Tainá Arruda de Aguilar
Aguilartaina7@gmail.com 08/08/2020

Dive into Deep Learning – Study Group

Deep Learning has shifted the programming paradigm. Before we had to know the rules of what we wanted to code or even to have a domain expertise that allowed us to manually distort our representation (aka feature engineering) in order to apply some traditional machine learning model. However, deep learning allows us to do this with even less domain knowledge by just feeding labeled data into a random initialized model and then letting the model adapt itself so that we get results that are more accurate, which can be refereed as "learning from experience". Even though seems easy or somewhat magical, the whole idea comes from comparing the predicted output with the ground truth and defining what is an improvement. We do this by defining a loss function that depends on out dataset and our goal; this is an important part of the deep learning process. The task will be to minimize this loss function, so that we can optimize our model. This idea can be achieved using a simple calculus concept: the chain rule. The backpropagation algorithm, which was derived from this, guides our steps through tuning the model's parameters in order to achieve the local minimum of loss function.

By now, we focused on two key components of deep learning, loss function and optimization algorithm. We also must address data and the model by itself. Data is an obvious component of data science and one of the main drives of deep learning growth in recent years has been the amount of data that is now available and that is generated by each second. Of course, the quality of a prediction relies on the quality of the data, if the labels are correct, if it is representative enough of our population and so on… One important area in deep learning now a day is about ethics issues regarding models and the way they can reproduce social prejudices. Finally, the models in machine learning perform sequential transformations in data. The intuition is to distort the inputs in a way that they became linearly separable in the last layer. To get from our input layer to the output layer we must pass first to some hidden layer, therefore the name deep learning.

In machine learning in general, we have two kinds of problems: supervised and unsupervised learning. The former consists in labeled data and our goal is to map from input x to output y. We feed our training data (inputs and labels) into a supervised learning algorithm that outputs a model, with which we can now make label predictions of data that the model has never seem. We have a few different supervised learning tasks. Regression means that we want to predict a quantitative value, while classification is a problem in which we want to predict the category from our feature's vector. In the case of our classification, we usually output the probability of each category in order to make our classification. Even though tagging is kind of a classification problem, it has is particulars features, since we do not have to classify into particular class, but rather we can classify each input into several labels. We can also work with sequential data, when the inputs are related with each other. The intuition is to not throw away the information present in the previous inputs, but rather use this to support our prediction. Other examples of unsupervised learning are search and ranking and recommender systems. Supervised learning has been the area in which machine learning has achieved the most progress. On the other hand, supervised learning is an area that still has much to grow. The idea is that you just have a bunch of data and you must do something with it, endless possibilities! We can find clusters, important features and even try to represent the data in an arbitrary structure.

Tainá Arruda de Aguilar
Aguilartaina7@gmail.com 08/08/2020

1. I have code that estimates the amount of cars that must be delivered to a city in order to achieve full income potential in a rent a car company. I made this with a linear regression model and some previous feature engineering.

2. One problem that my father has is to define the outcome of a legal process. With access to the documents, a model might be able to predict if the process proceeds or not. In my company, we altered the price of our product by looking at sales reports and the competition's price; this might be a candidate to be solved with deep learning.

3. Data, actually, the growth in data availability can be seem as one of the main responsible factors for this success of deep learning and consequentially of artificial intelligence. Similar what to coal is to steam engines, without data, an algorithm is has no practical value. Only when it is feed with that an algorithm can start to work and generate results and as coal quality influences on the amount of work generate by a steam machine, data quality is extremely important in the quality of the model generated by an algorithm. However, can be replaced by different fuels while data is yet the only fuel that an algorithm can have.

4. We can apply end-to-end in education (think about how Coursera could progress if it found a better way to evaluate students' answers, MOOCs could reach a new level). All companies could also benefit from a pricing model that would adjust price based on real-time supply and demand. In general, any area that we have a large amount of data might benefit from an end-to-end approach.