DYNAMIC CASCADE NEURAL NETWORK

Tofara Moyo

C.E.O.Mazusa
tofaramoyo@gmail.com

ABSTRACT

We introduce a type of cascade correlation network to predict the next word in a sentence, where the hidden layer is connected in such a way as to represent the topology of the conversation up to that point. There will be one input neuron that is fed the current word, and one output neuron that will output the predicted next word in the sentence. During training, we will be given a set of conversations to train on. That means that we shall know the input output pairs to train with, and the topology of the previous parts of the conversation correlating with each pair. We shall build the hidden layer in such a way that it is isomorphic to the previous parts of the conversation at the point here that input output pair coincide. During operation we will build the hidden layer of the network at each input word and in such a way as to be isomorphic to the previous parts of the conversation up to that point. Training will make the probability of the output n+1 word to be conditionally dependent on the previous words in that order and not just the nth word, conditional on the training conversations giving the chat-bot the ability to talk consistently on topics and themes found within the training conversations.

KEYWORDS

Dynamic, cascade network, Chatbot

1. INTRODUCTION

We will develop theory based of cascade correlation networks [1], where connections are added progressively during the training process. To that end we will label every node that could potentially be in the hidden layer of the DCNN with a word or token from the conversation.

Every node thus will have a list of connections which are labelled with a natural number from one to "s" that it could potentially use to connect to another hidden layer node.

To train the entire system we then perform the following algorithm

Take the hidden layer neuron labelled BOC, meaning beginning of conversation, and add it as the first node grown into the hidden layer of the network and connect it to the output and input neurons.

-Take the nth word as input and the n+1th word as output, and back propagate the error.

- take the hidden layer neuron labelled with the name of the nth word, and add it to the hidden layer.

-connect the BOC node laterally to the recently added node

-Take the n+1 word as input and the n+2 word as output and back propagate the error.

-Connect the hidden layer node labelled with the name of the n+1th word and add it to the hidden layer.

- connect the nth hidden layer node to the recently added n+1th node, using the nth nodes 1st connection from its list of s connections.

Repeat above adding the s+ 1 connection of a node to the network every time we get the same word being used as input as it was t an earlier part of the conversation rather than re adding the same node

So whenever we get the same word we used previously we do not duplicate its node within the hidden layer. Instead we make a recurrent connection using the next connection in its list of connections to connect it to the next incoming node.

Then we back propagate through time the entire network of nodes at each train example.

After training and during operation we simply grow networks hidden layer using the same rules we used to grow it during training after each input, using the predicted word as the input in the next time step until an EOU, end of utterance token is predicted. Then the first word in the live response to the chat is given, and instead of producing an output we grow the network with the relevant connections. Until the human signifies the EOU token, then the network begins prediction outputs staring with the output from EOU variable as input.

3. CONCLUSIONS

We have designed a neural network that can process time series data and generate new predictions. Here we used a chatbot as an example, giving the chatbot the attention to be aware of all the words that had been said before in a conversation and by whom and in what order.

REFERENCES

[1] Fahlman, Scott E.; Lebiere, Christian (August 29, 1991). "The Cascade-Correlation Learning Architecture" (PDF). Carnegie Mellon University. Retrieved 4 October 2014.

Authors

Tofara Moyo

is lead researcher at Mazusa

Photo