SHAR MARKET Case Study:

There are three conventional approaches for stock price prediction: technical analysis,traditional time series forecasting, and machine learning method. Earlier classical

regression methods such as linear regression, polynomial regression, etc. were used topredict stock trends. Also, traditional statistical models which include exponential

smoothing, moving average, and ARIMA makes their prediction linearly. Nowadays,Support Vector Machines (Cortes & Vapnik, 1995) (SVM) and Artificial Neural

Networks (ANN) are widely used for the prediction of stock price movements. Every algorithm has its way of learning patterns and then predicting. Artificial Neural Network

(ANN) is a popular and more recent method which also incorporate technical analysis for making predictions in financial markets. ANN includes a set of threshold functions.

These functions trained on historical data after connecting each other with adaptive weights and they are used to make future predictions. (Trippi & Turban, 1992;

Walczak, 2001; Shadbolt & Taylor, 2002) (Kuan & Liu, 1995) investigated the out-of-sample forecasting ability of recurrent and feedforward neural networks based

on empirical foreign exchange rate data (Kuan & Liu, 1995). In 2017, Mehdi Khasheiand Zahra Haji Rahimi evaluated the performance of series and parallel strategies to

determine a more accurate one using ARIMA and MLP (Multilayer Perceptron)(Mehdi & Zahra, 2017).

This dataset have the coloumn fields MARKET,SERIES,SYMBOL,SECURITY,PREV_CL_PR,OPEN_PRICE,HIGH_PRICE,LOW_PRICE,CLOSE_PRICE,NET_TRDVAL,

NET_TRDQTY,CORP_IND,TRADES,HI_52_WK,LO_52_WK.we can analyse this dataset according to the prices.

Data

MARKET,SERIES,SYMBOL,SECURITY,PREV_CL_PR,OPEN_PRICE,HIGH_PRICE,LOW_PRICE,CLOSE_PRICE,NET_TRDVAL,NET_TRDQTY,CORP_IND,TRADES,HI_52_WK,LO_52_WK

N,N1,IRFC,BOND 8.00% PA TAX FREE S1, 1086.00, 1085.00, 1085.00, 1084.54, 1085.00, 3371959.44,3108, ,8, 1194.00, 1050.00

N,N1,JNPT,BOND 6.82% PA TAX FREE S1, 1001.00, 1007.10, 1015.00, 1007.10, 1015.00, 51370.00,51, ,2, 1529.99, 1000.00

N,N1,NHAI,BOND 8.20% PA TAX FREE S1, 1091.38, 1092.00, 1092.00, 1087.00, 1089.98, 3861145.49,3544, ,39, 1140.00, 1045.00

N,N1,NTPC,8.41%S-R-NCD SERIES 1A, 1198.99, 1135.00, 1135.00, 1135.00, 1135.00, 113500.00,100, ,1, 1275.00, 1022.10

N,N1,RECLTD,TAXFREE SEC NCD TR1 S1, 1083.54, 1088.00, 1088.00, 1085.00, 1085.00, 108530.00,100, ,2, 1184.80, 1041.00

N,N2,BRITANNIA,8.00% SEC RED NCD, 33.07, 33.07, 33.20, 32.80, 33.14, 677372.39,20461, ,107, 36.00, 29.90

N,N2,HUDCO,8.20 NCD05MAR27 FV 1000, 1215.00, 1222.00, 1222.00, 1222.00, 1222.00, 3666.00,3, ,2, 1311.00, 1171.00

N,N2,IIFCL,8.66% TAX FREE NCD, 1135.00, 1137.00, 1137.00, 1137.00, 1137.00, 4548.00,4, ,1, 1390.00, 1101.10

N,N2,M&MFIN,9.00% UNSECURED NCD, 1104.00, 1101.98, 1101.98, 1080.05, 1089.43, 423116.50,388, ,12, 1229.00, 1049.50

N,N2,NABARD,7.64% TAX FREE TRI SRIIB, 1238.50, 1239.01, 1239.55, 1237.01, 1238.22, 125060.45,101, ,4, 1304.80, 1181.00

N,N2,NHAI,BOND 8.30% PA TAX FREE S2, 1258.50, 1255.50, 1257.00, 1253.10, 1256.45, 464590.50,370, ,6, 1294.90, 1176.00

N,N2,NTPC,8.48%S-R-NCD SERIES 2A, 1282.00, 1282.50, 1283.00, 1282.50, 1283.00, 128275.00,100, ,2, 1854.00, 1225.00

N,N2,PCHFL,SEC RE NCD 8.35% SR.I, 1000.00, 1000.00, 1004.67, 1000.00, 1001.68, 526800.91,526, ,17, 1028.00, 981.01

N,N2,RECLTD,TAXFREE SEC NCD TR1 S2, 1207.00, 1192.00, 1192.00, 1192.00, 1192.00, 238400.00,200, ,1, 1288.00, 1160.00

N,N2,TATACAPHSG,SEC RED NCD 8.10% SR.I, 1054.21, 1055.01, 1058.00, 1055.00, 1058.00, 1105942.50,1048, ,42, 1160.00, 950.00

N,N3,BRITANNIA,UNSEC RED NCD 5.50%, 29.43, 29.40, 29.70, 29.38, 29.47, 4233217.99,143810, ,787, 30.50, 29.00

N,N3,DHANILOANS,DHANI LOANS & SERV LTD, 1008.00, 1002.00, 1003.00, 1002.00, 1003.00, 78158.00,78, ,5, 1030.00, 810.00

N,N3,HUDCO,8.10 NCD 05MAR22 FV 1000, 1054.84, 1054.84, 1055.90, 1054.52, 1054.91, 1729096.92,1639, ,19, 1320.00, 1000.00

N,N4,DHANILOANS,DHANI LOANS & SERV LTD, 1020.80, 1022.01, 1022.01, 1022.01, 1022.01, 2044.02,2, ,1, 1098.75, 851.00

N,N4,ERFLNCDI,SEC RED NCD 9.00% SR. IV, 986.00, 988.22, 988.22, 988.22, 988.22, 29646.60,30, ,1, 1050.00, 829.95

N,N4,HUDCO,7.34 NCD 16FEB23 TR1 SR1, 1088.66, 1073.01, 1073.10, 1073.01, 1073.05, 118037.80,110, ,3, 1153.50, 1005.60

PROBLEMS :

1.Query to display the number of series present in the data.(using hive)

2.Display the series present in the data.(using hive)

3.Find the sumpof all the prices in the each series.(Using hive)

4.Display security,series with highest net trade value(use pyspark)

5.Display the series whose sum of all prices greater than the net trade value.(Using pyspark)

6.Display the series with highest net trade quantity.(Using pyspark)

7. Display the highest and lowest open price(Using sql)

8.Query to display the series which have trades more than 80.(Using SQL).

9.Display the difference between the net trade value net trade quantity for each series.(Using sql).