1. Research: big neural networks

The purpose of this question is to familiarize with the network Very large trained neural networks such as GPT-3 is during which this network and network Review similar large-scale studies and report on things like learning methods, data, and their applications.

2. Research: other learning method

Review the following learning methods and in addition to expressing their ideas, express the working method and application examples of each

- Progressive Learning
- Meta Learning
- Zero Shot Learning
- One Shot Learning

3. Sentiment analysis in text with LSTM network

In this question, it is supposed to useLSTMTrain a model that can identify the positive or negative sentiment associated with each comment in the data Recognize IMDBs. For this purpose, first need to prepare It is data creation and then model design and training and evaluation. The movie review dataset, commonly known as theIMDB dataset , contains positive and negative reviews for network training and a similar number for 25,000 evaluation. The problem is designed to distinguish between positive and negative feelings written in each comment. The desired data can be accessed from the link below. /http://ai.stanford.edu/~amaas/data/sentiment Before starting to work on and design the model, it is necessary to prepare and analyze the data. For this purpose, it is necessary to convert all the data into lowercase letters. Remove the punctuation marks. Make a list of comments. It is necessary to break down the text into separate units (unitization), for which you can use the space character and punctuation marks as the border between words.

It is necessary to convert each unit - word - into a numerical vector. For this, ready-made models and vectors such asVec2Word ,FastText Or useBert . UseBert s vectors in this ' task . Get an example in the following link: https://github.com/google-research/bert Now

it's time to teach the network. To do this from the network Use LSTM and build a model whose input is an opinion and whoseoutput is an agreement or disagreement. For this purpose, it is necessary to use theLSTM algorithm in general with the possibility of changing the number of hidden layers , the number of Neron of each layer , setting the activator function, changing the batch size , etc. you can from the library Use available tools such as tensorflow.

A- Create a one-wayLSTM network with batchsize = 50 a hidden layer with 64 memory , units and a sigmoid activation function and learn it at a rate of 0.01. The network was trained with theTrain part of the data and a diagram of the trend of MSE error changes of the network on this data draw Reportthe accuracy of the trained model on the test data .

B- Make the network of part A two-way and get the results in this state and compare it with the one-way state.