1.2 Objectives

The objectives of this project are:

- To translate English Nepali language to English in a better and efficient way.
- To implement this concept of machine translation for the Search Engine Optimizations
- To learn how encoder-decoder models with attention are used to learn the relationship between words in two different languages.
- To obtain a bleu score of at least 25% and interpret the results.

2. Literature Review

Statistical Machine Translation (SMT) has been the dominant translation paradigm for decades. Practical implementations of SMT are generally phrase-based systems (PBMT) which translate sequences of words or phrases where the lengths may differ. Even prior to the advent of direct Neural Machine Translation, neural networks have been used as a component within SMT systems with some success. Perhaps one of the most notable attempts involved the use of a joint language model to learn phrase representations which yielded an impressive improvement when combined with phrase-based translation. This approach, however, still makes use of phrase-based translation systems at its core, and therefore inherits their shortcomings. Other proposed approaches for learning phrase representations or learning end-to-end translation with neural networks offered encouraging hints, but ultimately delivered worse overall accuracy compared to standard phrase-based systems.

Neural machine translation is a recently proposed approach to machine translation. The models proposed recently for neural machine translation often belong to a family of encoder–decoders and encode a source sentence into a fixed-length vector from which a decoder generates a translation. In the paper by Dzmitry Bahdanau, they conjecture that the use of a fixed-length vector is a bottleneck in improving the performance of this basic encoder–decoder architecture, and propose to extend this by allowing a model to automatically (soft-)search for parts of a source sentence that are relevant to predicting a target word, without having to form these parts as a hard segment explicitly. With this new approach, they achieved a translation performance comparable to the existing state-of-the-art phrase-based system on the task of English-to-French translation.