Deep Learning

Final Exam

Name: Suryakiran George

1. What is the difference between machine learning and deep learning? Give an example of an application where you would use machine learning rather than deep learning and why?

Machine learning and deep learning are both subsets of artificial intelligence that deal with the problem of enabling computers to learn from data.

Machine learning algorithms work by using statistical models and algorithms to analyze and learn patterns from data. These algorithms typically involve selecting the appropriate features from the data and applying a learning algorithm to create a model that can make predictions or decisions based on new data. Machine learning algorithms are generally good at handling structured data, and can be used for a wide range of tasks, such as classification, regression, clustering, and recommendation.

Deep learning, on the other hand, is a subset of machine learning that involves training deep neural networks with multiple layers to automatically learn representations of the data. These deep neural networks are designed to automatically discover hierarchical representations of data by combining multiple layers of abstraction, allowing them to extract complex features from raw data. Deep learning algorithms are particularly useful for tasks that involve unstructured data, such as images, audio, and text, and can be used for tasks such as image classification, speech recognition, natural language processing, and autonomous driving.

Here are the main differences between them:

1) **Neural networks:** One of the biggest differences between machine learning and deep learning is that deep learning uses neural networks to model and solve complex problems, while traditional machine learning relies on more basic statistical models. Neural networks have multiple layers of interconnected nodes that allow them to learn and model complex relationships in data, making them well-suited for applications like image and speech recognition.

2) **Feature engineering:** Machine learning algorithms typically require that features (or input variables) be hand-engineered by domain experts before being fed into the algorithm. In contrast, deep learning algorithms are capable of automatically learning useful features from raw data, reducing the need for manual feature engineering. So machine learning is better with structured data while deep learning is better with unstructured data.

3) **Data requirements**: Deep learning algorithms typically require much more data to train effectively than traditional machine learning algorithms. This is because neural networks have a large number of parameters that need to be tuned, and more data is needed to achieve good generalization performance.

4) **Computation requirements**: Deep learning algorithms are typically more computationally expensive than traditional machine learning algorithms, requiring more powerful hardware and longer training times.

In summary, deep learning is a more advanced and powerful form of machine learning that uses neural networks to automatically learn features and model complex relationships in data. While traditional machine learning techniques can still be effective for certain tasks, deep learning has shown great success in many applications, especially in areas such as computer vision and natural language processing.

Examples:

Machine learning algorithms like linear regression, decision trees, and support vector machines can be used to make predictions on structured data with high accuracy and interpretability. These techniques are often more efficient and simpler to implement than deep learning algorithms.

Deep learning is best suited for unstructured data like images, audio, and natural language processing, where it can extract complex patterns and features. However, when dealing with structured data, machine learning techniques can often provide a more straightforward and interpretable solution.

Interpretability and explainability are very important in applications like health care and bioinformatics like predicting a person's susceptibility to certain diseases based on Gene mapping from their DNA information. They associate a model to its cause and effect. Whereas deep learning methods have a black box problem, where it is incapable of determining the cause and effect. If a model can take the inputs, and routinely get the same outputs, the model is interpretable.

High interpretable models equate to being able to hold another party liable. Interpretability is a requirement in applications where wrong decisions may lead to physical or financial harm. And when models are predicting whether a person has cancer or while making correlations between a patient's health data, people need to be held accountable for the decision that was made. So in such scenarios machine learning models that are deterministic are a safe bet over deep learning models that lack explainability/interpretability. Interpretable Machine learning methods such as decision trees, decision rules and linear regression are suitable for such scenarios.

2. Suppose you are building a deep classifier for cancer detection. There are two types of errors your system is going to make: predict cancer when it is not there and predict no-cancer when it is there. Obviously, these two errors do not have equal cost [That is the consequences in making these two errors are vastly different]. What modifications would you make to the cost function for training such a classifier? [You can assume any cost function of your choice]

In this scenario, the cost function should reflect the unequal costs associated with the two types of errors. A commonly used cost function for classification problems is the cross-entropy loss function, which measures the dissimilarity between the predicted probability distribution and the true probability distribution of the labels.

To modify the cost function to reflect the different costs of the two types of errors, we can use a weighted version of the cross-entropy loss, that assigns different weights to each type of error. Specifically, we can assign higher weights to the less desirable type of error. For example, if predicting a patient has cancer when they do not (false positive) is more costly than predicting they do not have cancer when they do (false negative), we can assign a higher weight to the false positive term in the cost function.

Let TP, FP, TN, and FN denote True Positive, False Positive, True Negative, and False Negative, respectively. Then, a possible modified cost function can be:

cost = - (w_p * TP * log(p) + w_n * FN * log(1-p) + w_p * FP * log(1-p) + w_n * TN * log(p))

where p is the predicted probability of cancer, and w_p and w_n are the weights assigned to false positive and false negative, respectively. The weights can be determined based on the specific costs associated with each type of error. For example, if a false positive is five times as costly as a false negative, we can set w_p = 5 and w_n = 1 or vice versa.

By using this modified cost function, the classifier can be trained to minimize the total cost associated with the errors, taking into account the different costs of false positives and false negatives.

3. What are the takeaways from this course and how these might help you in your further studies/job?

Takeaways:

It's a field of great interest nowadays and comes as a useful tool in several applications. Coming from a traditional engineering background where I routinely use physics based modeling to understand and interpret the world; deep learning methods add valuable skills and understanding and help me stay updated with the latest trends and technologies in the industry.

Deep learning involves experimentation, testing, and tuning to get the best results. Helps me with problem solving and is valuable in my studies and career.

Deep learning is a rapidly evolving field, with new techniques and applications being developed all the time. Studying deep learning provides me opportunities for research and contributes to the advancement of the field.

Nowadays deep learning is attempted in every field whether it's really required or not! So helps me with critical thinking; to know whether it's a bane or a boon in a given situation.

It was an intellectually stimulating experience and helped me develop a deeper understanding of the field, the tools and resources. When I started I had very limited idea about various open source or free tools like Jupyter notebook or pytorch, various open source libraries etc. (*actually struggled a bit coming up to speed with the basic set up and stuff!*) and did not know much outside of Deep learning using MATLAB which is a proprietary tool and lacks the amount of resources and options and examples that open source nowadays has.

It was a great beginning and Thank you!