shown to be very useful in Community Question and Answer Q&A as it reduces the dimensionality of data to more machine analyzable forms and it also features a predictive capability. Our model also optionally includes a temporal observer to capture the temporal states that are likely to occur in a real Question and Answer (Q&A) session and solve the temporal problem in existing design. The proposed architectural model is shown in Figure 1.

The system is similar to the proposed one (Liisa et al., 2017) except that the Auto-Encoder (AE) is replaced with LSTM blocks and a time delay unit (T_{delay}) is added to capture the temporal behavioral states. The idea is to learn the semantic Question and Answer (Q&A) using a predictive (recurrent) LSTM network and in the long run understand and interpret knowledge sharing context. The use case diagram captures the key details in a semantic web application. The community Question and Answer (Q&A) case study is used to gain insight into the possible use of semantic representational learning system in real world. The diagram is as shown in Figure 2.

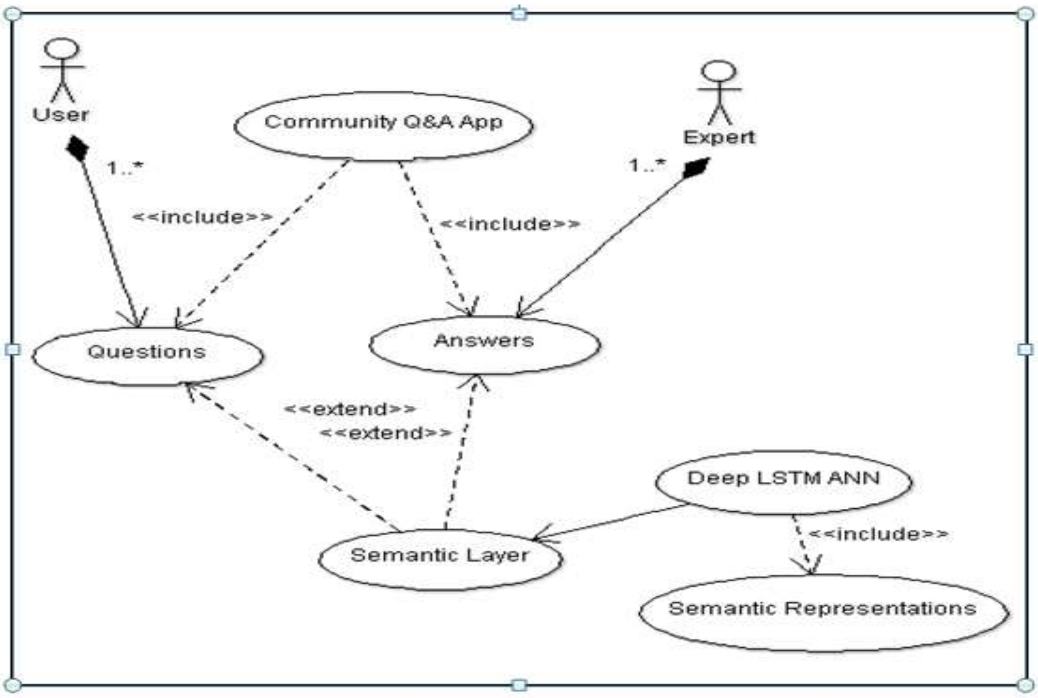

Figure 2: Use Case Design of the proposed System

In the use case model, two actors play a major role of query question and answering. While the user actor is dependent on the Question Use case model, the expert actor is fully associated with the Answer use case processing. To allow for semantic completeness, the semantic use case layer serves as an extension to the Questions and Answers use case which is conceptually part of the Semantic Layer. This layer allows the definition of semantic data forms and types including special embedding's that allow for proper meanings to be attached to the system. In addition, the deep LSTM ANN use case is dependent on the semantic layer and includes the learning of semantic representations.

The semantic Q&A system conforms to the specification expressing Dublin Core Metadata in HTML/XML documents and their use. The element name should be placed in the name attribute of the Meta tag, while the associated values (Title and Description fields for example) should be placed in the content attribute. In our case the name fields are really not necessary. Thus this system can be extended to Public domain using appropriate XML tags. The semantic Q&A systems application is hosted in the local directory in XAMPP root

htdocs folder which also serves as content repository with File Manager support.
In order to systematically model the relations that occur in a typical Q & A session, an implementation and ontology modelling tool (Protégé) is used. This is illustrated in Figure 3. This ontology captures diagrammatically the key operations in the Q & A system. These include the interface schema (or ontograph) showing the relation among class member entities, the interface display showing the class hierarchy, the description of class hierarchy and a sample class-object properties.

6. Training and Evaluating the Q&A System

The system is trained using the Long Short-Term Memory (LSTM) neural network. Data for training is obtained from the Q&A recent reports view and entered into the LSTM learning system (see Figure 4) to predict the most likely trending topic. The system is evaluated in terms of two recommended performance metrics for predictive recurrent neural networks. The performance metrics are the perplexity for evaluating the level of surprise seen by the predictive system and the Negative Log-Likelihood (NLL) estimates for evaluating the loss or cost of the predictive LSTM system.

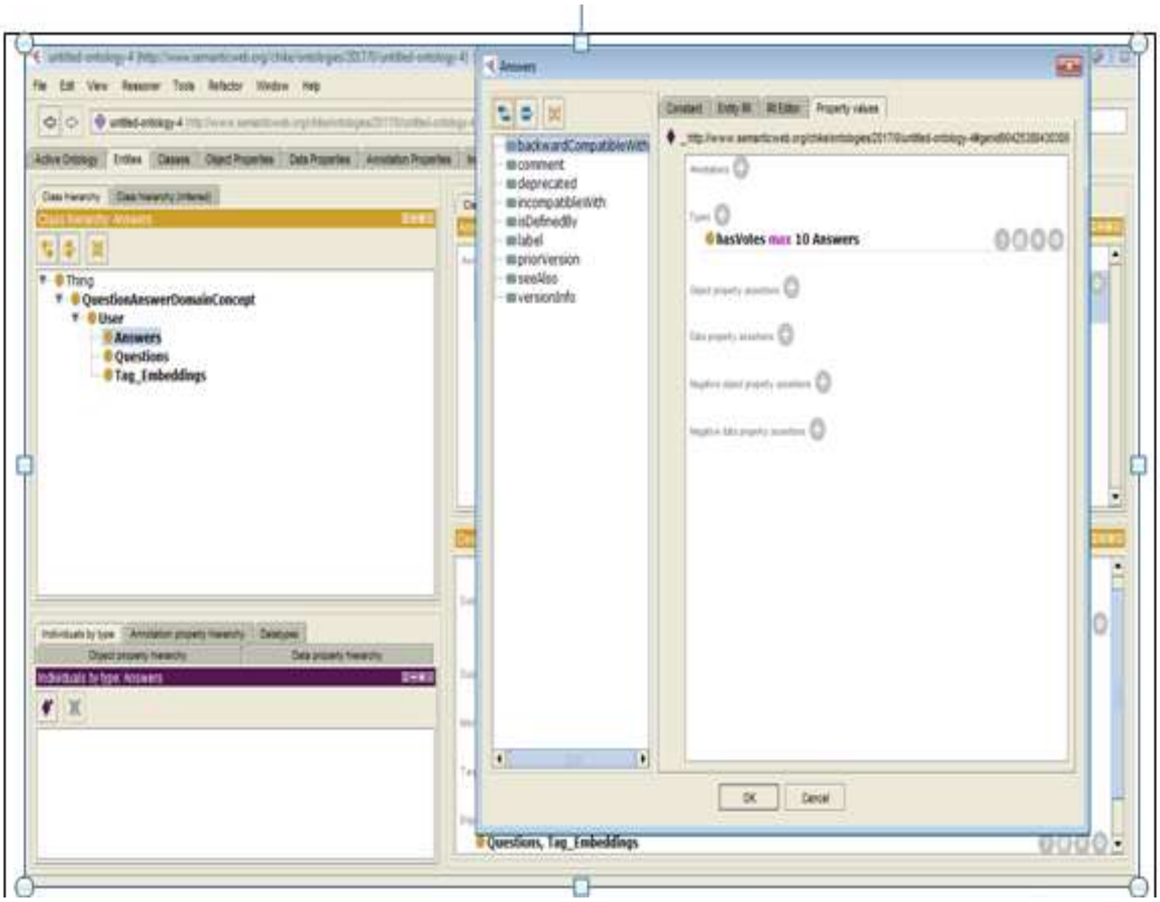

Figure 3: Interface display showing the Answers object properties

7. Results and Discussion

The system is trained using the Long Short-Term Memory (LSTM) neural network. Data for training is obtained from the Q&A recent reports view and entered into the LSTM learning system (see Figure 5) to predict the most likely trending topic. The system is evaluated in terms of two recommended performance metrics for predictive recurrent neural networks. The performance metrics are the perplexity for evaluating the level of surprise seen by the predictive system and the Negative Log-Likelihood (NLL) estimates for evaluating the loss or cost of the predictive LSTM system. The results showing the performances of the proposed system is provided in Table 1. From these results the training of the semantic web generated data by the LSTM shows the trending topic to be neatly close to "Secretary related to Senate". Thus, the LSTM can provide information on the most important topics in a community Q&A site. The perplexity and NLL estimates also indicate that these results can be achieved with reasonable accuracy or precision with higher hidden layer size (i.e. lower costs in prediction error). For instance, it was observed that increasing the number of hidden neurons (hidden layer size reduces the perplexity and NLL estimates and in turn increased the prediction performance. However, there is a limit to the number of hidden neurons that can be varied due to computational power requirements and running time issues. Figure 4 shows that the perplexity estimates degrades slightly while that of the NLL degrades gracefully.

Table 1: Results of Predictive Training Performances of the LSTM System on Semantic Data for 100 Training Epochs

LSTM Hidden Layer Size	Perplexity Estimate	NLL Estimate
400	1.10	3.49
225	1.21	4.84
100	1.27	7.33
25	1.98	17.5

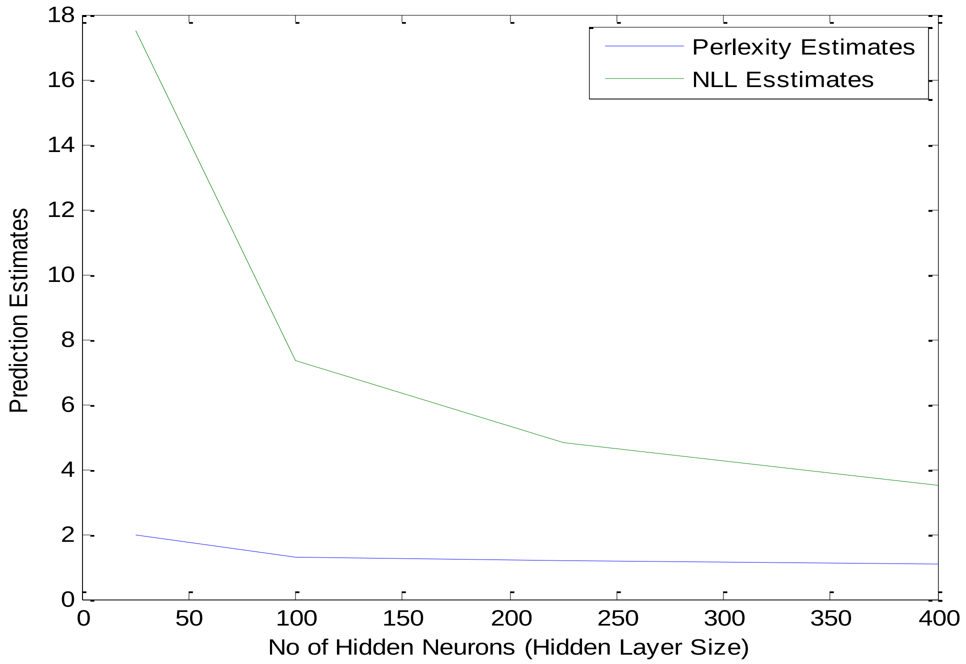

Figure 4: Prediction performance of the Semantic-LSTM system using Negative Log likelihood (NLL) and perplexity metric

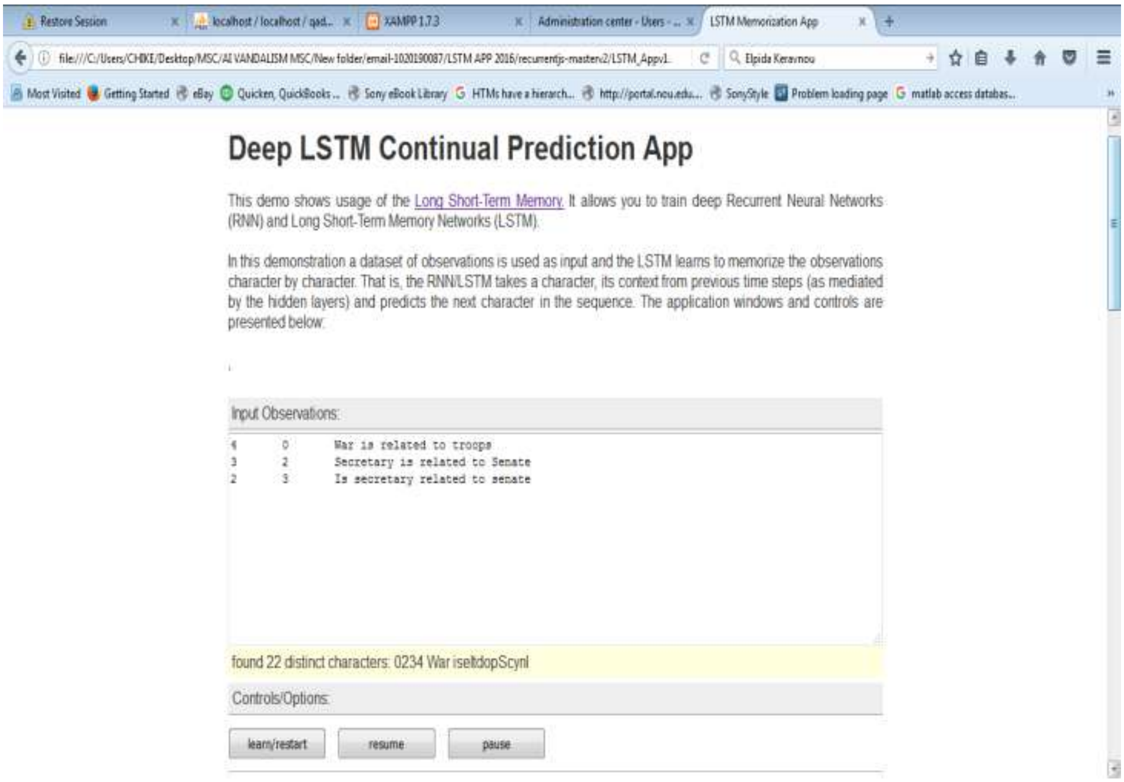

Fig 5: LSTM interface display for training generated semantic Q&A data

8. Conclusion

The world is evolving every day and the semantic web is the next new generation of web architecture in sight. It should therefore be a thing of great importance that the people follow the trend and not get lost in the search of in concise and linear direction. Obviously the semantic web presents a whole new world but it equals take for a generation of effortless computer scientist who are not afraid of its abstraction. Software model ontology artificial intelligent system has been developed to assist software engineers and ontology developers in the field of semantic web. This system is therefore a good starting point for researchers who are interested in developing ontology-compliant artificial intelligence products and services to the rapidly growing community of the semantic web and the internet

References

Bernstein, A. Hendler, J. & Noy, N. (2016). A new look at the semantic web. Communications of the ACM , 59(9), 35-37.

Da Silva, F. S. C. Vasconcelos, W. W. Robertson, D. S. Brilhante, V. De Melo, A. C. Finger, M. & Agust, J. (2002). On the insufficiency of ontologies: problems in knowledge sharing and alternative solutions. Knowledge-Based Systems. [209] P. W. Dalrymple. Knowledge15 (3), 147-167.

Hochreiter, S. & Schmidhuber, J. (1998). Long short-term memory. Journal of Machine Learning Research Neural computation, 9(8), 1735-1780.

Liisa, L. Juha-Matti, J. Sami, K. & Laura, P. (2017). The Constructive Research Approach: Problem Solving for Complex Projects. Chapter 8 of Design, Methods and practices for Research of Project Management (978-1-4094-4880-8) by Beverly pasian.

Mahmood, K. & Takahashi, H. (2015). Cloud based sports analytics using semantic web tools and technologies. IEEE 4th Global Conference on Consumer Electronics (GCCE) pp. 431-433.

Mikolov, T. Sutskever, I. Chen, K. Corrado, G. S. & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In: Advances in neural information processing systems pp. 3111-3119.

Minsky, M & Papert. S. (1969). Perceptron. Oxford, England: M.I.T. Press. http://dx.doi.org/, pp.334-435.

Najafabadi, M. M. Villanustre, F. Khoshgoftaar, T. M. Seliya, N. Wald, R. & Muharemagc. E. (2006). Deep Learning Techniques in Big Data Analytics. In Big Data Technologies and Applications. Springer International Publishing pp. 133-156.

Powers, S. & Practical, R. D. F. (2003). The Resource Description Framework (RDF). O'Reilly Media. Inc., Sebastopol, CA 1st edition; Language: English; ISBN-10: 0596002637; ISBN-13: 978-0596002633.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview Neural Networks. Elsevier Science Ltd. Oxford, UK, 61, 85-117.

Tang, D. Qin, B. & Liu, T. (2015). Deep learning for sentiment analysis: successful approaches and future challenges. Wiley Interdisciplinary Reviews in Data Mining and Knowledge Discovery, 5(6), 292-303.

Zhou, G. Zhou, Y. He, T. & Wu, W. (2016). Learning semantic representation with neural networks for community question answering retrieval Knowledge-Based Systems. Elsevier Science Publishers B. V. Amsterdam, 93, 75-83.