**A Machine Learning Approach for Automated Evaluation of Short Answers Using Text Similarity Based on WordNet Graphs**

PAPER LINK: [A Machine Learning Approach for Automated Evaluation of Short Answers Using Text Similarity Based on WordNet Graphs | Springer](https://link-springer-com.ezaccess.libraries.psu.edu/article/10.1007/s11277-019-06913-x)

Q. What the aim of the research?

A. This paper talks about the recent advances of deep learning methods on solving simple questions in two streams, the information extraction style and semantic parsing style. Later, the paper also introduces how to extend the neural architectures to answer more complex questions with iteration and decomposition techniques and summarize current research challenges.

Q. What are the scientific challenges?  
  
A. Here are some of the challenges the author faced during the research, which are:

Training data has been a long-standing problem in various machine learning based methods, especially for neural network models, which are generally assumed to require more training data than traditional methods. Particularly, in the KBQA scenario, it is especially expensive to collect question-answer pairs, let alone any fine-grained annotations.   
  
One of the drawbacks of Knowledge Bases (KB) is their inevitable incompleteness. It is almost impossible for a KB to contain all kinds of knowledge, restricted by its schema. The incompleteness brings an upper bound of the performance for KBQA models.   
  
Traditional semantic parsing based KBQA works usually rely on Combinatory Categorial Grammar (CCG) or Probabilistic CCG to derive the meaning representations from given questions, which is relatively hard to explicitly capture in a unified neural model without considerations.

Q. Describe the contribution of this research?

A. The research paper tries to implement two different types of paradigms to solve simple question, using the information extraction style and the semantic parsing style. We also understand that there is no strict difference between them, and most efforts benefit from both paradigms. We will be using these paradigms to help solve our research project of deep learning can be used in answering questions given a KB.

Q. How the proposed contribution differs from related work published in different venues?

A. The paper talks about three papers that focus on how the field of KBQA (Knowledge-based Question Answering) has seen various approaches to learn vector representations of words, entities, and relationships. Bordes et al. (2014) proposed a joint embedding framework that followed a retrieval-encoding-ranking pipeline. In the encoding step, the bag-of-words style was used, but this prevented the model from utilizing more sources of clues in natural language or the KB itself.  
  
Yih et al. (2014) improved the encoding step by using convolutional neural networks (CNN), which could automatically integrate local features within the convolution windows and capture fine-grained contextual information. Hao et al. (2017) introduced the attention mechanism to further improve the encoding step by representing the question differently according to different aspects of the answer.  
  
The memory network is another novel learning framework that has been applied to KBQA (Weston et al., 2015). Bordes et al. (1506b) demonstrated the potential of memory networks in managing large scale KB entries by reading and storing structured knowledge bases in a bag-of-symbols form. Miller et al. (2016) improved the idea by organizing KB triples in a Key-Value format, making the memory mechanism more flexible to store knowledge from heterogeneous resources. However, creating a memory for every knowledge requires large storage, and improving the space efficiency of memory networks may be a future research direction.

Bibliographical references:

* Bordes, A., Usunier, N., García-Durán, A., Weston, J., & Yakhnenko, “Translating embeddings for modeling multi-relational data”, Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014, pp. 2787-2795.
* Yih, W. T., Meek, C., & Harpale, A. “Semantic parsing for single relation questions”, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014, pp. 1533-1544.
* Weston, J., Bordes, A., Chopra, S., Mikolov, T. “Towards AI-complete question answering: A set of prerequisite toy tasks”, 2015, arXiv preprint arXiv:1502.05698.
* Miller, A., Fisch, A., Dodge, J., Karimi, F., Bordes, A., Laurent, T., “Key-Value Memory Networks for Directly Reading Documents”, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, pp. 1100-1109.

Q. Assess the weakness and the strength of the contribution?

A. One of the weaknesses in the research paper is that it focuses only on evaluating short answers, which may limit its applicability in other areas where more comprehensive evaluation is required. Another weakness is the heavy dependence on the WordNet database, and the accuracy of the results may vary depending on the quality and completeness of the database. The paper also lacks a comparison with other existing approaches to automated evaluation of short answers, which could provide a better understanding of the strengths and limitations of the proposed method.  
  
Strengths of the research paper is that they the use of WordNet Graphs provides a more comprehensive and accurate approach to evaluating text similarity. The results of this research could have practical applications in the field of education, where automated evaluation of short answers is a common requirement.  
  
Q. Explain how these papers are relevant to your project? What are the learned lessons and how do they help you in developing, training, and testing your project's models.

A. The research paper presents a machine learning approach for evaluating short answers in exams using text similarity based on WordNet graphs. The system uses a natural language processing technique to compare the answers with a set of predefined answers and evaluate the similarity based on the WordNet graph, which is a large lexical database of English words. The study found that this approach was effective in accurately evaluating short answers and providing consistent results. We will be taking inspiration from this paradigm to try and see if it can be implemented for evaluating more descriptive answers in exams.