Prediction of the Query of Search Engine Using Back Propagation Algorithm

Deepti Gupta[†] , Antima Puniya[††] and Komal Kumar Bhatia[†††]

†School of Computer Science and Information Technology, Shobhit University Meerut, INDIA
†† Shobhit Insitute of Technology, Gangoh, INDIA
†††Department of Computer Science, YMCA University of Science, Faridabad, INDIA

Abstract: The information user is depending on the Search Engine; therefore search engines are required as a prediction system to predict the next query hit by the user. Web mining techniques, like neural network can be used for this purpose. In this paper, a novel approach to predict the oncoming query for the search engine has been highlighted. This approach helps search engine to predict the oncoming query domain, by which search engine keeps most relevant web pages in the repository.

Keywords: Search Engine, Query, Neural Network, Backpropagtion Algorithm

1. Introduction:

Ten of Thousands of user surfers the internet through search engine [01, 02 and 06] to achieve their query results. A user enters the query into query interface [04] of search engine to find needed information. The search queries [17] are individual. It may be match from standard query languages. There are following four types of search queries:

- Informational queries – These queries are generally used and it covers a wide topic and gives thousands of relevant answers.
- Navigational queries – These queries are in the form of a single website.
- Transactional queries – These queries are referred to a particular action, like shopping or downloading a screen saver.
- Connectivity queries – These queries are based on the connectivity of the indexed web graph.

The act of a search engine [18, 19 and 20] is limited to the problem of "Information Overkill". Thus, there is a need to develop an approach for predicting the next query. In this paper, a novel approach for predicting the oncoming query on the search engine which uses neural network. This proposed prediction system helps to generate the oncoming query.

2. Related Work:

HweeHwa Pang et. al. [15, 16] formulated the properties of processing a text search query by adapting on existing threshold-based algorithms, and develop an authentication scheme for checking the validity of a result and also improve efficiency and practicality of the solution through an empirical evaluation with real documents and benchmark queries.

Ramon Lawrence [26] surveyed on the application of neural networks in forecasting stock market prices. With their ability to discover patterns in nonlinear and chaotic systems, neural networks [03, 05] offer the ability to predict market directions more accurately than current techniques. Common market analysis techniques such as technical analysis, fundamental analysis, and regression are discussed and compared with neural network performance. Also, the Efficient Market Hypothesis (EMH) is presented and contrasted with chaos theory and neural networks. He refutes that the EMH based on previous neural network work. Finally, future directions for applying neural networks for the financial markets are discussed.

Nakamura et.al. [23] introduced word category prediction system and it is used to implement an accurate word recognition system. Traditional statistical approaches require considerable training data to estimate the probabilities of word sequences, and many parameters to memorize the probabilities.

A critical look at the available literature indicates that till date only neural network [12, 13, 14 and 21] which was used for predication in any other applications like stock market, weather forecasting. Specifically a neural network's ability is to predict the future trends in Stock Market Indices which will be tested. Accuracy will be compared against a traditional forecasting method, using multiple linear regression analysis. In this paper the PMI (Point wise mutual information) has been used to make the predication system. It defines the maximum probability of each event. In the neural network each queries identified by their PMI value. With the help of artificial neural network the predication system is proposed for the oncoming queries.

3. Proposed Work:

This paper shows a novel approach to predict the oncoming query on the search engine [22] using neural network. This proposed prediction system helps us to generate the oncoming query. The proposed prediction system follows these steps:

Step-1 Selection of Domain Name:
In this present work four domains have been selected. These domains are follows:

- Education
- Entertainment
- Travel
- Sports

Step-2 Queries asked by the user from selected domain:
A survey has been conducted in Shobhit University, Meerut among final year students. They have been asked for 25 favorite queries in all the four domains. These queries are shown in section 4.

Step-3 Training the neural network:
Training and testing the artificial neural network for user queries based on frequency and PMI (Point Wise mutual Information) value. Finally 25 queries from all the domains have been selected for testing and training the neural network. The queries now have been triggered at Google search browser and frequency of these queries has been kept in the database for further processing. All the queries are further broken into the individual keywords and the frequency of individual keywords count has also been taken in account. The PMI (Point Wise mutual Information) of each query has been calculated, which basically defines the maximum probability of the event. In the neural network each queries identified by their PMI value.
In the neural network the PMI values are taken as inputs among twenty five queries. Twenty queries have been used for training the neural network and rest of all queries has been taken as testing inputs to test the efficiency of the neural network. Due to the usage of back propagation algorithm (supervised learning algorithm) the target values has been required. The target values have been tagged as 0,1,2,3 on the basis of PMI values of training dataset (maximum the PMI value tagged). Once the neural network [27 and 28] has been trained by the given training data, it can be used further for mining the large data sets.

Step-4 Predict the next query:
This last step shows oncoming query for next user. This proposed model has prediction factor which is easier to search to next query for the user.

The crawler [08, 09, 11 and 25] maintains a list of unvisited URLs called the frontier. The list is initialized with seed URLs which may be provided by a user or another program. Crawler crawls all web pages stored in the repository. Indexer indexes all the keywords stored in the local repository. The user sends a query through user interface and query processor processes the query and identifies the domain name. Find the PMI value of each query. The neural network is the tool which learns using some rules and conditions invented by incoming queries and work for oncoming queries. In the present work neural network is being used for prediction of the oncoming data on the bases of incoming queries. The proposed architecture is shown in figure 3.1 and architecture of Feed Forward neural network (4:5:1) [07, 22] is shown in figure 3.2.

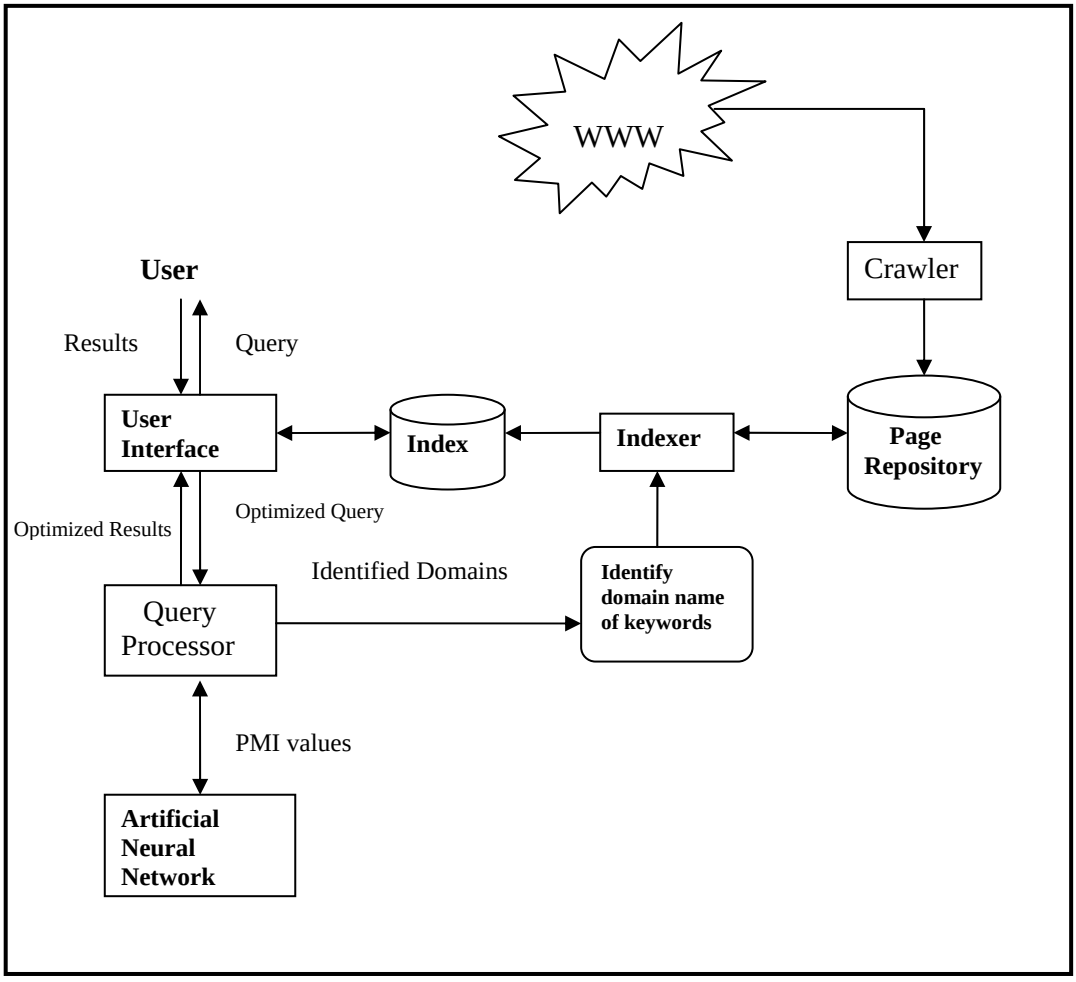

Figure 3.1: Block diagram of the Proposed Query Prediction System

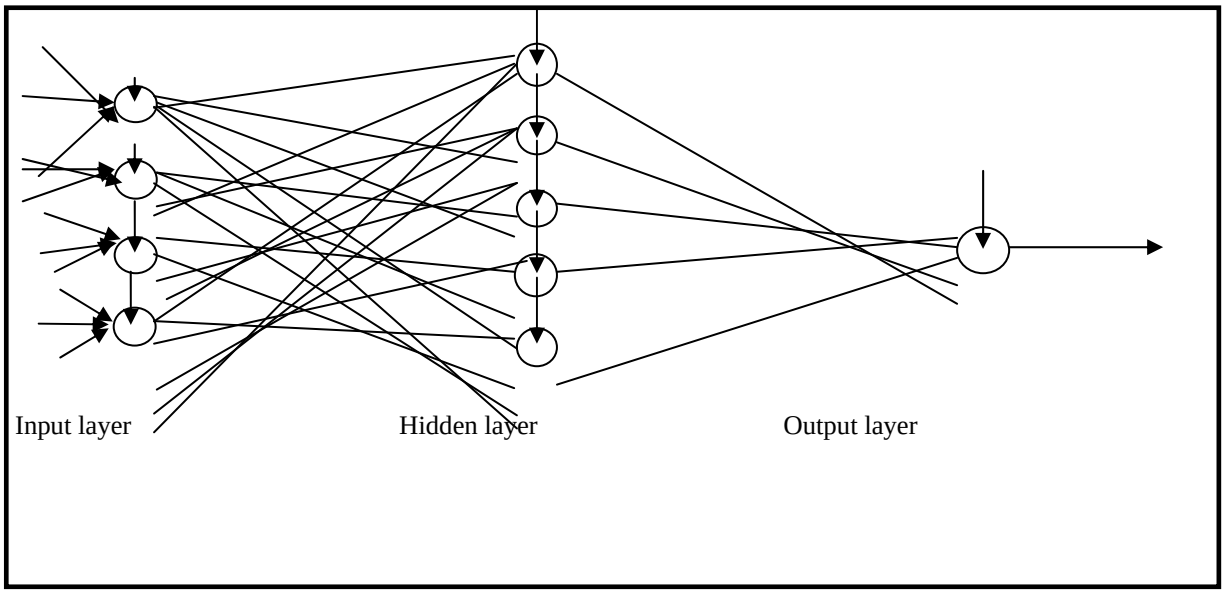

Figure: 3.2 Architecture of Feed Forward neural network [4:5:1] using in proposed Query Prediction system

This proposed prediction system has a methodology. The following the flow chart of the methodology is shown in figure 3.3.

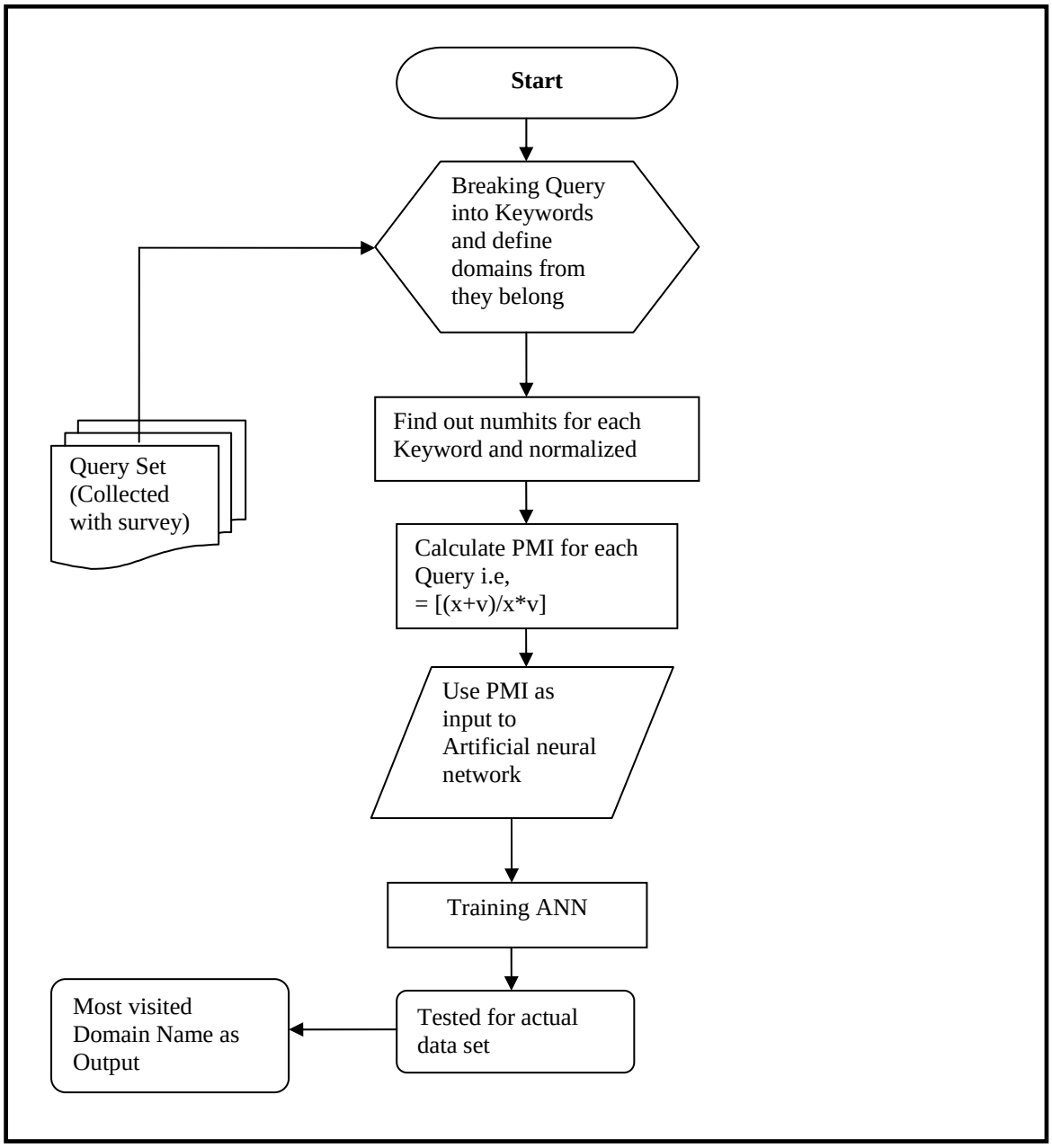

Figure 3.3: Flow chart of the proposed system

4. Example:

In this example there are four domains. Each domain has 25 queries. These are incoming queries and prediction of the oncoming queries. All domain names are shown below.

(i) Education:

Education Domain has been selected for query prediction due to its highest relevance. The following 25 queries have been selected from the database available from the students.

1	Technical Education
2	Master of Arts
3	Master of Technology

4	Information Technology
5	Bachelor Of Technology
6	Social Science
7	Master Of Business administration
8	Education Of Women
9	Primary Education
10	Higher Education
11	Adult Education
12	Education Counseling
13	Method of teaching
14	Research Publication
15	Teacher Training
16	Text Book Of History
17	Top 100 Universities
18	College Of Engineering
19	Business Schools
20	Teaching Faculty
21	Engineering Entrance
22	Course at IIT
23	Secondary Education
24	National Schools
25	Nursery training

(ii) **Entertainment:**

Entertainment Domain has been selected for query prediction due to its highest demand. The following 25 queries have been selected from the database available from the students.

1	Hindi Movies
2	Punjabi Movies
3	Gujrati Movies
4	English Movies
5	Hindi Songs
6	Punjabi Songs
7	Gujhrati Songs
8	English Songs
9	Old Songs
10	Latest Songs
11	Pop Music
12	Puzzle Games
13	Shahrukh Khan Movies
14	Salman Khan Movies
15	Latest movies
16	Latest Music
17	Pop Star
18	Amitabh Movies
19	Madhuri Songs
20	Lata Mangeshkar Songs

21	Asha Bhonsle Songs
22	Kishor Kumar songs
23	Sonu Nigam Songs
24	Classical Songs
25	Dance Beats

(iii) Travel:

Travel Domain has been selected for query prediction due to its highest significance in day to day life. The following 25 queries have been selected from the database available from the students.

1	Travel to Mumbai
2	Travel to Delhi
3	Travel to Srinagar
4	Travel to Orissa
5	Travel to Kerala
6	Travel to Gujarat
7	Travel to Jaipur
8	Weather of Mumbai
9	Weather of Delhi
10	Weather of Srinagar
11	Weather of Orissa
12	Weather of Kerala
13	Flights to Kerala
14	Flights to USA
15	Flights to Jaipur
16	Flights to Goa
17	Flights to Srinagar
18	Flights to Landon
19	Flights to Thailand
20	Trains to Mumbai
21	Trains to Kerala
22	Trains to Darjeeling
23	Flights to Canada
24	Hotels in Mumbai
25	Hotels in Kerala

(iv) Sports

Sports Domain has been selected for query prediction due to its highest popularity. The following 25 queries have been selected from the database available from the students.

1	Sports Magazine
2	Sports in India
3	National Games
4	Test Cricket
5	One Day Cricket
6	Common Wealth Games
7	Asian Games
8	Olympic Games

9	Children Games
10	Sports In Delhi
11	Hockey In India
12	Indian Cricket
13	Cricket Stadium
14	History Of Cricket
15	Baseball Game
16	Indoor Games
17	Outdoor Games
18	Racing Cars
19	Racing Boats
20	Long Jump
21	Swimming Sports
22	High Jump
23	Dangerous Games
24	Tennis Stadium
25	Football Match

Extract all the above mentioned queries into keywords and count the frequency of individual keywords. Calculate PMI (Point Wise Mutual Information) value of each query has been shown on the tables 4.1, 4.2, 4.3 and 4.4. There have selected twenty PMI values from each table for training and also selected five PMI values for testing data. There is need to structure of neural network and these specifications are shown below:

Neural Network Model Specifications
The neural network model used in this work has following specifications:
Number of Inputs=20x4
Number of Neurons in Input layer=4
Number of Neurons in Hidden layer=5
Number of Neurons in Output layer=1
Biases used
At input layers=4
At Hidden layer=5
At Output layer=1
Activation Functions used
For Input layer=Piece-wise linear
For Hidden layer=Sigmoid
For Output layer=Sigmoid
Error criteria used=Mean Square Error
Target accuracy=0.00000015

Table 4.1: Hit-ratio for Travel Query

S.No.	Main Query (x+v)	HR (x+v)	NHR (x+v/min(x+v)	sub-query1 (x)	HR1	NHR1	sub-query2 (v)	HR2	NHR2	PMI x+v/(x*v)
Training data										
1	Travel to Mumbai	33200000	33.2	Travel	796000000	796	Mumbai	63700000	63.7	0.65

2	Travel to Delhi	79300000	79.3	Travel	796000000	796	Delhi	82100000	82.1	1.21
3	Travel to Srinagar	1350000	1.4	Travel	796000000	796	Srinagar	5700000	5.7	0.30
4	Travel to Orissa	1520000	1.5	Travel	796000000	796	Orissa	82300	0.0823	23.20
5	Travel to Kerala	5120000	5.1	Travel	796000000	796	Kerala	218000	0.218	29.51
6	Travel to Gujarat	280000	0.3	Travel	796000000	796	Gujarat	1910000	1.91	0.18
7	Travel to Jaipur	3370000	3.4	Travel	796000000	796	Jaipur	10500000	10.5	0.40
8	Whether of Mumbai	6840000	6.8	Whether	2650000	2.65	Mumbai	63700000	63.7	40.52
9	Whether of Delhi	18000000	18.0	Whether	2650000	2.65	Delhi	82100000	82.1	82.73
10	Weather of Srinagar	873000	0.9	Whether	2650000	2.65	Srinagar	5700000	5.7	57.80
11	Weather of Orissa	603000	0.6	Whether	2650000	2.65	Orissa	82300	0.0823	2764.85
12	Weather of Kerala	2210000	2.2	Whether	2650000	2.65	Kerala	218000	0.218	3825.51
13	Flights to Kerala	1610000	1.6	Flights	161000000	161	Kerala	218000	0.218	45.87
14	Flights to USA	17800000	17.8	Flights	161000000	161	USA	875000000	875	0.13
15	Flights to Jaipur	945000	0.9	Flights	161000000	161	Jaipur	10500000	10.5	0.56
16	Flights to Goa	1850000	1.9	Flights	161000000	161	Goa	22400000	22.4	0.51
17	Flights to Srinagar	663000	0.7	Flights	161000000	161	Srinagar	5700000	5.7	0.72
18	Flights to Landon	18000000	18.0	Flights	161000000	161	Landon	900000000	900	0.12
19	Flights to Thailand	350000	0.4	Flights	161000000	161	Thailand	204000000	204	0.01
20	Trains to Mumbai	180000	0.2	Trains	201000000	201	Mumbai	63700000	63.7	0.01
							Testing data			
21	**Trains to Kerala**	**439000**	**0.4**	**Trains**	**201000000**	**201**	**Kerala**	**218000**	**0.218**	**10.02**
22	**Trains to Darjeeling**	**4260000**	**4.3**	**Trains**	**201000000**	**201**	**Darjeeling**	**3000000**	**3**	**7.06**
23	**Flights to Canada**	**69800000**	**69.8**	**Flights**	**161000000**	**161**	**Canada**	**354000000**	**354**	**1.22**
24	**Hotels in Mumbai**	**3000000**	**3.0**	**Hotels**	**658000000**	**658**	**Mumbai**	**63700000**	**63.7**	**0.07**
25	**Hotels in Kerala**	**2600000**	**2.6**	**Hotels**	**658000000**	**658**	**Kerala**	**218000**	**0.218**	**18.13**

Table 4.2 Hit-ratio for Entertainment Query

S.No.	Main Query (x+v)	HR (x+v)	NHR (x+v/min(x+v)	sub-query1 (x)	HR1	NHR1	sub-query2 (v)	HR2	NHR2	PMI x+v/(x*v)
							Training data			
1	Hindi Movies	69400000	69.4	Hindi	195000000	195	Movies	977000000	977	0.36
2	Punjabi Movies	3300000	3.3	Punjabi	40900000	40.9	Movies	977000000	977	0.08
3	Guajarati Movies	200000	0.2	Guajarati	34400000	34.4	Movies	977000000	977	0.01
4	English Movies	66000000	66.0	English	2200000000	2200	Movies	977000000	977	0.03
5	Hindi Songs	14700000	14.7	Hindi	2200000000	2200	Songs	466000000	466	0.01
6	Punjabi Songs	39900000	39.9	Punjabi	2200000000	2200	Songs	466000000	466	0.04

7	Guajarati Songs	3190000	3.2	Guajarati	2200000000	2200	Songs	466000000	466	0.00
8	English Songs	113000000	113.0	English	2200000000	2200	Songs	466000000	466	0.11
9	Old Songs	30200000	30.2	Old	6640000000	6640	Songs	466000000	466	0.01
10	Latest Songs	22700000	22.7	Latest	13700000	13.7	Music	466000000	466	3.56
11	Pop Music	74800000	74.8	Pop	36300000	36.3	Games	162000000	162	12.72
12	Puzzle Games	14500000	14.5	Puzzle	7900000	7.9	Movies	1480000000	1480	1.24
13	Shahrukh Khan Movies	2140000	2.1	Shahrukh Khan	4270000	4.27	Movies	977000000	977	0.51
14	Salman Khan Movies	4480000	4.5	Salman Khan	10500000	10.5	Movies	977000000	977	0.44
15	Latest movies	59100000	59.1	Latest	13700000	13.7	Movies	977000000	977	4.42
16	Latest Music	91000000	91.0	Latest	13700000	13.7	Music	162000000	162	41.00
17	Pop Star	16300000	16.3	Pop	363000000	363	Star	389000000	389	0.12
18	Amitabh Movies	9670000	9.7	Amitabh	4720000	4.72	Movies	977000000	977	2.10
19	Madhuri Songs	5260000	5.3	Madhuri	2760000	2.76	Songs	466000000	466	4.09
20	Lata Mangeshkar Songs	1600000	1.6	Lata Mangeshkar	3660000	3.66	Songs	466000000	466	0.94
							Testing data			
21	**Asha Bhonsle Songs**	**182000**	**0.2**	**Asha Bhonsle**	**276000**	**0.276**	**Songs**	**466000000**	**466**	**1.42**
22	**Kishor Kumar songs**	**1700000**	**1.7**	**Kishor Kumar**	**2240000**	**2.24**	**Songs**	**466000000**	**466**	**1.63**
23	**Sonu Nigam Songs**	**879000**	**0.9**	**Sonu Nigam**	**1580000**	**1.58**	**Songs**	**466000000**	**466**	**1.19**
24	**Classical Songs**	**20800000**	**20.8**	**Classical**	**119000000**	**119**	**Songs**	**466000000**	**466**	**0.38**
25	**Dance Beats**	**17900000**	**17.9**	**Dance**	**389000000**	**389**	**Beats**	**55300000**	**55.3**	**0.83**

Table 4.3 Hit-ratio for Education Query

S.No.	Main Query (x+v)	HR (x+v)	NHR (x+v/min(x+v)	sub-query1 (x)	HR1	NHR1	sub-query2 (v)	HR2	NHR2
					Training data				
1	Technical Education	24600000	24.6	Technical	377000000	377	Education	3020000000	3020
2	Master of Arts	204000000	204.0	Master	436000000	436	Arts	498000000	498
3	Master of Technology	212000000	212.0	Master	436000000	436	Technology	742000000	742
4	Information Technology	749000000	749.0	Information	3070000000	3070	Technology	742000000	742
5	Bachelor Of Technology	4130000	4.1	Bachelor	1690000	1.69	Technology	742000000	742
6	Social Science	75100000	75.1	Social	985000000	985	Science	1390000000	1390
7	MBA	4900000	4.9	Master	436000000	436	Business Admin.	28400000	28.4
8	Education Of Women	572000000	572.0	Education	3020000000	3020	Women	3010000000	3010

9	Primary Education	22600000	22.6	Primary	296000000	296	Education	3020000000	3020
10	Higher Education	62100000	62.1	Higher	3410000	3.41	Education	3020000000	3020
11	Adult Education	21300000	21.3	Adult	331000000	331	Education	3020000000	3020
12	Education Counseling	13300000	13.3	Education	3020000000	3020	Counseling	63100000	63.1
13	Method of teaching	45200000	45.2	Method	358000000	358	Teaching	16400000	16.4
14	Research Publication	185000000	185.0	Research	752000000	752	Publication	85800000	85.8
15	Teacher Training	28100000	28.1	Teacher	203000000	203	Training	45000000	45
16	Text Book Of History	80900000	80.9	Text Book	248000000	248	History	1170000000	1170
17	Top 100 Universities	2740000	2.7	Top 100	1010000000	1010	Universities	246000000	246
18	College Of Engineering	22800000	22.8	College	66400000	66.4	Engineering	74300000	74.3
19	Business Schools	85200000	85.2	Business	1610000000	1610	Schools	5050000000	5050
20	Teaching Faculty	32700000	32.7	Teaching	164000000	164	Faculty	83600000	83.6
								Testing data	
21	**Engineering Entrance**	**4750000**	**4.8**	**engineering**	**1340000000**	**1340**	**Entrance**	**78000000**	**78**
22	**Course at IIT**	**1910000**	**1.9**	**Course**	**504000000**	**504**	**IIT**	**95100000**	**95.1**
23	**Secondary Education**	**2970000**	**3.0**	**Secondary**	**13400000**	**13.4**	**Education**	**3020000000**	**3020**
24	**National Schools**	**534000000**	**534.0**	**National**	**112000000**	**112**	**Schools**	**5050000000**	**5050**
25	**Nursery Training**	**15700000**	**15.7**	**Nursery**	**251000000**	**251**	**Training**	**450000000**	**450**

Table 4.4 Hit-ratio for Sports Query

S.No.	Main Query (x+v)	HR (x+v)	NHR (x+v/min(x+v)	sub-query1 (x)	HR1	NHR1	sub-query2 (v)	HR2	NHR2	
									Training data	
1	Sports Magazine	53100000	53.1	Sports	656000000	656	Magazine	966000000	966	0.083
2	Sports in India	97700000	97.7	Sports	656000000	656	India	650000000	650	0.229
3	national Games	441000000	441.0	National	2390000000	2390	Games	1510000000	1510	0.122
4	Test Cricket	24100000	24.1	Test	748000000	748	Cricket	99500000	99.5	0.323
5	One Day Cricket	52400000	52.4	One Day	2160000000	2160	Cricket	99500000	99.5	0.243
6	Common Wealth Games	51500000	51.5	Common Wealth	194000000	194	Games	1510000000	1510	0.175
7	Asian Games	57700000	57.7	Asian	188000000	188	Games	1510000000	1510	0.203
8	Olympic Games	68500000	68.5	Olympic	69000000	69	Games	1510000000	1510	0.657
9	Children Games	209000000	209.0	Children	1340000000	1340	Games	656000000	656	0.237
10	Sports In Delhi	76500000	76.5	Sports	656000000	656	Delhi	124000000	124	0.940
11	Hockey In India	1090000	1.1	Hockey	4660000	4.66	India	650000000	650	0.359

12	Indian Cricket	33400000	33.4	Indian	270000000	270	Cricket	99500000	99.5	1.243
13	Cricket Stadium	64600000	64.6	Cricket	99500000	99.5	Stadium	686000000	686	0.946
14	History of Cricket	27400000	27.4	History	984000000	984	Cricket	99500000	99.5	0.279
15	Baseball Game	18300000	18.3	Baseball	119000000	119	Games	15100000	15.1	10.18
16	Indoor Games	5690000	5.7	Indoor	76600000	76.6	Games	1510000000	1510	0.049
17	Outdoor Games	15200000	15.2	Outdoor	206000000	206	Games	1510000000	1510	0.048
18	Racing Cars	14600000	14.6	Racing	196000000	196	Cars	72500000	72.5	1.027
19	Racing Boats	15850000	15.9	Racing	196000000	196	Boats	368000000	368	0.219
20	Long Jump	24600000	24.6	Long	1130000000	1130	Jump	656000000	656	0.033
Testing data										
21	**Swimming Sports**	**92400000**	**92.4**	**Swimming**	**115000000**	**115**	**Sports**	**368000000**	**368**	**2.183**
22	**High Jump**	**27100000**	**27.1**	**High**	**1540000000**	**1540**	**Jump**	**656000000**	**656**	**0.026**
23	**Dangerous Games**	**56600000**	**56.6**	**Dangerous**	**538000000**	**538**	**Games**	**1510000000**	**1510**	**0.069**
24	**Tennis Stadium**	**347000000**	**347.0**	**Tennis**	**165000000**	**165**	**Stadium**	**68600000**	**68.6**	**30.65**
25	**Football Match**	**28101000**	**28.1**	**Football**	**186000000**	**186**	**Match**	**348000000**	**348**	**0.434**

The extracted frequency of the data has been classified and necessary moderation has been carried out. The Point Wise Mutual Information (PMI) has been calculated from the moderated data. The obtained PMI for all hundred queries have been applied to the neural network model. The neural network model is trained to the following specifications:

Maximum results shown at a time=100;
Maximum number of iterations=100000;
Maximum allowed mean square error=0.00000150;

Number of training inputs = 20
Number of testing inputs = 5
The input matrix given to the model: (20 inputs for 4 commodities)

Columns 1 through 11

```
   (T)*  0.6500   1.2100   0.3000   3.2000   2.5100   0.1800   2.5200   2.7300   0.8000   0.8500   0.5100
   (E)*  0.3600   0.0800   0.0100   0.0300   0.0100   0.0400   0.0030   0.1100   0.0100   3.5600  12.7200
(ED)*0.0200   0.9400   0.6600   0.3300   3.2900   0.0500   0.4000   0.0600   0.0300   6.0300   0.0200
   (S)*  0.0800   0.2300   0.1200   0.3200   0.2400   0.2000   0.1800   0.6600   0.2400   0.9400   0.3600
```

Columns 12 through 20

```
   (T)*   0.8700   0.1300   0.5100   0.7200   0.1200   0.0100  10.0200   7.0600   1.2200
   (E)*   1.2000   0.5100   0.4400   4.4200   4.1000   0.1200   2.1000   4.0900   0.9400
   (ED)*0.0700   7.7000   2.8700   3.0800   0.2800   0.0100   4.6200   0.0100   2.3900
   (S)*   1.2400   0.9500   0.2800  10.1800   0.0500   0.0500   1.0300   0.2200   0.0300
```

Here:
(T)* = Travel
(E)*= Entertainment
(ED)*= Education
(S)* = Sport
The target matrix given to the model:
t = [0 0 2 0 2 3 0 0 0 2 1 3 2 2 3 1 1 0 0 2];
Here:
0= code for Travel
1= code for Entertainment

2= code for Education
3= code for Sports
The model took 1229 iterations to meet the target output. The training curve is shown in figure 4.1.

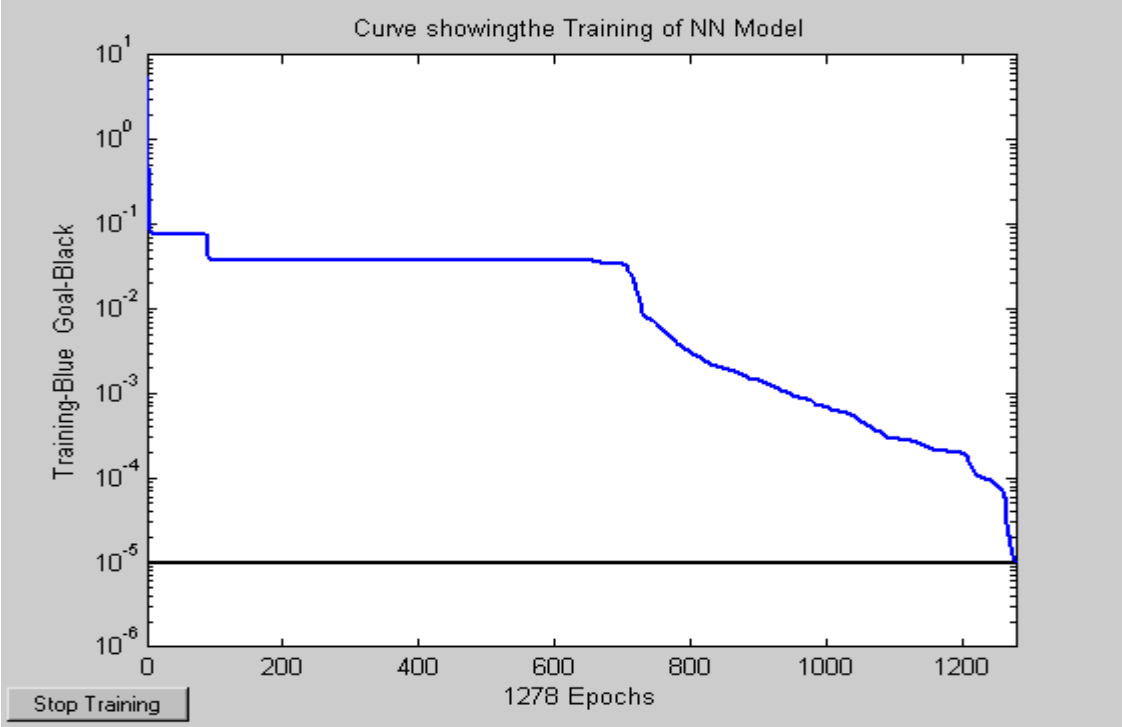

Figure 4.1: Training curve for Neural Network Model

5. Results

The trained neural network model has been tested for four selected testing data set.

The result obtained is

i) Entertainment

The obtained result has been found correct.
The testing curve for Neural Network is shown in figure 5.1.

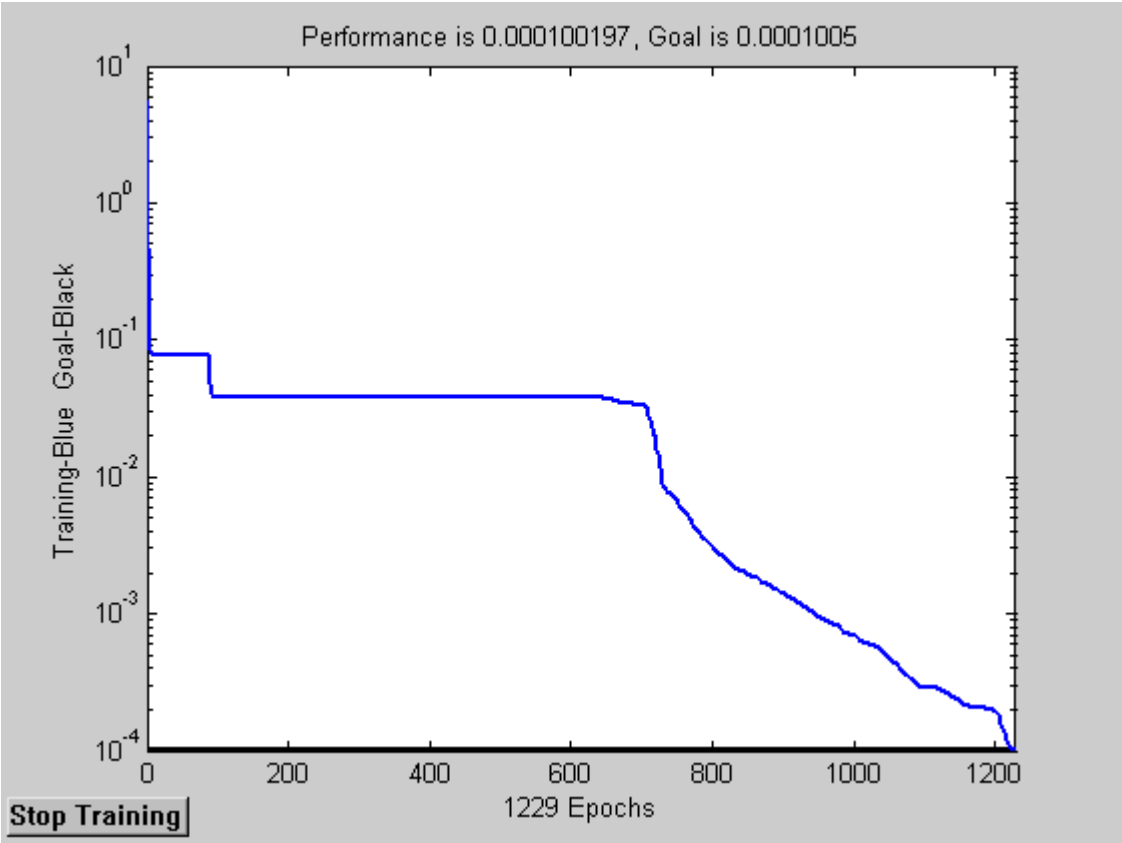

Figure 5.1: Testing curve for Neural Network Model

6. Conclusion and Future Work:

A predication model for the oncoming queries on the web was proposed and a neural network approach was used for training which provides relevant result in less time. This model helps us to predict the oncoming queries for a particular domain so that searching of documents becomes more efficient in terms of time complexity.

In future we can also develop predication model by using any other approach of neural network so that predication model will be become more efficient and scalable.

References

[1] A. Spink, D. Wolfram, B. Jansen, and T. Saracevic. "Searching the Web: The Public and Their Queries", Journal of the American Society for Information Science and Technology, 52(3):226–234, February 2001.

[2] Adar, Eytan, Teevan, Jaime, Dumais, Susan T., and Elsas, Jonathan L. , "The web changes everything understanding the dynamics of web content", In Proceedings of the Second ACM International Conference on Web Search and Data Mining, pp. 282-289, February 2009.

[3] Amir Abolfazl Suratgar, Mohammad Bagher Tavakoli, and Abbas Hoseinabadi, " Modified Levenberg-Marquardt Method for Neural Networks Training", World Academy of Science, Engineering and Technology 6 , pp. 46-48,2005.

[4] Anuradha1 & A.K Sharma, "A Novel Approach for Automatic Detection and Unification of Web Search Query Interfaces using Domain Ontology", International Journal of Information Technology and Knowledge Management, July-December 2010, Volume 2, No. 2, pp. 196-199.

[5] Anderson, J.A. and Rosenfeld, E. (eds) "Neurocomputing: Foundations of Research" , The MIT Press, Cambridge, MA, USA, 1988.

[6] B.D. Davison. "Topical locality in the web", In *Proc. 23rd Annual Intl. ACM SIGIR Conf. on Research and Development in Information Retrieval*, 2000.

[7] Bello, M. G. "Enhanced training algorithms, and integrated training architecture selection for multi layer perceptron networks," IEEE Trans. on Neural Net., vol 3, pp. 864-875, 1992.

[8] Berners-Lee, Tim, "The World Wide Web: Past, Present and Future", MIT USA, Aug 1996.

[9] Berners-Lee, Tim, and Cailliau, CN, R., "WorldWideWeb: Proposal for a Hypertext Project" CERN October 1990, available at: http://www.w3.org/Proposal.html.

[10] Brewington, B. and Cybenko, G., "How Dynamic is the Web", In Proceeding of 9th International World Wide Web Conference, pp.264-296, 2000.

[11] Brin, Sergey and Page, Larry ,"The Anatomy of a Large-Scale Hypertextual Web Search Engine", 1998.

[12] Caudill, Maureen and Butler, Charles, “Understanding Neural Networks, Volume 1:Basic Networks” , The MIT Press, Cambridge, MA, USA, 1993.
[13] Caudill, Maureen and Butler, Charles, “Understanding Neural Networks, Volume 2:Advanced Networks” , The MIT Press, Cambridge, MA, USA, 1993.
[14] Fausett, Laurene, “Fundamentals of Neural Networks: Architectures, Algorithms and Applications” , Prentice-Hall, New Jersey, USA, 1994.
[15] H. Pang and K.-L. Tan. “Authenticating Query Results in Edge Computing”. In IEEE ICDE, pages 560–571, 2004.
[16] H. Pang and K.-L. Tan. “Verifying Completeness of Relational Query Answers from Online Servers”. ACM Transactions on Information and System Security, 11(2):1–50, 2008.
[17] Hwee Hwa PANG and Kyriakos MOURATIDIS. "Authenticating the Query Results of Text Search Engines" *34th International Conference on Very Large Data Bases (VLDB)* (2008): 126-137.
[18] http://www.internetworldstats.com
[19] http://www.worldwidewebsize.com
[20] J. Zobel and A. Moffat. “Inverted Files for Text Search Engine”, ACM Computing Surveys, 38(2), July 2006.
[21] KARABOGA, D., and KALINLI, A.: ‘Training recurrent neural networks for dynamic system identification using parallel tabu search algorithm’. Proceedings of 12th IEEE international symposium on Intelligent control, Istanbul, Turkey, 1997.
[22] Kobayashi, M. and Takeda, K. (2000). "Information retrieval on the web", ACM Computing Surveys (ACM Press) 32 (2): 144–173. doi:10.1145/358923.358934.
[23] Masami NAKAMURA, Katsuteru MARUYAMA, Takeshi KAWABATA, Kiyohiro SHIKANO **“**Neural network approach to word category predication for English texts”, ATR Interpreting Telephony Research Laboratories**,** Seika-chou, Souraku-gun, Kyoto 619-02, JAPAN**.**www.aclweb.org/anthology/C/C90/C90-3038.pdf.
[24] M. T .Hagan and M. B. Menhaj, “Training feed forward network with the Marqual algorithm,” IEEE Trans. on Neural Net., vol. 5, no. 6, pp.989-993, 1994.
[25] M. Diligenti, F. Coetzee, S. Lawrence, C. L. Giles, and M. Gori. “Focused crawling using context graphs”, In *Proc. 26th International Conference on Very Large Databases (VLDB 2000)*, pages 527-534, Cairo, Egypt, 2000.
[26] Ramon Lawrence “Using Neural Networks to Forecast Stock Market Prices” Department of Computer Science University of Manitoba, December 12, 1997 https://people.ok.ubc.**ca**/rlawrenc/research/Papers/nn.pdf.
[27] Towell, G. and Shavlik, J. “The extraction of refined rules from knowledge based neural networks”. *Machine Learning*, 131, pp 71-101, 1993.
[28] VAI, M., PRASAD, S., and WANG, H.: ‘A smith chart represented by a neural network and its applications’. MIT-S international *Microwave* symposium Digest, pp. 1565-1568, 1992.