**Recurrent Neural Network Architectures for Time Series Prediction of Influenza Infections**

Mason Bursac

**Abstract:**

**Introduction:**

2023 spawned the term “tripledemic” in reference to the combination of RSV, Influenza, and COVID-19 infections causing abnormal stresses on the health care system. Facing these hard conditions, it would be advantageous for health officials to know the extent of infections prior to them happening. This research focuses on the implementation of Recurrent Neural Networks (RNN) which will train on previous flu seasons to predict 4 weeks ahead. Additionally, the hyperparameters of these networks which be changed to understand what is most effective in the this application.

The architectures to be tested will include a vanilla RNN, an LSTM, and a GRU on weekly data. The hyperparameters to be changed will include the activation function for the vanilla RNN, as well as the number of recurrent layers and number of hidden features. Some things to remain constant between all trials include input sequence length = 10, number of future predictions = 4, loss function = MSE, optimizer = Adam, learning rate = .001, and number of epochs = \_\_\_\_\_.

Some things that I will look for include the effect of gradient vanishing/exploding for the vanilla RNN. The input sequence length being equal to 10 causes me to believe this will not be an issue because the length must normally be much higher to begin to see this problem arise. Additionally, I predict there will be no noticeable improvement in the LSTM due to long term memory as new infections are much more influenced by more recent levels. Lastly, the GRU was made as a solution to gradient vanishing/exploding, but as stated before, I see this as being a non issue which makes me believe there will be minimal benefits.

As the predictions approach week 4, I believe that the LSTM model will perform the best because the long term memory will have the predictions be more influenced by the input and less by predictions of predictions.

**Technologies and Methods:**

The data used will be from the CDC FluView WHO NREVSS**.** This data begins at the end of 1997 and continues until week 39 of 2015. This is weekly data on the number of total influenza infections as reported to the CDC. The models are implemented using python notebooks and the PyTorch torch.nn and associated packages.

Following the order of the python notebook, after import all necessary python packages, a torch dataset class is defined which will read the data from the CSV into a tensor. This tensor is then sliced into the feature which includes infection data in the range [t-9,t]. The output is sliced to take on the range of data [t-8,t+4] where t+1 is the first week of predicted values. The dataset is then initialized on from the beginning of the data set up to and including 2010. The test data will be from 2011 until 2015 when the data ends. Additionally, all of the data points are normalized to allow for more efficient training of the models.

Next, the model class is defined. It utilizes the “cell” versions of each respective model and a for loop to iterate through each time index and update the hidden and prediction states. Each of the predictions [t-9,t+4] is captured and compared to the actual data and used to update the model parameters.

In addition to using training loss as a method to compare models, graphs of the data were implemented in order to visualize the difference between truth and prediction on the test data.

**Graphs and Figures:**

RNN GIF 1-4 weeks in the future:

numLayers = 1

hiddenSize = 32

GRU GIF 1-4 weeks in the future:

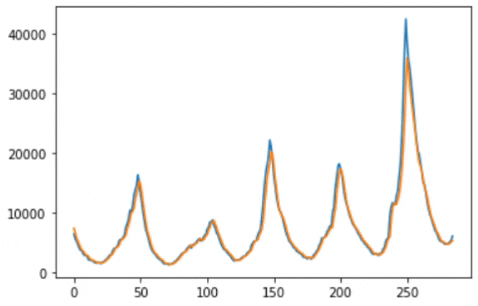
numLayers = 1

hiddenSize = 32

LSTM GIF 1-4 weeks in the future:

numLayers = 1

hiddenSize = 32



**Problems:**

**NNAR** learning the mean of the data because the lr was too high.

RNN would train slow and LSTM would only produce one value or only shape to the lower values of the data. Data normalization solved both of these.

used cell instead of full because needed the middle predictions to help train and predict further in the future

Predictions are much better with lower peak years as this is what was more common in the past.