Course : CS 598 Deep Learning for Healthcare

Name: Himangshu Das

Email: hdas4@illinois.edu

**General Questions**

1. **Please give a brief summary of the chapter?**

The challenge with conventional deep neural networks (DNNs) in modeling sequences arises from the growing input size over time, rendering traditional architectures incompatible. One proposed solution involves aggregating data over preceding weeks, but this may overlook weekly progression trends. To address this, Recurrent Neural Networks (RNNs) are employed, allowing the model to retain memory of prior results and seamlessly integrate new data for improved decision-making. RNNs account for present observations and persistent memory to generate predictions iteratively over sequential data inputs, ensuring continuity and adaptability in modeling dynamic sequences, such as weekly blood pressure observations in patients.

**Long Short-Term Memory (LSTM)**: LSTMs are a type of RNN that addresses the vanishing gradient problem, which is common in traditional RNNs and affects the model's ability to learn long-term dependencies. LSTMs use a memory cell with gates to regulate the flow of information, allowing them to capture long-term dependencies in sequential data

**Gated Recurrent Units (GRUs):** GRUs are another variation of RNNs that address the vanishing gradient problem while being computationally more efficient compared to LSTMs. GRUs use gating mechanisms to control the flow of information within the network, enabling them to capture dependencies in sequential data.

**Diverse RNN settings:** Compared to DNNs, RNNs offer more flexibility in handling various sequential dependencies. The key difference between type 1 and type 2 many-to-many relations lies in the length of input and output sequences.

**Training: back-propagation through time (BPTT):** In lengthy sequences, gradients can either explode or vanish over long distances. To mitigate this, researchers use truncated BPTT, partitioning long sequences into smaller segments for parallel gradient calculation. While this approach simplifies training, it may hinder the model's ability to capture long dependencies.

1. **What improvements do you want to see in this chapter? Please elaborate on them (50 Points)**

I would have loved to see the Heart Failure Dataset problem discussed in the slides instead of the textbook. Including real-life examples of healthcare applications in the weekly slides could have enhanced engagement significantly. For instance, presenting case studies that demonstrate the application of regression and cluster models and showcase their respective results would have been impactful.

In last week's weekly reflection submissions, I received 4 points for the section regarding "***What improvements do you want to see in this chapter.***" I feel this evaluation is unjust. It seems unreasonable to allocate such a significant portion of points to a section that may not necessarily warrant a lengthy response. Given the overall word limit of 200 and the weightage of 100 points for the entire reflection document, allocating 50 points solely to this section feels disproportionate. Are we supposed to write 100 words about suggesting improvements only? I suggest either revisiting the points distribution for this section or providing clearer guidelines to ensure fairness in evaluation.

1. **What are the typos in this chapter? (20 Points)**

I was not able to find any typo.

1. **Which part of the chapter do you like most? (10 Points)**

I like clustering topics. Clustering involves partitioning data points into homogeneous groups. K-means as one of the most popular clustering algorithms. With K-means algorithm being a popular method to minimize the distance between data points and their assigned cluster centers iteratively until convergence. For performance metrics in assessing the quality of the models. Accuracy is the ratio between the sum of true positives and true negatives. Another popular metric is called F1 score, which combines true positive rate and positive predictive value by taking the harmonic mean between them or more precisely, two times positive predictive value at times true positive rate divided by the sum of positive predictive value and true positive rate

1. **What are the most useful things you learned from this chapter? (10 Points)**

Machine learning basics covered in the chapter was very important- 1. Supervised learning, 2. Unsupervised learning, and 3. Evaluation metrics for machine learning.

We first started with how to build a predictive modeling pipeline. There are 6 different steps of Predictive Modelling Pipeline -

1. First, it will define a **prediction target**
2. Second we **construct a cohort** of relevant patients in the cohort construction step.
3. Third we define all the potentially relevant features in **feature construction**.
4. Fourth, we select which features are actually relevant to predicting the target variable in the **feature selection** step.
5. Fifth, we compute the **predictive model**, which can be a classification model or a regression model.
6. Sixth, we **evaluate the predictive model performance.**