Applied Soft Computing 38 (2016) 1060–1075

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Quantum neural network based machine translator for English to Hindi

Ravi Narayan [a,*], S. Chakraverty [b], V.P. Singh [a]

[a] Department of Computer Science, Thapar University, Patiala, Punjab, India
[b] Department of Mathematics, National Institute of Technology Rourkela, Odisha, India

ARTICLE INFO

Article history:
Received 30 November 2014
Received in revised form 9 July 2015
Accepted 19 August 2015
Available online 29 August 2015

Keywords:
Machine translation
Semantic translation
Syntactic translation
Quantum neural networks

ABSTRACT

This paper presents the machine translation system for English to Hindi which is based on the concept of machine learning of semantically correct corpus. The machine learning process is based on quantum neural network (QNN), which is a novel approach to recognize and learn the corpus pattern in a realistic way. It presents the structure of the system, machine translation system and the performance results. System performs the task of translation using its knowledge gained during learning by inputting pair of sentences from source to target language i.e. English and Hindi. Like a person, the system also acquires the necessary knowledge required for translation in implicit form by inputting pair of sentences. The effectiveness of the proposed approach has been analyzed by using 4600 sentences of news items from various newspapers and from Brown Cuprous. During simulations and evaluation, BLEU score achieved 0.9814 accuracy, NIST score achieved 7.3521, ROUGE-L score achieved 0.9887, METEOR score achieved 0.7254 and human based evaluation achieved 98.261%. The proposed system achieved significantly higher accuracy than AnglaMT, Anuvadaksh, Bing and Google Translation.

1. Introduction

Machine translation (MT) along with natural language processing (NLP) always remained an area of interest for researchers since the time when computers were invented. Many researchers have tried to build the system which can understand multiple languages to translate from one source language to another target language. They also searched the way how computers understand and generate the human languages with semantics and syntactic. However, they realized that still many languages have translation difficulties, grammatically and semantically.

Most of the machine translation researchers have done their work in the field of high exactitude in alignment of words in sentences (chunk). The translation model based on the structure, and the relationship of the two languages can be produced after aligning the parallel text upto the level of words and phrases. Since all the words in a sentence are related to the adjoining words. More refined models generally define their probability distributions over flanking phrases. Bilingual dictionaries may be used as a good source for directly pulling out the word correspondences. The attractive approach applied is to align higher level of syntactic or semantic structures, in addition to aligning individual words and contiguous strings or 'phrases'. Each language is having its own pattern to align the parts of speech in any sentence. Chandola et al. have also explained the alignment and reordering of words with the use of corpus pattern for English to Hindi machine translation [4].

For machine translation, the complex algorithms can be customized only by an intellectual who is an expert in expressing such rules or algorithms in the programming language of the computer. This clearly proves that still the machine translation systems are not capable of self-adopting the correct translation rules with higher accuracy. Some of the researchers have achieved the accuracy upto 96.22% during machine translation using neural network based approach [12].

Quantum neural network based machine translation system is a possible solution to this problem, as these have ability to learn from examples by recognizing their pattern. In this work, the main focus is to show the significant increase in accuracy achieved during our research done with machine translation by using new approach of quantum neural network for reordering of words for parts of speech tagging and their alignment during the machine translation. If the translation of a sentence of such a quantum neural network translation system is wrong, then it can be corrected and taught proper translation by a user without any expert technical knowledge of how the computer stores and represents rules. Quantum neural

* Corresponding author at: C-4/1499, Jaroda Gate, Near Old Grain Market, Jagadhri (Yamuna Nagar), Haryana 135003, India. Tel.: +91 931 558 0711.
E-mail address: ravi_n_sharma@hotmail.com (R. Narayan).

http://dx.doi.org/10.1016/j.asoc.2015.08.031
1568-4946/

network (QNN) is effective to classify indeterminate data, because QNN has inherently fuzzy properties which can encode the sample information into discrete levels of certainty or uncertainty. The transfer function of quantum neural network expressed as linear superposition of sigmoid function [13]. In this way, a hidden layer neural cell can express more states. Each sigmoid function has different quantum interval. Thus, QNN decreases indeterminacy and increases veracity of pattern recognition [6,14].

The motivation behind the study of QNN is its similarity with the living brain, as human performs successfully in complex environments where many ill-defined and uncertain factors are encountered. Human can successfully work in these unrealistic situations where classes of information are not bounded tightly or overlapped on each other and human thinking employs rules-of-thumb, experience, perception and other heuristics. Similarly QNN is able to independently detect the occurrence of uncertainty in the sample data and adaptively learns to quantify the existing uncertainty. QNN detects and quantify uncertainty in the sample data without any assumptions about the number of classes, the number of uncertainty levels in the data, the convexity of the classes. If the resulting dataset lies in the boundary between overlapping classes, the QNN will assign it partially to all classes whose overlapping boundaries include this resulting dataset. In realistic situation if there is no uncertainty regarding the classification of a certain feature vector, then QNN will assign it to the class indicated by the training set. The traditional neural network (NN) advocates the study of intracellular structures but it can only cover the realistic space in given data classes, having the nature of certainty [9,10,23].

Recently some researchers used the multi layer quantum neural network in other applications like robotics etc. A multi-layer quantum neural network (QNN) was considered in which qubit neurons were used as an information processing unit. Qubits were used to store the states circuits during quantum computation. A real coded genetic algorithm was used to improve the learning process rather than of back propagation algorithm for supervised training of the multi layer QNN. For evaluating the capabilities of the direct quantum neural network controller, the computational experiments were conducted for controlling a discrete time nonlinear system and a non holonomic system (a two wheeled robot) [32].

Liu et al. introduced the single hidden layer feed forward quantum neural network based on the Grover learning. The model was based on the concepts and principles of quantum theory. The quantum hidden neurons and the connected quantum weights were defined by joining the quantum mechanism with the feed forward neural network. These were used as a fundamental processing unit in a model. The quantum neuron and weights along with the Grover searching algorithm showed the result in an efficient neural network characteristic of reduced network, high efficient training and prospect application in future [19].

Li et al. introduced the hybrid quantum inspired neural networks with sequence inputs. The performance of classical neural network was increased with the use of Quantum Inspired Neural networks (QINN) model. The inputs in the model were the discrete sequences which were described by the matrix where number of input nodes represents the number of rows and sequence length represents the number of columns. The model has three layers, the hidden layer contained the quantum neurons and the output layer contained the classical neuron [17].

The major difference between conventional neural network and quantum neural network is the form of the nonlinear activation functions of their hidden units. In QNN, a multilevel activation function is used instead of the ordinary sigmoid functions. Every multilevel function contains the sum of sigmoid functions shifted by the quantum intervals. The transfer function of the quantum neuron in hidden layer applies the superstition of several traditional transfer functions [33]. The quantum computing can be used as a general framework for producing quantum analogs of well-known classical artificial neural networks.

In QNN, a neuron is regarded as a basic information unit of a quantum computer (qubit) and a synaptic connection corresponding to a qubit interaction. According to Li Fei et al., QNN has three layer architecture i.e. inputs, one layer of multilevel hidden units, and output units [16]. Consider a neuron with n inputs $|x_1\rangle \ldots |x_n\rangle$

$|x_j\rangle$ is a quantum bit (qubit) of the form

$$|x_j\rangle = a_j|0\rangle + b_j|1\rangle = (a_j, b_j)^t$$

In QNN model, the sigmoid function with various graded levels has been used as the activation function for each hidden neuron and is expressed as:

$$sgm(x) = \frac{1}{n_s}\sum_{r=1}^{n_s}(1/(1+\exp \quad \beta_h(x-\theta^r)))$$

where n_s denote the number of grades in the sigmoid transfer functions, and θ^r quantum interval of quantum level r and, where β_h is slope factor [9,10,23].

2. Machine translation systems

Machine translation (MT) is a field of natural language processing. It involves the complete linguistic analysis of sentence used for automatic translation from one language to another. The main challenging issues need to be addressed are word ambiguity, word order, word sense, idioms, pronoun resolution, syntactic ambiguity and structural ambiguity. The statistical machine translation approach was inspired by noisy channel model and this was introduced by research community as a statistical tools based on a noisy channel model which was originally developed for speech recognition [7,8,20,29].

Machine translation based on Markovian model, is based on phrases rather than words and coupled with a phrase-to-phrase translation table. In this approach, translating a text-amount to its most likely translation is based on its available model parameters. Inferring the parameters of this model from bilingual corpora is a matter of statistics. By model inference, the task of extracting is performed for all tables, parameters and functions from the corpus [21].

Rule based machine translation method for translating English sentences to Malayalam is proposed by Rajan et al. In this work the rules and the bilingual dictionaries are used for converting source language structures into target language structures. The rules used in this approach are based on the parts of speech (POS) tag and dependency information obtained from the parser. Two types of rules are used in this system, one is transfer link rule and the other is morphological rule. The transfer link rules are used for generating target structure. Morphological rules are used for assigning morphological features [28]. English to Malayalam machine translation system based on ANN-based transfer module has been introduced by Dam et al. In which the English to Malayalam translation is done by the ANN-based transfer module with the help of bilingual dictionary. According to author, the idea is to allow the ANN-based transfer module to automatically learn correspondences between source and target language structures using a large set of English sentence and their Malayalam translations [22].

Machine translation system based on multi-hierarchical approach for Chinese complex long sentence introduced by Yin et al., in which the translation of Chinese complex long sentence has been done through the analysis of syntax function, punctuation and conjunction words. This uses semantic characteristic of Chinese sentences including grammatical features, the length of