Indian Journal of Artificial Intelligence and Neural Networking (IJAINN)
ISSN: 2582-7626 (Online), Volume-1 Issue-6, October 2021

How to Choose VLSI IC from E-Commerce Sites?: Sentiment Analysis with the help of Python Tools

Biswajit Biswas, Tuhin Mukherjee

Abstract: ***Very Large Scale Integration (VLSI) dominates the digital technology in the present era. The VLSI based products have strong computing power in one hand & small in physical dimension due to its space minimization feature. Unfortunately, the people who use these ICs frequently are still not in a position to frequently purchase such VLSI IC in online mode & rely more on offline shopping from electronics shop in his or her locality. Due to the emergence of cashless economy, there is a paradigm shift in purchasing behavior of customers across the globe but it is an established fact that the ecommerce sites are still not matured for VLSI IC. This study is an attempt to use Natural Language Processing Tool Kit (NLTK) in Python & augment it with Valence Aware Dictionary and Sentiment Reasoner (VADER) analysis for development of a newly proposed App supposed to be used in the smart mobile phones for Sentiment Analysis of online feedbacks & reviews of customers who use VLSI IC frequently in their profession. The purpose is to create a confidence & prepare a convenient platform for those people towards their online purchasing behavior of VLSI IC. Uniqueness of this research is the use of word cloud for the textual review along with the star rating in the sentiment analysis through an automated system which is supposed to accept the Uniform Resource Locator (URL) of the concerned product only. The study demonstrates the validation on a set of specific product reviews from the amazon.in. & unfold the challenge of said App development successfully. The newly developed tools will be helpful for a customer to select such a VLSI IC before purchase.***

Keywords: Artificial Intelligence, VLSI IC, Feedback Analysis, Sentiment Analysis, and Online Product Reviews

I. INTRODUCTION

This paper is an attempt to explore the online review of two selected VLSI IC across e-commerce sites. The purpose is to develop & demonstrate a new App which will serve the end users by computing sentiment analysis metrics. The open source python platform is used and both numerical ratings & textual feed backs have been considered. The paper describes the algorithmic procedure of our newly developed App & demonstrates the use of that App with respect to two VLSI IC viz operational amplifier & timer over the e-commerce site of Amazon in Indian context. The date & time of our execution is also embedded within the output snapshots in the form of timestamp.

Manuscript received on 25 September 2021 | Revised Manuscript received on 06 October 2021 | Manuscript Accepted on 15 October 2021 | Manuscript published on 30 October 2021.
***** Correspondence Author(s)
Biswajit Biswas*, Department of Business Administration, University of Kalyani (West Bengal), India. E-mail: biswajit.biswas0012@gmail.com, ORCID ID: https://orcid.org/0000-0002-5302-7675
Dr. Tuhin Mukherjee, Department of Business Administration, University of Kalyani (West Bengal), India. E-mail: tm.ku.finance@gmail.com

II. OBJECTIVES OF THE STUDY

i. To evaluate the conceptualize framework of feedback analysis system in e-marketing.
ii. To develop an automated review collected system from e-marketing websites.
iii. To develop an automated system to extract the polarity & other features of the collected reviews.
iv. To find out the actual rating of a specific product, if customer's textual review contradicts with numerical reviews for relevant decision making.
v. To demonstrate the use of our newly developed App in the context of two VLSI IC viz Op Amp 741 & 555 Timer.

III. LITERATURE REVIEW

In this present age many researchers are doing their research work on sentiment analysis. Sentiment depends on domain where and how the words are used. In social science there is a vest area on sentiment analysis of people buying behavior. So it requires sentiment lexicons for specific domains. In this work authors' induction a domain specific algorithm for sentiment analysis (Hamilton, et al. 2016)[2]. Lee, J et al. observed in their research work that in this modern Industry 4.0 tools and machineries are working in a collaborative way. This evolution needs the updated prediction tools as the raw data are processed in a systematic way to remove the uncertainties and make more informative decision (Lee, et al. 2016)[4]. Jain et al said in their research work that marketers can find out many truthful compliments from the consumer data and it gives new opportunities to explore the market. However 4Vs of marketing helps the marketers to understand the success rate or failures rate in market. In this work authors frame work a visible GAP on the consumer feedback and it will help to design new product. It starts to identify the product features. The related data for this work collected from Amazon.com (Jain et al. 2016)[3]. Acceding to Pankaj et al. sentiment analysis is a rapid growing area of research in computer science. In this paper they introduced a review mining system on a particular product that will be helpful for the future customers. The reviews are collected from the Amazom.com and they have done a comparative analysis on retrieved reviews. Result of the work reflected on customers purchasing behavior (Pankaj et al. 2019)[5]. Gunawan et al. describe in their research work that to get opinion on a particular product or service, social network analysis is the best way. It gives the accurate result than other types of questionnaires analysis.

Data for this research may collect from social network sites like Twiter, blogging, e-commerce sites. In this study authors used the python library to extract and analysis the data (Gunawan et al. 2020)[1].

IV. ALGORITHMIC DESCRIPTION OF NEWLYDEVELOPED APP

A. Product URL to extract customer review

The first step is to collect data or to acquire data for the research purposes. In this study, we have developed web-scraping python program to extract customer feedbacks from the Amazon website by passing the URL in the program as parameter. In that program, python package Beautiful Soup is used for Web Scrapping activity from Amazon website. Amazon maintains separate unique URL for each of the products available in its online virtual store.Following two URLs are used in this paper to demonstrate the use of our newly developed App using python platform. The first one is for Operational Amplifier & the second one is Timer VLSI IC, which are most commonly used IC in daily use.
(i)https://www.amazon.in/dp/B0718X4LKW/ref=cm_sw_r_wa_apa_i_TRAM7M5YBTKT9B7S118Y &
(ii)https://www.amazon.in/Robo-India-555-Timer Pieces/dp/B00VTI9F3U/ref=sr_1_6?crid=LGSDYJZBZ5EX&keywords=timer+integrated+chip&qid=1653575672&sprefix=timer%2Caps%2C433&sr=8-6#customerReviews

B. Data extraction using Beautiful Soup API

Beautiful soup is a module of Python, which read and remove the data included within HTML and XML files. It allows researchers, data analysts to extract the non-critical business information using web scrapping programs.We have used it with our selected parser to explore, search, and alter the analysis tree language. It decreases the time a user needs to work.The files with extension .csv or .xls or .xlsx may be created to have customer reviews in excel files with row-column format.API serves as a conduit for communication between a server and a client.

C. Extract user's feedback and rating

Online e-commerce giants, Amazon, allows its customers to keep useful information such as reviewer identification, credibility for the reviewer, product evaluation, review time, assistance with other reviewers and the option to modify comments afterwards. Customer Feedbacks published on Amazon may be converted into useful driving tool to decide the usability and popularity of the product among its customers.

D. Data cleaning

Pre-processing is one of the key processes in working with text data. Data cleanup often means that any specific letters, null values or other terms that do not add any value to the analytical data are deleted. It also covers duplication of information and other anomalies. For cleansing data (pre-processing procedures) texts there are various standardized ways.

E. Word cloud

It identifies the most often occurring terms, which may help both consumers and manufacturers, retailers, or sellers understand how current customers feel about the product or what it's most important features are. A word cloud may show you how prevalent a word is throughout all of the extracted customer feedbacks. The size, weight, and color of the text all have a significant influence on how customers perceive it.

F. Sentiment analysis

Sentiment analysis determines if a sentence is good, negative, or neutral by recognizing the contextual polarity of the text. Because it seeks to understand people's attitudes, it is also known as opinion mining. The two fundamental approaches for sentiment analysis are machine learning-based techniques and lexicon-based techniques.

i. Valence Aware Dictionary for Sentiment Reasoning (VADER) Analysis

To do this Valence Aware Dictionary for Sentiment Reasoning (VADER) Analysis, this newly developed App used Natural Language Toolkit (NLTK) which is the API for Natural Language Processing (NLP) with python. The text classifier subjectivity in our App runs from 0.0 and 1.0 ranges. The 0.0 defines the text as objective, and the 1.0 rate the text as extremely subjective.

ii. Text blob analysis using NLTK

The data screened by our newly developed App use the tool set of Python Library after the extraction and cleaning of the data. TexBlob and Word Cloud tool are the Python tools used for this purpose. The polarity score shows positive text with a value greater than 0 and a negative text indicate below 0. The resulting emotion is then visualized. The findings of the sentiments are shown with a Matplotlib tool, a Python 2D graphic library and NumPy for anomalous maths.

V. RESULTS & FINDINGS

In this section some snap shots of our newly developed App have been included as an evidence of demonstration. Though we have done the entire experiment for both the VLSI IC across Amazon website, here we are selectively representing the outputs only for one of them. Following figure 1 is online review extraction part for operational amplifier (the first VLSI IC that is considered in this paper).

Retrieval Number: 100.1/ijainn.B39181212222
DOI:10.54105/ijainn.B3918.101621
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

Amazon_Product_sentiment_analysis_prog

```
In [3]: import pandas as pd
        import numpy as np
        import re
        import seaborn as sns
        import matplotlib.pyplot as plt
        import string
        from string import punctuation
        import nltk
        from nltk.corpus import stopwords
        from textblob import TextBlob, Word
        from nltk.stem import WordNetLemmatizer
        from nltk.corpus import wordnet
        from nltk import FreqDist
        from wordcloud import WordCloud

In [4]: df  = pd.read_csv("Review_All_Amazon_Productvlsi.csv")
        df.head()
```

Out[4]:

	Title	Rating	Date	Text	Found_Helpful
0	Agam	3.0 out of 5 stars	Reviewed in India on 5 May 2018	IC quality is good but packing isn't, all ICs ...	2 people found this helpful
1	Santosh Kumar	1.0 out of 5 stars	Reviewed in India on 22 May 2021	Not a single one is working. There is no outpu...	One person found this helpful
2	R.B.Suraj	4.0 out of 5 stars	Reviewed in India on 12 December 2020	Giving in a plastic case witth 5 free diodes i...	2 people found this helpful
3	usha	5.0 out of 5 stars	Reviewed in India on 22 June 2020	Good for home project	NaN

Figure 1: Online review extraction part for operational amplifier

Out[8]:

	Text	Rating
0	IC quality is good but packing isn't, all ICs ...	3
1	Not a single one is working. There is no outpu...	1
2	Giving in a plastic case witth 5 free diodes i...	4
3	Good for home project	5
4	All Op- Amp ICs came with Pins bent.	1

```
In [9]: #distribution of rating
        sns.set_style('whitegrid')
        sns.countplot(x='Rating',data=review_df, palette='YlGnBu_r')
```

Out[9]: <AxesSubplot:xlabel='Rating', ylabel='count'>

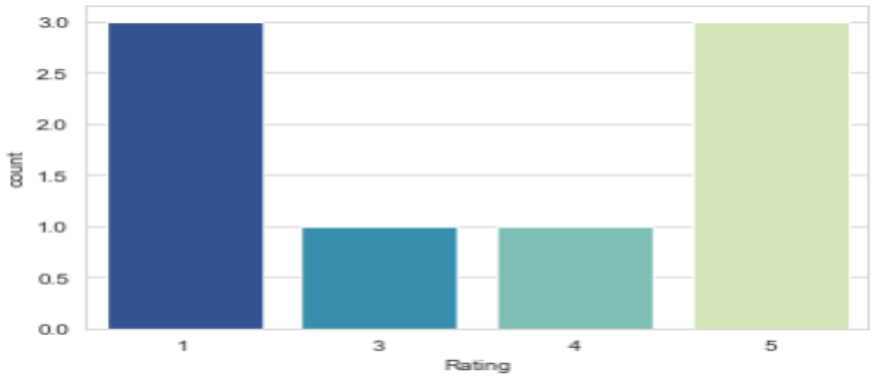

Figure 2: Online review (both text & numerical rating) extraction part for operational amplifier

Retrieval Number: 100.1/ijainn.B39181212222
DOI:10.54105/ijainn.B3918.101621
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

Out[9]:

	Text	Rating	Polarity_Rating
0	IC quality is good but packing isn't, all ICs ...	3	Neutral
1	Not a single one is working. There is no outpu...	1	Negative
2	Giving in a plastic case witth 5 free diodes i...	4	Positive
3	Good for home project	5	Positive
4	All Op- Amp ICs came with Pins bent.	1	Negative
5	Pin 6 is dead inspite of inputs. Defective ICs...	1	Negative
6	Excellent	5	Positive
7	Very very good and best quality	5	Positive

Figure 3 polarity computation part of our newly developed App for operational amplifier

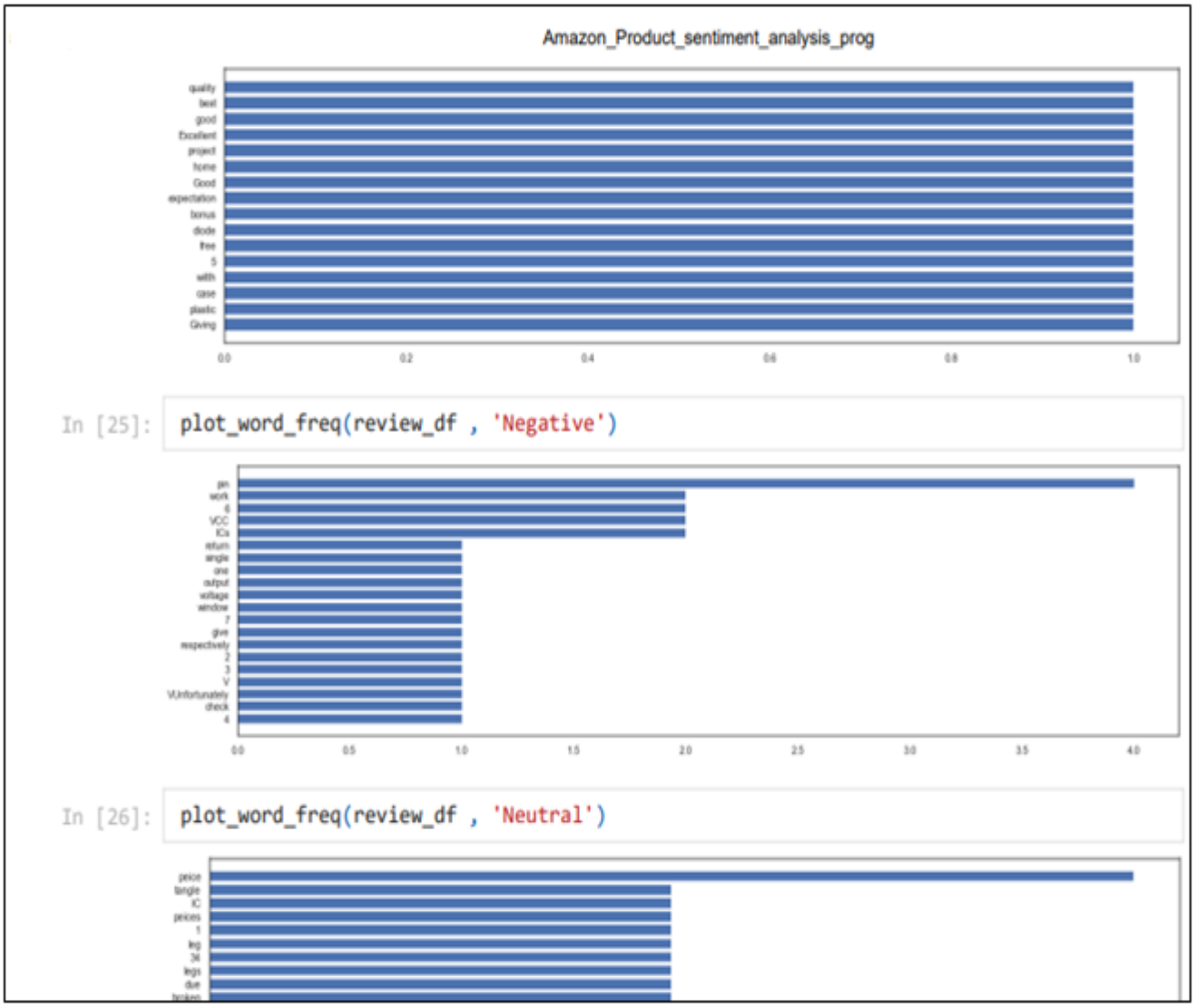

Figure 4: polarity frequency part of our newly developed App for operational amplifier

Retrieval Number: 100.1/ijainn.B39181212222
DOI:10.54105/ijainn.B3918.101621
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

Figure 5: positive word cloud part of our newly developed App for operational amplifier

```
In [31]: review_df[['Polarity_Rating' , 'Review_Processed' , 'score' ]].head()
```

Out[31]:

	Polarity_Rating	Review_Processed	score
0	Neutral	IC quality good packing isnt ICs normal small ...	{'neg': 0.135, 'neu': 0.739, 'pos': 0.126, 'co...
1	Negative	single one work output voltage pin 6 pin 4 7 g...	{'neg': 0.0, 'neu': 0.888, 'pos': 0.112, 'comp...
2	Positive	Giving plastic case witth 5 free diode bonus e...	{'neg': 0.0, 'neu': 0.352, 'pos': 0.648, 'comp...
3	Positive	Good home project	{'neg': 0.0, 'neu': 0.408, 'pos': 0.592, 'comp...
4	Negative	Op Amp ICs come Pins bent	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound...

Figure 6: score computation part of our newly developed App for operational amplifier

```
In [37]: from sklearn.metrics import accuracy_score , classification_report, confusion_matrix

In [38]: accuracy_score(rev_df_m['Polarity_Rating'] , rev_df_m['comp_rating'] )
Out[38]: 0.7142857142857143

In [39]: print(classification_report(rev_df_m['Polarity_Rating'] , rev_df_m['comp_rating'] ))
                  precision    recall  f1-score   support

        Negative       1.00      0.33      0.50         3
        Positive       0.67      1.00      0.80         4

        accuracy                           0.71         7
       macro avg       0.83      0.67      0.65         7
    weighted avg       0.81      0.71      0.67         7

In [40]: print(confusion_matrix(rev_df_m['Polarity_Rating'] , rev_df_m['comp_rating'] ))
         [[1 2]
          [0 4]]
```

Figure 7: Computations of f1-score for Texblob analysis of our newly developed App for operational amplifier

Retrieval Number: 100.1/ijainn.B39181212222
DOI:10.54105/ijainn.B3918.101621
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

Text rating and numerical rating both have been collected by our newly developed App as shown in figure 2. In the Figure 3, the polarity computation of online review has been demonstrated. Figure 4 depicts the positive polarity, negative polarity & neutral polarity of the words as present in the online review. Figure 5 demonstrates the positive word cloud formation. Similarly we formulated negative cloud too though it is not included here. The Figure 6 describes the computation of polarity scores required for VADER analysis. The final Figure 7 depicts the f1- score computation to be used for Texblob sentiment analysis.

VI. CONCLUSSION

The developed model is helpful for searching truthful information for a product from the previous user's reviews in e-commerce sites. Here authors take into account VLSI IC which plays an important role in digital era but people are not well aware about this product. It is not sufficiently available in common offline shop so people are basically dependent on e-commerce. This developed model easily analyse the previous users reviews and bring out their sentiment on that product.

MANAGERIAL IMPLICATIONS OF FINDINGS

According to the proposed Business Intelligence (BI) model, market research of VLSI IC as a product on ecommerce sites based on online reviews is a worthwhile endeavor for the manufacturers, retailers and whole sellers of VLSI IC across the country & abroad also. Customer satisfaction dimensions (CSDs) derived from the e-marketing sites reviews with respect to VLSI ICs are the most important for a potential customer. In addition to that knowledge about customers' perception of a specific electronic characteristic of VLSI IC will be useful for product up gradation.

LIMITATIONS

The online market for VLSI IC is in embryonic stage due to many factors. Consequently this study is restricted to few VLSI ICs only which are presently available (i.e. 555 Timer & Operational Amplifier) in the Amazon website. The consequence of sentiment analysis for the reviews on VLSI ICs in Amazon website could also be analyzed by modeling consumer happiness based on the sentiment strengths but due to paucity of time, this study is restricted to sentiment analysis part only.

FUTURE SCOPES

This study was done based on actual consumer comments about the VLSI ICs, but demographic profile of the consumers & technical specifications are not considered for the analysis. In future, the relation of these two factors with the findings may also be explored. This research is based on empirical work done on Amazon's review site. Customer review systems at other e-commerce firms (like flipkart, eBay, snapdeal) are not considered. It can also be explored by the future researchers. Moreover, other AI based techniques like Nave Bayes machine-learning and SVM may provide more statistical insight in future.

DECLARATION ON CONFLICT OF INTEREST

In this research work authors used publicly available data across Ecommerce websites. This work does not contain any studies with human participants or animals performed by any of the authors. Further, all three authors are attached in the Department of Business Administration, University of Kalyani and consequently they used the infrastructure of their University to carry on this research activity. Consequently there is no conflict of interest involved in this case. There is no internal and external funding agency to complete this research work.

Funding/ Grants/ Financial Support	No, did not receive fund from any resources.
Conflicts of Interest/ Competing Interests	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material/ Data Access Statement	In this research work authors used publicly available data across Ecommerce websites.
Authors Contributions	All authors having equal participation in the article.

REFERENCES

1. Gunawan.T.S; Abdullah.N.A.J; and et al.(2020) "Social Networking Analysis using Python Data Mining" 8th International Conference on Cyber and IT Service Management (CITSM)",**DOI:** 10.1109/CITSM50537.2020.9268866@ IEEE Xplore [CrossRef]
2. Hamilton.W.L, Clark.K, Leskovec.J, and Jurafsky.D(2016), "Inducing domain-specific sentiment lexicons from unlabeled corpora," in Proceedings of the Conference on Empirical Methods in NaturalLanguage Processing. Conference on Empirical Methods in NaturalLanguage Processing, 2016, vol. 2016: NIH Public Access, p. 595. [CrossRef]
3. Jin, J., Liu, Y., Ji, P., Liu, H., 2016, Understanding Big Consumer Opinion Data for Market-Driven Product Design. International Journal of Production Research, 54/10: 3019–3041 [CrossRef]
4. Lee, J., Kao, H., Yang, S., 2016, Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. Procedia CIRP, 16:3–8. 2–14. [CrossRef]
5. Pankag,Pandey.P, Muskan and Soni.N(2019),"Sentiment Analysis on Customer Feedback Data: Amazon Product Reviews", International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (Com-IT-Con), India, IEEE;ISBN:978-1-7281-0211-5/19/$31.00 2019 [CrossRef]

AUTHOR PROFILE

Biswajit Biswas, Assistant Professor in the Department of Business Administration, University of Kalyani. India. He has six years research and teaching experience in Post Graduate and Under Graduate level in the domain of Business Administration. His expertise is about Business Intelligence, Web-data mining, Soft computing, Principle of Management, Strategic Management and Digital Marketing. He has published many research articles in international journal and conferences. He also visited many prestigious academic Institute like IIMA, IIMB, IIMC, IISC, AMITY University, KIIT, DUYTUN University Vietnam and many more in India and aboard. He has attached several teaching institutes including undergraduate & postgraduate levels as an examiner.

Dr. Tuhin Mukherjee, Assistant Professor in The Department of Business Administration, University of Kalyani has a research experience of more than one decade and published many papers in national and international journals. He use to teach in Post Graduate level in the domain of Business Administration. His expertise is about Business Analytics, Statistical Methods, Machine Intelligence, and Financial Management & Option Pricing. He already served several teaching institutes including undergraduate & postgraduate levels