Introduction to Deep Learning and its Applications

Theodore Austin Dyer

Abstract—A high level overview of machine learning and deep learning concept fundamentals, with examples of their applications.

Index Terms—CMPE185, Deep Learning, Machine Learning, Supervised learning, Unsupervised Learning, Reinforcement Leaning, Feature Learning, Algorithms, Models, Predictions

1 Introduction

THE purpose of this document is to provide a very high level explanation of the basic concepts behind deep learning, and the encapsulating subject of machine learning and some of their applications. Irrespective of your educational background, the reader should be able to obtain valuable information and learn something about this revolutionary technology after reading my work.

Machine learning is a buzzword that you have likely heard thrown around in the context of computer and data science, or even in household conversation if you do not have any connection to these fields. That being said: not many people, especially those without computing backgrounds, actually understand what machine learning is on any sort of surface level. I believe that this subject matter is extremely important for people to understand at a base level in order to understand the ramifications of things that deep learning and machine learning may change in the future, and how they may affect everyone's personal lives.

2 Concept Introduction

2.1 Machine vs. Deep Learning, what do they mean?

Before you can begin understand the concept of deep learning, you need at least some knowledge of the broader encapsulating topic: machine learning. Machine learning (ML) can be thought of as a branch off of artificial intelligence. Common artificial intelligence applications that you may be able to think of likely employ machine learning solutions in some of their functions. Machine learning algorithms attempt to achieve the following goal: automatic improvements in accuracy and reliability through experience with data. These solutions are able to execute tasks without being told step-by-step how to do so.

Machine learning Algorithms begin by forming some model structure based on the statistics of some (relatively) small input data set, referred to as a 'training' data set. This model is then used to make inferences or predictions about any new data and that might be passed into the model, automatically making decisions based on the predetermined rule set that has been developed within the model. There are many ways in which a machine learning algorithm can approach developing its model, some examples of these employed in both machine learning and more specifically deep learning are: supervised learning, semi-supervised learning, or unsupervised learning. All of these different types of learning can then further be applied to different types of models, which will be explained later.

2.2 Different Learning Algorithms

Machine learning can be employed to solve many different types of tasks: from making predictions about stock prices to learning how to beat the snake game. As you might imagine, the data that we need to pass to an algorithm, and the data that we can interpret from output are completely different for each of these types of tasks. This creates the need for a number of different types of learning algorithms.

2.2.1 Supervised Learning

Algorithms utilizing supervised learning build their models based on data that includes a set of inputs and a set of outputs that these inputs should produce. The goal here is to train the model with the correct solutions in order for it to predict the correct solutions for any future input data that it does not have the specified output for. Imagine studying for an exam by reading an answer key to a previous exam: you are given questions (input) and answers (correct output) with the hope that after learning from this data, you will be able to predict the correct answers for future questions on your own exam (unknown input). Supervised learning is often used for problems such as predicting some output where you already know all of the possible outputs: regression - where some input falls along an already known range, or classification - determining which already known category an input falls into.

2.2.2 Unsupervised Learning

The approach of unsupervised learning differs from supervised in that the training data set does not include a predetermined set of outputs that the inputs should map to. Instead of learning by memorizing what 'should' be done, unsupervised approaches have the goal of learning how to interpret groupings or relationships within the input data. An example employment of an unsupervised learning approach would be that the model splits each of the data inputs in the training set into a 'cluster' or group based on some similar feature that they all include. If a new data point that is passed in shares the same data feature as others in a cluster, you would know where to group it among the entire data set.

2.2.3 Semi-supervised Learning

The semi-supervised approach is simply a combination of both of the above approaches, and is used when neither of the other can be used exclusively. The advantages to this approach include flexibility of input data sets, you do not need everything to be labeled but it is helpful when it does happen to be labeled.

2.2.4 Feature Learning

The learning approach utilized by deep learning is referred to as representation learning, or 'feature' learning. This learning approach aims to redefine or re-represent input data in a way that provides more useful information in a different format, while preserving all of the original data.

2.3 Machine Learning Models

Above it was briefly mentioned that 'models' are trained based on some input data set, now we will begin to explore how these models are structured.

2.3.1 Neural Networks

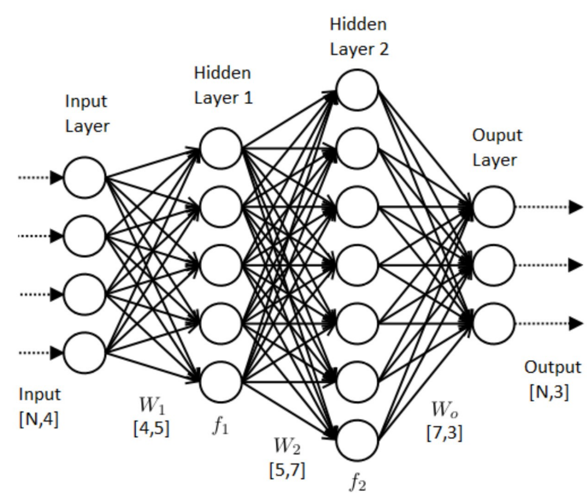

Fig. 1. Neural Network Example

Often referred to as Artificial Neural Networks (ANNs), these models are meant to in a way reproduce the way that a brain operates when processing data as it travels through neurons. ANNs map inputs to possible outputs through a series of hidden 'nodes' illustrated above by the circles in between the input and output. The connections between these nodes represent ways in which the logic of processing the data may flow, certain inputs provide intermediate results, which will ultimately lead the model to some final output.

2.3.2 Decision Trees

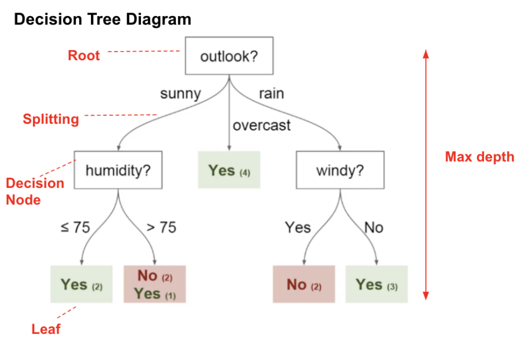

Fig. 2. Decision Tree Example

The decision tree model comes to conclusions about data field observations by following a set of rules determined from its training. This results in a flow of logic down the tree and to a single result or "leaf" (Bottom row for Fig. 2).

Suppose the question to be answered in this circumstance is, "Should I go on a run today?" This decision is then determined by which bottom node of the tree that is produced after following the tree's rules starting from the top. For example: it is rainy and windy, the question would have the answer of no, you should not go on a run today - but if it is sunny with a humidity of less than or equal to 75, you should.

3 Deep Learning Introduction

Deep learning exists as a branch of machine learning algorithms that utilize a combination of the learning methods and models described above.

3.1 How is Deep Learning Different?

Now that you have an understanding of the above terminology, we can begin to understand where deep learning fits into picture and how it can be used to solve some of today's pressing problems.

3.1.1 Deep Learning Composition

Deep learning combines a number of the subtopics detailed above. These algorithms utilize what is referred to as a deep neural network (DNN) which is essentially an ANN as described above, except it contains many intermediate layers of hidden nodes in between input and output - or in other words, multiple layers of abstraction. This abstraction allows for analysis and decision making at many different levels and based on intermediate stages of data representation. Additionally, the deep learning approach uses representation/feature learning that can be composed of a combination of supervised or unsupervised learning (Quiz: what is this combination called?).

4 Applications of Deep Learning

Now that you understand on a fundamental level how the technology works, the following practical application examples will hopefully begin to illustrate just how large the impact of deep learning techniques will be on the world moving forward.

4.1 Deep Learning For Healthcare

We are seeking to solve the problem of providing the correct medical treatment to the right patient at the right time. Due to the fact that healthcare professionals now have the technology at their disposal to track patient data including molecular traits, predispositions to certain medication, etc. there is now a large amount of biomedical data available for use in research. Without some more sophisticated technology or methods it is extremely difficult to detect relationships within this vast amount of data. Previous attempts to analyze and compile this data have taken the approach of linking data sources to form larger pools and the use of some application of predictive tools based on machine learning techniques, however machine learning up until this point has still not had the revolutionary impact that it may some day achieve. This situation has set the stage for a new methodology to step in and make a larger change within the field.

4.1.1 Challenges

Now that the baseline goal has been determined: to discover more meaningful relationships within the available data to achieve advances in medical treatment (application and recognition), it is important to understand a few of the challenges that our current models are facing.

These include: Data that is available is not uniform, many different sources of data (although they may be referring to the same subject) use different terminologies. This results in the data being much more complex to compile. Next, there is a scarcity for data related to certain subcategories of healthcare, for example data related to rarer diseases. Lastly, the sheer volume of data can be hard to sort through without a more sophisticated approach, and it is too difficult to perform fully supervised learning (a healthcare professional specifying the phenotypes to use) due to scaling issues.

4.1.2 Machine and Deep Learning

You might be asking yourself: what is the solution to the aforementioned challenges and issues in the space of analyzing and drawing conclusions from biomedical data? The answer to that question is constantly evolving and changing, but deep learning has provided a noteworthy advancement.

The traditional machine learning approach answers the question of: how do we learn relationships from the data without the need to define it prior? This applies directly to the issue at hand because biomedical data can be difficult to interpret given variances and noise in the data, thus being able to keep analysis flexible without the need for in depth understanding about the rules the data follows can be extremely beneficial. This is achieved via a four step process: "data harmonization, representation learning, model fitting and evaluation," (DLFH p. 1237).

Deep learning veers from the technique of base machine learning in the way that is learns representations from the underlying data. Traditional artificial neural network (ANN) approaches typically include three hidden processing layers with supervised representations optimized for a specific task. Deep learning on the other hand produces a data representation at each layer, and layers are not designed by human engineers, they are instead learned (derived) from the data.

4.1.3 More Medical Applications

Deep learning is still relatively new in regards to its applications and exploration of use in the medical field, but the following are a few examples in which it has been useful.

Clinical Imaging: Deep learning has been utilized in image processing of MRI scans in order to predict/detect diseases (Alzheimer, MS, etc.) and has been proven stable over larger data sets.

Electronic Health Records: Deep learning has been applied for the analysis/processing of collections of electronic health records containing diagnoses, tests, notes, etc. and has been used in some cases to predict diseases from this data and patient statuses.

4.2 Predicting Flash Floods

Deep Learning has now been used as a new method of predicting a location's susceptibility to flash floods, an application which would not be immediately obvious to most. This illustrates the technology's flexibility in being able to solve a wide range of problems.

4.2.1 Resources

This experiment was made possible via an input set of data including elevation, slope, curvature, aspect, stream density, NDVI, soil type, lithology, and rainfall used for training. The neural network structure in use is comprised of 192 neurons in 3 hidden layers, and creates an inference model for predicting flash flood danger level.

4.2.2 Algorithm Success

The DLNN used for this prediction achieved the following and is classified as a 'good' predictor: 92.05 Classification Accuracy Rate, 94.55

Positive Predictive Value, 89.55 Negative Predictive Value.

The DLNN was benchmarked against two other models: Multilayer Perceptron Network, Support Vector Machine. The result was that the trained DLNN performed much better as a predictor of flash flood susceptibility.

4.2.3 What Does This Suggest?

It may be hard to imagine the impact of this technology as not all areas of the world are susceptible to flash floods, however for those that are, this research could prove extremely beneficial if applied further. Further analysis could result in the avoidance of structural damage, agricultural damage for those planning to utilize areas that they do not realize are high risk, et cetera.

4.2.4 Who Benefits?

This technology could see use across a large number of different industries and has an extremely high potential if it were applied further. Governments may be interested in using it to ensure the safety of the general public in issuing warnings, individuals and businesses who rely on the use of natural resources such as farm land would also be extremely interested in knowing whether or not their resources are going to be compromised as a result of a flash flood.

4.2.5 Conclusion

This example serves to illustrate one of the many ways in which deep learning has applications far beyond what is immediately obvious.

4.3 Breast Cancer Risk Prediction

There currently exist a few very rudimentary models for predicting breast cancer risk based on data from radiologists. Additionally there are a large number of factors at play in complicating advances in the field, and deep learning may be a key to more accurate predictions moving forward.

4.3.1 Resources

A model used for predicting breast cancer risk was built on and trained with data from approximately 90 thousand screening mammograms over the period of January 2009 to December 2012. This data was then filtered and categorized to be used as 72 Thousand for training, 9 Thousand for validation, and 9 Thousand for tests.

4.3.2 Success

The model was designed with the goal in mind of predicting risk within the next five years of a patient after receiving the exam, and achieved the following success statistics:

Where: RF-LR refers to risk factor based logistic regression DL refers to Deep Learning – the upcoming model. TC refers to Tyrer-Cuzick model – the existing model. And performance is measured using area under the receiver operating characteristic curve (AUC).

Results: The following test set included 3937 women, aged 56.20 years 6 10.04

Hybrid DL and image-only DL showed AUCs of 0.70 (95 percent confidence interval [CI]: 0.66, 0.75) and 0.68 (95 percent CI: 0.64, 0.73), respectively.

RF-LR and TC showed AUCs of 0.67 (95 percent CI: 0.62, 0.72) and 0.62 (95 percent CI: 0.57, 0.66), respectively.

Hybrid DL showed a significantly higher AUC (0.70) than TC (0.62; P ¡ .001) and RF-LR (0.67; P =.01).

The conclusion to be drawn from this data is that deep learning approach models yield substantially better risk analysis compared to the Tyrer-Cuzick model that is currently in place, and more research and application could be beneficial for future risk predictions.

4.3.3 Conclusions

With diseases such as cancer where early recognition and early treatment and action are so critical, deep learning advancements and applications such as these in review have the potential to make sweeping changes in preventing the number of deaths that occur globally from breast cancer. With the ability to determine risk before risk is already high, especially without an extensive and expensive test, I would imagine that many deaths could be prevented.

5 Conclusion

Now that you have some idea of how powerful deep learning can be when applied in the right environments, hopefully you have some idea about things that may change in your life moving forward with machine and deep learning applications evolving even further. These concepts could prove entirely revolutionary for a wide range of fields including industry and production automation, medical data processing, the limits here are truly incomprehensible. This may ultimately prove to have a large negative impact on many people's lives as further automation can threaten a large portion of the job market, so it is hard to determine whether or not there will be a net positive impact on society.

References

[1] Deep learning for healthcare: review, opportunities and challenges By: Miotto, Riccardo; Wang, Fei; Wang, Shuang; et al. BRIEFINGS IN BIOINFORMATICS Volume:19 Issue:6 Pages 1236-1246 Published NOV 2018

[2] A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area By: Dieu Tien Bui; Nhat-Duc Hoang; Martinez-Alvarez, Francisco; et al. SCIENCE OF THE TOTAL ENVIRONMENT Volume: 701 Article Number: 134413 Published: JAN 20 2020

[3] A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction By: Yala, Adam; Lehman, Constance; Schuster, Tal; et al. RADIOLOGY Volume: 292 Issue: 1 Pages: 60-66 Published: JUL 2019