Dialogue System using Long Short-Termed Memory

Muhammad Husain Toding Bunga[1], **Suyanto**[2]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[4]Divisi Digital Service PT Telekomunikasi Indonesia
[1]muhammadhsn@students.telkomuniversity.ac.id, [2]suyanto@telkomuniversity.ac.id,

Abstract

As the technology of natural language understanding and language generation improve, there is increasing human interest towards human-computer interaction, which can be used for various applications such as customer services, travel, and much more. Well known examples of conversational AI are Apple's Siri, Amazon's Alexa, Microsoft's Cortana and Google's Google Assistant.

Most work related on this field are emphasizing on single sentence or speaker turn. While sometimes a conversation has their own context according to previous conversation. Designing this kind of conversational system is challenging, most of the time conversational agent are built based on knowledge based system and rule based system. By building a conversational agent with data driven approaches, which learn from a corpus we could improve the amount of time and effort needed to create a rule based system.

Keywords : natural language understanding, dialogue system, conversational agent, LSTM

1. Introduction

Natural language application are getting popular as the technology of natural language understanding and language generation improve. Many big companies are building natural language based application or AI conversational agent, these applications are popular to customers since they give a unique experience of being served by human-like applications. Many start-ups and business are also starting to adopt and use AI for economical customer relation and customer service. Besides that, the behavior of users now have changed, people are using internet more often, they expect something fast and consistent, which are the main usage and ability of computer.

Background

Conversational AI are designed to imitate how human talk, creating an experience of being served by human, while giving consistent and fast response which might not be able to be provided by regular applications. Most conversational agent are built with knowledge based system and rule based system, but on this research, the dialogue model will be implemented with rule based model but with data-driven approach. Many other work has accomplished data-driven approach, the work of Ritter [1], for example demonstrates that a response generation system can be constructed from Twitter conversations using statistical machine translation techniques, where a status post by a Twitter user is "translated" into a plausible looking response. However, the approach that is done by Ritter et al. (2011) does not address the challenge of generating responses that are sensitive to the context of the conversation. The ability to take into account previous utterances is key to building dialog systems that can keep conversations active and engaging. A neural network architecture is used to address sparsity issues that arise when integrating contextual information into classic statistical models, allowing the system to take into account previous dialog utterances[2].

One of the advantages of RNNs is that they might be able to connect previous information to the present task, such previous conversation might inform something about current conversation. But as the gap of the context grows , RNNs became unable to connect the information. This problem is called vanishing gradient problem[12]. In this work, we are using an LSTM cell, that is designed to avoid the long-term dependency or vanishing gradient problem.

Problem Identification

A conversational system with a context based on dialogue history can be a challenge to design, given the context of a dialogue and number of possible outcome in a dialogue. To solve the problem, this research will try to build a dialogue system using rule based with data-driven approach from a corpus by utilizing an LSTM network. In this case LSTM neural network is chosen because it has the ability to remember past observations arbitrarily long, and has been shown to yield superior performance in many domains[17].

To be able to use the conversation model to state the track of a conversation, the ability to understand a text is needed. To create this module, this research uses GloVe (Global Vectors) for word representations, the GloVe model weight statistical information by training only on the nonzero elements in word occurrence matrix rather than whole sparse matrix[13]. Using the vectors, a model will be train with a Multiclass Support Vector classifier to classify a sentence's intent.

Goals

Based on the problem defined, the main goals of this research is to build a complete dialogue system or a dialogue agent that manages to understand the intent of a text and give response based on the state of dialogue.

Writing Organization

In this sub-section the following sections (after Introduction) are written in this Final Task journal, accompanied by a very brief explanation.

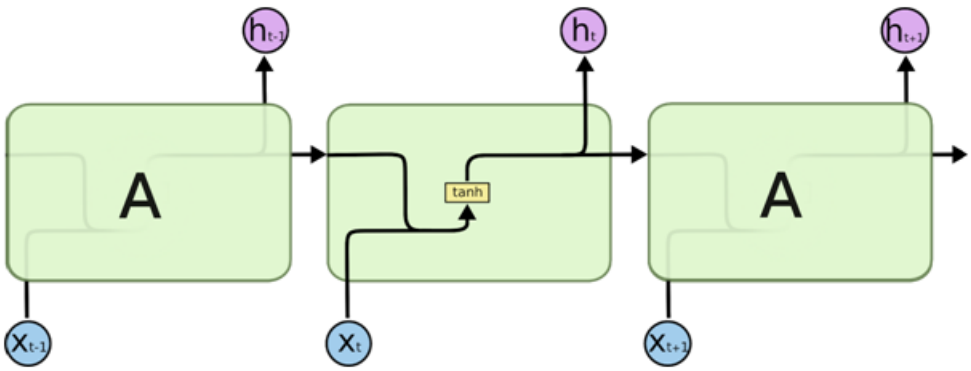
2. Literature Review

Data driven methods are now widely used for building conversation system since the popularity of social medias such as Facebook, Twitter, Instagram and online forums are making it easier to collect data[11]. Microsoft's team research, for example has build a dialogue system with Convolutional Neural Network as their product entity detection even though their dialogue state tracker is using their own custom engine, resulting 0.898 precision, 0.677 recall and 0.772 F1[19].

Another similar work is also done by Jason [17] with LSTM combined reinforcement learning. Their LSTM itself has 0.91 accuracy with 20 training dialogue used.

Support vector machine is a machine learning algorithm for two group classification problems, SVM basically implement the following idea : the input vectors are non-linearly mapped to a very high dimension feature space. In this feature space a linear decision surface is constructed. Special properties of the decision surface ensures high generalization ability of the learning machine[5]. However, svm is not good approach for multiclass problems. By using Multiclass SVM, where two-class classifier over a feature vector derived from the pair consisting of the input features and the class of the datum. At test time, the classifier chooses the class . The margin during training is the gap between this value for the correct class and for the nearest other class, and so the quadratic program formulation will require that . This general method can be extended to give a multiclass formulation of various kinds of linear classifiers. It is also a simple instance of a generalization of classification where the classes are not just a set of independent, categorical labels, but may be arbitrary structured objects with relationships defined between them[3].

LSTM is one of the Recurrent Neural Networks, which capable of learning long-term dependencies, designed to avoid long-term dependency problems, remembering information for long periods of time. LSTM have the form of a chain repeating modules, where each modules have 4 gates. These gates are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a pointwise multiplication operation. These gates are called input gate, output gate ,forget gate and input modulation gate. Input gates decides which information that should be remembered by the network, output gate handles the attention [4] mechanism, which data should it focus on. And forget gates decides which information to forget [6]



Gambar 1. **Structure of LSTM cell**