CHAPTER-4

4. DATA ANALYSIS AND RESULTS

4.1. Introduction

The empirical research in the field of developing EWSs focuses on exploring various factors, variables and tools and techniques utilized in its construction. Research carried out in this study aims at contributing to this ongoing empirical research in the field. This has been done by exploring various macroeconomic and financial variables that can act as potential early warning indicators for predicting the probability of a banking/stock market crisis in Indian context. The work is divided into two sections. First section focuses on examining the significance of the potential indicators for an approaching banking crisis. It also explores the effectiveness of machinelearning techniques in comparison to the traditional signal extraction and limited dependent variable models. The second section focuses on examining the significance of the potential indicators for predicting the probability of a stock market crisis. It also explores the role of varioussentiment variables related to domestic, developed and aggregate emerging market sentiment in predicting the likelihood of a stock market crisis in India. This is followed by investigating and comparing the traditional approach of Logit models with the ANNs in their predictive power.

4.2. Analysis on Indian Banking Crisis

This section deals with the two major aspects related to comparison of alternate tools and techniques and the indicators which have got a direct bearing on the probability of crisis in Indian context. The crisis variable has been defined as a binary dummy variable based on an index constructed using different components. As mentioned earlier, BSF is a composite index constituting Aggregate Time Deposits, Foreign Currency Borrowing, Net Bank Reserves and

Domestic Credit as proxies for Credit risk, Liquidity risk, and Interest rate risk. The crisis variable has been defined as a binary dummy variable based on an index constructed using different components. The Figure 4-1 shows the banking sector fragility considering four components between 2001 and 2017. The components constitute a group of banking system related variables such as real deposits, real foreign currency borrowing, real credit, and real bank reserves, to assess the fragile conditions experienced by the banking system.

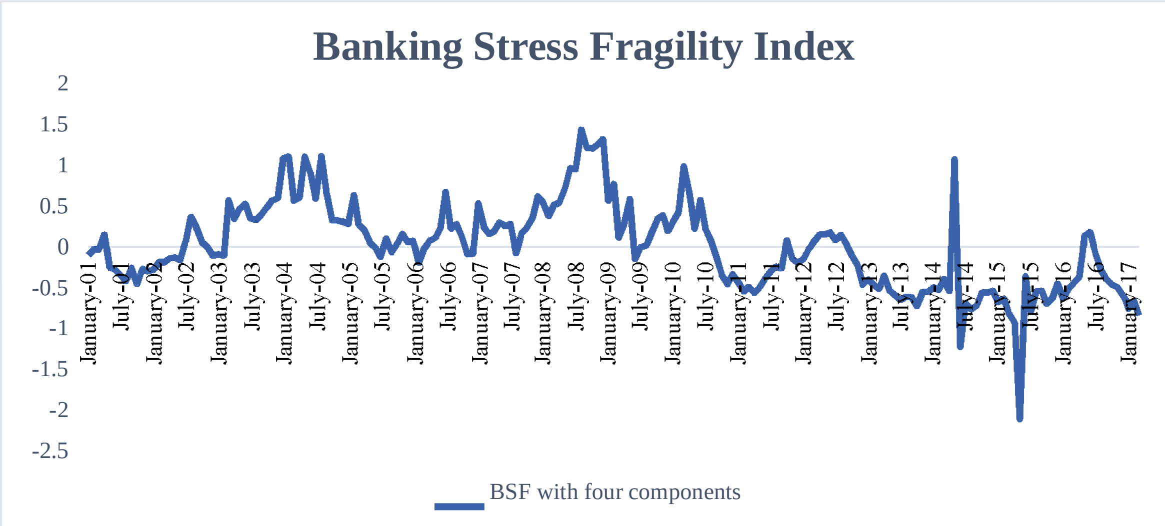
The episodes of high banking fragility can be observed from the trend followed by the BSF index. For most of the time, the index is positive showing a period of stability, while the period from 2010-2012 and from 2013 to 2017 depicts the zones of high fragility in the system. During 2008-09 sub-prime crisis, the direct impact of the crisis was relatively muted however, the economy did go through a significant slowdown as compared to the last five years' robust growth. The stress in international markets was reflected in domestic financial markets through widening credit spreads and higher liquidity crunch during Q3 of 2008 and Q1 of 2009. The global economy was in a recovery period when the sovereign debt crisis emerged in May 2010. As compared to the passive growth in 2009-2010, the economy witnessed a robust revival in credit off take during 2010-11. However, the revival in credit growth signaled caution as this could happen due to less stringent credit risk assessments and could lead to impairment of assets during the downturn of the credit cycle. On the other hand, the deposit growth slacked which resulted in continued dependence on borrowing and other short term means to fund the expansion. Thus, the liquidity risk increased owing to the increasing structural mismatch in the maturity profile of advances and the deposits. The deterioration of asset quality became a major concern for SCBs during 2013. The distress in the banking system increased marginally in the first half of 2014. SCBs' credit growth on a y-o-y basis declined significantly from 17.1% in September 2013 and 15.1% in March 2013 to 13.6% in

March 2014, while the deposit growth declined from 14.4% to 13.9%. The business of SCBs slowed significantly during 2015-16 which was reflected in decline in deposit and credit growth. Between March and September 2015, restructured standard advances ratio declined and the Gross Non-Performing Advances (GNPAs) ratio registered an increase in its value. Public sector banks (PSBs) continued to record the lowest CRAR among the bank groups and the Capital to Risk-Weighted Assets ratio (CRAR) of SCBs deteriorated as well. During the first half of 2015-16, among other bank groups, the asset quality of both Scheduled Urban Co-operative Banks (SUCBs) and non-banking financial companies (NBFCs) also registered decline. The second half of 2015-16 continued to witness decline in credit growth of all SCBs, on a year-on-year basis, as it declined from 9.7% to 9.4%. The growth in deposits also declined from 10.7% to 9.9% between March 2015 and September 2015. During first half of 2016, the subdued performance of Public Sector Banks (PSBs) resulted in single digit numbers of overall credit and deposit growth of SCBs. During March 2016, all SCBs recorded a decline in credit growth, on a year-on-year basis, from 9.4% to 8.8% in September 2015 while the growth in deposit declined from 9.9% to 8.1%.

4.2.1. Identification of banking crisis indicators -Signal Extraction approach

This section deals with the identification of early warning indicators for prediction of the banking crisis. The first technique employed for identification of warning indicators is the signal extraction approach. Thresholds for each variable has been defined relative to the percentiles of the distribution of the observations of the indicators. A grid search has been performed to find an optimal threshold by calculating NSR for a range of potential threshold values. The optimal threshold selected is the one where the NSR value for each warning indicator is minimized.

Figure 4-1:Banking Stress Fragility Index



Source: Based on Author's calculations

In case an indicator is positively related to the probability of a banking crisis, upper tail distribution (>50th percentile) threshold has been chosen. The 50h percentile is chosen because it represents the score, at or below which 50 percent of the observations fall dividing the dataset equally. On the other hand, when an indicator is negatively related to the probability of a banking crisis, the lower tail distribution threshold has been chosen. Table 4-1 presents the expected relation between the indicator and the probability of occurrence of a banking crisis. For example, a surplus in the current account balance would diminish the probability while a high current account deficit would disrupt generation of foreign exchange to finance the balance of payments deficit, thereby putting pressure on the exchange rate and hence, increasing banking sector problems. Therefore, for the

crisis, the optimal threshold value (-0.140) has been chosen at the 40th percentile of its distribution. Similarly, a sudden stop or a sharp decline in capital flows would increase the probability of a currency crisis further deepening the problems in banking sector while a higher amount of FDI would imply attractive economic policies and a low share of current account being financed by volatile capital inflows. This lowers the probability of a currency attack thereby lowering the probability of a banking crisis. Therefore, the optimal threshold value (-0.379) for the indicator "FDI_GDP" has been selected 10^{th} percentile of its distribution. The relationship between the overvaluation of Exchange rate and the probability of occurrence of a banking crisis is expected to be negative as well. This is because, an overvalued exchange rate would slow down the exports which can lead to loss of competitiveness and business failures of domestic enterprises. This would lead to increased bank loan defaults and increased imports, hence increasing the pressure on the banking sector. A positive relation is expected between the interest rates and the probability of a banking crisis as high domestic interest rates result in liquidity crunch which in turn brings a slowdown in the economy and puts pressure on the banks making them fragile. Therefore, for indicator "YTM" the optimal threshold value (7.519) has been selected at the 70^{th} percentile of its distribution. Table 4-1 also presents the calculated minimum NSR values for each of the indicator at which their respective thresholds have been selected.

Table 4-1:Optimal thresholds, NSR and Persistence of Univariate Potential Lead Indicators

Potential Variable Indicator	Optimal Threshold	Threshold Value	Relation of variable with the banking crisis	NSR	$\frac{A}{A+B}-\frac{A+C}{A+B+C+D}$
YTM	70	7.519	Positive	0.185	0.358
SPREAD	90	1.592	Positive	0.341	0.252
GGFD_GDP	90	0.230	Uncertain	0.380	0.230
M3_FEX	70	0.033	Positive	0.341	0.252
GWPI	60	0.059	Positive	0.747	0.073
GCAB_GDP	40	-0.140	Negative	1.229	-0.051
GSP	30	0.002	Negative	0.901	0.026
REER_DEV	30	-1.532	Negative	1.112	-0.026
GCDR	60	0.016	Positive	1.274	-0.060
GFDI_GDP	10	-0.379	Negative	1.062	-0.015

GOILP	60	0.139	Positive	1.543	-0.105
GIIP	10	-0.007	Negative	1.126	-0.029
GRM	60	0.061	Positive	1.182	-0.041
GSTD	60	0.272	Positive	1.498	-0.098
CMR	70	7.477	Positive	0.302	0.276

As discussed in the Chapter 3, the indicators having NSR greater than 1 are the ones which have low predictive power as they produce more false alarms than the good signals. Also, the lower the NSR value than 1, the better the indicator is at predicting the occurrence of a banking crisis. It may be observed from analysis presented in Table 4-1 that out of 15 variables, only 7 indicators namely YTM (0.185), SPREAD (0.341), GGFD_GDP (0.38), M3_FEX (0.341), CMR (0.302), GWPI (0.747) and GSP (0.901) have NSR<1. This suggests that a large number of selected indicators (8 out of 15) give rise to more false alarms. The best indicator with minimum NSR is YTM as evident from the lowest NSR value of 0.185 while the worst indicator is growth in Oil prices with NSR value of 1.543 among all the indicators. This implies that the 91day T-bills yield to maturity provides a good signal for an impending banking crisis. The Indian economy imports 70% of its oil requirements from international markets thus, making it vulnerable to any fluctuations in oil prices. However, it may be mentioned that in the present study oil prices are not effective in signaling banking crisis in Indian scenario as the NSR is greater than 1. This may be mainly because of the government's administered pricing policies that diffuse the hikes by raising subsidies. The last column of the Table 4-1 represents the difference between conditional (A/A+B) and unconditional probability (A+C/A+B+C+D) of the respective indicators. For the indicators with NSR greater than 1, the difference in the conditional and unconditional probability is negative, while for the indicators having NSR less than 1, the respective differences are positive. Thus, the indicator is useful and has good predictive power when the conditional probability is greater than the unconditional probability as this is equivalent to indicator having NSR<1.

In general, the leading indicators may not be of great use when considered individually as different variables act differently at a certain point of time. Therefore, the leading indicators found in the above analysis have been compressed into two composite indicators CI^1 and CI^2. The composite indicator namely CI^1 is based on the summation of the number of all the selected leading indicators while the second composite indicator namely CI^2 weighs more the signals issued by the indicators having high predictive power i.e. NSR<1. The index CI^2 is the weighted sum of the indicators using inverse of indicator's NSR at its respective weight. The threshold value for CI^1 is set at 75% percentile of the distribution. Therefore, signals are issued by CI^1 when the value of composite indicator crosses the threshold.

In case of CI^2, the weighted composite indicator, the highest weight is given to the indicator with the best performance i.e. minimum NSR. The persistence i.e. inverse of NSR and weights used for each of the indicators for the construction of the CI^2 are presented in Table 4-2. It is can be observed from the Table 4-2, that indicators namely YTM, SPREAD, CMR, GGFD_GDP, M3_FEX, GWPI and GSP have persistence greater than 1 as these indicators have NSR lesser than 1. Also, the highest weightage of 0.207 has been allotted to the indicator "YTM" which has the highest predictive power while the lowest weight of 0.025 has been given to the indicator "GOILP" which is the worst performing indicator in the above analysis.

Table 4-2:Weights of the indicators used in CI^2

Early Warning Indicator	Persistence	Weights
YTM	5.414	0.207
SPREAD	2.933	0.112
GGFD_GDP	2.632	0.101
M3_FEX	2.933	0.112
GWPI	1.340	0.051
GCAB_GDP	0.814	0.031
GSP	1.110	0.042
REER_DEV	0.899	0.034
GCDR	0.785	0.030
GFDI_GDP	0.942	0.036
GOILP	0.648	0.025
GIIP	0.888	0.034
GRM	0.846	0.032
GSTD	0.668	0.026
CMR	3.309	0.126

Table 4-3 compares the performance of the two composite indicators in terms of their forecasting abilities. It may be observed from Table 4-3 that the weighted composite indicator CI^2 is superior to the unweighted composite indicator CI^1 in terms of the forecasting ability. The percentage share of bad signals decreased from 20.78% to 12.37% for CI^2 while the percentage of good signals increased from 35.53% to 67.44% which is almost twice of the value for CI^1. The gain in efficiency in terms of lower NSR of CI^2 compared to CI^1 is a result of combining the univariate leading indicators and their respective NSR values. The QPS and GSB values also indicate superior performance of CI^2 as compared to CI^1. The QPS value for CI^1 is 0.554 while for CI^2, it is 0.463. In terms of predictive performance, the closer the value of QPS to 0, the better the indicator. Hence,

CI^2 outperforms CI^1 in terms of accuracy. Similarly, for GSB, measuring calibration, the value for CI^2 is 0.0017 which is closer to 0 in comparison to 0.0034 of CI^1, again confirming that CI^2 is better than the CI^1 in terms of predictive power.

Table 4-3:Comparison of Performance of CI^1 and CI^2

Composite Indicator	% share of bad signals in total signals (B/B+D)	% share of good signals in total signals A/(A+C)	NSR= $\frac{B/(B+D)}{A/(A+C)}$	A/(A+B) (%)	$\frac{A}{A+B} - \frac{A+C}{A+B+C+D}$	QPS	GSB
CI^1	20.78	35.53	0.585	62.79	13.12	0.554	0.0034
CI^2	12.37	67.44	0.183	82.86	35.86	0.463	0.0017

The conditional probabilities of a banking crisis associated with different values of the weighted and un-weighted composite indicators are presented in Table 4-4. The conditional probability of occurrence of a banking crisis has been defined as the probability of a banking crisis within 12 months under the condition that the composite index ranges between the allotted lower and upper limits.

As mentioned above, for CI^1, the 75% percentile has been set as the threshold for identifying the signals for an approaching banking crisis. On computation, the 75^{th} percentile comes out to be “8” and this has been chosen as the cut off probability. For CI^2, various upper and lower limits have been tested to find out the cut-off probability. It can be observed that the conditional probability of occurrence of a future banking crisis increases alarmingly once the value of CI^2 exceeds 0.405; midpoint of two intervals [0.21-0.40] and [0.41-0.60]. Therefore, the probability threshold of 0.263 (conditional probability of banking crisis in the interval of 0.21 – 0.40) has been chosen as the cut off probability.

Table 4-4:Distribution of conditional probabilities of banking crisis based on CI[1] and CI[2]

CI^1 intervals	Conditional Probability	CI^2 intervals	Conditional Probability
[0-2]	0	0.01-0.20	0.1667
[2-4]	0.0714	0.21-0.40	0.2631
[4-6]	0.4412	0.41-0.60	0.8491
[6-8]	0.4520	0.61-0.80	0.75
>8	0.5068		

4.2.2. Identification of banking crisis indicators -multivariate Logit model

The second approach concerns with the use of a Logit model. The identified 15 leading indicators in Chapter -3 are used as explanatory variables and the crisis dummy variable is entered as the dependent variable. Before the model is constructed, the correlation is checked for the explanatory variables to take care of multi-collinearity.

There is no evidence of strong correlation among the variables hence all the variables have been used for the analysis. The maximum correlation exists between YTM and CMR of 88%. However, most of the variables are lightly correlated as the value does not exceed 50% in most of the cases. The correlation matrix has been presented as Table 4-5.

The 15 indicators are included as explanatory variables in the Binomial Multivariate Logit regression model as there is no problem of correlation among the variables. Table 4-6 presents the results of the Logit model.

Table 4-5:Correlation Matrix of leading indicators

Variable	GSP	GRM	GCDR	M3FEX	FDIGDP	GFDGDP	CABGDP	GSTD	GOILP	SPREAD	GIIP	GWPI	YTM	CMR	REER_DEV
GSP	1	0.29	-0.04	-0.03	0.71	-0.52	0.03	0.07	0.42	0.21	0.57	0.06	-0.28	-0.34	0.22
GRM		1	0.27	-0.15	0.19	-0.18	0.06	0.14	0.27	-0.42	0.37	0.26	0.05	0.005	0.35
GCDR			1	0.06	-0.42	-0.04	-0.07	0.48	0.23	-0.18	0.13	0.19	0.13	0.12	-0.03
M3FEX				1	-0.13	0.23	0.04	-0.19	-0.18	0.11	-0.10	0.08	-0.00	-0.07	-0.03
FDIGDP					1	-0.36	0.04	-0.16	0.13	0.24	0.44	0.001	-0.26	-0.28	0.30
GFDGDP						1	0.07	-0.11	-0.28	0.06	-0.19	0.001	-0.15	-0.03	-0.51
CABGDP							1	-0.05	0.07	0.01	0.02	0.068	-0.04	-0.03	-0.06
GSTD								1	0.40	0.03	0.38	0.20	-0.19	-0.21	-0.02
GOILP									1	-0.26	0.41	0.60	-0.11	-0.15	0.23
SPREAD										1	0.13	-0.37	-0.68	-0.64	-0.33
GIIP											1	0.16	-0.49	-0.44	0.09
GWPI												1	0.12	0.13	0.07
YTM													1	0.88	0.14
CMR														1	0.03
REER															1

Table 4-6:Estimates of Multivariate Binary Logit model

	Coefficient	t-statistics
GSP	-0.0107	-0.65
GRM	0.141**	3.27
GCDR	0.0110	0.14
M3_FEX	0.0853**	3.25
FDIGDP	-0.00472	-0.44
GFDGDP	-0.0103	-0.96
CABGDP	-0.00108	-1.05
GSTD	0.101	0.24
GOILP	0.0123	0.99
SPREAD	3.691***	5.54
GIIP	-0.0789	1.00
GWPI	0.312	2.69
YTM	2.323***	3.85
CMR	0.608*	2.05
REER_DEV	-0.0298	-0.23
cons	-22.23***	5.35
Pseudo R^2	0.492	
Number of obs	194	
Hosmer- Lemeshow statistic	0.31	
Probabiltiy	(0.5785)	

(*) p<0.1 ;(**) p<0.05; (***)p <0.01

It may be observed from Table 4-6 that five out of 15 indicators are found to be significant at 95% confidence level namely growth in reserve money (GRM), growth in ratio of M3 to Foreign Exchange Reserves (M3FEX), weighted average call money rate (CMR), yield to maturity on 91 day T-bills (YTM), and the spread between bank rate and YTM (SPREAD). Growth in stock prices, growth in gross fiscal deficit, growth in FDI to GDP, growth in industrial production and growth in current account balance to GDP are found to be negatively related to rise in the probability of a banking crisis while growth in Credit-Deposit Ratio, growth in Money Supply relative to Foreign Exchange Reserves, growth in short term debt, s growth in spread between bank rate and YTM and growth in Call Money Rate are found to be positively related to the rise in the probability of a banking crisis. This indicates that an increase in the annual growth of Stock prices and Industrial Production leads to a decrease in the probability of an occurrence of a banking crisis while an increase in annual growth of C-D ratio, Money Supply relative to Foreign Exchange Reserves and Call Money rate results in increased probability of an occurrence of a banking crisis. All the indicators are found to be having expected signs. The most significant warning indicators are found to be YTM, SPREAD, M3_FEX, CMR and GRM in predicting the probability of a banking crisis.

The model has a Pseudo R^2 equal to 49.2% and the Hosmer-Lemeshow (H-L) test for goodness of fit indicates a good Logit regression model fit since the p-value comes out to be 0.5785 which is greater than 0.05. The null hypothesis for the H-L goodness of fit test is that the observed and expected proportions are the same across all population subgroups. It is basically a test which tells how well the model fits the given data. The test produces p-value, which, if low (<0.05), leadsto rejection of the estimated model and if high, leads to acceptance of the estimated model. Thus,the value of 0.5785 greater than 0.05 clearly leads to acceptance of the model.

It is necessary to estimate the power of the model and its forecasting ability to use the model for forecasting probability of occurrence of a banking crisis. The standard method is to compare the estimated probability with the actual occurrences. The study analyses different probability thresholds to test the model. The results with different probability thresholds are presented in Table 4-7.

Table 4-7:Goodness of fit at different cut-off probabilities

Probability cut off (0.2)	
% of observations correctly called	79.38%
Probability of an alarm conditional on a crisis	92.47%
Probability of a crisis following an alarm	72.27%
Probability of a tranquil period following no alarm	90.67%
Probability of tranquil alarm conditional on tranquil period	67.33%
Probability cut off (0.25)	
% of observations correctly called	81.44%
Probability of an alarm conditional on a crisis	91.40%
Probability of a crisis following an alarm	75.22%
Probability of a tranquil period following no alarm	90.12%
Probability of no alarm conditional on tranquil period	72.28%
Probability cut off (0.5)	
% of observations correctly called	87.63%
Probability of an alarm conditional on a crisis	81.72%
Probability of a crisis following an alarm	91.57%
Probability of a tranquil period following no alarm	93.07%
Probability of no alarm conditional on tranquil period	84.68%

The results presented in Table 4-7 reveal that as the cut off probability is increased from 0.2 to 0.5, the percentage of observations correctly classified rises from 79.38% to 87.63%. However, on further increasing the cut off value, the percentage of observations correctly classifies falls. Thus, the critical probability threshold is set at 0.5. The value is approximately closer to the probability threshold of 0.405 which is determined for the Weighted Composite index (CI^2) in the Signal Extraction approach.

The model is tested for it's out of sample performance by estimating the coefficients for the period 2001 to 2014 and then forecasting the probabilities for the period from 2015 to 2017. The performance is measured in terms of QPS and GSB. The in-sample estimates of QPS and GSB are found to be 0.418 and 0.0044 respectively while the out of sample estimates of QPS and GSB are found to be 0.523 and 0.0243 respectively. As expected, the out of sample estimates are found to be greater than in-sample estimates. As explained in Chapter 3, the closer the value of QPS and GSB to 0, the more accurate is the prediction of the estimated model. As within sample and out of sample values for QPS and GSB are very close to 0, the performance is found to be satisfactory suggesting the model to be a good fit.

4.2.3. Identification of banking crisis indicators – Artificial Neural Networks

The third approach utilized in present study is ANN. The two neural networks used are Elman recurrent and Multilayered Feedforward Back Propagation neural network. For experimental purpose, different structures with different number of neurons and combinations of hidden layers and transfer functions have been tested. Two network structures, RNN1 corresponding to MLFN and RNN2 corresponding to Elman Neural Network used in the study are presented below. The construction of both ANNs has been presented in Fig.4-2 and Fig. 4-3.

RNN1:

1) Input layer: 15 input units/neuron (for fifteen indicators).

2) Two Hidden layers: 30 neurons in the first layer and 15 neurons in the second layer.

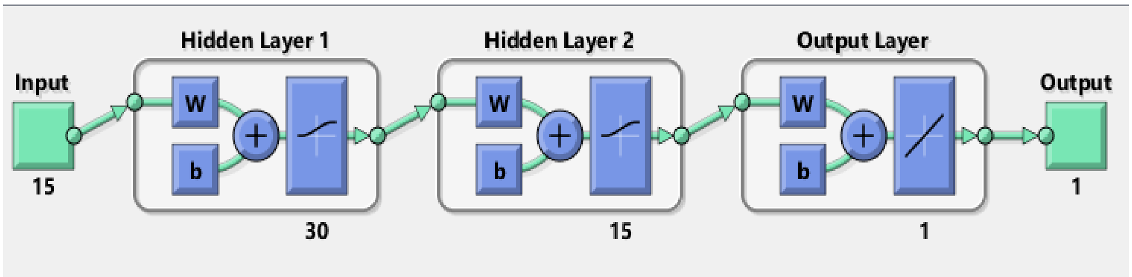
3) One Output layer with a targeted value equal to 1 for crisis periods and 0 for tranquil periods.

4) Training function: Gradient descent with momentum and adaptive learning rate backpropagation.

5) The ‘Pure-linear’ function is applied to the network output while ‘Log-Sigmoid’ is applied to both the hidden layers.

6) Mean Square Error is taken to be the performance function.

Figure 4-2:Feedforward Backpropagation Neural Network

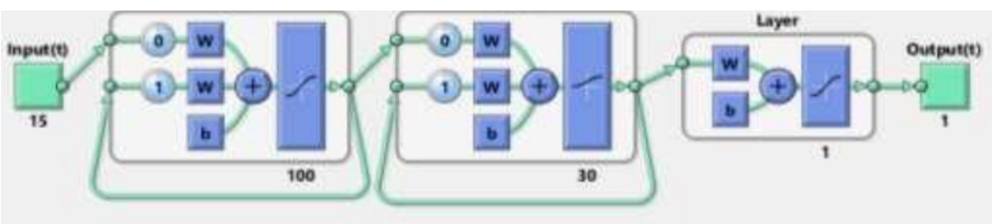


Source: Based on Author's calculations

RNN2:

1) Input layer: 15 input units/neuron (for fifteen indicators)
2) Two Hidden layers: 100 neurons in the first layer and 30 neurons in the second layer.
3) One Output layer with a targeted value equal to 1 for crisis periods and 0 for tranquil periods.
4) Training function: Conjugate gradient backpropagation with Fletcher-Reeves updates.
5) The 'Pure-Linear' function is applied to the network output while 'Tan-Sigmoid' is applied to the first hidden layer and 'Log-Sigmoid' is applied to the second hidden layer.
6) The transfer function at all layers is such that it simply reproduces the value of its own argument.
7) Mean Square Error is taken to be the performance function.

Figure 4-3:Elman Recurrent Neural Network



Source: Based on Author's calculations

Table 4-8:Calibration Scores – within the sample and out of sample for Neural Networks

Neural Network	In sample		Out of Sample	
	QPS	GSB	QPS	GSB
Feedforward BP	0.135	$4.3*10^{-5}$	0.156	0.035
Elman Recurrent	0.068	0.0002	0.148	0.011

The calibration scores from within the sample and out of sample are reported in Table 4-8. It may be observed from Table 4-8, that Quadratic Probability Score (QPS) for Elman recurrent network is lesser than the QPS for MLFN for in sample data. However, the Global Squared Bias (GSB) for MLFN is lesser than the GSB for Elman recurrent neural network. As mentioned earlier, a QPS value lies between 0 and 2. A score of zero implies perfect accuracy while a value nearer to 2 implies that the indicator is not accurate at prediction. The results presented show that Elman recurrent neural network is more accurate than MLFN as the QPS for Elman recurrent network is less than QPS for MLFN. The calibration score measured by GSB shows that the overall forecast calibration for MLFN is better than that of Elman recurrent neural network. Similar to QPS, the GSB also lies between zero and two. The value equal to zero corresponds to perfect global calibration. These results correspond to within-sample prediction. A study by Roy (2009) aiming to develop an EWS for twenty three countries using ANNs reports QPS of 0.926 and 0.846 and GSB of 0.20 and 0.117 for the two constructed Elman neural networks for in-sample data. The values of QPS and GSB in the present study are 0.068 and 0.0002 for in sample analysis which are lesser than the ones reported by Roy (2009).

The QPS and GSB scores for out of sample data were also calculated. The neural networks were trained on data spanning from 2001 to 2014 and using the trained networks the probabilities were simulated for two years spanning from 2015 to 2017. The results show that the out of sample QPS scores were greater than in sample QPS scores respectively which is desirable. The QPS for Elman network rose from 0.068 to 0.148 while for MLFN, it rose from 0.135 to 0.156. The GSB scores also increased for both the networks, however, the performance for the Elman network was found to be superior to MLFN. This was in contrast to in sample results where MLFN performed better than Elman in calibration. The study conforms with the study by Roy (2009), which reports QPS scores of 0.560 and 0.599 and GSB scores of 0.124 and 0.137 for the two Elman networks constructed. The estimated values of QPS and GSB scores arrived at in the present study indicate satisfactory performance in a single country context.

The comparative results of signals extraction, Logit regression and ANNs arrived at based on QPS and GSB score are presented in Table 4-9.

Table 4-9:Forecasting estimates of Signals approach, Logit regression and ANNs

	QPS	GSB	QPS	GSB
Composite Indicator	**In- sample**		**Out of sample**	
CI^1	0.554	0.0034	--	--
CI^2	0.463	0.0017	--	--
Logit				
Static Logit	0.423	0.0044	0.541	0.024
Neural Network				
Feedforward BP	0.136	4.3E-05	0.156	0.035
Elman BP	0.068	0.000201	0.148	0.011

It is evident from the results presented in Table 4-9 that ANNs are superior to Signaling Approach and Binary Multivariate Logit Regression. On comparing the in-sample results, Logit model and the weighted composite indicator have more or less similar QPS scores of 0.423 and 0.463 respectively. However, in terms of GSB, CI^2 fared well in comparison to Logit model. The out of sample statistics suggest Logit model outperforming MLFN in terms of GSB, though MLFN has a QPS significantly lesser than Logit's QPS score, thus better in terms of accuracy. The Elman recurrent neural network outperformed all the conventional techniques in terms of both QPS and

GSB, therefore standing out as the best crisis prediction technique. The in sample QPS and GSB scores for Elman Network are 0.068 and 0.0002 respectively. In comparison with Logit model's QPS score of 0.423 and GSB score of 0.0044 and Signal Approach's QPS score of 0.463 and GSB score of 0.0017, Elman network performs best in anticipating banking sector stress events. The out of sample QPS and GSB scores for Elman network are 0.148 and 0.011 which are closer to 0 as compared to out of sample QPS (0.541) and GSB (0.024) scores of out of sample Logit model.

4.2.4. Critical variables using feature selection in ANNs

The important input variables identified based on stepwise selection using ANNs are: growth in inflation, growth in oil prices, call money rate, growth in stock prices, deviation from the Real Effective Exchange Rate, spread between bank rate and yield to maturity on 91 day T-bills, and growth in credit to deposit Ratio. Table 4-10 and Table 4-11 presents the R^2 values and MSE values obtained for forward and backward stepwise selection procedures respectively. The variables in Table 4-10 are in order of addition i.e. using only call money rate gives R^2 = 49.15% and when growth in inflation is combined with CMR, the R^2 value goes to 63.87%. Similarly, when Spread is added to the previous two input variables, the R^2 rises to 73.59%. This procedure goes on till all 15 variables are added which result in value of R^2 = 81.81%.

Table 4-10:Forward stepwise feature selection

Variable	MSE	R^2
Call Money Rate	0.193	49.15
Growth in Inflation	0.153	63.87
Spread	0.111	73.59
Growth in Oil Prices	0.0974	80.76
Growth in Reserve Money	0.0992	78.16
Growth in Industrial Production	0.0962	77.57

Growth in ratio of Gross Fiscal Deficit to GDP	0.0770	83.42
Growth in Short term Debt	0.0781	81.60
Growth in ratio of Broad Money to Foreign Exchange Reserves	0.0941	80.25
Growth in stock prices	0.0873	80.49
Growth in ratio of Foreign Direct Investment to GDP	0.0813	82.25
Growth in ratio of Current Account Balance to GDP	0.0754	82.23
Yield to Maturity on 91 days T-Bills	0.0868	79.34
Deviation from Real Effective Exchange Rate	0.0930	80.04
Growth in Credit to Deposit Ratio	0.0878	81.81

In Table 4-11, the ANN is trained with all the 15 variables and one variable is removed each time. For example, the value of R^2 decreases to 68.62% from 81.81% when the variable M3/FEX is removed from the model leaving only 14 variables. When the next variable (growth in reserve money) is removed from the 14 variable model, the R^2 value drops to 56.95% for the model with 13 variables. The process is carried out till only variable is left in the model i.e. FDI_GDP, which gives R^2 value equal to 39.10%.

Table 4-11: Backward stepwise feature selection

Variable	**MSE**	**R^2**
Growth in ratio of Broad Money to Foreign Exchange Reserves	0.123	68.62
Growth in Reserve Money	0166	56.95
Growth in ratio of Current Account Balance to GDP	0.146	67.51
Growth in Inflation	0.134	65.28
Growth in Short term Debt	0.174	59.66
Growth in Oil Prices	0.153	58.73
Deviation from Real Effective Exchange Rate	0.162	56.57
Spread	0.184	50.21
Yield to Maturity on 91 days T-Bills	0.182	55.41

Growth in Credit to Deposit Ratio	0.166	53.26
Growth in stock prices	0.213	40.68
Growth in Industrial Production	0.165	58.82
Growth in ratio of Gross Fiscal Deficit to GDP	0.193	39.46
Call Money Rate	0.252	20.84
Growth in ratio of Foreign Direct Investment to GDP	0.215	39.10

Based on the performance of R^2 values, 9 variables are found to contribute in increasing R^2 values in forward selection from 49.15% to 81.81% and 11 variables are found to be contributing in decreasing R^2 values from 81.81% to 20.84% as presented in Table 4-12 and Table 4-13. Consequently, a set of seven common variables out of those 9 and 11 variables are selected and their respective performance was tested on both the MLFN and Elman Recurrent network (ERN) against the variables identified by the forward and backward procedures. The results are presented in Table 4-12 and 4-13.

Table 4-12:Performance comparison of variables on ANNs with max_fails = 6

Network Construct	Feature Selection	MSE	R^2 value
Feed Forward Network	Forward Stepwise	0.151	68.76
	Backward Stepwise	0.120	72.62
	Common important variables	0.126	74.07
Elman Recurrent Network	Forward Stepwise	0.174	60.57
	Backward Stepwise	0.109	74.43
	Common important variables	0.097	78.68

Table 4-13:Performance comparison of variables on ANNs with max_fails = 1000

Network Construct	Feature Selection	MSE	R^2 value
Feed Forward Network	Forward Stepwise	0.0776	85.04
	Backward Stepwise	0.0772	82.41

	Common important variables	0.0359	91.94
Elman Recurrent Network	Forward Stepwise	0.0698	85.28
	Backward Stepwise	0.0462	89.26
	Common important variables	0.0396	93.25

As can be seen from the Tables 4-12 and 4-13, the common set of seven variables namely growth in oil prices, growth in stock prices, growth in inflation, Real Effective Exchange Rate, call money rate, growth in credit to deposit ratio and spread between bank Rate and YTM perform better than the set of the variables found in forward and backward selection. The R^2 value for both the MLFN and ERN containing only important variables is greater than the R^2 values of the networks containing variables obtained solely by forward and backward selection methods even when the parameter "max_fails" is changed. In terms of R^2, most important variables result in 74.07% and 78.68% for MLFN and ERN respectively when the max_fails parameter is equal to 6 and 91.94% and 93.25% when the max_fails parameter is equal to 1000. In order to test the robustness of the common variables, a parameter was changed in ANNs namely "max_fails" signifying maximum number of validation checks allowed to reach an optimal solution. This is done in order to take care of overfitting. Thus, the robustness of the most important variables is evident from the above analysis.

One of the limitation of the study lies in identification of the relationships among the input variables and the probability of banking fragility. The positive and negative relationships, and their relative weightages are difficult to determine due to the black box critique. Therefore, in this study relationships have been discussed based on logic and prior studies carried out in the same domain.

Inflation, interest rates and rapid lending growth have been found significantly contributing to banking sector vulnerability. The findings are in line with the existing literatureas suggested by Lambregts and Ottens (2006) who investigated the leading indicators of banking crisis in emerging economies using multivariate logit model and found the above mentioned important determinants of a banking crisis. Hardy and Pazarbaşıoğlu (1998), while examining the bankingcrisis episodes in 38 countries, also suggested that banking distress is associated with boom bust cycles in Inflation, credit expansion, and capital inflows; rising real interest rates and a sharp decline in real exchange rate; and adverse trade shocks which could include oil price shocks. Thisstudy also indicates that variables like Inflation, credit growth and interest rates like call money rate, spread between bank rate and YTM on 91-days T-bill, oil prices, and real exchange rate playa major role in indicating an approaching crisis situation. In the case of banking crisis, in addition,this study suggests that growth in equity stock prices have also been found as one of the indicativeEWS which is supported by the outcome of study by Allen and Gale (1998) that related the assetprice declines and banking crisis. According to Marshall (1998), if investors perceive asset pricesas a function of the probability of occurrence of future crisis, a decline in asset prices may lead toa banking crisis. Peter (2009) has linked banking system distress with asset prices in a monetary macroeconomic model highlighting that the effect of falling asset prices is indirect, non-linear, andinvolves feedback from the banking system in the form of credit contraction. In contrast, it was suggested that sharp declines in equity prices do not cause problems in the banking sector when the linkages between the banking crisis and equity market crisis were examined for 14 developedcountries during 1970-99 (Vila, 2000).

Thus, above variables could be used by policy planners for taking proactive measures to combat the adverse implications arising out of possible happening of a banking crisis. It may be mentioned that it has to be a dynamic process to identify significant variables and their possible implications on banking crisis. The indicators constituting EWS provide an insight as to which macro variables need to be given greater cognizance to minimize adverse implications by taking proactive policy measures during a specific time period. Hence, aiding the policy planners in deciding a balanced mix of policies wary of the situations present in a country. It is evident from the study that the performance of neural networks as a tool to predict early warning signal for banking fragility in India is effective and promising alternative to conventional techniques. It should be combined with other techniques to so as to get a clearer picture and take better and necessary corrective measures to prevent a future crisis and/ or safeguard from adverse implications on the economy.

4.2.5. Comparison of predictive power of dynamic and static Logit models

This section has been divided into two parts. The first analysis deals with the estimation of logit models using various lagged independent variables. The second analysis deals with the estimation of logit models employing lagged values of dependent variables. Both of the analyses examine the predictive power of estimated models using ROC analysis. Based on the work by Candelon et. Al. (2014) the present section also attempted to analyze the dynamic specifications of the logit models for development of EWSs. The study proposed a new generation of EWSs which considered the persistence dimension of the crisis process. As it has been suggested by Tudela,2004, the longer a country has stayed in a crisis period, the higher has been the probability of exit for that country irrespective of the political reaction. Therefore, negligence of the endogenous crisis persistence may have led to incorrect evaluation of the likelihood prediction by earlier EWSs.

4.2.5.1. Comparison of dynamic and static models using lagged independent variables

The following section presents the results for within-sample and out-of-sample analysis. The in-sample results, reported in Table 4-14, estimate the static logit models with lagged (t, t-3, t-6 and t-12) and without lag early warning indicators. The final dynamic model has been estimated using the indicators which were found to be significant and correctly signed for static logit models. The forecast performance of the models has been evaluated on the basis of their goodness of fit for different probability thresholds which has been presented in Table 4-15 and the AUROC curves which have been presented in Figure 4-4. Also, all the five models have been compared for the goodness of fit using ROC curves.

Table 4-14:Comparison of static Logit models with t, t-3, t-6, t-12 variables, and dynamic Logit model- In Sample

	Model(1)	Model (2)	Model (3)	Model (4)	Model (5)
YTM	2.323***				-1.296
	(3.85)				(-1.52)
CMR	0.608*				1.201*
	(2.05)				(2.54)
GIIP	-0.0789				
	(-1.00)				
GWPI	0.312				
	(2.69)				
GOILP	0.0123				
	(0.99)				
GCDR	0.0110				
	(0.14)				
GRM	0.141**				0.167*
	(3.27)				(2.09)
M3FEX	0.0853**				0.152
	(3.25)				(1.84)
GFDGDP	-0.0103				
	(-0.96)				
FDIGDP	-0.00472				
	(-0.44)				
CABGDP	-0.00108				
	(-1.05)				
GSP	-0.0107				
	(-0.65)				
GSTD	0.101				
	(0.24)				
SPREAD	3.691***				0.873
	(5.54)				(1.15)
REER_DEV	-0.0298				
	(-0.23)				

L3REERDV	-0.320**			-0.179
	(-2.66)			(-0.72)
L3GSTD	0.0356			
	(0.09)			
L3GWPI	0.154			
	(1.46)			
L3GRM	0.0520			
	(1.44)			
L3GCDR	-0.0925			
	(-1.12)			
L3YTM	2.778***			2.767
	(4.84)			(1.12)
L3GSP	-0.0194			
	(-1.22)			
L3CMR	-0.346			
	(-1.32)			
L3M3FEX	0.0956***			0.0226
	(3.75)			(0.27)
L3FDIGDP	0.00327			
	(0.33)			
L3GFDGDP	-0.0217			-0.0138
	(-1.87)			(-0.44)
L3CABGDP	-0.00151			
	(-1.70)			
L3GOILP	0.0159			
	(1.39)			
L3SPREAD	2.406***			2.190
	(4.42)			(0.87)
L3GIIP	0.0492			
	(0.70)			
L6REERDV		-0.0811		
		(-0.72)		
L6GSTD		0.359		
		(0.88)		
L6GWPI		-0.00760		
		(-0.08)		
L6GRM		-0.0343		
		(-0.66)		
L6GIIP		-0.00637		
		(-0.09)		
L6YTM		1.628***		0.191
		(3.51)		(0.08)
L6GSP		0.00396		

	(0.27)		
L6CMR	0.0980		
	(0.37)		
L6M3FEX	0.102***		0.117
	(4.18)		(1.17)
L6FDIGDP	-0.00395		
	(-0.43)		
L6GFDGDP	-0.00612		
	(-0.54)		
L6CABGDP	-0.000440		
	(-0.50)		
L6GOILP	0.0285*		0.0723**
	(2.42)		(2.65)
L6SPREAD	1.786***		1.037
	(3.78)		(0.43)
L6GCDR	-0.211*		-0.438*
	(-2.53)		(-2.06)
L12GIIP		-0.0362	
		(-0.43)	
L12GCDR		-0.318**	-0.243
		(-3.13)	(-1.84)
L12YTM		1.101*	0.208
		(2.40)	(0.32)
L12GSP		0.0351*	0.000191
		(2.07)	(0.01)
L12CMR		-0.342	
		(-0.83)	
L12M3FEX		0.125***	0.0771
		(4.05)	(0.78)
L12FDIGDP		-0.00883	
		(-0.87)	
L12GFDGDP		0.0193	
		(1.61)	
L12CABGDP		0.000476	
		(0.50)	
L12GOILP		0.00850	
		(0.61)	
L12SPREAD		0.132	
		(0.29)	
L12GRM		-0.333***	-0.580***
		(-4.28)	(-3.31)
L12REERDEV		0.236*	0.955***
		(1.94)	(3.56)

L12GWPI				0.0477	
				(0.46)	
L12GSTD				1.115*	1.278
				(2.26)	(1.27)
_cons	-22.23***	-17.58***	-11.11***	-1.050	-15.08*
	(-5.35)	(-4.75)	(-3.65)	(-0.36)	(-2.39)
N	194	191	188	182	183
pseudo R-sq	0.492	0.404	0.376	0.475	0.685

t statistics in parentheses
* p<0.1 ** p<0.05 *** p<0.01

Table 4-14 illustrates the coefficients estimated by the Logit Model with different lagged variables (static and dynamic). The variables which are found to be significant at 5% significance level for Model 1 with no lagged variables constitute Reserve Money (GRM), ratio of M3 to Foreign Exchange Reserves (M3FEX), Weighted Average Call Money Rate (CMR), Yield to Maturity on 91 day T-bills (YTM), and the spread between Bank rate and YTM (SPREAD). GRM, M3FEX, YTM, CMR, and SPREAD are found to be positively related i.e. the growths in the above variables are more likely to cause a crisis. The pseudo R^2 of Model 1 is equal to 49.23%. The prediction evaluation at 0.5, 0.25 and 0.2 presents the total percentage of observations correctly classified. The probability of a crisis following an alarm decreases from 91.57% to 72.27% as the cut-off probability decreases from 0.5 to 0.2. The ROC curve for Model 1 is illustrated in **Figure 4-4 (i)**. The AUROC is found to be 92.90% which suggests that the estimated model is a good fit.

The coefficients for Logit model with explanatory variables which are lagged by 3 months (Model 2) are also presented in Table 4-14. YTM, M3FEX, SPREAD, REER_DEV are found to be significant variables at 5% significance level. GFD_GDP is significant at 10% significance level. YTM, M3FEX, and SPREAD are found to be positively related while REER_DEV and GFD_GDP are found to less likely increase the probability of a crisis. The pseudo R^2 Model 2 is equal to 40.36%. The prediction evaluation presented in Table 4-15 suggests the total percentage of correctly classified observations to be 82.72% at 0.5, 76.44% at 0.25 and 70.68% at 0.2 cut-off probabilities. The AUROC curve comes out to be 89.58%. as reported in **Figure 4-4 (ii)**.

Table 4-14 accounts for the estimated Logit model with t-6 lagged early warning indicators (Model 3). YTM, GCDR, M3FEX, OILP, and SPREAD are found to be significant indicators with a 6-month lag for each variable at 5% significance level. YTM, M3FEX, OILP, and SPREAD are found to be positively related to the occurrence of a crisis while GCDR is found to be negatively related to the incidence of a crisis. The pseudo R^2 Model 3 is equal to 37.56%. The prediction evaluation in Table 4-15 reveals that the total percentage of correctly classified observations decreases from 80.32% to 69.51% as the probability threshold decreases from 0.5 to 0.2. However, the probability of a tranquil period following no signal increases from 77.59% to 89.09%. The AUROC curve for Model 3 is illustrated in **Figure 4-4 (iii).** The AUROC is 87.72% which is less than the other models. The model hence performs relatively bad in terms of discrimination ability with respect to the other models.

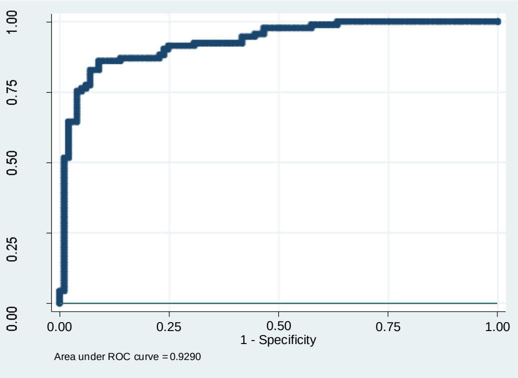
The results for the Logit estimation using 12 months lagged variables (Model 4) are also presented in Table 4-14. The significant variables at 5% significance level are GCDR, YTM, GSP, M3FEX, GRM, and GSTD while at 10% significance level REER_DEV becomes statistically significant as well. GSP, M3FEX, REER_DEV, YTM, and GSTD are more likely to cause a crisis while GCDR and GRM are less likely to cause a crisis. The pseudo R^2 for Model 4 comes out to be 47.52%. The prediction evaluation in Table 4-15 reveals the total percentage of correctly classified observations to be 86.26% at 0.5, 79.67% at 0.25 and 76.37% at 0.2 cut-off probabilities. The AUROC comes out to be 91.76%. as reported in **Figure 4-4 (iv)**.

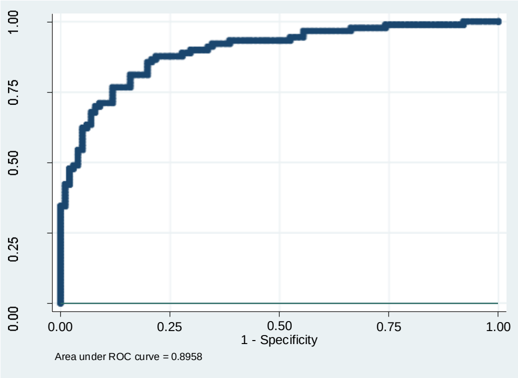
The coefficients and their standard errors for the final Dynamic Logit model (Model 5) incorporating all the significant variables from the prior static models are presented in Table 4-14. The R^2 comes out to be 68.55% which is the highest among all the tested models. The total percentage of correctly classified observations, presented in Table 4-15, is also the highest at 0.5 cut-off probability which is equal to 90.71% and the AUROC curve is also the highest i.e. 97.15% hence depicting the highest discrimination ability. The AUROC for Model 5 is presented in **Figure 4-4 (v)**.

Table 4-15:Prediction Evaluation for Static and Dynamic Logit Models- In Sample

Probability cut off (0.50)	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**
Total % of observations correctly classified	87.63%	82.72%	80.32%	86.26%	90.71%
Probability of an alarm conditional on a crisis	81.72%	76.67%	70.11%	82.56%	89.53%
Probability of a crisis following an alarm	91.57%	85.19%	84.72%	87.65%	90.59%
Probability of a tranquil period following no signal	93.07%	80.91%	77.59%	85.15%	90.82%
Probability of no signal conditional on tranquil period	84.68%	88.12%	89.11%	89.58%	91.75%
Probability cut off (0.25)					
Total % of observations correctly classified	81.44%	76.44%	74.47%	79.67%	88.52%
Probability of an alarm conditional on a crisis	91.40%	92.22%	89.66%	90.70%	94.19%
Probability of a crisis following an alarm	75.22%	68.60%	66.67%	72.90%	83.51%
Probability of a tranquil period following no signal	90.12%	90%	87.32%	89.33%	94.19%
Probability of no signal conditional on tranquil period	72.28%	62.38%	61.39%	69.79%	83.51%
Probability cut off (0.20)					
Total % of observations correctly classified	79.38%	70.68%	69.15%	76.37%	87.98%
Probability of an alarm conditional on a crisis	92.47%	93.33%	93.10%	91.86%	96.51%
Probability of a crisis following an alarm	72.27%	62.69%	60.90%	68.70%	81.37%
Probability of a tranquil period following no signal	90.67%	89.47%	89.09%	89.55%	96.30%
Probability of no signal conditional on tranquil period	67.33%	50.50%	48.51%	62.50%	80.41%

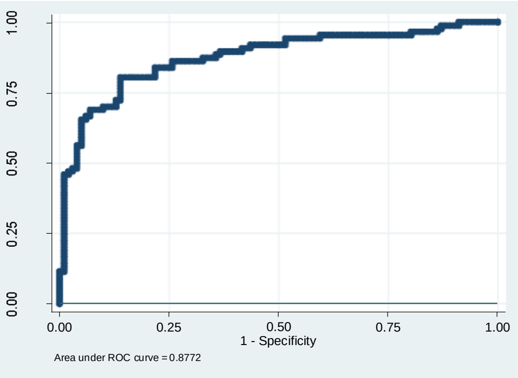
Figure 4-4:Area under the ROC curves

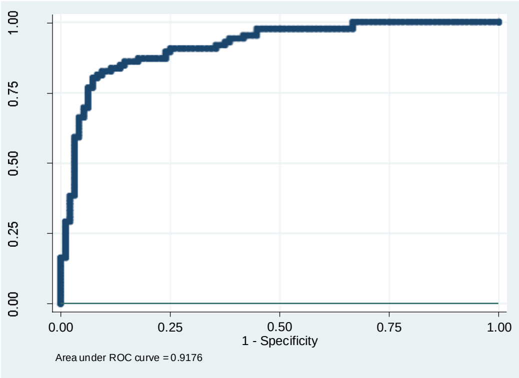




(i)

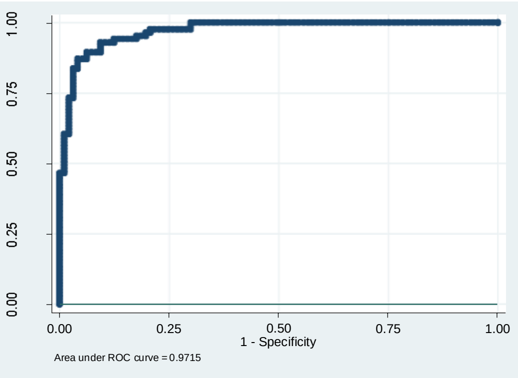
(ii)





(iii)

(iv)



(v)

Source: Based on Author's calculations

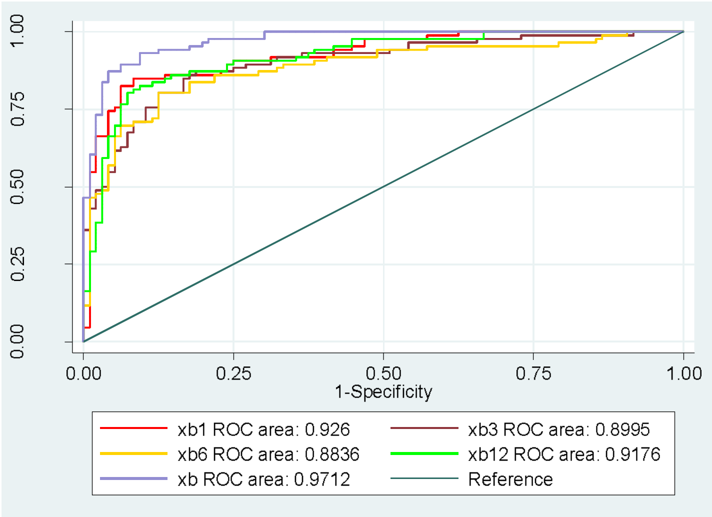
Table 4-16 presents the comparison of area under ROC curves for static and dynamic models. The AUC for Model 5 is highest (97.12%) which denotes the curve under the dynamic model incorporating variables at their level and selected lags. The p-value is 0.0001 which is less than 0.05, hence failing to accept the null hypothesis of AUC being equal for all the ROC curves. **Figure 4-5** presents the graphical representation of the comparison of AUROCs for all the estimated models for in sample analysis. It can be inferred from the graph that the area (97.12%) for the dynamic model is greater than all the static logit models at all the probability threshold levels.

Table 4-16:Comparison of AUC for ROC curve- In sample

	Obs	ROC Area	Std. Err	Asymptotic Normal [95% Conf. Interval]	
Model 1	182	0.926	0.0196	0.88756	0.96443
Model 2	182	0.8995	0.0232	0.85397	0.94497
Model 3	182	0.8836	0.0262	0.83229	0.93491
Model 4	182	0.9176	0.0208	0.87695	0.95832
Model 5	182	0.9712	0.0101	0.95132	0.99103
H0: Area (Model 1) = Area (Model 2) = Area(Model 3) = Area(Model 4) = Area(Model 5)					
chi2(4)		24.44	**Prob>chi2**	0.0001	

Table 4-17 reports the Hosmer-Lemeshow statistics. Considering the H-L test, if the associated p-value is significant ($p<0.05$), it might be an indication that the model doesn't fit the data. Since, for all the static and dynamic models, the H-L goodness- of- fit test statistic is much greater than 0.05, the null hypothesis that there is no difference between the model-predicted values and the observed values of the dependent variable is not rejected, implying that the model's estimates fit the data at an acceptable level.

Figure 4-5:Comparison of AUC for ROC curves



Source: Based on Author's calculations

Table 4-17:Hosmer-Lemeshow statistics for static and dynamic logit models

Model	Hosmer-Lemeshow chi2(1)	Prob>chi2
Model 1	0.08	0.7750
Model 2	0.69	0.4078
Model 3	0.91	0.3392
Model 4	0.29	0.5932
Model 5	1.60	0.2059

Out of sample results

The out of sample results are presented in Table 4-18 for static and dynamic logit models. The sample has been limited to December 2014 and the probabilities for the time period of January 2015 to March 2017 are predicted on the basis of the estimated model using data limited to December 2014. The pseudo R^2 of the Model 1 (t-0 lagged) with no lagged explanatory variables comes out to be 50.90%. The pseudo R^2 for Model 2 (t-3 lagged) is 53.1%, for Model 3 (t-6 lagged) is 45.3%, for Model 4 (t-12 lagged) is 52.7% and for Dynamic model, the corresponding R^2 is

72.9%. It can be inferred from the table that the dynamic model outperforms the static models for out of sample analysis as well.

Table 4-18:Comparison of static Logit models with t, t-3, t-6, t-12 variables, and dynamic Logit model- Out of Sample

	Model(1)	Model (2)	Model (3)	Model (4)	Model (5)
YTM	1.224*				-1.995
	(2.25)				(-1.65)
CMR	0.740*				1.273*
	(2.20)				(2.43)
GIIP1	-0.0808				
	(-0.83)				
GWPI	0.384**				0.133
	(3.00)				(0.47)
GOILP	0.0251				
	(1.68)				
GCDR	0.00597				
	(0.06)				
GRM	0.0132*				-0.148
	(0.20)				(-0.82)
M3FEX	0.0973***				0.154
	(3.41)				(1.43)
GFDGDP	-0.00372				
	(-0.31)				
FDIGDP	0.0152				
	(1.10)				
CABGDP	-0.00142				
	(-1.28)				
GSP	-0.0196				
	(-1.05)				
GSTD	0.262				
	(0.58)				
SPREAD	2.523***				0.889
	(4.02)				(0.73)
REER_DEV	-0.108				
	(-0.76)				
L3REERDV		-0.518**			-0.312
		(-3.08)			(-0.92)
L3GSTD		0.222			
		(0.48)			
L3GWPI		0.420			

	(2.99)		
L3GRM	-0.138		
	(-1.81)		
L3GCDR	-0.125		
	(-1.11)		
L3YTM	1.542**		3.425
	(2.78)		(0.92)
L3GSP	-0.0378		
	(-1.89)		
L3CMR	-0.296		
	(-0.97)		
L3M3FEX	0.167***		-0.163
	(4.37)		(-1.26)
L3FDIGDP	0.0428		
	(2.51)		
L3GFDGDP	-0.0301*		-0.00273
	(-1.89)		(-0.06)
L3CABGDP	-0.00294**		-0.00166
	(-2.63)		(-1.62)
L3GOILP	0.0465		
	(2.61)		
L3SPREAD	0.821***		2.567
	(1.37)		(0.66)
L3GIIP	-0.0220		
	(-0.23)		
L6REERDV		-0.251	
		(-1.80)	
L6GSTD		0.567	
		(1.21)	
L6GWPI		0.186	
		(1.67)	
L6GRM		-0.114	
		(-1.85)	
L6GIIP		-0.117	
		(-1.29)	
L6YTM		0.927***	-0.462
		(1.79)	(-0.12)
L6GSP		0.000180	
		(0.01)	
L6CMR		0.195	
		(0.63)	
L6M3FEX		0.157***	0.163
		(4.56)	(1.17)

L6FDIGDP			0.0207		
			(1.53)		
L6GFDGDP			-0.00920		
			(-0.67)		
L6CABGDP			-0.00158		
			(-1.59)		
L6GOILP			0.0482**		0.0812*
			(2.77)		(2.30)
L6SPREAD			1.050*		1.251
			(2.02)		(0.31)
L6GCDR			-0.216*		-0.674*
			(-2.28)		(-2.53)
L12GIIP				-0.249	
				(-2.02)	
L12GCDR				-0.284**	-0.384*
				(-2.61)	(-2.22)
L12YTM				0.464**	0.345
				(0.81)	(0.43)
L12GSP				0.0395*	0.0575
				(2.08)	(1.82)
L12CMR				-0.436	
				(-1.04)	
L12M3FEX				0.140***	0.411*
				(4.06)	(2.29)
L12FDIGDP				-0.00235	
				(-0.17)	
L12GFDGDP				0.0126	
				(0.94)	
L12CABGDP				-0.0000607	
				(-0.06)	
L12GOILP				0.0143	
				(0.93)	
L12SPREAD				-0.405	
				(-0.69)	
L12GRM				-0.394***	-0.904**
				(-3.92)	(-3.19)
L12REERDEV				0.175	1.226**
				(1.17)	(3.13)
L12GWPI				0.241	
				(1.67)	
L12GSTD				1.219*	4.017*
				(2.05)	(2.51)
_cons	-15.25***	-9.148*	-7.150*	4.296	-4.348

	(-4.24)	(-2.57)	(-2.22)	(1.09)	(-0.57)
N	167	164	161	155	156
pseudo R-sq	0.509	0.531	0.453	0.527	0.729

t statistics in parentheses
* p<0.1 ** p<0.05 *** p<0.01"

The prediction performances for the static and dynamic models for different probability cutoffs for out of sample analysis are presented in Table 4-19. The estimates for Model 5 (Dynamic) showsthe superior performance of the dynamic model over other static models. The total percentage of observations correctly classified for Dynamic model is 91.67% when the cut off probability is 0.50, 91.03% when the cut off probability is 0.25 and 89.10% when the cut off probability is 0.20.

Table 4-19:Prediction Evaluation for static and dynamic Logit Models- Out of Sample

Probability cut off (0.50)	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**
Total % of observations correctly classified	86.83%	85.98%	84.47%	87.10%	91.67%
Probability of an alarm conditional on a crisis	79.45%	75.71%	76.12%	80.30%	87.88%
Probability of a crisis following an alarm	89.23%	89.83%	85.00%	88.33%	92.06%
Probability of a tranquil period following no signal	85.29%	83.81%	84.16%	86.32%	91.40%
Probability of no signal conditional on tranquil period	92.55%	93.62%	90.43%	92.13%	94.44%
Probability cut off (0.25)					
Total % of observations correctly classified	83.23%	81.10%	75.78%	80.65%	91.03%
Probability of an alarm conditional on a crisis	89.04%	87.14%	86.57%	87.88%	93.94%
Probability of a crisis following an alarm	76.47%	73.49%	65.91%	72.50%	86.11%
Probability of a tranquil period following no signal	90.24%	88.89%	87.67%	89.33%	95.24%
Probability of no signal conditional on tranquil period	78.72%	76.60%	68.09%	75.28%	88.89%
Probability cut off (0.20)					
Total % of observations correctly classified	79.04%	79.27%	74.53%	79.35%	89.10%
Probability of an alarm conditional on a crisis	90.41%	91.43%	91.04%	92.42%	93.94%
Probability of a crisis following an alarm	70.21%	69.57%	63.54%	69.32%	82.67%
Probability of a tranquil period following no signal	90.41%	91.67%	90.77%	92.54%	95.06%
Probability of no signal conditional on tranquil period	70.21%	70.21%	62.77%	69.66%	85.56%

For assessment of the prediction ability of the models, QPS scores are compared. The out of sample QPS scores for static and dynamic models are presented in Table 4-20. The QPS score for dynamic model is least in comparison of static models. The QPS for dynamic model is 0.1935 which is closer to 0 than any other estimated model, hence depicting the highest predictive ability among the estimated models.

Table 4-20:QPS Scores for Static and Dynamic Logit Models- Out of Sample

Model	QPS Score
Model 1	0.3197
Model 2	0.4721
Model 3	0.4604
Model 4	0.5496
Model 5	0.1935

The outcome of the present study reveals that yield to maturity on 91 days T- Bills, the spread between bank Rate and YTM and growth in the ratio of broad money supply relative to foreign exchange reserves are significant in all the static models (t, t-3, t-6 and t-12) within sample, hence it is concluded that these three explanatory indicators are the most crucial ones for predicting a banking stress situation in Indian context. The variable namely Overvaluation of Real Effective Exchange Rate become significant at their third lag. Growth in credit to deposit ratio and growth in oil prices come into play at their sixth lag. Growth in stock Prices, growth in short term debt, and growth in reserve money become significant at their twelfth lag. Corroborating the superior performance of Elman recurrent neural network, it is evident from the outcomes of this study that macroeconomic and financial indicators contribute to the prediction of banking stress withdifferent time lags and hence including their lagged values into the development of EWS helps inimproving the predictive power of the EWS.

The above analysis identifies and compares the potential indicators for predicting an occurrence of a banking crisis in Indian context, and further compares the predictive performance of traditional methodologies namely signal extraction and limited dependent variable models like Logit models with those of new techniques like ANNs. The study also compares the predictive power of variables at their level and at their lags for both limited dependent variable models and ANN models. The analysis reveals that Neural networks perform better in comparison to the traditional techniques as it can learn itself from the fed data and hence, can forecast better by evolving itself through training from the new data. The process is dynamic in nature, and it requires no assumption of the functional fit of the inputs and outputs. Whereas, the probit/logit models assume the functional form prior to the data is fed to the model. This results into a major limitation as the true nature about the relationships among the variables remains unraveled. However, ANNs suffer from a major limitation of "black box critique" which renders its usefulness in identification of the warning indicators and their positive/negative relationships with the other variables. The first limitation can be overcome by using feature selection techniques which have been utilized in the present analysis. The stepwise (forward and backward) selection helps in identification of the significant variables which are found to be contributing in predicting the probability of a banking crisis in Indian context. To overcome the second limitation of unknown relationships among the variables, it becomes imperative to combine these ANN models with the traditional models like logit/probit to get a sense of the dynamics undergoing the whole process.

Secondly, the analysis reveals the importance of process being dynamic in nature. As the study employs lagged variables for ANNs using Elman recurrent network and for Logit models by adding specific lagged values of the input indicators, the predictive power of these models improves. This suggests that the lagged values of the potential warning indicators can be useful in anticipating an approaching crisis.

4.2.5.1. Lagged dependent variables

The benchmark model i.e., the static EWS model (labeled Model 1), as well as three types of dynamic EWSs: one including the lagged binary dependent variable (labeled Model 2), a dynamic one including the lagged index (Model 3), and a dynamic model which includes both the lagged binary dependent variable and the lagged index (labeled as Model 4) have been estimated in Table 4-21. It may be observed that the co-efficient of the lagged binary dependent variable is significant in both the models (Model 2 and Model 4) and has a positive sign. This clearly indicates that the probability of being in a crisis increases if a crisis regime prevailed in the previous period. The selection of the most parsimonious dynamic specification has been carried out by relying on the Schwarz Information Criterion (BIC). For purpose of comparison, the values of BIC for the static model have also been reported. The BIC values have been used because it penalizes the model complexity more heavily and helps in finding true model among the sets of candidates. This goodness of fit indicator revealed that the right – hand- side variables have important explanatory power, especially when the lagged dependent variable or/and the lagged index are present in the model. It may be observed that the lowest BIC value corresponds to Model (3) which has 191.1 as its BIC value. To put it another way, the goodness-of- fit indicator provides evidence of the fact that in-sample dynamic specifications generally outperform the static one.

Following estimation, the static and dynamic models are statistically tested for the in-sample and out-of-sample forecasting abilities in prediction of a banking crisis.

Table 4-21:Estimation Results (time-series logit models) using lagged dependent variable

Indicator	Model (1)	Model (2)	Model (3)	Model (4)
GSP	-0.0107	-0.0442	-0.01521	-0.0142
	(-0.65)	(-0.24)	(-0.76)	(-0.71)
GRM	0.141**	0.131**	0.209**	0.204**
	(3.27)	(2.72)	(3.19)	(3.10)
GCDR	0.011	0.045	0.093	0.092
	(-0.14)	(-0.53)	(-1.01)	(-0.99)
M3_FEX	0.0853**	0.0548	0.0463	0.0451
	(3.25)	(1.92)	(1.41)	(1.38)
FDIGDP	-0.00472	-0.00269	0.00343	0.00301
	(-0.44)	(-0.23)	(0.26)	(0.23)
GFDGDP	-0.0103	-0.0066	-0.0126	-0.0127
	(-0.96)	(-0.55)	(-0.90)	(-0.90)
CABGDP	-0.00108	-0.00051	-0.00056	-0.00048
	(-1.05)	(-0.40)	(-0.40)	(-0.35)
GSTD	0.101	0.253	0.211	0.219
	(0.24)	(0.55)	(0.42)	(0.44)
GOILP	0.0123	0.0111	0.0244	0.0236
	(0.99)	(0.81)	(1.54)	(1.49)
SPREAD	3.691***	2.355***	2.283***	2.204**
	(5.54)	(3.54)	(3.41)	(3.26)
GIIP	-0.0789	-0.1326	-0.1353	-0.1406
	(-1.00)	(-1.50)	(-1.42)	(-1.47)
GWPI	0.312	0.178	0.075	0.0736
	(2.69)	(1.53)	(0.56)	(0.55)
YTM	2.323***	1.128*	1.111*	1.053*
	(3.85)	(2.11)	(2.31)	(2.19)
CMR	0.608*	0.556	0.637*	0.613
	(2.05)	(1.82)	(2.05)	(1.95)
REER_DEV	-0.0298	-0.0221	-0.105	-0.105
	(-0.23)	(-0.16)	(-0.74)	(-0.73)
Lagged Binary Variable		2.479***		0.493
		(4.41)		(0.65)
Banking Stress Fragility Index			-4.728***	-4.167**
			(-4.68)	(-3.24)
_cons	-22.23***	-14.37***	-13.86***	-13.47***
	(-5.35)	(-3.77)	(-3.70)	(-3.59)

N	194	194	194	194
Pseudo R-sq	0.492	0.569	0.622	0.623
BIC	220.7	205.2	191.1	196.0

t statistics in parentheses
* p<0.1** p<0.05*** p<0.01"

In-sample analysis

The forecasting abilities using ROC comparison of the static and dynamic time series models for the complete data has been checked for in-sample and presented in Table 4-22.

Table 4-22:ROC comparison test- in sample analysis

	Obs	ROC Area	Std. Error	Asymptotic Normal [95% Confidence Interval]	
omodel1	194	0.9290	0.0185	0.89265	0.96533
omodel2	194	0.9439	0.0169	0.91073	0.97706
omodel3	194	0.9637	0.0119	0.94032	0.98707
omodel4	194	0.9626	0.0121	0.93889	0.98637

Ho: area(model1) = area(model2) = area(model3) = area(model4)
chi2(3) = 8.44 Prob>chi2 = 0.0377

The forecasting abilities of the static and dynamic models have been assessed by considering both AUC evaluation criterion and the ROC comparison test. Table 4-22 reports the results of AUC performance assessment criterion, the ROC test statistic and the p-values for all the models. The higher the AUC, the better the model. For further comparison, ROC test has been used. The significance in differences among the models' performances can be checked using this test. Since the p- value is less than 0.05, the ROC test fails to accept the null hypothesis of equal forecasting abilities thus corroborating the finding that accounting for the endogenous persistence of the crisis by including lagged index is important for the forecasting abilities of banking crisis EWSs.

Out of sample analysis

Table 4-23 reports the results of AUC performance assessment criterion, the ROC test statistic and the p-values for all the models. The forecast has been made for two years using the data from 2001 – 2014. The results indicate the superior performance of dynamic logit models over static logit models. Since the p- value is less than 0.05, the ROC test fails to accept the null hypothesis of equal forecasting abilities. Thus, it can be inferred that dynamic models fare well than static models.

Table 4-23:ROC comparison test- out of sample analysis

	Obs	ROC Area	Std. Error	Asymptotic Normal [95% Confidence Interval]	
omodel1	194	0.8508	0.0277	0.79664	0.90506
omodel2	194	0.9094	0.0216	0.86714	0.95167
omodel3	194	0.9574	0.0128	0.93234	0.98249
omodel4	194	0.9571	0.0127	0.93223	0.98196
Ho: area(omodel1) = area(omodel2) = area(omodel3) = area(omodel4)					
chi2(3) = 23.75				Prob>chi2 = 0.0000	

The in-sample and the out of sample ROC curves confirm that the dynamic model outperforms static model estimated in the present study. It can also be concluded from the present analysis that the dynamic model with lagged crisis index performs better than all other dynamic models.

4.3. Analysis on Indian Stock Market Crisis

This section focuses on the identification, selection and comparison of effectiveness of warning indicators for predicting a stock market crisis in Indian context. The study further compares the predictive power of traditional tools with that of techniques like ANNs. The analysis firstly identifies the episode of stock market crisis using an index called CMAX based on stock prices from Nifty 50. It then carries out identification of the potential warning indicators followed by testing the predictive power of the sentiment variables in addition to the selected macro and financial variables in predicting a stock market crisis. This is followed by the comparison of different techniques like Logit models and ANNs to gauge their respective predictive power in estimating the probability of a stock market crisis in India.

The following section discusses the empirical results of the above analysis. Table 4-24 presents the results of the estimated logit models where the dependent variable is a binary indicator of a crisis period. The independent variables include the real interest rate (RIRR), weighted average call money rates (CMR), Index of industrial production (IIP), Real effective exchange rate (REER), Credit to Deposit ratio (CDR), the Foreign Institutional net inflows (FII), the U.S. sentiment (MCSI), U.K. sentiment (EURO), Emerging countries sentiment (EMERSENT) and the Indian sentiment (SENT). In Table 4-24, Model 1 represents the estimated logit model with only macroeconomic and financial variables. Model 2 introduces the Indian sentiment in Model 1, while Model 3, Model 4, and Model 5 introduce U.S. sentiment, U.K. sentiment and aggregate emerging market sentiment respectively in Model 1. The statistics tabulated in parentheses are p values. The sample period includes monthly data from June 2001 to December 2018.

Table 4-24:Logit models using sentiment variables – In sample

	-1	-2	-3	-4	-5
CMR	0.178	0.183	0.395	0.581*	0.143
	(0.305)	(0.318)	(0.063)	(0.015)	(0.478)
RIRR	0.978**	1.499***	0.487	1.267**	1.294***
	(0.001)	(0)	(0.192)	(0.006)	(0.001)
CDR	0.0118	-0.0667	-0.168	-0.102	0.0933
	(0.863)	(0.384)	(0.054)	(0.268)	(0.298)
FII	-0.000136**	-0.000167**	-0.000128*	-0.000160**	-0.000121*
	(0.002)	(0.004)	(0.033)	(0.003)	(0.02)

REER	0.0149	0.0425	0.0198	0.111	0.0228
	(0.726)	(0.396)	(0.741)	(0.09)	(0.626)
IIP	-0.00699	-0.00011	-0.00128	-0.00643	-0.0206
	(0.594)	(0.993)	(0.919)	(0.679)	(0.25)
SENT		-1.090***			
		(0.0004)			
MCSI			-0.158***		
			(0.00001)		
EURO				-0.231***	
				(0.00005)	
EMERSENT					-1.151***
					(-0.001)
_cons	-12.51*	-16.97*	13.27	-23.24*	-18.57**
	(0.041)	(0.017)	(0.181)	(0.013)	(0.009)
N	211	210	211	211	211
Pseudo R^2	0.323	0.444	0.482	0.471	0.42

Following Table 4-25 reports the Goodness of fit test namely Hosmer-Lemeshow test. All the models are tested with different number of groups. It can be observed all the models, except Model 5 which constitutes the aggregate emerging market sentiment fails to pass the model specification test with 10 groups at 10% significance level.

4.3.1. Predictive power of macroeconomic and financial variables

Model 1 in the above tables represents the estimated logit model using macroeconomic and financial variables only. The variables which are found to be significant are net Foreign Institutional Investments flow (FII), and real interest rates (RIRR). With exception of the CDR in Model 2, 3 and 4, all variables are having expected signs.

Table 4-25:Hosmer Lemeshow Goodness of Fit tests

Hosmer Lemeshow Test	Number of groups	Hosmer-Lemeshow chi2(1)	Prob > chi2
MACRO	3	0.6	0.4386
	10	1.98	0.9817
SENT	3	0.37	0.5415
	10	5.23	0.7327
USENT	3	4.75	0.1293
	10	8.49	0.3867
UKSENT	3	1.78	0.1824
	10	3.45	0.9027
EMERSENT	3	2.24	0.1349
	10	15.02	0.0587

The net FIIs inflows and Industrial production are found to be negatively related to the probability of a stock market crisis while variables namely Call money rate, real interest rate, real effective exchange rate are found to be positively related to the probability of a stock market crisis in Indian context.

The goodness of fit test in Table 4-25 confirms the specification of the overall fit of the model. The present study shows that an increase in real interest rates is likely to increase the probability of a stock market crisis. This could be due to reason that the increase in interest rates lead to fall in stock prices as both businesses and consumers start cutting their spending and profits of corporate fall. An interest hike generally impacts the banks directly as banks increase their rates for consumer loans. Also, increased cost of borrowing impacts businesses by adversely affecting their earnings which eventually results in drop in stock market prices. The literature on the relationship between interest rates and stock prices include many studies which have inferred a negative relationship between them. These include studies by Pearce and Roley (1985), Hafer (1986), who documented that equity prices react negatively to the changes in discount rates; Mukherjee and Naka (1995) and Muktadir-Al-Mukit (2013) found that the long run interest rates have a negative impact on the stock market; Alam and Uddin (2009) examined the relationship between stock prices and interest rates in 15 developed and developing countries and reported a negative association between the two variables. This observation is in contrast to the other studies like Zouaoui et al. (2011) that reports a negative relationship between the interest rates and the probability of a stock market crisis in a panel of countries. According to the study, the monetary authorities cut the rates to stabilize the lagging economy in order to increase the consumer spending. This makes the credit cheaper and leads to banks taking excessive risk which can result in a stock market crisis.

The net FIIs inflows have emerged as a significant variable and is having negative relationship with the probability of a stock market crisis. This is expected as FIIs play a major role in Indian economy. There are many studies such as Srikanth and Kishore, (2012) , Shrivastav, (2013), Rajput and Thaker (2008), Jayaraj et al. (2009), Karthikeyan and Mohanasundaram (2012) and Gupta and Kumar, (2020) which have confirmed a positive relationship between the FIIs and Indian stock market. Hence, an increase in net inflow of FIIs is likely to decrease the likelihood of a stock market crisis as it boosts the investors' confidence and is indicative of a positive outlook of the economy.

4.3.2. Incremental predictive power of the sentiment variables on stock markets

In the present analysis, all the sentiment variables are found to be highly significant and negatively related to the stock market crisis probability. Results of Models 2, 3, 4 and 5 show that investor sentiment is significant even after controlling for financial and economic variables. The negative relationship between sentiment variables and crisis probability indicate that the likelihood of a stock market crisis increases with the decrease in the value of investor sentiment. This is at odds with the study by Zouaoui et al. (2011) which favors the fundamental hypothesis that investor sentiment in contrarian in nature i.e. when the sentiment is low, the future returns are expected to grow and when the sentiment is high, the future returns are expected to fall as a high value of sentiment would mean that stocks are overpriced and will experience a decline in the value to correct itself. The Mc-Fadden R^2 increases from 32.3% to 44.4 % when Indian investor sentiment is introduced in the Model-1 constituting only macroeconomic and financial variables. Introduction of U.S. sentiment in the Model-1 increases the R^2 to 48.2% while with Eurozone sentiment, the Mc-Fadden R^2 increases to 47.1%. The model-5 reports the Mc-Fadden-R^2 of 42% with aggregate emerging market sentiment.

Tables 4-26 presents the predictions at 50% and 25% probability cut offs for each of the model when full sample is considered. At 50% and 25% probability cut off, the predictive performance improves as the variables related to sentiment are included in the model. In terms of classifying correctly, the total proportion of cases increased from 88.15% (Model 1) to 90.95% (Model 2) to 91.94% (Model 3) to 92.89% (Model-5) and 93.36% (Model 4) for 50% cutoff. At 0.25 cutoff, again inclusion of sentiment variables increases the predictive power of the models. However, the predictive ability for Model-3 (U.S. sentiment) is found to be better than Model -4 (Euro sentiment) i.e. 90.05% (Model-3) and 87.20% (Model-4).

Table 4-26:Prediction Evaluation for Individual Sentiment Logit Models- In Sample

At 0.5 cutoff	**Model-1**	**Model-2**	**Model-3**	**Model-4**	**Model-5**
Sensitivity Pr(+| D)	22.22%	38.46%	55.56%	55.56%	51.85%
Specificity Pr(-|~D)	97.83%	98.37%	97.28%	98.91%	98.91%
Positive predictive value Pr(D| +)	60.00%	76.92%	75.00%	88.24%	87.50%
Negative predictive value Pr(~D| -)	89.55%	91.88%	93.72%	93.81%	93.33%
False + rate for true ~D Pr(+|~D)	2.17%	1.63%	2.72%	1.09%	1.09%
False - rate for true D Pr(-| D)	77.78%	61.54%	44.44%	44.44%	48.15%
False + rate for classified + Pr(~D| +)	40.00%	23.08%	25.00%	11.76%	12.50%
False - rate for classified - Pr(D| -)	10.45%	8.12%	6.28%	6.19%	6.67%
Correctly classified	88.15%	90.95%	91.94%	93.36%	92.89%

At 0.25 cutoff	**Model-1**	**Model-2**	**Model-3**	**Model-4**	**Model-5**
Sensitivity Pr(+| D)	66.67%	73.08%	81.48%	77.78%	66.67%
Specificity Pr(-|~D)	89.13%	89.13%	91.30%	88.59%	91.30%
Positive predictive value Pr(D| +)	47.37%	48.72%	57.89%	50.00%	52.94%
Negative predictive value Pr(~D| -)	94.80%	95.91%	97.11%	96.45%	94.92%
False + rate for true ~D Pr(+|~D)	10.87%	10.87%	8.70%	11.41%	8.70%
False - rate for true D Pr(-| D)	33.33%	26.92%	18.52%	22.22%	33.33%
False + rate for classified + Pr(~D| +)	52.63%	51.28%	42.11%	50.00%	47.06%
False - rate for classified - Pr(D| -)	5.20%	4.09%	2.89%	3.55%	5.08%
Correctly classified	86.26%	87.14%	90.05%	87.20%	88.15%

To confirm the models' predictive power, AuROC has been used. The area under the curve is a measure of the usefulness of a test. The greater the AUC, the more useful is a test. The results in Table 4-27 show that USENT has the highest value of 93.1% for AUC followed by UKSENT (92.85%), SENT (92.66%), EMERSENT (89.07%) and MACRO (88.38%) in decreasing order.

This suggests that USENT outperforms all the other models in terms of prediction. Apart from that, it is necessary to compare the ROCs to check the significance of difference among the prediction ability of different models. Table 4-27 reports these statistics and suggests that models differ in their predictive ability, but it varies with different threshold points. As observed from the Table 4-26, at 0.5 cutoff, UKSENT outperforms USENT while at 0.25, USENT outperforms UKSENT.

Table 4-27:ROC comparison for Individual Sentiment Logit Models- In Sample

Model	Obs	Area	Std. Err.	Asymptotic Normal [95% Conf. Interval]	
MACRO	210	0.8838	0.0297	0.82563	0.94193
SENT	210	0.9266	0.0246	0.87837	0.97489
USENT	210	0.931	0.0285	0.87514	0.9869
UKSENT	210	0.9285	0.0258	0.87801	0.97901
EMERSENT	210	0.8907	0.034	0.8241	0.95725
Ho: area(macro) = area(SENT)= area(USENT)=area(UKSENT)=area(EMERSENT)					
chi2(4) = 6.31				Prob>chi2 = 0.1774	

4.3.3. Effectiveness of alternate models using sentiment variables

The following section deals with the results of combining different sentiment variables and checking the respective predictive abilities.

After estimating the above models, Hosmer-Lemeshow test is used to check the specification of the models. Table 4-28 presents the results for the Goodness of Fit test. It can be observed that Models 1 (INDUK) and 2 (INDUS) are misspecified as the values 0.0004 for Model 1 and 0.0002 for Model 2 show the presence of misspecification of the models. As mentioned earlier, values less than 0.05 (5% significance level) leads to acceptance of the alternate hypothesis of H-L test that the model does not fit the data well for different number of subgroups.

Table 4-28:Hosmer Lemeshow Goodness of Fit tests

	Number of groups	Hosmer-Lemeshow chi2(1)	Prob > chi2
INDUK	3	12.33	0.0004
	10	9.67	0.2891
INDUS	3	13.44	0.0002
	10	10.97	0.2033
EMERUK	3	0.23	0.6314
	10	5.58	0.694
EMERUS	3	0.37	0.5427
	10	6.28	0.6157
EMERIND	3	0.07	0.7958
	10	5.99	0.6484
EMERINDUK	3	0.09	0.7629
	10	11.13	0.1944
EMERINDUS	3	0.09	0.7583
	10	4.78	0.781

4.3.3.1.Predictive power of the alternate models using sentiment-In sample analysis

Table 4-29 reports the predictions at 50% and 25% probability cut-offs for each of the combined models for the full data sample. Models 1 and 2 have not been reported as the concerned models are misspecified. On comparing Table 4-26 with the Table 4-29, it can be observed that the predictive ability of the models is superior to those which include only a single sentiment variable. At 50% cutoff level, both EMERUK and EMERUS are found to be classifying 93.84% of the cases correctly. While models EMERIND and EMERINDUS are found to be classifying 93.81% of the cases correctly. EMERINDUK which includes emerging market aggregate sentiment and Eurozone sentiment is found to be outperforming all other models with highest percent (95.25%).

Table 4-29:Prediction Evaluation for Combination Sentiment Logit Models- In Sample

At 0.5 cutoff	EMERUK	EMERUS	EMERIND	EMERINDUK	EMERINDUS
Sensitivity Pr(+| D)	59.26%	59.26%	57.69%	65.38%	61.54%
Specificity Pr(-|~D)	98.91%	98.91%	98.91%	99.46%	98.37%
Positive predictive value Pr(D| +)	88.89%	88.89%	88.24%	94.44%	84.21%
Negative predictive value Pr(~D| -)	94.30%	94.30%	94.30%	95.31%	94.76%
False + rate for true ~D Pr(+|~D)	1.09%	1.09%	1.09%	0.54%	1.63%
False - rate for true D Pr(-| D)	40.74%	40.74%	42.31%	34.62%	38.46%
False + rate for classified + Pr(~D| +)	11.11%	11.11%	11.76%	5.56%	15.79%
False - rate for classified - Pr(D| -)	5.70%	5.70%	5.70%	4.69%	5.24%
Correctly classified	93.84%	93.84%	93.81%	95.24%	93.81%

At 0.25 cutoff	EMERUK	EMERUS	EMERIND	EMERINDUK	EMERINDUS
Sensitivity Pr(+| D)	62.96%	77.78%	73.08%	69.23%	76.92%
Specificity Pr(-|~D)	92.39%	94.57%	92.39%	92.93%	94.02%
Positive predictive value Pr(D| +)	54.84%	67.74%	57.58%	58.06%	64.52%
Negative predictive value Pr(~D| -)	94.44%	96.67%	96.05%	95.53%	96.65%
False + rate for true ~D Pr(+|~D)	7.61%	5.43%	7.61%	7.07%	5.98%
False - rate for true D Pr(-| D)	37.04%	22.22%	26.92%	30.77%	23.08%
False + rate for classified + Pr(~D| +)	45.16%	32.26%	42.42%	41.94%	35.48%
False - rate for classified - Pr(D| -)	5.56%	3.33%	3.95%	4.47%	3.35%
Correctly classified	88.63%	92.42%	90.00%	90.00%	91.90%

At 25% probability cutoff, model EMERUS constituting the aggregate emerging market sentiment and U.S. sentiment is found to be outperforming all the other models with 92.42% of the total cases correctly classified. EMERINDUS which includes the Indian sentiment also performs fairlywell in classifying 91.90% of the total cases correctly. This is followed by models EMERIND and EMERINDUK with 90% of the total cases correctly classified. This shows that inclusion of sentiment variables definitely improves upon the predictive ability of the logit models. And, at different probability thresholds, different sentiment variables fare slightly better than the others.

To confirm the performance of the models ROC under AUC is employed for combined models as well and they are compared with the single sentiment variable model. Table 4-30 presents the results of the comparison among all the models for the full sample data. Even though Models 1 and 2 namely (INDUK and INDUS) are misspecified as already proved and stated in the above section, they have been included in ROC comparison to check their respective predictive abilities in comparison to the other models. It can be observed that EMERINDUS (95.23%) outperforms all the other models followed by EMERINDUK (94.75%) and EMERUS (94.8%). The value of 0.0568 indicates that the areas under all these model curves differ at 10% significance level.

Table 4-30:ROC comparison for Combination Sentiment Logit Models- In Sample

Model	Obs	Area	Std. Err.	[95% Conf. Interval]
MACRO	210	0.8838	0.0297	0.82563
SENT	210	0.9266	0.0246	0.87837
USENT	210	0.931	0.0285	0.87514
UKSENT	210	0.9285	0.0258	0.87801
EMERSENT	210	0.8907	0.034	0.8241
INDUK	210	0.9342	0.0271	0.88102
EMERINDUS	210	0.9523	0.0218	0.90966
EMERINDUK	210	0.9475	0.0209	0.90654
EMERIND	210	0.9442	0.0217	0.90163
EMERUS	210	0.948	0.0221	0.90471
EMERUK	210	0.9327	0.0228	0.88804
INDUS	210	0.9385	0.0278	0.88413
Ho:area(macro)=area(SENT)=area(USENT)=area(UKSENT)=area(EMERSENT)=area(INDUK)=area(EMERINDUS)=area(EMERINDUK)=area(EMERIND)=area(EMERUS)=area(EMERUK)=area(INDUS)				
chi2(11)= 19.25			Prob>chi2= 0.0568	

4.3.3.2. Predictive power of the alternate models using sentiment- Out of sample analysis

Following section reports the out of sample analysis. The models have been tested for 3 horizons i.e. one year, two years and three years. The logit models are estimated using data from June 2001 to December 2015, then till December 2016 and then till December 2017. Table 4-31 and 4-32 reports the out of sample analysis for 50% and 25% probability cutoffs for the logit model estimated from June 2001 to December 2015. At 0.5 cutoff, the models EMERINDUK and EMERIND outperform all the other models with 93.68% of the total cases correctly classified. This is followed by the models EMERINDUS and EMERUS identifying 93.1% of the cases correctly. However, at 0.25 probability cutoff, the model EMERUS outperforms all the other models classifying 92% of the total cases correctly. The models EMERIND and EMERINDUS, classifying 90.23% of the cases correctly confirms the high predictive power of Indian and US sentiment in comparison to the Eurozone sentiment. The models including Euro sentiment i.e. EMERINDUK and EMERUK identify 88.51% and 86.86% of the cases correctly respectively which is lesser than EMERUS and EMERINDUS.

Table 4-33 and 4-34 reports the results for the out of sample analysis for the sample 2001-2016. At 0.5 probability cut off the model EMERINDUK outperforms all the other models with classification accuracy of 94.62% followed by EMERUS and EMERUK with 93.05% and EMERIND and EMERINDUS with 93.01%.

At 0.25 probability cutoff, EMERINDUS outperforms all the other models with 91.4% classification accuracy followed by EMERUS classifying 90.91% of the total observations correctly. Models namely EMERIND and EMERINDUK perform equally in terms of classifying the observations correctly with 88.71% accuracy.

Table.4-35 and 4-36 presents the out of sample analysis for the period 2001-2017. Again at 0.5 cutoff probability, EMERINDUK is found to be outperforming all the other models with 94.95% classification accuracy, followed by EMERUS and EMERUK classifying 93.47% and EMERINDUS and EMERIND classifying 93.43% of the observations correctly. While at 0.25 probability cutoff, EMERUS outperforms all other models with 91.96% observations correctly identified followed by the EMERIND and EMERINDUK classifying 89.39% of the total observations correctly.

Table 4-31:Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2015 at 0.5 cutoff

Table 4-31: Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2015 At 0.5 cutoff

At 0.5 cutoff	EMERIND UK	EMERIND US	EMER US	EMERUK	EMERIND	SENT	UKSENT	MACRO	USENT
Sensitivity Pr(+| D)	68	64	61.54	61.54	68	48	57.69	23.08	57.69
Specificity Pr(-|~D)	97.99	97.99	98.66	97.99	97.99	97.32	97.99	97.32	96.64
Positive predictive value Pr(D| +)	85	84.21	88.89	84.21	85	75	83.33	60	75
Negative predictive value Pr(~D| -)	94.81	94.19	93.63	93.59	94.81	91.77	92.99	87.88	92.9
False + rate for true ~D Pr(+|~D)	2.01	2.01	1.34	2.01	2.01	2.68	2.01	2.68	3.36
False - rate for true D Pr(-| D)	2	6	8.46	8.46	2	52	42.31	76.92	42.31
False + rate for classified + Pr(~D| +)	15	15.79	11.11	15.79	15	25	16.67	40	25
False - rate for classified - Pr(D| -)	5.19	5.81	6.37	6.41	5.19	8.23	7.01	12.12	7.1
Correctly classified	93.68	93.1	93.14	92.57	93.68	90.23	92	86.29	90.86

Table 4-32:Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2015 at 0.25 cutoff

Table 4-32: Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2015 At 0.25 cutoff

At 0.25 cutoff	**EMERINDUK**	**EMERINDUS**	**EMERUS**	**EMERUK**	**EMERIND**	**SENT**	**UKSENT**	**MACRO**	**USENT**
Sensitivity Pr(+| D)	72.00	72.00	84.62	65.38	76.00	72.00	76.92	69.23	69.23
Specificity Pr(-|~D)	91.28	93.29	93.29	90.60	92.62	86.58	87.92	87.25	85.23
Positive predictive value Pr(D| +)	58.06	64.29	68.75	54.84	63.33	47.37	52.63	48.65	45.00
Negative predictive value Pr(~D| -)	95.10	95.21	97.20	93.75	95.83	94.85	95.62	94.20	94.07
False + rate for true ~D Pr(+|~D)	8.72	6.71	6.71	9.40	7.38	13.42	12.08	12.75	14.77
False - rate for true D Pr(-| D)	8.00	8.00	5.38	4.62	4.00	28.00	23.08	30.77	30.77
False + rate for classified + Pr(~D| +)	41.94	35.71	31.25	45.16	36.67	52.63	47.37	51.35	55.00
False - rate for classified - Pr(D| -)	4.90	4.79	2.80	6.25	4.17	5.15	4.38	5.80	5.93
Correctly classified	88.51	90.23	92.00	86.86	90.23	84.48	86.29	84.57	82.86

Table 4-33:Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2016 at 0.5 cutoff

Table 4-33: Prediction Evaluation for Sentiment Logit Models- - Out of Sample for 2001-2016 at 0.5 cutoff

At 0.5 cutoff	EMERIN DUK	EMERIN DUS	EMER US	EMERUK	EMERIN D	SENT	UKSEN T	MACRO	USENT
Sensitivity Pr(+| D)	65.38	61.54	59.26	59.26	61.54	38.46	55.56	22.22	55.56
Specificity Pr(- |~D)	99.38	98.13	98.75	98.75	98.13	97.50	98.75	98.13	96.88
Positive predictive value	94.44	84.21	88.89	88.89	84.21	71.43	88.24	66.67	75.00
Negative predictive value	94.64	94.01	93.49	93.49	94.01	90.70	92.94	88.20	92.81
False + rate for true ~D Pr(+|~D)	0.63	1.88	1.25	1.25	1.88	2.50	1.25	1.88	3.13
False - rate for true D Pr(-| D)	4.62	8.46	0.74	0.74	8.46	61.54	44.44	77.78	4.44
False + rate for classified +	5.56	15.79	11.11	11.11	15.79	28.57	11.76	33.33	25.00
False - rate for classified - Pr(D|	5.36	5.99	6.51	6.51	5.99	9.30	7.06	11.80	7.19
Correctly classified	94.62	93.01	93.05	93.05	93.01	89.25	92.51	87.17	90.91

Table 4-34:Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2016 at 0.25 cutoff

***Table** Error! No text of specified style in document.-34: **Prediction Evaluation for Sentiment Logit Models-** - **Out of Sample for 2001-2016** at 0.25 cutoff*

At 0.25 cutoff	EMERIN DUK	EMERIN DUS	EMER US	EMERU K	EMERI ND	SENT	UKSEN T	MACRO	USENT
Sensitivity Pr(+| D)	73.08	76.92	77.78	62.96	73.08	73.08	74.07	66.67	81.48
Specificity Pr(-|~D)	91.25	93.75	93.13	90.63	91.25	86.88	86.88	86.88	90.00
Positive predictive value Pr(D| +)	57.58	66.67	65.63	53.13	57.58	47.50	48.78	46.15	57.89
Negative predictive value Pr(~D| -)	95.42	96.15	96.13	93.55	95.42	95.21	95.21	93.92	96.64
False + rate for true ~D Pr(+|~D)	8.75	6.25	6.88	9.38	8.75	13.13	13.13	13.13	10.00
False - rate for true D Pr(-| D)	6.92	3.08	2.22	7.04	6.92	26.92	25.93	33.33	18.52
False + rate for classified + Pr(~D| +)	42.42	33.33	34.38	46.88	42.42	52.50	51.22	53.85	42.11
False - rate for classified - Pr(D| -)	4.58	3.85	3.87	6.45	4.58	4.79	4.79	6.08	3.36
Correctly classified	88.71	91.40	90.91	86.63	88.71	84.95	85.03	83.96	88.77

Table 4-35:Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2017 at 0.5 cutoff

Table 4-35: Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2017 at 0.5 cutoff

At 0.5 cutoff	EMERIND UK	EMERIN DUS	EMERU S	EMERU K	EMERIN D	SENT	UKSENT	MACR O	USENT
Sensitivity Pr(+| D)	65.38	61.54	59.26	59.26	57.69	38.46	55.56	22.22	55.56
Specificity Pr(-|~D)	99.42	98.26	98.84	98.84	98.84	98.26	98.84	98.26	97.09
Positive predictive value Pr(D| +)	94.44	84.21	88.89	88.89	88.24	76.92	88.24	66.67	75.00
Negative predictive value Pr(~D| -)	95.00	94.41	93.92	93.92	93.92	91.35	93.41	88.95	93.30
False + rate for true ~D Pr(+|~D)	0.58	1.74	1.16	1.16	1.16	1.74	1.16	1.74	2.91
False - rate for true D Pr(-| D)	4.62	8.46	0.74	0.74	2.31	61.54	44.44	77.78	44.44
False + rate for classified + Pr(~D|)	5.56	15.79	11.11	11.11	11.76	23.08	11.76	33.33	25.00
False - rate for classified - Pr(D| -)	5.00	5.59	6.08	6.08	6.08	8.65	6.59	11.05	6.70
Correctly classified	94.95	93.43	93.47	93.47	93.43	90.40	92.96	87.94	91.46

Table 4-36: Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2017 at 0.25 cutoff

Table 4-36: Prediction Evaluation for Sentiment Logit Models- Out of Sample for 2001-2017 at 0.25 cutoff

At 0.25 cutoff	EMERINDUK	EMERINDUS	EMERUS	EMERUK	EMERIND	SENT	UKSENT	MACRO	USENT
Sensitivity Pr(+| D)	69.23	76.92	77.78	62.96	73.08	73.08	74.07	66.67	81.48
Specificity Pr(- |~D)	92.44	93.60	94.19	91.86	91.86	88.37	87.79	87.79	90.70
Positive predictive value Pr(D| +)	58.06	64.52	67.74	54.84	57.58	48.72	48.78	46.15	57.89
Negative predictive value	95.21	96.41	96.43	94.05	95.76	95.60	95.57	94.38	96.89
False + rate for true ~D Pr(+|~D)	7.56	6.40	5.81	8.14	8.14	11.63	12.21	12.21	9.30
False - rate for true D Pr(-| D)	0.77	3.08	2.22	7.04	6.92	26.92	25.93	33.33	18.52
False + rate for classified + Pr(~D|	41.94	35.48	32.26	45.16	42.42	51.28	51.22	53.85	42.11
False - rate for classified - Pr(D| -	4.79	3.59	3.57	5.95	4.24	4.40	4.43	5.63	3.11
Correctly classified	89.39	91.41	91.96	87.94	89.39	86.36	85.93	84.92	89.45

4.3.3.3. Comparison of AuROC for estimated Logit models- Out of sample analysis

Similar to the in-sample analysis, the AUROC for the estimated logit models for different out of sample periods have been compared and results are presented in Table4-37, 4-38 and 4-39. It is found that area-wise, model EMERINDUS outperforms all the other models in out of sample analysis. The model reports AUC of 96.67% for sample 2001-2015, 94.59% for sample 2001-2016, and 94.9% for sample 2001-2017. Following EMERINDUS is the model EMERINDUK with 96.13%, 94.21%, and 94.39% for the samples 2001-2015, 2001-2016, and 2001-2017 respectively. However, the ROC comparison among the models suggests that with different probability cutoffs, different models perform better. This is in accordance with the results found in the above section using 0.5 and 0.25 probability cutoffs. This means that even though the area covered by the model EMERINDUS is largest among all the models, indicating best performance, models other that EMERINUS may perform better conditional on the probability cutoff chosen.

Table 4-37:ROC comparison for Sentiment Logit Models- Out of Sample (2001-2015)

Model	Observation	Area	Std. Error	Asymptotic Normal [95% conf. Interval]	
osMACRO16	174	0.8792	0.033	0.81452	0.94387
osSENT16	174	0.9332	0.0198	0.89441	0.9719
osUKSENT16	174	0.9353	0.022	0.89223	0.97837
osUSENT16	174	0.9423	0.0221	0.89894	0.98562
osEMERSENT16	174	0.8851	0.0392	0.80835	0.96185
osEMERIND16	174	0.9592	0.0154	0.92898	0.98941
osEMERUK16	174	0.9385	0.0217	0.89608	0.98097
osEMERUS16	174	0.9552	0.0192	0.91762	0.99272
osEMERINDUS16	174	0.9667	0.0136	0.94008	0.99334
osEMERINDUK16	174	0.9613	0.0147	0.93254	0.99014

Ho: area(osMACRO16) = area(osSENT16) = area(osUKSENT16) = area(osUSENT16) = area(osEMERSENT16) = area(osEMERIND16) = area(osEMERUK16) = area(osEMERUS16) = area(osEMERINDUS16) = area(osEMERINDUK16)
chi2(9) = 13.62 Prob>chi2 = 0.1367

Table 4-38:ROC comparison for Sentiment Logit Models- Out of Sample (2001-2016)

Model	Observation	Area	Std. Err.	Asymptotic Normal[95% Conf. Interval]	
OsMACRO17	186	0.8748	0.0321	0.81179	0.93773
osSENT17	186	0.9159	0.028	0.86093	0.9708
osUKSENT17	186	0.9171	0.0292	0.85986	0.97427
osUSENT17	186	0.9216	0.032	0.85898	0.98429
osEMERSENT17	186	0.8791	0.0376	0.80532	0.95285
osEMERIND17	186	0.9389	0.0234	0.89299	0.98489
osEMERUK17	186	0.9224	0.0261	0.87129	0.97342
osEMERUS17	186	0.9401	0.0249	0.89131	0.98898
osEMERINDUS17	186	0.9459	0.0245	0.89781	0.99402
osEMERINDUK17	186	0.9421	0.023	0.89706	0.98708

Ho: area(os17_macro) = area(osSENT17) = area(osUKSENT17) = area(osUSENT17) = area(osEMERSENT17) = area(osEMERIND17) = area(osEMERUK17) = area(osEMERUS17) = area(osEMERINDUS17) = area(osEMERINDUK17)

chi2(9) = 13.07 Prob>chi2 = 0.1594

Table 4-39:ROC comparison for Sentiment Logit Models- Out of Sample (2001-2017)

Model	Observation	Area	Std. Err.	Asymptotic Normal[95% Conf. Interval]	
osMACRO18	198	0.8813	0.0305	0.82143	0.94109
osSENT18	198	0.9215	0.0263	0.86995	0.97308
osUKSENT18	198	0.9226	0.0273	0.8692	0.97606
osUSENT18	198	0.9264	0.0303	0.86706	0.98581
osEMERSENT18	198	0.8833	0.0361	0.81256	0.95399
osEMERIND18	198	0.9403	0.0232	0.89483	0.98577
osEMERUK18	198	0.928	0.0243	0.88031	0.97569
osEMERUS18	198	0.9443	0.0236	0.89811	0.99053
osEMERINDUS18	198	0.949	0.0233	0.9034	0.99464
osEMERINDUK18	198	0.9439	0.0223	0.90007	0.98767

Ho: area(os18_macro) = area(osSENT18) = area(osUKSENT18) = area(osUSENT18) = area(osEMERSENT18) = area(osEMERIND18) = area(osEMERUK18) = area(osEMERUS18) = area(osEMERIND~18) = area(osEMERIND~18)

chi2(9) = 14.59 Prob>chi2 = 0.1027

The best performing models at each probability cut-off have been presented in Table 4-40 and Table 4-41. Considering single sentiment variable models, the U.K. sentiment is found to outperform others at 0.5 probability cutoff while at 0.25 cutoff, the U.S. sentiment is found to be outperforming the other sentiment variables. Similarly, considering models with more than one sentiment variable, the analysis again finds that the models containing U.S. sentiment variables are superior in their predictive performance in comparison to all the other models at probability 0.25 cut off. While at 0.5 cutoff, the models containing U.K. sentiment variables have better predictive performance compared to other alternate models. This indicates that there is a sentiment spillover from other developed and emerging economies which does affect the Indian stock market.

Table 4-40: Best performing models-In sample

Probability cut off	**In sample**	**Out 2001-2015**	**Out 2001-2016**	**Out 2001-2017**
0.5	UKSENT	EMERSENT	UKSENT	UKSENT
0.25	USENT	EMERSENT	USENT	USENT

Table 4-41:Best performing combination models-In sample

Probability cut off	**In sample**	**Out 2001-2015**	**Out 2001-2016**	**Out 2001-2017**
0.5	EMERINDUK	EMERINDUK/EMERIND	EMERINDUK	EMERINDUK
0.25	EMERUS	EMERUS	EMERINDUS	EMERUS

Finally, the ROC analysis reveals the AuROC is highest for the models containing the U.S. sentiment i.e. USENT for individual sentiment models and EMERINDUS for models including more than one sentiment.

A growing number of studies have substantiated the idea that international investor sentiment matters in home stocks' valuation. A study by Hwang (2011) found that American investor sentiment affected the demand for securities from a specific country and caused deviation of security prices from their fundamental values. Another study Aissia (2016), investigated the impact of both foreign and home investor sentiment of stock returns for French stock market. Thestudy found strong evidence that both the home and foreign investor sentiment affect stock returns. Specifically, to the context of the U.S. sentiment, a number of studies have examined the impact of shocks originating from the U.S. on the international stock markets. These shocks include macroeconomic news announcements, monetary policy, volatility, election cycles and sentiment originating from the U.S. The study (Soydemir, 2000) examined the transmission patterns of stock market movements between the U.S. and emerging market economies and found a significant impact of the U.S. stock market on emerging stock markets at varying degrees. There exists a mixed evidence for the integration of the Indian stock market with the developed nations. The study by Wong et al. (2005) found that the Indian stock market was integrated with the U.S., the U.K. and Japan for the post liberalization period. The study by Tripathi and Sethi (2010) examined the integration of the Indian stock market with four major stock markets namely, the U.S., the U.K., Japan and China from 1998-2008. The results indicated that Indian stock markets (NIFTY) is not integrated with any of the developed markets analyses except for the U.S. The study by Wu et al. (2015) also found the U.S. stock market co-integrated with the Asian stock markets (including India) during the pre- and post- 2008 financial crisis periods. On the contrary, Mukherjee and Mishra (2005) has found that there is no integration between Indian stock market and thedeveloped nations. The present study is in line with the studies supporting the integration of Indianstock markets with the U.S. stock markets.

Predictive power of ANN models

The second approach utilized in present study is ANN. The optimal number of neurons in the single layer feedforward network is decided based on the classification accuracy obtained using k-fold cross validation. With k=5, the optimal number of neurons has been found to be 36. The network is tested from 1 neuron to 40 neurons in the hidden layer so as to avoid the problem of overfitting. Table. 4-42 presents the classification accuracies obtained for different number of neurons with k= 5. The procedure is repeated for five times and the average of the classification accuracies is computed for each number of neuron. This is done to ensure the choice of the best performance neural network. The maximum average classification accuracy is observed when number of neurons is 36.

Table 4-42:Selection of best performance ANN using K-fold cross validation

Number of Neurons	Classification accuracy-1	Classification accuracy-2	Classification accuracy-3	Classification accuracy-4	Classification accuracy-5	Average Classification accuracy
1	88.63	88.15	88.62	89.58	89.57	88.91
2	87.67	90.03	90.99	88.16	87.21	88.81
3	94.31	86.27	88.62	90.53	85.78	89.10
4	88.64	88.64	90.03	89.57	87.21	88.82
5	91.02	92.91	88.64	90.04	91.46	90.82
6	91.46	92.88	90.08	88.14	89.13	90.34
7	91.02	90.51	92.89	94.32	90.52	91.85
8	94.31	90.10	95.73	86.71	92.89	91.95
9	87.19	89.13	92.40	89.11	84.83	88.53
10	90.54	89.08	94.32	90.54	90.06	90.91
11	93.84	93.84	90.52	89.09	91.94	91.85
12	95.75	91.94	89.58	90.09	89.09	91.29
13	90.53	93.83	90.04	91.03	93.37	91.76
14	93.37	90.04	89.11	89.56	94.31	91.28
15	92.44	89.58	91.48	89.59	89.59	90.54
16	92.87	91.47	89.10	90.06	89.56	90.61
17	87.69	92.88	88.14	91.00	91.02	90.14
18	94.31	91.45	94.32	92.43	91.94	92.89
19	90.52	95.74	87.66	92.43	94.32	92.13

20	92.41	93.37	91.94	89.13	90.97	91.56
21	91.96	91.45	92.41	94.32	88.16	91.66
22	93.82	88.64	87.69	90.52	90.07	90.15
23	92.89	94.78	92.89	93.36	87.21	92.23
24	91.50	91.47	91.46	90.99	91.01	91.28
25	89.10	91.95	89.11	95.25	93.82	91.85
26	91.45	91.93	92.90	92.90	92.90	92.42
27	94.33	91.01	92.92	90.06	91.48	91.96
28	93.37	89.11	95.75	95.26	90.99	92.89
29	92.90	93.85	92.44	92.90	90.03	92.43
30	92.91	92.91	89.09	93.36	93.83	92.42
31	91.01	92.48	95.24	95.25	90.07	92.81
32	92.89	89.58	91.02	91.96	91.93	91.48
33	93.37	90.52	90.97	93.83	90.51	91.84
34	94.30	90.55	90.51	92.87	95.27	92.70
35	88.68	92.43	94.31	90.52	94.32	92.05
36	94.33	91.45	96.22	91.95	91.48	93.09
37	93.36	87.24	93.84	86.74	91.98	90.63
38	90.99	93.37	90.51	88.18	91.47	90.90
39	92.43	90.53	91.46	93.37	94.80	92.52
40	91.51	88.21	93.36	91.96	87.21	90.45

The construction of the two-layer feedforward pattern classification neural network is as follows:

PNN:

1) Input layer: 6 input units/neuron (for six macroeconomic indicators).

2) One Hidden layer: 36 neurons in the hidden layer.

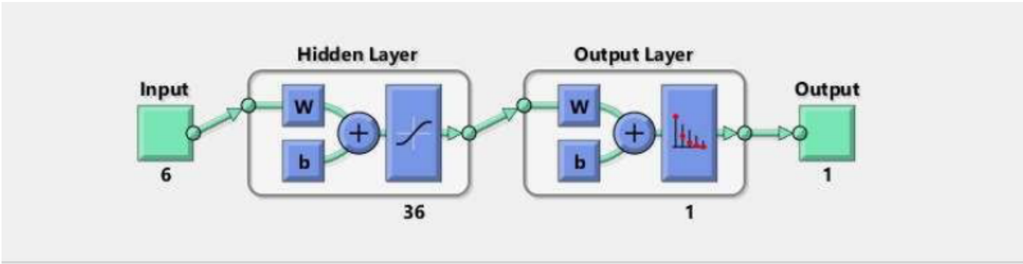
3) One Output layer with a targeted value equal to 1 for crisis periods and 0 for tranquil periods with 1 neuron.

4) Training function: Scaled Conjugate Gradient

5) Transfer function: Sigmoid transfer function in the hidden layer and a softmax transfer function in the output layer.

6) Cross entropy is taken to be the performance function.

Figure 4-6:Construct for the ANN



Source: Based on Author's calculations

The calibration scores from within sample and out of sample for both Logit and ANN models are reported in Table 4-43. The comparison among the estimated Logit models shows that the models EMERINDUK and EMERINDUS outperform all the other models for in sample analysis with lowest QPS values of 0.0898 and 0.089 respectively. The highest QPS value is associated with the model MACRO which consists of only the macroeconomic variables. This proves that sentiment does carry a predictive ability in estimating the probability of a stock market crisis. The QPS values are also compared for out of samples with three different samples i.e. 2001-2015, 2001-2016, and 2001-2017. It can be observed that the QPS values increase for the out of sample analysis which is expected as the accuracy decreases when the model is estimated without full sample and tested on the hold out sample. For the hold out periods of one year and two years, the QPS values do not vary much. For most of the models, the QPS is found to be almost same for both one year and two year hold out periods. The model EMERINDUS still outperforms all the other models with lowest QPS values of 0.0786 for three year hold out, 0.0889 for two year hold out, and 0.089 for the one year hold out period for out of sample analysis. Looking at the results for ANN models, the difference is clear in the QPS values. In terms of accuracy, ANN models outperform the Logit models as evident from the values for in-sample and out-of-sample analysis. Comparing the in-sample results, it can be seen that the QPS values range from 0.0412 (EMERINDUS) to 0.1138 (SENT) while the range for Logit models is between 0.089 and 0.1584. Since, the QPS value is inversely proportional to the degree of accuracy of the model, the results indicate the superior performance of ANN models. Again, the model EMERINUS outperforms all the other models with the lowest QPS value of 0.0464 for one year hold out period analysis.

Table 4-43:Calibration Scores- Logit vs ANNs- In sample and Out of sample

	ANN				**LOGIT**			
	OS_18	**OS_17**	**OS_16**	**IN_SAMPLE**	**OS_18**	**OS_17**	**OS_16**	**IN_SAMPLE**
MACRO	0.1492	0.0826	0.1030	0.1106	0.1577	0.1575	0.1483	0.1584
SENT	0.1100	0.0978	0.0835	0.1138	0.1267	0.1268	0.1156	0.1267
USENT	0.0881	0.0651	0.1256	0.0738	0.1101	0.1101	0.0986	0.1100
UKSENT	0.0718	0.0907	0.0682	0.0826	0.1151	0.1152	0.1037	0.1150
EMERSENT	0.0537	0.0778	0.0567	0.1136	0.1203	0.1201	0.1096	0.1201
EMERIND	0.0748	0.1237	0.0549	0.0415	0.1009	0.0992	0.0842	0.1009
INDUS	0.0780	0.0715	0.0543	0.0831	0.1026	0.1027	0.0898	0.1026
INDUK	0.0750	0.0748	0.0836	0.0570	0.1058	0.1058	0.0934	0.1058
EMERUK	0.0983	0.0445	0.0351	0.0721	0.1024	0.1024	0.0927	0.1024
EMERUS	0.0871	0.0852	0.0907	0.0656	0.0948	0.0948	0.0856	0.0948
EMERINDUS	0.0464	0.0730	0.0876	0.0412	0.0890	0.0889	0.0786	0.0890
EMERINDUK	0.0676	0.0637	0.0904	0.0636	0.0898	0.0900	0.0796	0.0898

The present section on the analysis of the investor sentiment as a potential early warning indicator reveals its high significance in predicting the probability of a stock market crisis in Indian context. The study finds FIIs and real interest rates as highly significant indicators in anticipating a stock market crisis. In addition to the domestic investor sentiment, the foreign investor sentiment, specifically the U.S. sentiment, if also found to be playing a significant role in developing an EWS model for Indian stock market. Development of EWS using logit and ANNs also suggests the improved prediction by ANNs in comparison to the traditional techniques like logit models.

4.4. Hypotheses Testing

The results of the formulated hypotheses are presented below:

Section I: Analysis of banking crisis

The outcome of data analysis carried out in this section reveals that the selected macroeconomic and financial variables such as Yield to Maturity, spread between bank rate Yield to Maturity, growth in Broad money to foreign exchange reserves do act as warning indicators and have significant impact on the likelihood of a banking crisis. Therefore, the null hypothesis H_{01} is rejected. The findings on comparison of traditional techniques and ML techniques indicate that ANNs outperform the limited dependent variable models in terms of predicting the probability of a banking crisis. Therefore, the null hypothesis H_{02} is rejected.

H_{01}: The selected macroeconomic variables have no impact on the likelihood of a banking crisis. - Rejected

H_{02}: The predictive performance of logit models is better than the predictive performance of ANNs. - Rejected

Section II: Analysis of stock market crisis

The findings of section II suggest that the selected macroeconomic and financial variables have significant impact on the likelihood of a stock market crisis. Specifically, variables like FIIs and real interest rates came out to be highly significant in predicting the probability of a stock market crisis. Therefore, the hypothesis H_{03} has been rejected. The analysis also examines the role of sentiment variables in contributing to the predictive ability of the developed EWS. The findings indicate a high significance of sentiment variables representing domestic (Indian), U.S., European and emerging market sentiment. Therefore, the hypotheses $H_{04,}$ $H_{05,}$ and H_{06} have been rejected. The results of the comparison between the limited dependent variable approach and ANNs indicate that ANNs outperform the Logit approach in terms of predicting the probability of a stock market crisis. Therefore, the hypothesis H_{07} has been rejected as well.

H_{03}: The selected macroeconomic variables have no impact on the likelihood of a stock market crisis. - Rejected

H_{04}: The domestic investor sentiment plays no role in predicting a stock market crisis in Indian context. - Rejected

H_{05}: The emerging market sentiment plays no role in predicting a stock market crisis in Indian context. - Rejected

H_{06}: The developed market sentiment plays no role in predicting a stock market crisis in Indian context. - Rejected

H_{07}: The predictive performance of logit models is better than the predictive performance of ANNs. – Rejected