ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 7, September 2012

F-Measure Metric for English to Hindi Language Machine Translation

Neeraj Tomer [1] Deepa Sinha [2] Piyush Kant Rai [3]

Abstract: The main objective of MT is to break the language barrier in a multilingual nation like India. Evaluation of MT is required for Indian languages because the same MT is not works in Indian language as in European languages due to the language structure. So, there is a great need to develop appropriate evaluation metric for the Indian language MT. The present research work aims at studying the Evaluation of Machine Translation Evaluation's F-Measure Metric for English to Hindi for tourism domain. This work will help to give the feedback of the MT engines. We may make the changes in the MT engines and further we may revise the study. We see that as we increase the number of references there is improvement in our results.

Keywords: MTE- Machine Translation Evaluation, MT- Machine Translation, SVO- Subject-Verb-Object, SOV-Subject-Object-Verb, Tr.-Tourism.

INTRODUCTION

India is a highly multilingual country with 22 constitutionally recognized languages. Still, English is understood by less than 3% of Indian population. Hindi, which is official language of the country, is used by more than 400 million people. Therefore, Machine Translation (MT) assumes a much greater significance in breaking the language barrier within the country's sociological structure. The main objective of MT is to break the language barrier in a multilingual nation like India. English is a highly positional language with rudimentary morphology, and default sentence structure as Subject-Verb-Object. Indian languages are highly inflectional, with a rich morphology, relatively free word order, and default sentence structure as Subject-Object-Verb. In addition, there are many stylistic differences. So the evaluation of MT is required for Indian languages because the same MT is not works in Indian language as in European languages. The same tools are not used directly because of the language structure. So, there is a great need to develop appropriate evaluation metric for the Indian language MT.

The present research work aims at studying the "Evaluation of Machine Translation Evaluation's F-Measure metric for English to Hindi" for tourism domain. The present research work is the study of statistical evaluation of machine translation evaluation for English to Hindi. The research aims to study the correlation between automatic and human assessment of MT quality for English to Hindi. The main goal of our experiment is to determine how well a variety of automatic evaluation metric correlated with human judgment.

MATERIALS AND METHODS

In the present work we propose to work with corpora in the tourism domain and limit the study to English – Hindi language pair. It may be assumed that the inferences drawn from the results will be largely applicable to translation for English to other Indian Languages. Our test data consisted of a set of English sentences that have been translated from expert and non-expert translators. The English source sentences were randomly selected from the corpus of tourism domain. These samples are taken randomly from the tourism domain in which we have the 15200 sentences and from the health care domain in which we have the 10000 sentences. Each output sentence was score by Hindi speaking human evaluators who were also familiar with English. It may be assumed that the inferences drawn from the results will be largely applicable to translation for English to other Indian Languages, as assumption which will have to be tested for validity. We intend to be consider the following MT engine in our study-

- Anuvadaksh

OBJECTIVE

The main goal of this work is to determine how well a variety of automatic evaluation metrics correlated with human judges. A secondary goal is to determine for which the correlation of automatic and human evaluation is particularly good or bad. The other specific objectives of the present work are as follows.

1. To design and develop the parallel corpora for deployment in automatic evaluation of English to Hindi machine translation systems.
2. Assessing how good the existing automatic evaluation metric F-Measure, will be as MT evaluating strategy for evaluation of Indian language machine translation (EILMT) systems by comparing the results obtained by this with human evaluator's scores by correlation study.
3. To study the statistical significance of the evaluation results as above, in particular is the effect of-
   - size of corpus
   - sample size variations
   - increase in number of reference translations

Creation of parallel corpora: Corpus quality plays a significant role in automatic evaluation. Automatic

metrics can be expected to correlate very highly with human judgments only if the reference texts used are of high quality, or rather, can be expected to be judged high quality by the human evaluators. The procedure for creation of parallel corpora is as under:

1. Collect English corpus from the domain from various resources.
2. Generate multiple references (we limit it to three) for each sentence by getting the source sentence translated by different expert translators.
3. XMLise the source and translated references for use in automatic evaluation.

Description of Corpus

Domain	Tourism
Source Language	English
Target Language	Hindi
No. of Sentences	1000
No. of Words	23000
No. of Human Translation	3
No. of MT Engine	1

For the corpus collection our first motive was to collect as possible to get better translation quality and a wide range vocabulary. For this purpose the first corpus we selected to use in our study is collected from different sources. We have manually aligned the sentence pairs.

In our study for tourism domain we take 1000 sentences. When the text has been collected, we distributed this collected text in the form of Word File. Each word files having the 100 sentences of the particular domain. In this work our calculation will be based on four files- source file and three reference files. Reference files are translated by the language experts. We give the file a different identification. For e.g. our first file name is Tr_0001_En where Tr_ for tourism 0001 means this is the first file and En means this is the Candidate file. We treat this as the candidate file. In the same way our identification for the Hindi File is Tr_0001_Hi, in this Hi is for the Hindi file and we have called this a reference file. As we already mention that we are taking the three references we named them reference 1(R1), reference 2(R2), reference 3(R3). In the study we take the candidate sentence and the reference sentences, as shown below. For e.g.

Source Sentence: Internal problems dominated the agenda for the next 30 years, as a series of governments struggled to keep the economy, which was almost completely destroyed as a result of the Pacific War, from disintegrating.

Candidate Sentence:

उष्ण कटिबंधीय निम्नभूमियाँ वन्य जीवन का बहुत बड़ी क्रम छुपाती हैं, घनी वनस्पति और समान देश के अंतर्गत अभी तक नाटकीय जलप्रपातों ट्रेककरों औरपर्वतारोहियों को चुनौती देकर ऊँची एन्डियन चोटियाँ में है

Reference Sentences:

R1- अगले 30 वर्षों तक एजेण्डा में अन्दरूनि समस्याएं सर्वोच्च रही, जबकि एक के बाद एक सरकारें संघर्षरत थी उस अर्थव्यवस्था को विघटित होने से बचानेमें जो कि पैसिफिक युध्द के परिणाम स्वरुप लगभग पूरी तरह नष्ट हो चुकी थी।

R2- जिस तरह अर्थव्यस्था को बनाए रखने के लिए सरकारों की श्रेणियों मेंसंघर्ष किया, आने वाले 30 वर्षों के लिए आन्तरिक समस्याओं ने कार्य सूचि को अधिकार में रखा जो विघटित होने से Pacific War के परिणामस्वरूप पूर्णतया समाप्त हो गयी थी।

R3- अगले तीस वर्षों के लिए आन्तरिक समस्याएं कार्यक्रम पर प्रभावी रहीं, क्योंकि सरकारों की एक श्रृंखला ने अर्थव्यवस्था, जो पेसिफिक युद्ध के परिणामस्वरूप पूर्णतः नष्ट हो गयी थी,विखण्डित होने से बचाए रखने के लिए संघर्ष किया।

HUMAN EVALUATION

Human evaluation is always best choice for the evaluation of MT but it is impractical in many cases, since it might take weeks or even months (though the results are required within days). It is also costly, due to the necessity of having a well trained personnel who is fluent in both the languages, source and targeted. While using human evaluation one should take care for maintaining objectivity. Due to these problems, interest in automatic evaluation has grown in recent years. Every sentence was assigned a grade in accordance with the following four point scale for adequacy.

	Score
• Ideal	1
• Acceptable	.5
• Not Acceptable	.25
• If a criterion does not apply to the translation	0

AUTOMATIC EVALUATION BY F-MEASURE METRIC

We used F-Measure evaluation metric for this study. This metric is specially designed for English to Hindi. F-Measure metric, designed for evaluating MT quality, scores candidate sentences by counting the number of n-gram matches between candidate and reference sentences. F-Measure metric is probably known as the best known automatic evaluation for MT. To check how close a candidate translation is to a reference translation, an n-gram comparison is done between both. Metric is designed from matching of candidate translation and reference translations. We have chosen correlation analysis to evaluate the similarity between automatic MT

evaluations and human evaluation. Next, we obtain scores of evaluation of every translated sentence from both MT engines. The outputs from both MT systems were scored by human judges. We used this human scoring as the benchmark to judge the automatic evaluations. The same MT output was then evaluated using both the automatic scoring systems. The automatically scored segments were analyzed for Spearman's Rank Correlation with the ranking defined by the categorical scores assigned by the human judges. Increases in correlation indicate that the automatic systems are more similar to a human in ranking the MT output.

Statistical significance is an estimate of the degree, to which the true translation quality lays within a confidence interval around the measurement on the test sets. A commonly used level of reliability of the result is 95%. To reach at decision, we have to set up a hypothesis and compute p-value to get final conclusion.

The present research is the study of statistical evaluation of machine translation evaluation's F-Measure metric. The research aims to study the correlation between automatic and human assessment of MT quality for English to Hindi. While most studies report the correlation between human evaluation and automatic evaluation at corpus level, our study examines their correlation at sentence level. The focus in this work is to examine the correlation between human evaluation and automatic evaluation and its significance value, not to discuss the translation quality. In short we can say that this research is the study of statistical significance of the evaluated results, in particular the effect of sample size variations.

So, firstly we take source sentences and then get these sentences translated by our MT engine, here we consider the Anuvadaksh. We have the different references of these sentences. After doing this we do the evaluations of these sentences human as well as the automatic evaluations and we collect the individual scores of the given sentences considering all the three references one by one. The following table shows the individual scores of the five sentences (particular sentences can be seen at the end of the paper) using different no. of references.

Table 1: Human Evaluation and F-Measure Evaluation scores

S. No.	F-Measure Score			
	Human Eval.	one no. of reference	two no. of references	three no. of references
1.	0.75	0.1286	0.1575	0.1575
2.	0.25	0.1814	0.1689	0.1865
3.	0.75	0.1917	0.2073	0.2472
4.	0.75	0.1352	0.1505	0.1679
5.	1	0.1336	0.1713	0.1868

In this way we also collect the individual scores of all the sample sizes like 20, 60,100,200,300,500 and 1000 sentences. After this we do the correlation analysis of these values. In order to calculate the correlation with human judgements during evaluation, we use all English–Hindi human rankings distributed during this shared evaluation task for estimating the correlation of automatic metrics to human judgements of translation quality, were used for our experiments. In our study the rank is provided at the sentence level.

For correlation analysis we calculate the correlation between human evaluation and automatic evaluations one by one by the Spearman's Rank Correlation method. The Spearman's rank correlation coefficient is given as (when ranks are not repeated)-

$$\rho = 1 - \left(\frac{6\sum_{i=1}^{n} d^2}{n(n^2 - 1)} \right)$$

where d is the difference between corresponding values in rankings and n is the length of the rankings. An automatic evaluation metric with a higher correlation value is considered to make predictions that are more similar to the human judgements than a metric with a lower value. Firstly, we calculate the correlation value in between the human evaluation and automatic evaluation F-Measure metric means human evaluation with F-Measure for sample size 20, 60, 100, 200, 300, 500 and 1000.

Table 2: Correlation (ρ) values

Sample Size	ρ values		
	one no. of reference	two no. of references	three no. of references
20	.384	.399	.410
60	.141	.151	.204
100	.071	.092	.106
200	.219	.212	.260
300	.216	.199	.232
500	.116	.256	.256
1000	.176	.256	.176

After calculating the correlation, we need to find out which type of correlation is there between the variables and of which degree and whether the values of the correlation are significant.

ANALYSIS OF STATISTICAL SIGNIFICANCE TEST FOR HUMAN EVALUATION AND AUTOMATIC EVALUATION

Statistical significance is an estimate of the degree, to which the true translation quality lays within a confidence interval around the measurement on the test sets. A commonly used level of reliability of the result is 95%, for e.g. if, say, 100 sentence translations are evaluated, and 30 are found correct, what can we say about the true translation quality of the system? To reach at decision, we have to set up a hypothesis and compute p-value to get final conclusion that whether there is any correlation

between the human evaluations and automatic evaluations. If yes, then what is the type and degree of correlation? Also what is the significance of the correlation value? In this work we set the hypothesis that there is no correlation between the values of human and automatic evaluation. The p-value will provide the answer about the significance of the correlation value.

A Z-test is a statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. For each significance level, the Z-test has a single critical value (for example, 1.96 for 5% two tailed) which makes it more convenient than the Student's t-test which has separate critical values for each sample size. The test statistic is calculated as:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means, s_1^2 and s_2^2 are the sample variances, n_1 and n_2 are the sample sizes and z is a quartile from the standard normal distribution.

Table 3 : p-values of output of Anuvadaksh using different no. of references

Sample Size	p-values		
	one no. of reference	two no. of references	three no. of references
20	0.0001	0.0001	0.0001
60	0.0001	0.320	0.0001
100	0.0001	0.0001	0.450
200	0.0001	0.0001	0.0001
300	0.0001	0.0202	0.0162
500	0.2296	0.0069	0.0069
1000	0.0764	0.069	0.0754

Now on the basis of these values we conclude our results like which type and degree of correlation is there between the given variables and whether the correlation results are significant. In the above example we have done all the calculations by considering the single reference sentence and in tourism domain using 5 numbers of sentences.

But in our research work we consider the different references like 1, 2, 3 and we use the different sample sizes like 20, 60, 100, 200, 300, 500, and 1000. We see whether the results remains uniform for different sample sizes and different number of references in particular domains.

For above calculation we used following sentences:
English Sentences:

1. Little wonder, then, that its environment is so rich.
2. French Guiana is also home to a colorful blend of different cultural backgrounds.
3. All this combined with rich historical and archaeological past and enduring indigenous cultures.
4. The country enjoys 500km of fine sandy beaches on the Atlantic and the Río de la Plata, woods, hills, hot springs, hotels, casinos, art festivals and numerous opportunities for sport and entertainment.
5. First occupied by the French in 1764, the islands were quickly ceded to Spain, which then ruled the adjacent territory in Latin America.

Candidate Sentences:

1. थोड़ा सा विस्मय, तब जो अपनी पर्यावरण इसलिए समृद्ध है
2. फ्रेंज गुआना विभिन्न सांस्कृतिक पृष्ठभूमियाँ का रंगीन मिश्रण को आवास भी है
3. सभी यह समृद्ध ऐतिहासिक और पुरातात्विक भूतकाल और चिरस्थाई स्वदेशी संस्कृतियाँ के साथ सुमेलित किया
4. देश खेल और मनोरंजन के लिए ला प्लाह्वकड़ियाँ, पहाड़ों, गर्म वसंतों, होटलों, नृत्यशालाएँ, कला उत्सवों और अनेक अवसरों द अटलांटिक और आर. ओ पर अच्छे बालुकामय समुद्र तटों क500 किमी को आनन्द प्राप्त करता है
5. 1764 में फ्रेंच द्वारा पहला अधिकार किया गया द्वीपों स्पेन को शीघ्रतापूर्वक सौंप दिये गये थे जो लैटिन अमेरिका में संलग्न प्रांत को तब शासन किया

RESULTS

In the domain tourism there is significance difference between the average evaluation score of human with F-Measure at 5% level of significance and for the sample size 20. There is highest correlation between the evaluation score of human and F-Measure for MT engine Anuvadaksh.

We see that as we increase the number of references there is improvement in our results. In Table 2 (Correlation (ρ) values) correlation value for F-Measure is .399 and .410 these values are for sample size 20 and for two and three number of references which is significant at 5% level of significance. A similar result is seen in the case of sample size 60, 100 and 200 for two and three no. of references. But for the sample sizes 500 and 1000 value of correlation is .256 for two no. of references which is insignificant on the given level of significance. From the analysis on the basis of z-test used for the significance test of human evaluation and automatic evaluation we obtain the following important point; in the domain tourism there is significance difference between the average evaluation score of human with F- Measure at 5% level of significance and for the some sample sizes.

CONCLUSION

This work will help to give the feedback of the MT

engines. In this way we may make the changes in the MT engines and further we may revise the study. Corpus quality plays a significant role in automatic evaluation. Automatic metrics can be expected to correlate highly with human judgments only if the reference texts used are of high quality, or rather, can be expected to be judged of high quality by human evaluators. In this evaluating MT evaluation metrics for English to Indian Language machine translation work, we developed the tool SEMTE which is good enough according to the results obtained.

ACKNOWLEDGEMENT

The present research work was carried under the research project "English to Indian Languages Machine Translation System (EILMT)", sponsored by TDIL, Ministry of Communications and Information Technology, Government of India. With stupendous ecstasy and profundity of complacency, we pronounce utmost of gratitude to Late Prof. Rekha Govil, Vice Chancellor, Jyoti Vidyapith, Jaipur Rajasthan.

REFERENCES

1. Akiba, Yasuhir Taro Watanabe, Eiichiro Sumita, (2002): "Using Language and Translation Models to Select the Best among Outputs from Multiple {MT} System", Proceeding of Colong, 8-14.
2. Babych, Bogdan and Anthony Hartley, (2004): "Extending the BLEU MT Evaluation Method with Frequency Weightings", Proceedings of ACL, 621-628.
3. Bradley Efron and Robert Tibshirani, (1986): "Bootstrap Methods for Standards Errors, Confidence Intervals, and other Measures of Statistical Accuracy", Statistical Science, 54-77.
4. Calison-Burch, Chris and Raymond S. Flournoy, (2001): "A Program for Automatically Selecting the Best Output from Multiple Machine Translation Engines", Proceedings of MT Summit VIII, 63-66.
5. Calison-Burch, Chris and Raymond S. Flournoy, (2001): "A Program for Automatically Selecting the Best Output from Multiple Machine Translation Engines", Proceedings of MT Summit VIII, 63-66.
6. Deborah Coughlin, (2003): "Correlating Automated and Human Assessments of Machine Translation Quality", In Proceedings of MT Summit IX. New Orleans, 63-70.
7. Deborah Coughlin, (2003): "Correlating Automated and Human Assessments of Machine Translation Quality", In Proceedings of MT Summit IX. New Orleans, 63-70.
8. Doddington G., (2002): "Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrences Statistics", Human Language Technology: Notebook Proceedings, San Diego, USA, 128-132.
9. Donaway, R. L., Drummey, K. W., Mather, L. A. (2000): "A Comparison of Rankings Produced by Summarization Evaluation Measures", Proceedings of the Workshop on Automatic Summarization, 69-78.
10. Donaway, R.L., Drummey, K.W., and Mather, L.A., (2000): "A Comparison of Rankings Produced by Summarization Evaluation Measures", In Proceedings of the Workshop on Automatic Summarization, 69-78.
11. Dunning, Td. (1993): "Accurate Methods for the Statistics of Surprise and Coincidence", Computational Linguistics. 61-74.
12. http://en.wikipedia.org/wiki/Evaluation_of_machine_translation
13. http://en.wikipedia.org/wiki/History_of_machine_translation
14. Joseph P. Turrian, Luke Shen and I. Dan Melamed, (2003): "Evaluation of Machine Translation and its Evaluation", In Proceedings of MT Summit IX. New Orleans, 63-70.
15. K.-Y.Su, M.-W.Wu, and J.-S.Chang., (1992): "A New Quantitative Quality Measures for Machine Translation System", In Proceedings of the 15th International Conference on Computational Linguistics (COLING) Nantes, France. 433-439.
16. Lin C. Y., E. hovy. (2002): "Manual and Automatic Evaluations of Summaries", In Proceedings of the Workshop on Automatic Summarization, post-Conference Workshop of ACL. 45-52.
17. Liu Ding and Gildea Daniel, "Syntactic Features for Evaluation of Machine Translation", Department of Computer Science University of Rochester, Rochester, NY 14627.
18. Meng, X., R. Rosenthal, and D. Rubim., (1992): "Comparing Correlated Correlation Coefficients", Psychological Bulletin-11, 172-175.
19. Rajman M., Hartley A. (2001): "Automatically predicting MT system rankings Compatible with Fluency, Accuracy or Informativeness scores", In Proceedings of the 4th workshop on MT Evaluation, MT Summit VIII, Santigo de Compostela, 29-34.
20. Rao, Durgesh (2001): "Machine Translation in India: A Brief Survey", National Centre for Software Technology Gulmohar Road 9, Juhu, Mumbai 400049, India, 21-23.
21. S.Niessen, F.J.Och, G. Levsch, and H.Ney. (2000): "An Evaluation Tool for Machine Translation: Fast Evaluation for Machine Translation Research", In Proceedings of the Second International Conference on Language Resources and Evaluation. Athens, Greece, 39-45.
22. T. Neeraj (2012): "Evaluating Machine Translation (MT) Evaluation Metrics for English to Indian Language Machine Translation", Ph.D. Thesis, Banasthali University, Banasthali.
23. T. Neeraj and Sinha Deepa (2012): "Evaluating Machine Translation Evaluation's BLEU Metric for English to Hindi Language Machine Translation", in The International Journal of Computer Science & Application , Vol-01-NO-06-Aug-12, 48-58.
24. T. Neeraj and Sinha Deepa (2012): "Evaluating Machine Translation Evaluation's NIST Metric for English to Hindi Language Machine Translation", paper accepted in The International Journal of Multidisciplinary Research Academy –IJMRA.
25. www.amadeus.net/home/destinations/en/guides/.../intro.htm
26. www.escapelets.com/en-GB/countryproperties/PE.aspx
27. www.iexplore.com/travel.../**french-guiana**/overview
28. www.noble-caledonia.co.uk/.../countries_information.asp?section
29. www.thetravelwebsite.co.uk/uruguay-c446.html
30. Y. Akiba, K.Imamura, and E. Sumita., (2001): "Using Multiple Edit Distances to Automatically Rank Machine Translation Output", In Proceedings of MT Summit VIII. Santiago de Compostela, Spain, 15-20.

Author 1

Neeraj Tomer
tneeraj12@rediffmail.com
9460762117

Area of Interest:

- Machine Translation and Indian Language Technology
- Theoretical Computer Science and related technologies

Academic Qualification:
Ph.D (thesis submitted) in Computer Science, Banasthali University, Banasthali.
MCA, Maharishi Dayanand University, Rohtak 2005.
Master of Economics, Kurukshetra University Kurukshetra 1999.
Bachelor of Economics, Kurukshetra University Kurukshetra 1997.

Employment History:
Post graduate and graduate teaching at Mahatma Gandhi Institute of Applied Sciences, Jaipur as a lecturer from July 2003 to August 2006. As a Research Associate at Banasthali University Banasthali in 2007. As a lecturer at LCRT College of Education Panipat from August 2007 to July 2010. As an Assistant Professor at SINE International Institute of Technology, Jaipur from August 2010 to March 2012.

- Papers Published : 1
- In Press : 3
- Communicated : 3

Seminar and Conferences Attended: 5

Research Guidance:
Guided 3 students for their dissertation work at PG (M.Sc) level.

Future Plans: To grow academically

Author 2

Deepa Sinha
Associate Professor
Department of Mathematics
South Asian University
Akbar Bhawan
Chanakyapuri, New Delhi 110021 (India)
Cell No: 08744022273
deepasinha2001@gmail.com

Research Area: **Graph Structures**
Academic Qualification:
M.Sc., Ph. D. (University of Delhi), CSIR-NET (twice)
Future Plans: To grow academically
Achievements: CSIR_NET (qualified Twice)
Publications:
(a) Books: one
(b) Research Papers: 27
Conference/workshop/symposium attended: 39
Invited talks delivered: Ten
Papers presented: 23
Work experience: 16 years
Served several Institutions for graduate and post graduate courses, particularly Khalsa Girls Degree College (1996-1999), Delhi College of Engineering (1999-2004), Banasthali University (2004-2012).

Seven students got awarded their M. Phil. Degree under her supervision
.
Three students got awarded their Ph.D. in the year 2011-2012.

Have refereed several research papers for National and international Journal of high impact factor like Discrete Applied Mathematics, Graphs and Combinatorics, International Journal of Physical Sciences etc.

Sessions chaired in the National/ International conferences: four

Author 3

Name: Piyush Kant Rai
Date of Birth: February 5, 1980
E-Mail: raipiyush5@gmail.com
Current Affiliation:
Working as Assistant Professor (Senior-Scale) (Statistics) in the Department of Mathematics & Statistics (AIM & ACT), Banasthali University, Rajasthan.

Working Experience: Eight years of teaching and research experience including one year experience as a Project Assistant at Banaras Hindu University (BHU), U.P. and five months as a Statistical Fellow at Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS) Lucknow, U.P.

UGC Major/Minor Project: One Minor Research Project as Principal Investigator and one Major Research Project as Co-Principal Investigator.
Academic Qualification:

- Ph.D. (Statistics) from Banasthali University, Rajasthan.
- M.Phil. (Statistics), 2008, with First Class from Periyar Institute of Distance Education (PRIDE), Periyar University, Salem Tamilnadu.
- M.Sc. (Statistics), 2003, with First Rank (Gold Medal) from Banaras Hindu University (BHU), Varanasi, U.P.
- B.Sc. (Hons., Statistics), 2001, with First Class (5^{th} Rank in Faculty of Science) from Banaras Hindu University (BHU), Varanasi, U.P.
- Intermediate (Physics, Chemistry, Math, Hindi and English), 1996, with First Class (1^{st} Rank in College) from U.P. Board.
- High School (Hindi, English, Science, Math, Social Science and Biology), 1994, with first-Honours (1^{st} Rank in School) from U.P. Board.

Qualified National Eligibility Test:

- UGC NET, June 2004, UPSLET, August 2004 and ISI-JRF, May 2005.

Scholarships:

- 10^{th} to 12^{th} standard, 8^{th} to 10^{th} standard and 6^{th} to 8^{th} standard scholarship by Education Department U.P.

Publication: (National/International):

- Papers Published : 8
- In Press : 1
- Communicated : 3
- Books Published : Two Book Chapters are published and two are in Press.

Supervisor of 2 Ph.D. Scholars and 5 M.Phil. Scholars
Workshop and Training Program Attended /Organized: 17
Seminar and Conference Attended/Organized: 12

Membership:
Membership of Actuarial Society of India and Indian Bayesian Society.