Jenish Maharjan **Week 1**

Paper: Deep Learning

Summary:

In the paper "Deep Learning", the authors LeCun, Bengio and Hinton give various aspects of deep learning – a brief history and introduction to deep learning, the learning procedure of backpropagation that enhanced revived the research in deep learning, convolutional neural networks, recurrent neural networks, application of deep learning in computer vision and language processing and the future of deep learning thus providing a comprehensive overview of the field of deep learning. Unlike conventional machine learning techniques that required a carefully engineered feature extractor specific to the task domain that would convert raw data to a representation that could be used to classify or detect patterns in the input, deep learning techniques methods of representation learning that allow automatic discovery of the internal representations from the raw inputs required for classification or detection of patterns in the input. Deep learning methods have multiple levels of representations that transform representations from each level into a more abstract level as they move up the various levels. What these features represent may differ based on the task at hand.

In a supervised learning setup like image classification, the machine is shown labelled images during training. An objective function is then computed which measures the distance between the actual and the desired outputs. The weights of the connections between units are then modified based on that function. The learning procedure tries to minimize this error using gradient descent. To make the process efficient, another method called stochastic gradient descent (SGD) is implemented where the error is computed for small subsets of the input and their gradient is averaged. Deep learning architectures employ the learning procedure of backpropagation to compute the gradient of an objective function in each level of the multilayer stack of modules that constitute the architecture. Backpropagation basically implements the chain rule of derivative to propagate errors from higher level units to lower level units. The backpropagation method is used with feed forward neural networks which learn to map a set of inputs to a set of outputs. Such feedforward networks compute a linear combination of the inputs and the weights which is then passed through a non linear function like the rectified linear unit (ReLU) or the sigmoid function.

The authors of the paper describe convolutional neural networks (CNN) as a type of "deep, feedforward network" that has achieved "practical success during the period when neural networks were out of favor". A typical CNN architecture can be understood as a "series of stages" – each stage with convolutional and pooling layers. As we move up the "stages", higher and more abstract features are extracted from the input data. The role of the convolutional layer is to detect local features from the previous layer and the role of the pooling layer is to merge the results of the convolutional layers based on semantic similarity. With the development and availability of powerful hardware and software, training deeper architectures with billions of parameters became possible thus making them great for domains like computer vision and natural language processing. Deep learning architectures also have an advantage over classical machine learning

methods in such domains because they are able to learn distributed representations that makes predicting the desired outputs easier.

Recurrent neural network (RNN) is another type of neural network that "process an input sequence one element at a time" while maintaining a 'state vector' in their hidden units that store information about the history of the sequence. With developments in the architecture and training techniques for RNNs, their performance in predicting characters or words in a sequence increased. They're also used in more complex tasks such as analyzing sentiments from a text or language translation. After studies showed that it is difficult to learn while storing information for very long, another kind of networks called long short-term memory (LSTM) networks were introduced with special hidden units to support storing inputs for a long time which have outperformed conventional RNNs. Many other networks such as the Neural Turing Machine have been proposed for augmenting RNNs with a memory module which have shown increased performances.

The authors conclude that although supervised deep learning methods have had great successes, unsupervised learning has had effects on deep learning of higher magnitudes. The effects are justifiable because most learning in humans and animals is unsupervised. There have been systems that implement a combination of different learning techniques to achieve better results. The authors predict that major improvements in the field of artificial intelligence will be driven by systems that merge representation learning and complex reasoning.

Strengths:

- The content is well explained using clear and simple language. They have not used unnecessary jargons.
- The paper does a great job of providing a comprehensive overview about the field of deep learning. This paper can definitely be good material for people who're new to this field and want to learn about it.

Weaknesses:

- The explanatory nature of the paper sometimes makes it verbose and it becomes difficult to narrow down the gist of the content. I found it difficult to understand how the authors related supervised learning to deep learning in the first few reads.
- Restructuring the content would definitely help increase the readability of the paper. For instance, the section "Distributed representation and language processing" – although it talks about both these topics very clearly – could have been presented as applications of CNNs, and it could have then included the discussion on computer vision applications as well.

Confusions:

- I'm confused about what "distributed representations" actually mean. Is it different from "internal representations" that are extracted by the hidden units?
- Are conventional RNNs still used? Or only augmented RNNs in use?

Discussion Questions:

- CNNs are usually used for image processing and they're known for it. What are some other domains that CNNs could be used for and how do you know that a certain problem could benefit from CNNs? (Same questions with LSTMs and RNNs)
- How has unsupervised learning been "catalytic" in the field of deep learning?
- How are deep learning and reasoning being used to drive progress in artificial intelligence? What are the most recent breakthroughs in this area of research?