CSC 215-01 Artificial Intelligence

Project 2: Time Series Forecasting using NN, LSTM and CNN

Name	Meghana Rao Kanneganti, Nikitha Mandhana
ID	302804645, 303315350
Course Title	CSC- 215, Artificial Intelligence
Assignment ID	Project 2
Due Date	March 06, 2023

Problem Statement :
Time series data is a succession of data points that measure the same item across time and are recorded in chronological order. Time is frequently the independent variable in a time series, and the purpose is generally to anticipate the future. Time series can help us understand past behavior and forecast future behavior. In this project, we use time series data to forecast stock prices.

Methodology :
Initially, the dataset comprises seven columns. They are Date, Open, High, Low, Close, Adj_Close, and Volume. We began the project with data preprocessing step. We first tested for null values and special characters and discovered that the dataset is clean. We then removed the ‘date’ column and the ‘Adj_Close’ column, which were no longer necessary. Because the value to be predicted is 'close,' we produced a copy of 'close' variable to use as the output variable. Next, we went on to the data transformation step, where we used z_score to normalize all of the input characteristics. We next translated x and y into a format suitable for use as input to neural networks, LSTM and CNN, respectively. Finally, we divided the dataset into 70% training and 30% testing.

Fully connected neural networks :
To forecast the y value, we used the function ‘to_sequences’ to prepare an input with 7 * 5 = 35 input characteristics. It is then reshaped (rows, 35).

LSTM Model :
Here, we molded the data into a 3D form to feed into an LSTM model. Each record may be thought of as a series of seven vectors, each with five dimensions.

CNN Model :
In CNN, x is represented as a 4D vector. Here its reshaped to (rows, 1, 7, 5)

The models were then trained using x train and y train data with various activation functions and optimizers. During training the model, we also employed ModelCheckpoint, Early Stopping and plotted the training and test loss using Tensorboard while dealing with TensorFlow.

Experimental Results and Analysis :

Fully Connected Neural Networks :

Number of layers and neuron count	Activation	Optimizer	Score train (RMSE)	Score test (RMSE)
1 hidden layer has been used here. First hidden layer has 64 neurons, and second layer has 32 neurons respectively. Finally, the output layer has 1 neuron.	Adam	Relu	0.963	0.877
		Sigmoid	1.111	1.062
		Tanh	0.967	0.899
	Sgd	Relu	116.580	112.513
		Sigmoid	12.603	12.547
		Tanh	18.084	17.205
2 hidden layers have been used here. First hidden layer has 64 neurons, second layer has 32 and third layer also has 32 neurons respectively. Finally, the output layer has 1 neuron.	Adam	Relu	0.944	0.872
		Sigmoid	0.950	0.872
		Tanh	0.932	0.842
	Sgd	Relu	246.054	244.430
		Sigmoid	7.464	7.534
		Tanh	23.081	21.984

LSTM :

Number of layers and neuron count	Activation	Optimizer	Score train (RMSE)	Score test (RMSE)
1 LSTM layer with 64 neuron count and 2 Dense layers with 32 neuron count each. Finally, one output layer with 1 neuron	Adam	Relu	1.251	1.185
		Sigmoid	1.327	1.323
		Tanh	1.270	1.259
	Sgd	Relu	18.241	18.384
		Sigmoid	1.529	1.541
		Tanh	6.983	7.007

1 LSTM layer with 64 neuron count and a Dense layer with 32 neurons. Finally, one output layer with 1 neuron	Adam	Relu	1.389	1.360
		Sigmoid	1.281	1.263
		Tanh	1.248	1.220
	Sgd	Relu	13.586	13.641
		Sigmoid	1.252	1.230
		Tanh	2.381	2.395
2 LSTM layers with 128 and 64 neurons count respectively and 2 Dense layers with 32 neurons each. Finally, one output layer with 1 neuron	Adam	Relu	1.587	1.546
		Sigmoid	1.336	1.328
		Tanh	1.429	1.378
	Sgd	Relu	36.246	35.018
		Sigmoid	1.486	1.492
		Tanh	36.252	35.020

CNN :

Kernel size	Activation	Optimizer	Score train (RMSE)	Score test (RMSE)
(1, 2)	Adam	Relu	1.377	1.357
		Sigmoid	2.295	2.256
		Tanh	2.225	2.273
	Sgd	Relu	33.756	33.207
		Sigmoid	2.993	2.895
		Tanh	12.813	12.281
(3, 3)	Adam	Relu	1.377	1.340
		Sigmoid	2.726	2.632
		Tanh	2.111	2.090
	Sgd	Relu	34.353	33.095
		Sigmoid	5.646	5.549
		Tanh	17.246	16.391

Task Division and Project Reflection :

Initially we both started working together on Fully connected neural networks and achieved really good RMSE values. Later we divided the tasks where in Meghana worked on the LSTM model and Nikitha took up the CNN model. We also made sure out of the 4 additional features; 2 features are worked by each one of us.

We faced challenges in order to figure out how to convert or reshape into required format such as 2D, 3D. While dealing with CNN model, we faced issues while adding different layers and understand them. We learned how to handle time series data and to give input to the models. Also, while performing the additional features, we learnt how to add different Attention and Bidirectional layers and use Transformer model.

Additional Features :

1) In the project, we have found the best N value (number of the days we should consider in the past) that yields the most accurate model:
   N Best Value = 2
2) We built a multi-regression model to predict the next 5 continuous ‘Close’ values based on past 30 days inputs using LSTM. Taking the sequence length 30, using 3 hidden layers. We obtained a **RMSE score of 0.9097.**
3) We used a transformer model with a sequence length of 2 and number of features as 1 which when compared to the other models, fully connected Neural Networks, LSTM and CNN did not give those great results.
   The best **RMSE score for Transformer model is 1.926** using adam optimizer and relu activation function.
4) We built a mode using bidirectional LSTM and Attention layers to see if we can beat our baseline NN/CNN/LSTM models, however we noticed that we got almost the same results for this model when compared to LSTM and CNN.
   The best **RMSE score for the model is 1.956** using adam optimizer and relu activation function.