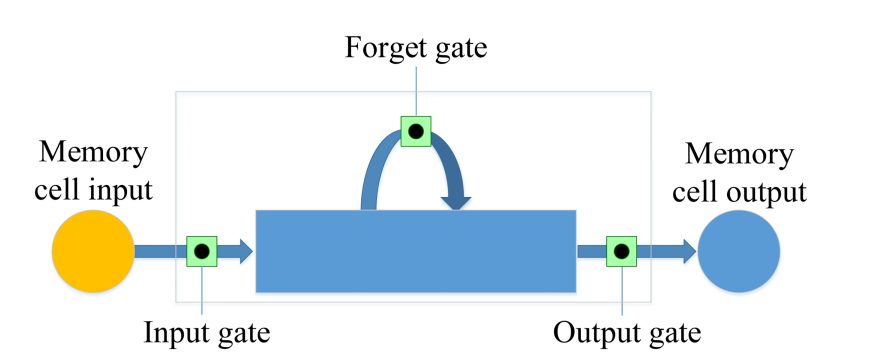
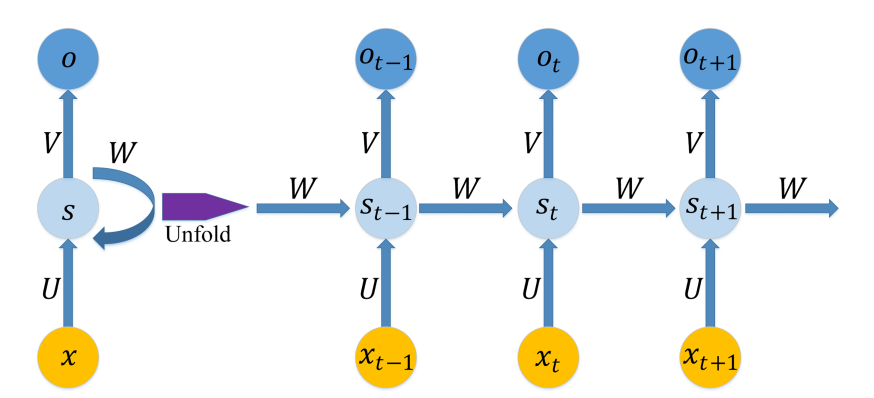
In Question 3, we adopted LSTM model, with the data processed from Wavelet Transform and Stacked Autoencoder. LSTM is employed as a key component of the model, primarily used for predicting stock prices based on high-level features generated by the SAEs. It's a type of recurrent neural network (RNN) that excels in learning from experiences to predict time series data, particularly adept at handling time steps of arbitrary sizes. The LSTM, in this model, contributes to improving the predictive accuracy by retaining and processing time-related information, addressing issues like vanishing gradients which are common in traditional RNNs. This inclusion of LSTM in the proposed model helps achieve better performance in terms of predictive accuracy and profitability, as demonstrated through tests on various individual stocks.

**Below is the basic structure of LSTM:**



**Advantages of using LSTM model rather than Decision Trees, etc. in Time-Series forecast:**

1. Handling Time-series Data: They are specifically designed for time-series forecasting, allowing them to inherently manage sequential data more effectively than decision trees.
2. Avoiding Overfitting: LSTMs can generalize better to unseen data, reducing the risk of overfitting that is often a concern with decision trees, especially in complex or noisy datasets.
3. Dynamic Adaptation: They can adapt to changes in the trend of time-series data, which decision trees might struggle with without retraining.
4. Managing Temporal Dynamics: LSTMs are superior in capturing temporal dynamics, which is essential for time-series analysis where the sequence and timing of data points are critical.

**How the data is processed in the final model?**

1. Process the initial data, in our final model, I used 24 features for X, the independent variables. To be specific, we used 'r\_m', 'total\_MV', 'PB', 'DivYield', 'PS', 'open', 'high', 'low', 'close', 'volume', 'turnover', 'BIAS', 'CCI', 'EMV', 'MA10', 'MA20', 'MACD', 'MTM6', 'MTM12', 'RSI', 'TRIX', 'VOSC','VRSI', 'VWAP'. By the way, the factors are normalized using StandardScaler provided by Sklearn, and the normalized factors are used in both Q1, Q2 and Q3.
2. Then, the 24 factors are the inputs for the Wavelet Transform process. After WT, I use SAE to denoise it. SAE has 3 layers, with hidden\_size = [20, 16, 8], which means that after the dimensionality reduction process, the number of higher-level features is 8, which is the input for our final LSTM model.
3. Then, I defined the LSTM model, along with the hyperparameters. Window\_size, input\_size, hidden\_size, num\_layers, etc. (You can choose whether to present these hyperparameters, should be optional) Then, the Y, dependent variable is the cumulative return of certain stocks. The batch\_size is 512 since our whole data size is relatively small, so quite a large batch can also efficiently update the neural network. The max\_epoch is set to be 30 to avoid overfitting.
4. For the train and test set, because there’re in total 1703 entries over 7 years for each stock, I set the first 80% to be the trainset, and the remaining 20% to be the testset. The graph in Wechat group is the cumulative return of several stocks in the test period, also the metrics are conducted based on the testset statistics.