

# Project Proposal: Building a Question Answering model

**Noul Singla**

Northeastern University,  
Boston, Massachusetts  
singla.n@husky.neu.edu

**Vipul Sharma**

Northeastern University,  
Boston, Massachusetts  
sharma.vip@husky.neu.edu

## Abstract

This project will be based on building a machine learning model which can read an article and then based on the article, answer a few questions. Dataset used in this project is taken from the Stanford: Stanford Question Answering Dataset (SQuAD).

## 1 Introduction

We all have interacted with Wikipedia articles for our research or just for casual reading about a topic, and many times thought if there would be a better option to skim through it to get what we want. Even in our daily life reading an article, a newsfeed or some social media post, we want to just quickly scroll through all the stuff and get our curiosities answered.

The main problem we face is that we have questions in our heads and simple search through the article doesn't yield any useful information we are searching. Search results on a page can help us search a word but there is no way that we can ask a question which is answered automatically. With the recent progress in Machine learning and deep learning methods in the Natural Language Processing (NLP), there is a belief that things that needed a manual intervention can be now automated using these advanced the concepts. NLP with these machine learning models gives the machine the ability to understand human language and semantically parsing the language into questions for machine to understand and respond in natural language. In this project an exploration of such a functionality is being considered with usage of Neural network or deep learning framework to carry out this complex task. Using this technology will not only save time but can enhance the capability of the digital assistants.

## 2 Related Work

SQuAD dataset has been used as a benchmark for question answering automation in NLP. Several University researchers as well as researches from organizations like Google, Microsoft, Facebook have all worked on this dataset. None of the models till date have been close to the human efficiency. Results of all those models are available at this squad website. [1]

## 3 Approach

We solve the problem at hand using the Machine Comprehension Model (MC)[2], that answer natural language questions by selecting an answer span within an evident text. The application starts by mapping words or phrases from the vocabulary, here the training dataset, to vectors of real numbers. For achieving this, we will be using the Word Embeddings like GloVe vectors [3]. The main purpose of using Word Embeddings over simple vectors is that, they are much better at capturing the context around the words than using a single hot vector for each word, boosting the performance of syntactic parsing. After processing the dataset, we need to work on the contextual binding of these vectors. We take the hidden vectors for the context and the hidden vector for question and bind them with context using the Attention Mechanism. With an attention mechanism we no longer try to encode the full source sentence into a fixed-length vector. Rather, we allow the decoder to "attend" to different parts of the source sentence at each step of the output generation. We will start with the implementation of simplest possible attention model, the encoder decoder using the Attention Score [4] and implement BiDAF [5] if time allows. Since the performance of Dot Product Attention will be too low, we are planning to dive into Context-to-Question (C2Q) Attention and Question-to-Context(Q2C) Attention, to find the context position for each context in our dataset.

## 4 Dataset and Evaluation

The evaluation of the SQuAD dataset models are done conventionally using the F1 and EM score and we plan on using the same evaluation metrics. The dataset has is already divided into learning and dev test dataset which would be used for model learning and model evaluations at a local level. There is also a provision to utilize an online test set to submit the model which can be explored if needed for further validation.

## References

- [1] <https://rajpurkar.github.io/SQuAD-explorer/>
- [2] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In EMNLP, 2014.

- [3] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. Squad: 100,000+ questions for machine comprehension of text. In EMNLP, 2016.
- [4] [phontron.com/class/nn4nlp2017/assets/slides/nn4nlp-09-attention.pdf](http://phontron.com/class/nn4nlp2017/assets/slides/nn4nlp-09-attention.pdf)
- [5] Minjoon Seo, Aniruddha Kembhavi, Ali Farhadi, Hannaneh Hajishirzi; Bidirectional Attention Flow for Machine Comprehension.
- [6] [CoRR, abs/1606.05250, 2016]. Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. *SQuAD: 100,000+ Questions for Machine Comprehension of Text*.
- [7] [ACL, 2017]. Wenhui Wang, Nan Yang, Furu Wei, Baobao Chang, and Ming Zhou. *Gated Self-Matching Networks for Reading Comprehension and Question Answering*.
- [8] [Stanford. 2014]. Jeffrey Pennington, Richard Socher, Christopher D. Manning. *GloVe: Global Vectors for Word Representation*.