Proceedings of the SMART–2018, IEEE Conference ID: 44078
2018 International Conference on System Modeling & Advancement in Research Trends, 23rd–24th November, 2018
College of Computing Sciences & Information Technology, Teerthanker Mahaveer University, Moradabad, India

Designing and Developing a Chatbot Using Machine Learning

Praveen Kumar[1], Mayank Sharma[2], Seema Rawat[3] and Tanupriya Choudhury[4]
[1,2,3,4]*Amity University, Noida, Uttar Pradesh, India, UPES Dehradun*
E-mail: [1]*pkumar3@amity.edu,* [2]mayank.sharma28@amity.edu,
[3]*srawat1@amity.edu,* [4]tanupriya@ddn.upes.ac.in

***Abstract*—Deep Learning is a new area of Machine Learning and research, which was introduced with the aim of taking the Machine learning closer to one of its original and main goal which is: AI (Artificial intelligence). If we speak of automatic learning algorithm, they tend to be linear, the DL (Deep Learning) algorithms are configured to increase complexity and abstraction. To learn in depth, imagine a 5-year-old whose first word is cat. The child Keeps on learning what a cat is by showing objects and saying the word cat. The mother says: "Yes, it's a cat" or "No, it's not a cat". As the child continues to point objects, he becomes more aware of the characteristics and features of all cats. What the child does, without even Knowing what he is doing, it clarifies a very complex level of abstraction by constructing a scale in which each level of abstraction is created with the help of Knowledge acquired the layer which was before on the scale.**

The programs that run on the system which uses deep learning go through the same process. Each algorithm in the way that applies a non-linear approach on its input and it uses what it learns to make a statistical model from the given output. Repetition continues until the output has reached an acceptable point of correctness. The number of layers in which the process is passing at each stage is what gave the name a prefix as deep.

Keywords: Deep Learning, Chabot, Machine learning

I. Introduction

Any business designed to entice a first-class customer to experience requires a proper and comprehensive customer advantage office. Already, offering a strong phone number with an informed customer benefit aggregate was enough to satisfy customers with requests and issues. In any case, in today's world, customers tend to expect another level[5]of simplicity and speed, no matter when they need to contact our organization to solve a problem. This resolute attitude leads to an impressive disappointment regarding the information lines and the maneuvers with the authorities who are indifferent or disconnected from the problems and needs of the client.

In view of these shifting needs, associations are beginning to introduce an optional visit as a technique of close contact with their clients. The decision of the discussion allows a greater adaptability, allowing our customers to start a discussion at any moment and at any time, whatever the period[2] necessary. Ideally, customers look for open discussions throughout the day, consistently and provide information about each type of customer problem. Communication associations are turning to this other option to deal with the strange condition of the customer asking once a day. Our clear work experience with telecom[6] associations on this issue has given our association a cautious perception of how to use chatbots more cost-effectively while overseeing improved measures of chat volumes. development behind chatbots to meet the needs of our customers. Although we are doing a powerful chatbot, we will most likely automate the system to limit the precondition for human organization, as this would be normal. To achieve this computerization, there are two options of directives. The first decision is to draw from our guides the response proposals given by AI to improve the customer benefit. This optional course always requires the use and regulation of customer service meetings. The second option is to completely replace all specialists with a chatbot, eliminating the precondition for the benefit of the client. We focus here on the second decision, which also asks anyway to give a more remarkable result over time that we characterize: we will probably accumulate a Chabot who will speak tastefully with customers of our telecommunications company. In all cases, including machine learning, the underlying progression is to[2][1] prepare the data for the machine to decode accurately. Because of the creation of a Chabot, the task includes the contribution of innumerable associations between clients and customer benefit representatives in order to teach the machine which words and phrases are business-sensitive. The next step in the process is called pre-preparation. It is hoped that this movement is consolidated in the understanding of the machine and makes it possible to recognize the spelling errors in the compositions.

II. Neural Networks

It all started when Warren McCulloch and Walter Pitts created the first model of an NN in 1943. Their model was purely based on mathematics and algorithms and couldn't be tested due to the lack of computational resources. Later on, in 1958, Frank Rosenblatt created the first ever model that could do pattern recognition. This would change it all. The Perceptron. However, he only gave the notation and the model. The actual model still could not be tested. There were relatively minor researches done before this.

The first NNs that could be tested and had many layers were published by Alexey Ivakhnenko and Lapa in 1965.Therefore, the first layer of the frame implies the "incorporation of the card" resulting from the convolution forms, with a tendency to an additional substance and perhaps a weight or the institutionalization of the characteristics. This fundamental step is followed by sub-sampling that further reduces the dimensionality and provides a bit of life for the development of the space. The subsampled integrated design is weighted and the driving trend multiplied by an authorization job. There are several varieties with a single guide for each layer or sums of several cards.

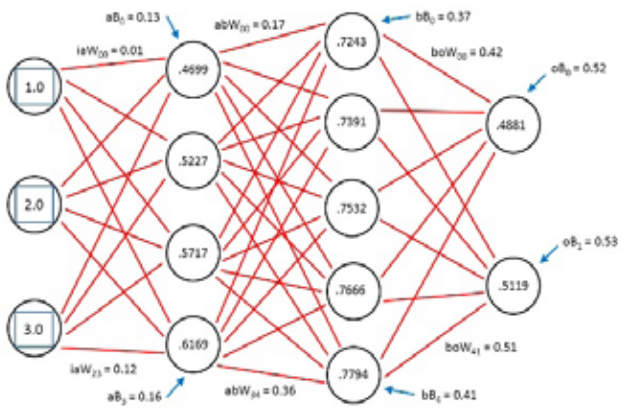

Fig. 1: Neural Network

After these, the research on NNs stagnated due to high feasibility of Machine Learning[4] models. This was done by Marvin Minsky and Seymour Papert in 1969.

Max-pooling was later introduced in 1992 which helped with 3D object recognition as it helped with least shift invariance and tolerance to deformation. Between 2009 and 2012, Recurrent NNs and Deep Feed Forward NNs created by Jürgen Schmidhuber's research group went on to win 8 international competitions in pattern recognition and machine learning.

In 2011, Deep NNs started incorporating convolutional layers with max-pooling layers whose output was then passed to several fully connected layers which[8] were followed by an output layer. These are called Convolutional Neural Networks. There have been some more researches done after these but these are the main topics one should know about.

A. What really is a Neural Network?

A good way to think of an NN is as a composite function. You give it some input and it gives you some output.

There are 3 parts that make up the architecture of a basic NN. These are:

- Units / Neurons
- Connections / Weights / Parameters
- Biases

All of the things mentioned above are what you need to construct the bare bones architecture of an NN.You can think of these as the building blocks/bricks of a building. Depending on how you want the building to function, you will arrange the bricks and vice versa. The cement can be thought of as the weights. No matter how strong your weights are, if you don't have a good amount of bricks for the problem at hand, the building will crumble to the ground. However, you can just get the building to function with minimal accuracy(using the least amount of bricks) and then, progressively build upon that architecture to solve a problem.

B. Units / Neurons

Being the least important out of the three parts of an NNs architectures, these are functions which contain weights and biases in them and wait for the data to come them. After the data arrives, they, perform some computations and then use an activation function to restrict the data to a range(mostly).

Think of these units as a box containing the weights and the biases. The box is open from 2 ends. One end receives data, the other end outputs the modified data. The data then starts to come into the box, the box then multiplies the weights with the data and then adds a bias to the multiplied data. This is a single unit which can also be thought of as a function. This function is similar to this, which is the function template for a straight line:

$y = mx + b$

Imagine having multiple of these. Having more than 2 of these could promote non-linearity in an NN. Since now, you will be computing multiple outputs for the same data-point(input). These outputs then get sent to another unit as well which then computes the final output of the NN.If all of this flew past you then, keep reading and you should be able to understand more.

C. Weights / Parameters / Connections

Being the most important part of an NN, these(and the biases) are the numbers the NN has to learn in order to generalize to a problem. That is all you need to know at this point.

D. Biases

These numbers represent what the NN "thinks" it should add after multiplying the weights with the data. Of course, these are always wrong but the NN then learns the optimal biases as well.

E. Hyper-Parameters

These are the values which you have to manually set. If you think of an NN as a machine, the nobs that change the behavior of the machine would be the hyper-parameters of the NN.

III. Activation Functions

These are also known as mapping functions. They take some input on the x-axis and output a value in a restricted range(mostly). They are used to convert large outputs from the units into a smaller value, most of the times. Your choice of an activation function can drastically improve or hinder the performance of your NN. You can choose different activation functions for different units if you like.

Here are some common activation functions:

A. Sigmoid

The Sigmoid function

B. Tanh

The tanh function
ReLU: Rectified Linear Unit
The ReLU function
Leaky ReLU
The Leaky ReLU function

C. Layers

These are what help an NN gain complexity in any problem. Increasing layers(with units) can increase the non-linearity of the output of an NN.Each layer contains some amount of Units. The amount in most cases is entirely up to the creator. However, having too many layers[5] for a simple task can unnecessarily increase its complexity and in most cases decrease its accuracy. The opposite also holds true.

There are 2 layers which every NN has. Those are the input and output layers. Any layer in between those is called a hidden layer. The NN in the pictu An NN with 2 or more hidden layers with each layer containing a large amount of units is called a Deep Neural Network which has spawned a new field of learning called Deep Learning. The NN shown in the picture is one such example.

IV. Deep Neural Net

A. What happens when a Neural Network Learns?

The most common way to teach an NN to generalize to a problem is to use Gradient Descent. Since I have already written an elaborate article on this topic which you can read to fully understand GD(Gradient Descent), I will not be explaining GD in this article. Here's the GD article: Gradient Descent: All YOU Need to Know.

Coupled with GD another common way to teach an NN is to use Back-Propagation. Using this, the error at the output layer of the NN is propagated backwards using the chain rule from calculus. This for a beginner can be very

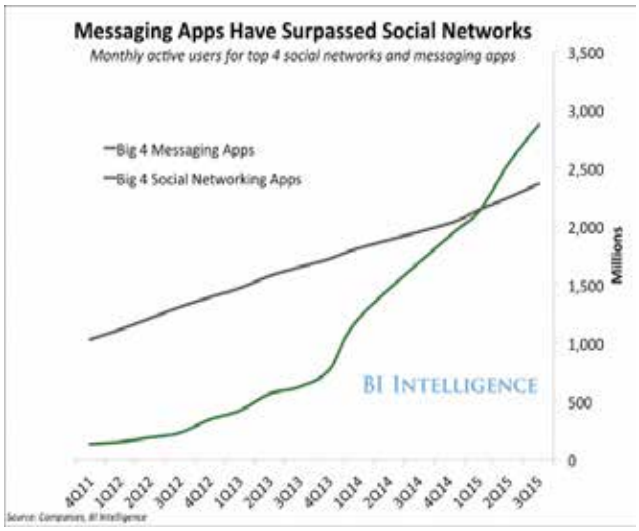

Fig. 2: Surpassed Social Networks

Challenging to understand without a good grasp on calculus so doesn't get overwhelmed by it. Click here to view an article that really helped me when I was struggling with Back-Propagation. It took me over a day and a half to figure out what was going on when the errors were being propagated backwards.

There are many different caveats[7] in training an NN. However, going over them in an article meant for beginners would be highly tedious and unnecessarily overwhelming for the beginners.

B. Implementation Details (How Everything is Manged in a Project)

To explain how everything is managed in a project, I have created a JupyterNotebook containing a small NN which learns the XOR logic gate. Click here to view the notebook.

After viewing and understand what is happening in the notebook, you should have a general idea of how a basic NN is constructed. The training data in the NN created in the notebook is arranged in a matrix. This is how data is generally arranged in. The dimensions of the matrices shown in different projects might vary.

Usually with large amounts of data, the data gets split into 2 categories: the training data(60%) and the test data(40%). The NN then trains on the training data and then tests its accuracy on the test data.

V. Deep Learning Applications

There have been a couple of tests that demonstrate the suitability of significant learning procedures in a grouping of use spaces. Despite handwriting challenge, there are applications to defend the area, the affirmation of the conversation and the recognizable evidence, the affirmation of the general disagreement, the typical vernacular preparation and the application of self-government. A case If Amazon decides to put a bot, there will be no need for we

to browse the website. Instead we can message the bot about wer requirement and it will give we all the information and options about the item. This will also mirror the behavior that we get when we visit a retailer store.

Couple of other examples:

A chatbot is an affiliation, made by rules and by all the academic capacity made by man, with which it is connected by strategies for a conversation interface. The affiliation could be different, from utilitarian to fun, and could live in any thing of affirmed conversation (Facebook Messenger, Slack, Telegram, texts, etc.).

To imagine a Chabot we take an event from a site:

We should take Amazon, for example, Amazon is a segment of web shopping that allows customers to handle things on the web and transmit them particularly to wer home merges continuous work in small coding where the characteristic high dimensionality[7]of the data is reduced by using the filled-in identification theory allowing a careful representation of the signs with small amounts of preface vectors. Another delineation is the semi-supervised complex learning [8] where the dimensionality of the data is diminished by evaluating the similarity between the preparation data tests, when visualizing these proximity estimates to reduce the dimensional spaces. In addition, driving inspiration and frameworks can be found from transformative programming approaches where hypothetically adaptable learning and basic changes of focus can be academics with unimportant tracing attempts. A segment of the media request that requires incitement thinking includes While meaningful learning has been appropriately associated with the test configuration construction assignments, the objective of the field is far from the specific applications of the missions. This grade can make the examination of several methodologies dynamically confusing and presumably will require a synergistic effort on the part of the research system to address it. Similarly, it must be seen that by paying little attention to the titanic perspective offered by meaningful learning progressions, some specific messages from the space can not be directly improved by those plans. A representation is perceiving and scrutinizing the numbers they direct at the base of bank checks.

A. What is a Chatbot?

A chatbot is an organization, created by rules and by all the intellectual competence created by man, with which it is associated by methods for a discussion interface. The organization could be a number of things, from utilitarian to fun, and could live in anything of genuine conversation (Facebook Messenger, Slack, Telegram, text messages, etc.).

As necessary, chatbots are a worthy experience. It is perhaps an open business entry for anyone who wants to jump without[4] thinking and create something that people require.

A chatbot is updated over time. An untrained instance of chatbot begins without data on how to confer. Each time a client enters a statement, the library stores the substance[10] they entered and the substance to which the declaration was referred. As chatbot gets more data, it increases the number of responses it can respond and the accuracy of each response in association with the explanation of the information. The program chooses the closest organizational response by searching for the closest Known clarification of planning that matches the data, and then restores the answer without question to that statement in perspective of how frequently conceivable each response is issued by the all-inclusive community with which the bot speaks.

B. Limited Capability of a Chatbot

The possibility of speech is critical for a chatbot. A Chabot can and jokes with a human anyway as determined as now their sKills are bound. Saying this does not imply that, in any case, that extraordinarily restricts the ways in which the substance or speech based on the voice can be as extraordinary or better than talking to a certified human being. Chatbots[4][6] can have biases on human administrators. They are available throughout the day, consistently and come very close to a wide screen of information and convenience. In addition, they can outperform individuals to the extent of speed and accuracy in a limited space. The problem in any case is to ensure that final consumers are aware of these obstructions.

While chatbots can exchange people for particular messages, they can be used in the same way to develop what human managers can offer their clients. The chatbot can, for example, give prescribed answers to the human administrator or bring the appropriate information in a useful way that the human teacher could have the ability to catch up. How chatbots are used clearly in a correspondence channel infers that participation between the bot and a human administrator is much less difficult to achieve. This is another way that chatbots are isolated from applications.

VI. Chatbot Using Wat Son Assist

A. Influence a Conversation to Benefit Before it use the IBM Wat Son Conversation

API'S, we must influence a conversation to make a profit in the IBM Bluemix organization and get sign-in capabilities for it. To do all that is considered, log in at the comfort of Bluemix, investigate Services[9]- Wat son and press the Create Wat son Advantage option select Conversation of the

summary of open organizations. In the outline that is indicated immediately, write a reasonable name for the organization and press Create get..

B. Make a Conversation Workstation

A Conversation organization can work only if we have a Conversation workstation related to it. To the extent that

it is noticed, we can consider a work space as a social event of indicators and intriguing purposes of the plan, which portrays the limits and character of the conversational UI.

The comfort of Bluemix[2] has an easy-to-use instrument that allows us to create and monitor work spaces. To send it, press the Get launch button. On the attached screen, press Create catch to create another workstation. In the business that flies, give a great name to the workstation and choose a language for it.

VII. Conclusion and Future Work

A Wat son chatbot which shows us and performs the tasks like “on headlamps” or “Turn on wipers” has been created. A user may input commands while driving through voice assist easily without any distraction from the road and the bot will perform those tasks for him. When designed right these digital assistants are able to send texts, images, audios, links and messages in real time as responses to queries from users.

Depending on the use case of the chatbots, they are often used to find information about the services a company[1] provides, consult their products availability, make reservations or bookings, and even assess the customer’s experience with the service the company has provided. These conversational agents have proposed a new scenario for customer service, pushing the bar to better services where costumers are cared for and assisted on a 24/7 basis. In a world where mobile and digital are first and musts it makes little sense for companies looking to succeed to no harness the benefits of deploying chatbots as a key tool to strengthen their relationships with their clients booth internally and externally.

References

[1] Rafael E. Banchs and Haizhou Li. IRIS: a chat-oriented dialogue system based on the vector space model. In The 50th Annual Meeting of the Association for Computational Linguistics, Proceedings of the System Demonstrations, July 10, 2012, Jeju Island, Korea, pages 37–42, 2012.

[2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, pages 6000–6010, 2017.

[3] Oriol Vinyals and Quoc V. Le. A neural conversational model. CoRR, abs/1506.05869, 2015.

[4] I. Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In Advances in Neural Information Processing Systems 27:Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada, pages 3104–3112, 2014.

[5] Oriol Vinyals and Quoc V. Le. A neural conversational model. CoRR, abs/1506.05869, 2015.

[6] Aria, Massimo, Cuccurullo, Corrado, “bibliometrix: an R tool for comprehensive bibliometric analysis of scientific literature”, University of Naples Federico II, 2016.

[7] J. O’Shea, Z. Bandar, K. Crockett, A. Tolk, L. C.Jain, “Systems Engineering and Conversational Agents” in Intelligence-Based Systems Engineering, Berlin Heidelberg: Springer, vol. 10, pp. 201-232, 2011.

[8] Kumar, Praveen, Tanupriya Choudhury, Seema Rawat, and Shobhna Jayaraman. “Analysis of various machine learning algorithms for enhanced opinion mining using twitter data streams.” In Micro-Electronics and Telecommunication Engineering (ICMETE), 2016 International Conference on, pp. 265-270. IEEE, 2016.

[9] Rawat, S., Gupta, P. and Kumar, P., 2014, November. Digital life assistant using automated speech recognition. In Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), 2014 Innovative Applications of (pp. 43-47). IEEE.

[10] Yadav, S., Haq, A.U., Rawat, S. and Kumar, P., 2018. Analysis and Implementation of Business Intelligence Software for Report Bursting. In Smart Computing and Informatics (pp. 461-473). Springer, Singapore.