Volume 6 Issue 1 ISSN: 2321-1857

International Journal of Business Analytics & Intelligence

April 2018

A Publication of Publishing India Group

International Journal of Business Analytics & Intelligence

Editorial Message

Greetings from team IJBAI!!!

It's quite a mission to retaining the core value and essence of IJBAI after completion of 5th year in a row.
We are continuously aiming in bringing the crux of Data science practice as it is experiencing a rapid evolution path. The journal is a platform for exchanging the thoughts of "data to decision" journey, understanding eco system, visions from various research studies, computational techniques and technology in various applications and thereby we try to emphasis to connect between "information" and "intelligence". To make our mission successful, we warmly welcome Favio Vázquez, Principal Data Scientist at OXXO and Chief Data Scientist at Iron AI, as the new member of our coveted editorial board.

In this issue of IJBAI, we are delighted to bring out six application and technology-oriented domain focused papers.
Readers will get a great perspective on distribution tail through a different lens by Prof. Arnab Laha of IIM Ahmedabad in his column "Analytically Yours". An IBMer Ms. Madhumita Ghosh's outlook on the 'Contextual Customer Support – Cognitive Way' throws light on a new edge analytics way to tap customer experience, which is beyond predictive analytics. This is in line with Favio Vazque's paper 'Data Science and Intelligence'

In soaring e-commerce age, the lower penetration is a business concern and media impact on customer's buying habit is a crux to increase the coverage as well as enhance the penetration. Text mining in 'News Analysis' and Machine Learning Algo for impact scoring helps determining the News influence on buyers. The study and methodology is depicted in a paper meticulously.

Talent retention is key to business which has direct impact on both topline as well as bottom line. One paper aims on determining the relationship between the talent-management crunch which the institutions are facing and the strategies that are devised to hold on to their talent. Readers will get a step by step understanding to devise a structural equation model to ensure winning and retaining the talent. While mentioning about talent, it is obvious that enhancing power and quality of mankind through technology is a great vision and when an innovation takes place to fulfill, we feel our mission is in a right path. Readers will be delighted to read the paper which talks about 'Wise/Smart Cane' that alerts visually impaired people about obstacles before might facilitate them in walking with less accident

In Digital era, software plays a crucial role not only in business and human application as well as in daily lifestyle of an individual. An error free in nature; however, developing a reliable software is a major challenge faced by the software-developing industry. While we spoke about new edge analytics and intelligence, in that line, this paper presents a software reliability growth model which is used to solve the problem under fuzzy environment.

The last two papers deal with the financial analytics. The first research paper on financial analytics presented the performance of foreign exchange trade in both India and China. The aim of this study is to understand trends of currency trade to predict how likely these countries are going to emerge as best in the region. The usefulness of Random Forest forecasting technique can be witnessed here. The second research paper uses nine models for estimating optimal hedge ratio, of which six are constant hedging models and three are time-varying hedging models to suggest the use of conventional hedging tools for estimating optimal hedge ratio.

We wish to create this IJBAI, a leading repository of knowledge in data science by constantly improve the quality of the journal and thereby delight our esteemed readers. We are sure that our readers will appreciate and learn a lot from the present issue. Do let us know your wish, suggestions and views to enrich our journal. Therefore, it would be great to have valuable feedback from our learned readers about the enriched version of IJBAI. We would like to thank all the researchers and renowned data science practitioners who have honored us by selecting our journal to publish some of their research cases. At the end, we extend our heartfelt thanks to all our esteemed readers who continued to support us for the last five plus years.

Sincerely Yours,
Madhumita Ghosh
Joint Editor-in-Chief

&

Tuhin Chattopadhyay
Editor-in-Chief

International Journal of Business Analytics and Intelligence

Volume 6 Issue 1 April 2018

ISSN: 2321-1857

Journal is also available online at www.publishingindia.com

International Journal of Business Analytics and Intelligence
6 (1) 2018, 01-03
http://publishingindia.com/ijbai/

Analytically Yours:

On Tails

Arnab Kumar Laha*

This article is about tails. Not only animals have tails, but probability distributions have tails too. We will be discussing the tails of probability distributions in this article. In recent times, we have heard about different kinds of tails of probability distributions such as light tail, heavy tail, and long tail. What are these and what should we know about them? This article aims to give a glimpse.

A typical course in probability and statistics begins with definitions of measures of central tendency and variation. One of the most popular measures of central tendency of a given dataset $\{x_1,\ldots, x_n\}$ is the arithmetic mean $(\bar{x})$ which is defined as $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$. When considering a random variable X having a cumulative distribution function (cdf) F_X, we define the analogue of the arithmetic mean called the Expectation of as $E(X) = \int_{-\infty}^{\infty} x dF_X(x)$. For discrete random variables with probability distribution $\{(x_i, p_i): i = 1,\ldots\}$ where $p_i = P(X = x_i)$ and $\sum_{i=1}^{\infty} p_i$ the above expression of $E(X)$ simplifies to give $E(X) = \sum_{i=1}^{\infty} x_i p_i$. Again for absolutely continuous random variables having probability density function (pdf) f_X, we get $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$. Now, suppose we have information about $E(X)$, can we say something useful about the cumulative distribution function (cdf) F_X (recall that $F_X(t) = P(X \le t)$? Markov's inequality provides an answer to this question for positive random variables.

Suppose, the random variable is positive. Then, the Markov inequality states that for any $t > 0$, $P(X > t) \le \frac{E(X)}{t}$ i.e. $F(t) = P(X \le t) \ge 1 - \frac{E(X)}{t}$. As an example suppose $E(X) = 4$; then, we can say that $P(X > 16) \le 0.25$ and $P(X \le 100) \ge 0.96$. More generally, we can say that $P(X \le c.E(X)) \ge 1 - \frac{1}{c}$ for any $c \ge 1$.

The most widely used measure of variation of a dataset is standard deviation (sd) which is defined as $s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$. It is often convenient to work with s^2 (called the sample variance) instead of s. For a random variable X, its variance is defined analogously as:

$$\mathrm{Var}(X) = E((X - E(X))^2) = \int_{-\infty}^{\infty} (x - E(X))^2 dF_X(x)$$

In the discrete case, the above is interpreted $\mathrm{Var}(X) = \sum_{i=1}^{\infty}(x_i - E(X))^2 p_i$; whereas in the absolutely continuous case, this is interpreted as $Var(X)$ $\int_{-\infty}^{\infty} (x - E(X))^2 f_X(x) dx$. What can we say about F_X if we know both $E(X)$ and $Var(X)$? Can we do better than the case when we only knew $E(X)$?

The Chebyshev's inequality tells us that indeed we can generally do much better if we have information about both the expectation and variance of the random variable X. Noting $P(|X - E(X)| > t) = P((X - E(X))^2 > t^2)$ and then applying the Markov's inequality we get:

$$P(|X - E(X)| > t) \le \frac{E((X - E(X))^2)}{t^2} = \frac{Var(X)}{t^2}$$

Assume as before $E(X) = 4$ and suppose that we have the additional information that $sd(X) = 2$, i.e. $Var(X) = 4$ Now, $P(X > 16) = P(X - 4 > 12) \le P(|X - 4| > 12) \le \frac{2^2}{12^2}$ $= 0.028$ which is a vast improvement on the upper bound of 0.25 we obtained earlier. Taking $t = c.sd(X)$, we get, $P(|X - E(X)| > c.sd(X)) \le \frac{1}{c^2}$

* Indian Institute of Management, Ahmedabad, Gujarat, India. Email: arnab@iima.ac.in

Next we look at Cantelli's inequality which is a generalization of Chebyshev's inequality. Let us denote $E(X) = \mu$ and $sd(X) = \sigma$. Hence, $Y = X - \mu$ has expectation 0 and variance σ^2. Then, for $t, u \geq 0$, we get

$$\begin{aligned} P(X-\mu > t) &= P(Y > t) \\ &= P(Y+u > t+u) \\ &= P((Y+u)^2 > (t+u)^2 \\ &\leq \frac{E((Y+u)^2)}{(t+u)^2} \text{ (by Markov inequality)} \\ &\leq \frac{\sigma^2+u^2}{(t+u)^2} \end{aligned}$$

Since the above inequality is true for every value of $u \geq 0$, we get,

$$P(X-\mu > t) \leq \inf_{u\geq 0} \frac{\sigma^2+u^2}{(t+u)^2)}$$

Using calculus, it is easy to show that $\inf_{u\geq 0} \frac{\sigma^2+u^2}{(t+u)^2} = \frac{\sigma^2}{\sigma^2+t^2}$ yielding the Cantelli's inequality:

$$P(X-\mu > t) \leq \frac{\sigma^2}{\sigma^2+t^2}.$$

Now, if we apply Cantelli's inequality to find an upper bound of $P(X > 16)$ where the random variable X has $E(X) = 4$ and $sd(X) = 2$, we get,

$$P(X > 16) = P(X-4 > 12) \leq \frac{2^2}{2^2+12^2} = 0.$$

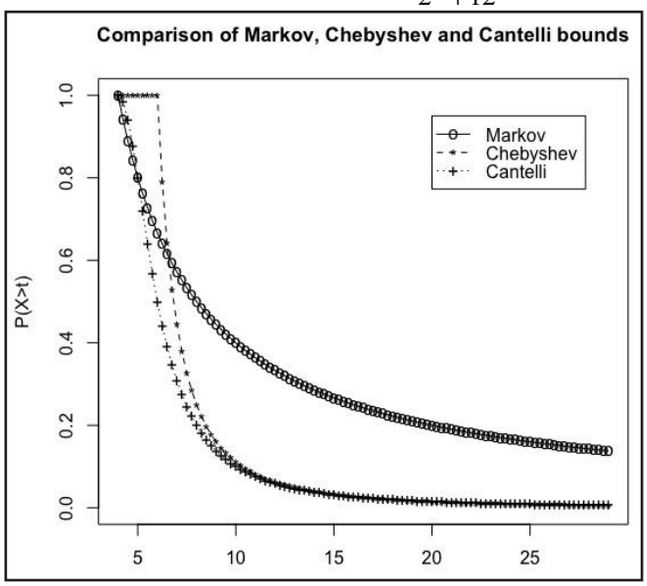


Fig. 1: Comparison of Markov, Chebyshev and Cantelli Bounds

In Figure 1, we give a comparison of the three bounds for $P(X > t)$ discussed above for different values of t.

You may recall that the moment generating function (mgf) of the random variable X is defined as $M_X(\lambda) = E(e^{\lambda X})$. The distributions for which the mgf exists are called *light-tailed*. Normal, Exponential, Uniform, and Gamma are some commonly occurring light-tailed distributions. Since the exponential function is strictly increasing, for $\lambda > 0$ we get using Markov inequality:

$$P(X > t) = P(\lambda X > \lambda t) = P(e^{\lambda X} > e^{\lambda t}) \leq \frac{E(e^{\lambda X})}{e^{\lambda t}} = \frac{M_X(\lambda)}{e^{\lambda t}}.$$

Since this inequality holds for every $\lambda > 0$, we get,

$$\begin{aligned} P(X > t) &\leq \inf_{\lambda>0} \frac{M_x(\lambda)}{e^{\lambda t}} = \inf_{\lambda>0} e^{\ln M_X(\lambda) - \lambda t} \\ &= e^{\inf_{\lambda>0}(\ln M_X(\lambda) - \lambda t)} \end{aligned}.$$

The above inequality is referred to as the Chernoff's inequality.

As an example suppose $X \sim N(0, \sigma)$. Then, $M_X(\lambda) = e^{\frac{1}{2}\lambda^2\sigma^2}$. Now using calculus, it can be easily seen that $e^{\inf_{\lambda>0}(\ln M_X(\lambda)-\lambda t)} = e^{-\frac{t^2}{2\sigma^2}}$ which implies $F(t) = P(X \leq t) \geq 1 - e^{-\frac{t^2}{2\sigma^2}}$

A distribution is said to be *heavy-tailed* if the mgf $M_X(t)$ does not exist (i.e., is not finite for any $t > 0$). Pareto, Cauchy, and Log-normal distributions are some examples of heavy-tailed distributions that occur frequently in Finance as distribution of returns of stocks. Note that a distribution which is not heavy-tailed is light-tailed. Table 1 gives exceedance probabilities for both normal and Cauchy distributions with Median = 0 and MAD = 1 (recall MAD is the acronym for Median Absolute Deviation about median). For normal distribution, the standard deviation is related to MAD through the relation $\sigma = 1.4826MAD$. In Table 1, we observe that the exceedance probabilities $P(X > x)$ for the Cauchy distribution are much larger than those of the normal distribution. While the chance of observing a departure of magnitude greater than 15 MAD from the median is practically 0 for the normal distribution, it is approximately 2% for the Cauchy distribution. This suggests that a careful study of the tail of a probability distribution is very important in many business applications such as determination of Value-at-Risk of market investments.

Now suppose, $X \sim F$ satisfies the condition that for any fixed $y > 0$, $P(X > x + y | X > x) \to 1$ as $x \to \infty$, i.e., $\frac{1-F(x+y)}{1-F(x)} \to 1$ as $x \to \infty$. Then, the distribution F is said

to be long-tailed. It can be proved that every long-tailed distribution is heavy-tailed. However, the converse is not true in general. Some interesting properties of random variables having long-tailed distributions are as follows:

(a) Let $X \sim F$ and $Y \sim G$ be two independent random variables and suppose that the distributions F and G are both long-tailed. Then, the distribution of $X + Y$ is also long-tailed.

(b) Suppose $X_1,\ldots, X_n$ are independent identically distributed random variables having a long-tailed distribution then the distributions of (i) $max\{X_1,\ldots, X_n\}$, (ii) $min\{X_1,\ldots, X_n\}$, and (iii) $X_1 + \ldots + X_n$ are all long-tailed.

Table 1: Exceedance probabilities for Normal and Cauchy distributions with median = 0 and MAD = 1

x	*P(X >x) for Normal*	*P(X > x) for Cauchy*
3	2.1×10^{-2}	0.102
4	3.5×10^{-3}	0.078
5	3.7×10^{-4}	0.063
6	2.6×10^{-5}	0.053
8	3.4×10^{-8}	0.040
10	7.7×10^{-12}	0.032
15	0	0.021
20	0	0.016

International Journal of Business Analytics and Intelligence
6 (1) 2018, 04-06
http://publishingindia.com/ijbai/

Contextual Customer Support – An Outlook

Madhumita Ghosh*

Introduction

Machine learning and Artificial Intelligence (AI) developments are happening at a prompt speed! Post the soaring usage of AI in customer care and support, the most apparent reason behind why users seek contextual help is – 'Convenience'. Support teams can conveniently solve their queries without waiting for any assistance or referring to a user manual and different system support while saving on a lot of time. According to a research report, approximately 65% of customers say that valuing their time is the most important thing a company can do to provide good and swift service. However, apart from this, there are many more reasons as to why the users direly want to seek contextual help. Tech giants are investing heavily in both applications and R&D, with an objective to stay ahead of the curve of what many believe to be an inevitable AI-led paradigm shift. At the forefront of this resurgence are the fields of conversational interactions (personal assistants or chatbots) and computer vision and autonomous navigation (in IVR (Interactive Voice Response)), with advancement in hardware, data availability, accessibility, and radical machine learning techniques that have enjoyed tremendous progress within the span of just a few years.

With the advent of AI-powered 'Customer Responses' in IVR, chatbots, and email-bots, the long-lasting problem has a possible solution. Despite a few disagreements on this notion, chatbots possess the potential to provide more natural, smoother, and better customer experience (CX).

Responsive and good customer services are all about bringing customers back as they are the lifeline for business. As the number of calls / emails / chats on contact centre increase, organization deploys IVR to reduce the human-cost and deal with the primary customer service functions.

Notwithstanding the above strategy of deploying IVR, everyone understands the exasperation upon interacting with them. One can come up with the simple question and then routed to the endless options until you press "0" to talk to someone who could understand your query in a precise way. With the advent of AI-powered IVR, this long-standing problem has a possible solution. Despite a few disagreements on this notion, AI-powered contextual IVR possess the potential to provide more natural, smoother, and better customer experience (CX).

IVR systems work in a directed manner. They navigate the static options to stream the conversation with customers and do not allow to deviate from the actual flow. This flow does not influence the customer as compared to the natural one. Once the contextual IVR is developed with business rules, then upon collecting the further subscriber information on journey path by context or mix of contexts AI can be applied. This objective of context-aware IVR with an aim to provide a first-time resolution to customer queries will reduce repeat calls and improve customer satisfaction. The purpose of context-aware IVR is to collect relevant information available pertaining to the customer and provide a contextually aware and personalized self-care experience to end users calling to the organization's IVR.

Getting the correct context and determining the proper goal of customer's query is the primary goal of any customer service. Real-life conversations are not always linear. Chatbots are flexible and natural. If customer switched the context of discussion from meeting to the timing of the branch, BOT kept the context and responded back accurately. Later on, when the customer just wrote 6 PM right before the confirmation so the BOT need not to require going back and come again with all the questions in a linear way. This is another unique feature of chatbots compared to the traditional IVR.

* Practice Leader - Big Data & Advanced Analysis BA & Strategy - Global Business Services IBM, India.

Likely Smart-Sales with Predictive Analysis

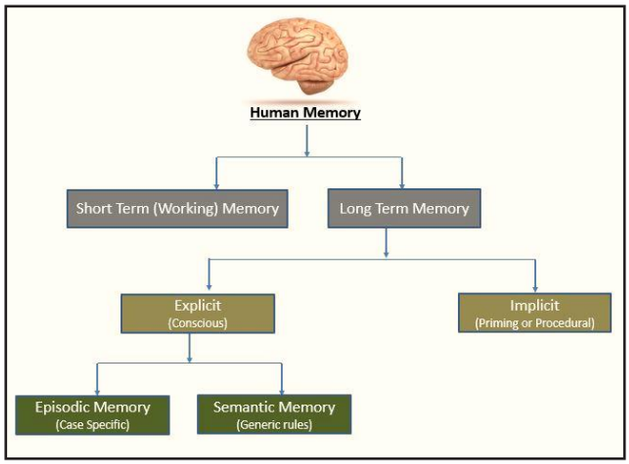
IVR navigates the user to the point that is already there in its option set and takes up the conversation from thereon. IVR systems do not use any AI or machine learning to suggest something else based upon the customer's objective or discussion.

In contrast, based upon having the context, non-linear dialogue communication capability and AI-powered chatbots have a feature to recommend something useful or informative to the end users. For example, if you're trying to book a movie in a theater, it suggests you some good dinner deals based upon your taste in nearby food outlets. As they is power of machine learning as a backend system, they do not only suggest or recommend the best possible options, but also learn and improve themselves continuously for forthcoming communications.

Generally, chatbots are flexible and give the consumer a feel of real person communicating with them because of the power of natural language understanding, contextual information management, and so on. However, IVRs are obstinate and do not allow customer to deviate from its context which do not maintain a natural composition during conversation. Despite the slow adoption of this change among customer services platforms, chatbots are still growing and getting improved by the time with easy handling capability of the first encounters of the customers.

Combinations of Multiple Disciplines

We are still in the early stages of the AI-powered conversational revolution, and it is fair to assume some problems that seem undefeatable today will likely be solved in the forthcoming years. We are quickly moving towards a ecosphere in which one will be able to have long and multifaceted interactions with AI support, which will not only understand what one wants to say, but will also know your preferences and tailor your experiences accordingly. To do so, the need is to combine multiple disciplines, including deep learning, applied statistics, and others, building technology that blends consumer preferences, environment, and language into one piece of intelligent, and flexible software; however, today's date system architects are already gearing towards that integrated approach.



One of the main drivers behind this wave of novel AI applications is deep learning, an area of machine learning that, despite existing for many decades, has recently revolutionized the fields such as computer vision and natural language processing (NLP). Nonetheless, despite its incredible performance, deep learning alone is not sufficient to solve the challenges faced by chatbots. Understanding context, disambiguating between subtle differences in language that can lead to wildly different meanings, employing logical reasoning, and most crucially, understanding the "preferences and intent of the consumer", are just a few of the many challenging tasks a system must be able to perform to sustain conversation with a human behind the machine. If a modern conversation engine hopes to go beyond answering simple, one-level questions, it must blend the most prominent techniques emerging from the field of deep learning with solid statistics, linguistics, other machine-learning techniques, and more structured classical techniques, such as semantic parsing and program induction.

The first building block in building an intelligent conversational system is information richness. In particular, deep learning is disreputable for needing vast amounts of high-quality data before it can unleash its true potential. However, while we live in an era where endless streams of data are constantly being generated, most of it is too raw to be of immediate use for machine learning algorithms. Unsupervised Learning, the subfield of machine learning dedicated to extracting information from raw data unaided by humans, is likely a promising alternative. Among its many uses, it can be utilized to build

an embedding model, which allows data to be represented in a less complex form, allowing patterns to be discovered more easily. While unsupervised learning is already omnipresent in machine learning, deep learning offers additional innovative ways to build — such embedding models — providing the state-of-the-art performance. Optimization of these techniques can alleviate the need for a lot of high quality and expensive labelled data, which is essential in getting AI chatbots to perform well.

Scientists are constantly (and usually subconsciously) checking every new piece of information we receive from our surroundings against an internal model of the world — a model of what is normal and what is not, of how entities are related, how we can make logical inferences involving said entities, and so on.

Finally, the ability to put it all together is yet another edge waiting for a solution. Unlike a search engine where the user is content with being presented a list of matches ordered by relevance, a conversation engine must be more specific. Simply using NLP to identify a set of relevant information is insufficient. It should be able to parse the input, break it down, and present a response to the user that is not only clear and concise, but is also highly relevant to their taste — solution and repeat.

As one embarks on the transition from multichannel to omnichannel, contextual with consistent connections with customer approach would help getting customers the experiences they're seeking as they go about their busy lives. It can also put the employees with a central set of cross-channel tools that gives them the integrated customer history, joined-up processes, and consistent answers they need. With these, employees will emerge more enabled, more empowered, and well positioned to deliver efficient customer service that meets and exceeds customer expectations.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 07-08
http://publishingindia.com/ijbai/

Data Science and Intelligence

Favio Vázquez*

Abstract

I show in this article that Data Science, Big Data, and Artificial Intelligence (AI) are serious areas of research and development. We need all of them to get to Artificial General Intelligence (AGI), and also how Data Science is crucial to this endeavour.

What is Intelligence?

Not an easy question to answer. I struggled a while ago trying to define what it was, but I came across a single-phrase definition that I liked.

So let's define intelligence as: The ability to accomplish complex goals.

But what is complex here? How are we defining something complex? If you look in the Internet, you will find several different definitions; but, I think the "main" one is close to what I think the definition of intelligence is talking about.

If we think that something complex is something having many parts related to each other in ways that may be difficult to understand, then we can say that something complex is a mix of things or parts, which combine together to form a bigger thing, and the way those parts are related is not very easy to understand.

For example, something complex that has many parts which work together in a way that is not that easy to understand is a car. But, if we take a look, the individual parts are not that hard to understand. I'm not saying that they are easy to build or see exactly what they do; however, it is easier to grasp what they do.

So, we can say now that intelligence is:

The ability to accomplish difficult goals by understanding the parts that form the main goal.

These goals will be defined in the context we want, but now we want to focus on the field of Artificial Intelligence (AI). So as AI wants to build intelligence using machines and computation trying to mimic the ways we as human see, hear, learn, and more, these goals will be seeing, learning, hearing, moving, understanding, and more.

What is Understanding?

Other important concept that we need to look at is understanding. We have used that word several times by now, so let us define it:

Understanding is the ability to turn complex information into simple, useful information.

We need this, and we talked about it when we saw the parts of the car. When we are trying to understand, we decode the parts that form this complex thing, and transform the raw data we collected in the beginning to something useful and simple to see.

We do this by modeling. This is the process of understanding the "reality", the world around us, but creating a higher-level prototype that will describe the things we are seeing, hearing, and feeling; however, it's a representative thing, not the "actual" or "real" thing.

So, how we humans create intelligence: by modeling the world around us, understanding its parts, transforming the raw data we collected into useful and simple information in order to see how these parts form more complex things, accomplishing in the end goals, i.e., "difficult" goals.

Why Do We Need to Create Intelligence with AI?

I think the recipe to create intelligence is not that hard at a high level. This is why I propose we need to do it:

* Data Scientist, BBVA Data and Analytics and Chief Data Scientist at Iron AI.

Big Data + AI + Data Science = Artificial General Intelligence

I'm talking about Artificial General Intelligence (AGI) as the main goal of this revolution. AGIs are general-purpose systems with intelligence comparable to that of the human mind (or maybe beyond humans).

We need Big Data as a Catalyst to get to AGI, because with more data, in addition to new ways of analysing data and better software and hardware, we can create better models and better understanding. We need the current state of AI, very close to Deep Learning, Deep Reinforcement Learning, and its surroundings, and then we need Data Science as the controller and science behind this revolution.

What is Data Science?

This definition may cause controversy for some people, but this is something which I think is very close to what the leaders (both theoretical and in the business) are saying right now. Thus,

Data Science is the resolution of business/organization's problems through mathematics, programming, and the scientific method that involve the creation of hypotheses, experiments, and tests through the analysis of data and the generation of predictive models. It is responsible for transforming these problems into well-posed questions that can also respond to the initial hypothesis in a creative way. It must also include the effective communication of the results obtained and how the solution adds value to the business/organization.

And with this definition, we can define who a Data Scientist is. Thus,

A Data Scientist is a person (or system?) in charge of analysing business/organization's problems and gives a structured solution starting by converting this problem into a valid and complete question, then using programming and computational tools to develop codes that prepare, clean, and analyse the data to create models and answer the initial question.

What I'm saying here is that Data Science is very much linked to the business, but it is a science in the end, or in the process of becoming one, or maybe not. I think, it could be very useful that Data Science is a Science because if that's the case, every project in Data Science should be at least:

Reproducible: Necessary for making easy to test other's work and analysis.

Fallible: Data Science and Science are not meant to look for the truth; they are meant to look for knowledge, so that every project can be substituted or improved in the future, as no solution is the ultimate solution.

Collaborative: The data scientist doesn't exist alone; he needs a team. This team will make things possible for creating intelligence and solutions. Collaboration is a big part of science, and data science should not be an exception.

Creative: Most of what data scientists do is new research, new approaches, or takes on different solutions, so their environment should be very creative and easy to work. Creativity is crucial in Science, and is the only way we can find solutions to hard and complex problems.

Compliant to Regulations: Right now, there are a lot of regulations on Science, not that much on Data Science, but there will be more in the future. Is important that the projects we are building can be aware of these different types of regulations so we can create a clean and acceptable solution to the problems.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 09-13
http://publishingindia.com/ijbai/

Content Analysis Concerning Online Shopping in UAE: Evaluation of Impact Score from News

Jitendra Singh Rathore*, Rajita Srivastava**

Abstract

The news in newspaper and magazine plays an imperative role to apprise public. People get influenced by the contents (either news or articles), and for businesses, this subjects to spontaneous decisions. It also corroborates greater defy for businesses owing to mammoth amount of data aggregated. The research conducts content analysis of "online shopping news" from two prominent newspapers (Gulf News and The National) and other protuberant sources (including Khaleej Times and Arabian Business) in UAE, summing to 49 news published during 2016-2017. The aim of collecting these documents is to find their impact (in terms of polarity using Plutchik emotional model) on public. The tests consist of impact identification phase, which associates news with impact aggregation and scoring.

Keywords: Content Analysis, Gulf News, Impact Score, Plutchik Emotional Model, The National

Introduction

The study is grounded on two vital observations, wherein, on one side the competition in online shopping is getting tougher in UAE with Amazon acquired Souq (Turner & Wang, 2017), and brick-and-mortar retail company, Emaar, procured two online retail companies, Namshi and JadoPado (Reuters, 2017; Scott, n.d.). Secondly, these acquisitions are happening in a place where online shopping accounted for just 1% of total shopping value as compared to international average of 15-20% elsewhere (Euromonitor, 2017). Both these annotations are contrary, as companies see potential in the market and customers instead of trusting of local online shopping companies, do cross-border shopping. Almost 46% of UAE online shoppers purchased online from websites outside of UAE (Jayakumar, 2016). This disintegration draws a major attention — does the news regarding local online shopping in UAE newspapers and magazine have any impact on consumer buying behaviour?

Newspapers and magazines do play an imperative role in informing public about the happenings; predominantly, in those areas in which audiences do not possess unswerving knowledge or experience (Happer & Philo, 2013). There is also an argument that newspapers and articles not only publicize opinion, but also generate stimulus for these opinions (Livingstone, 1994). Researchers focus on news articles and are mainly concerned about the polarity of these articles, whether they convey a positive or negative attitude towards the subject of discussion (Haider-Markel, Allen, & Johansen, 2006). These polarities generate shoppers' mindset, referred to as "specific cognitive orientation", resulting in different modes of emotions and information processing (Wind & Mahajan, 2001). The large linguistic miscellany in these documents makes these newsworthy, but challenging the cradle of information (Domann & Lommatzsch, 2017). Newspaper and magazine articles can be precarious, as customer can endorse or criticize a product through it. This information is of prodigious significance for a business because the existence of business depends on it (Lommatzsch, Bütow, Ploch, & Albayrak, 2017). As the reader of these news is emotional, the overall impact analysis can be evaluated using the eight different types of polarities (Plutchik, 2001).

In this paper we present a supervised methodology to compute impact analysis for news articles. Using text mining, the texts can be categorized inevitably with their polarities; thereby, overall impact score of each news can be assessed independently. In the course of this paper,

* Assistant Professor, Faculty of Management Studies, WISDOM, Banasthali University, Rajasthan, India.
** Research Scholar, Faculty of Management Studies, Rajasthan, WISDOM, Banasthali University, Rajashthan, India.
Email: rajitariktesh@gmail.com

we describe the development of "News Analysis Model" that takes advantage of a machine learning approach to accomplish this objective. We first collect the news from three types of sources, and then conduct the evaluation of emotions using Plutchick Emotional Model. Furthermore, we compare our findings; the results are presented on a front-end. Finally, we draw an inference in which we also recommend points for additional research and augmentation of our current approach.

The complete paper is divided into four sections. Section 2 particularizes the proposed News Analysis model, which comprises of four steps. Section 3 expresses the results analysis for 49 news, contributed from three types of sources, exploring the comprehensive set of outcomes. Section 4 accomplishes the paper with inferences and conclusion.

News Analysis Model

There are four steps in news analysis model, as mentioned in Figure 1. Step 1 is collection of news about pertinent topic of interest. For the study, we selected three types of sources (*Gulf News*, *The National,* and Others). Step 2 is preprocessing, which is done to remove common words and URLs from the texts. Step 3 does tokenization, which chops the document into individual words. Step 4 conducts the impact of these words using Plutchik Emotional Model.

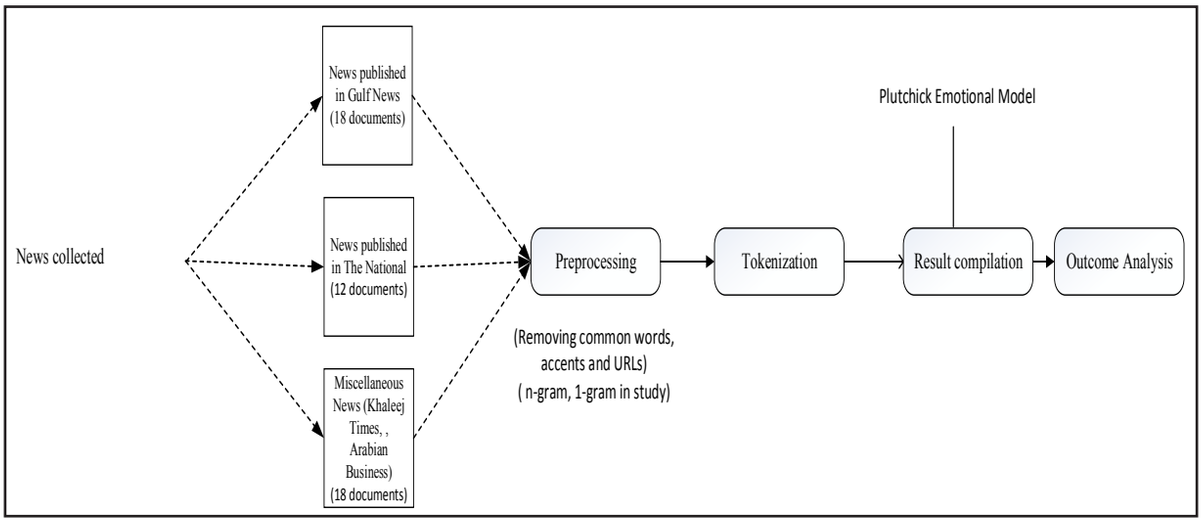


Fig. 1: News Analysis Model

Preprocessing

The documents collected from newspapers and magazines are in raw format. Preprocessing converts this raw format into more meaningful format for analysis. The steps for transformation are:

- Convert all text into lowercase.
- Remove all diacritics or accents in text, as héllo → hello.
- Documents also include links. These tags act as a single word and results in wrong identification of emotions. <a href> i love online shopping gets converted into → i love online shopping.
- Remove all URLs from documents.

The steps can be represented mathematically by equation (1) as:

$$\sum_{i=1}^{n}\prod_{j=1}^{m} d_{ij} = \sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(d_{11})\\(d_{12})\\\vdots\\\vdots\\\vdots\\(d_{nm})\end{bmatrix}\right] = \sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(Word_{11}Word_{12}\ldots Word_{1j})\\(Word_{21}Word_{22}\ldots Word_{2j})\\\vdots\\\vdots\\\vdots\\(Word_{n1}Word_{n2}\ldots Word_{nm})\end{bmatrix}\right] \quad (1)$$

where, is number of documents where each documents contain words. Notice from equation (1) that each word of documents is not disjointed and embodied as sentences.

Tokenization

Tokenization is a process of breaking documents into words (as stated in equation (1)). These words are smaller components called tokens. To do so, broad rules of splitting each word in the documents by:

- Punctuations are also considered as tokens separately from words. These tokens are represented by either whitespace or full stop.
- Emoticons are also used as tokens. The impact of emoticons is relatively substantial in categorizing impact analysis.

Equation (1) can now be defined as in equation (2) as:

$$\sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(d_{11})\\(d_{12})\\\vdots\\\vdots\\\vdots\\(d_{nm})\end{bmatrix}\right]=\sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(Word_{11})(Word_{12})\ldots\ldots(Word_{1j})\\(Word_{21})(Word_{22})\ldots\ldots(Word_{2j})\\\vdots\\\vdots\\\vdots\\(Word_{n1})(Word_{n1})\ldots\ldots(Word_{nm})\end{bmatrix}\right] \tag{2}$$

Notice, that each word from collected documents is now separated into expressive tokens. Based on equation (2), the number of tokens for the collected documents is mentioned in Table 1.

Table 1: Tokens Collected from Documents

Document Types	*Total number of documents*	*Number of words*	*Number of Tokens*
Gulf News	18	10093	5146
The National	13	7243	3705
Miscellaneous	18	7731	4166

Impact Analysis

The impact analysis is conducted using Plutchik model (Plutchik, 2001). Plutchik model divides the words into polarities of eight categories – Joy, Surprise, Trust, Anticipation, Anger, Disgust, Fear, and Sadness. The first four categories give positive impact score and later four give negative impact score, as mentioned in equations (3) and (4), respectively.

$$\sum_{j=1}^{m} s1_j = \sum_{j=1}^{m}\left[\prod_{j=1}^{m} Joy_j + \prod_{j=1}^{m} Surprise_j + \prod_{j=1}^{m} Trust_j + \prod_{j=1}^{m} Anticipation_j\right] \tag{3}$$

$$\sum_{j=1}^{m} s2_j = \sum_{j=1}^{m}\left[\prod_{j=1}^{m} Anger_j + \prod_{j=1}^{m} Disgust_j + \prod_{j=1}^{m} Fear_j + \prod_{j=1}^{m} Sadness_j\right] \tag{4}$$

Based on equations (3) and (4), the overall impact score for each document is calculated as mentioned in equation (5) below.

$$IS_j = \left[\sum_{j=1}^{m} s1_j - \sum_{j=1}^{m} s2_j\right] \tag{5}$$

Grounded from equation (5), if , the document is quantified to have positive impact, else if , the document provides negative impact, neutral otherwise. Overall, the impact score is revealed in equation (6).

$$\sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(IS_1)\\(IS_2)\\\vdots\\\vdots\\\vdots\\(IS_j)\end{bmatrix}\right]=\sum_{i=1}^{n}\left[\prod_{j=1}^{m}\begin{bmatrix}(Word_{11})\to\langle s1_j,s2_j\rangle\,(Word_{12})\to\langle s1_j,s2_j\rangle\ldots\ldots(Word_{1j})\to\langle s1_j,s2_j\rangle\\(Word_{21})\to\langle s1_j,s2_j\rangle(Word_{22})\to\langle s1_j,s2_j\rangle\ldots\ldots(Word_{2j})\to\langle s1_j,s2_j\rangle\\\vdots\\\vdots\\\vdots\\(Word_{n1})\to\langle s1_j,s2_j\rangle(Word_{n1})\to\langle s1_j,s2_j\rangle\ldots\ldots(Word_{nm})\to\langle s1_j,s2_j\rangle\end{bmatrix}\right] \tag{6}$$

Research Analysis

The research analysis includes the independent content analysis for three types (18, 13, 18) of documents collected through varied sources. Figures 2, 3, and 4 depict the content analysis of 49 news published concerning online shopping in UAE during 2016–2017. The selection is purely on news published (without revisions or public

feedback), as we intend to observe the direct impact of news and evaluate impact score as mentioned in equation (6).

Gulf News Observations

The three distinguishable factors for *Gulf News* are. The outcome, as mentioned in Figure 2 clearly indicates.

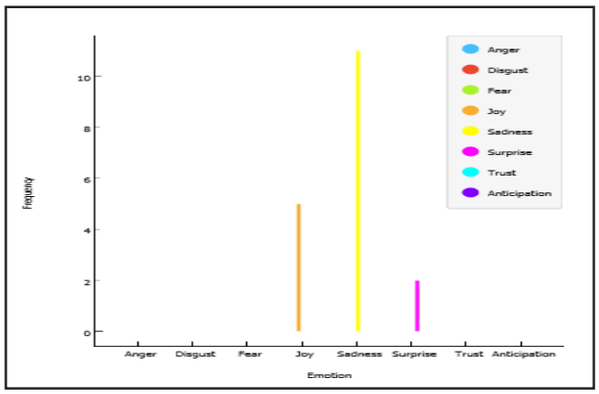


Fig. 2: Gulf News Impact Analysis

The National Observations

Again, the three distinguishable factors for *The National* are. The outcomes, as mentioned in Figure 3, illustrate.

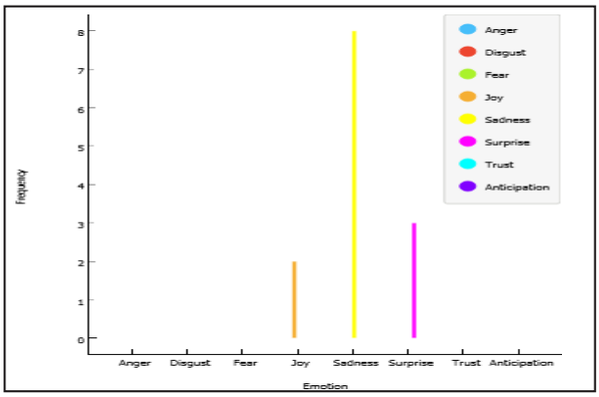


Fig. 3: The National Impact Analysis

Miscellaneous Observations

The miscellaneous collection of news includes collection from other important sources including *Khaleej Times* and *Arabian Business* and again three distinguishable factors are. The outcome of analysis clearly indicates.

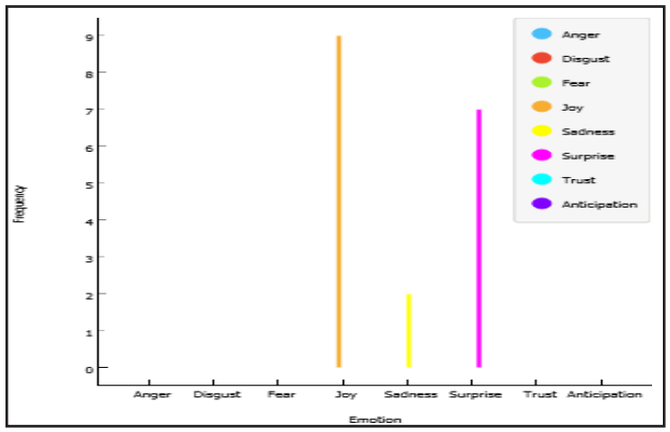


Fig. 4: Miscellaneous News Impact Analysis

Inferences and Conclusions

Usage of published news in newspapers and magazines with the computational conduct of subjectivity of text becomes vital part of direct spurt on the curiosity of consumers. The results obtained through analysis of 49 such news are quite upsetting. Implementation of algorithm provides the results, as specified in Table 2. The three polarities in all three news analysis are, other polarities are set to 0 and thus ignored. Astoundingly, overall impact score of *Gulf News* and *The National* (two leading newspapers in UAE) are negative; whereas, impact of news from miscellaneous sources is exceedingly positive.

Table 2: Overall Impact Score

News Source	*Documents collected*	*Impact --> Joy*	*Impact --> Sadness*	*Impact --> Surprise*	*Impact --> +ve*	*Impact --> -ve*	*Overall Impact Score*
Gulf News	18	27.8	61.1	11.1	38.9	61.1	-22.2
The National	13	15.38	61.54	23.08	38.46	61.54	-23.08
Misc.	18	50	11.11	38.89	88.89	11.11	77.78

As outcomes show, news published in newspaper may also be one of the reason that UAE has highest cross-border online shopping rate. The study indicates that foremost drive for UAE online shoppers to do cross-border online shopping is – 47% regarding secure payment, 44% because of free shipping, and 43% revealed proof of product genuineness (Jayakumar, 2016). The electronic commerce companies in UAE should focus that news published in newspaper should mention the positivity regarding the above three factors to boost confidence among consumers. Similarly, for magazines and other sources of information, these companies should lay stress on advertisements and product information.

The research will be further enhanced by evaluating opinions with UAE-based electronic commerce companies through social media analysis (especially on Facebook and Twitter). Combination of these two sources of information will provide a greater insight to companies and communication strategies can be deliberated and implemented accordingly.

References

Domann, J., & Lommatzsch, A. (2017). A Highly Available Real-time News Recommender based on Apache Spark. *CLEF' 17: Proceedings of the 8th International Conference of the (CLEF) Initiative*.

Euromonitor. (2017, January 1). Retailing in the United Arab Emirates. Euromonitor. Retrieved from http://www.euromonitor.com/retailing-in-the-united-arab-emirates/report

Haider-Markel, D. P., Allen, M. D., & Johansen, M. (2006). Understanding variations in media coverage of u.s. supreme court decisions comparing media outlets in their coverage of Lawrence v. Texas. *The Harvard International Journal of Press/Politics*.

Happer, C., & Philo, G. (2013). The role of the media in the construction of public belief and social change. *Journal of Social and Political Psychology*, *1*(1), 321-336. Retrieved from https://doi.org/10.5964/jspp.v1i1.96

Jayakumar. (2016, December 7). Infographic: UAE Online Consumers Shopping Trends Revealed. Retrieved December 1, 2017, from https://arabiangazette.com/uae-online-consumers-shopping-trends-revealed/

Livingstone, S. (1994). Watching talk: Gender and engagement in the viewing of audience discussion programmes. *Media, Culture and Society*, *16*(3), 429-447.

Lommatzsch, A., Bütow, F., Ploch, D., & Albayrak, S. (2017). Towards the automatic sentiment analysis of German news and forum documents. *Springer*, 1-16.

Plutchik, R. (2001). The nature of emotions. *American Scientist*, *89*, 344-350.

Reuters. (2017, May 11). Fund led by Dubai billionaire Alabbar buys UAE website JadoPado. Reuters. Retrieved from https://www.reuters.com/article/emirates-technology-jadopado/fund-led-by-dubai-billionaire-alabbar-buys-uae-website-jadopado-idUSL8N1ID1GK

Scott, A. (n.d.). Emaar Malls takes 51 per cent stake in online fashion retailer Namshi - The National. Retrieved December 1, 2017, from https://www.thenational.ae/business/emaar-malls-takes-51-per-cent-stake-in-online-fashion-retailer-namshi-1.60869

Turner, G., & Wang, S. (2017, March 28). Amazon Buys Souq.com as Middle East Online Market Takes Off. *Bloomberg.Com*. Retrieved from https://www.bloomberg.com/news/articles/2017-03-28/amazon-wins-battle-to-buy-middle-east-e-commerce-firm-souq-com

Wind, Y., & Mahajan, V. (2001). *Digital marketing: Global strategies from the world's leading experts* (1st ed.). Wiley.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 14-24
http://publishingindia.com/ijbai/

A Study on the Impact of Employee Engagement and Organizational Commitment on Talent Retention: A Structural Equation Modeling Approach

Srividya Prathiba C.S.*

Abstract

This paper aims at devising a structural equation model on employee engagement, organizational commitment, and talent retention in new generation private sector banks. The study analyses the relationship between the talent-management crunch which the banks are facing and the strategies that are devised in order to hold on to their talent. This article aims to review and discuss existing talent-management practices in new generation private sector banks and the strategies adopted by them through which, they retain their indispensable talent. Thus, this study aims to devise a structural equation model to ensure winning and retaining the talent. For this purpose, a conceptual framework is formulated and tested with a sample of 500 employees working with private sector banks. The results reveal that employee engagement and organizational commitment positively lead to talent retention.

Keywords: Employee Engagement, Human Capital, Impending Gap, Organizational Commitment, Talent Management

Introduction

Employees are said to be the assets of an organization. Research tells us that great talents outdo normal ones in many ways, including greater productivity, lower employee turnover, better client services, and greater employee morale and motivation. Given the probable talent crunch in the forthcoming years in the banking sector as employees are promoted faster to higher roles, the challenge lies in finding requisite talent and retaining them. Banks do not escape this reality and in fact, face some thoughtful and unique challenges in finding potential talent. This paper aims at identifying the strategies devised by banks to retain talent. It explores to understand the link between talent retention and employee engagement.

Theoretical Background and Literature Survey

Given the industry's significant financial crisis, most of the banks anticipate increased turnover of managers (Hymowitz, Carol 2008). The limited strength that exists in most of the banks, coupled with fast-tracked career growth, intensifies this impending gap. Banking sector is not the only industry confronting talent crunch challenges; yet, its challenges are unique and expected to amplify sooner than they do in other sectors (Brousseau, Kenneth R., Driver, Michael J., Hourihan, Gary & Larsson, Rikard 2006). Strategic talent management is a part of 21st century's essential changes as well as a complement to establishment of change in the organizations (Grossman R.J 2007). Talent management is the foundation for success in this economy . There are several reasons as to why recognizing, developing, and retaining talent is becoming more popular within organizations.

"To meet the challenge, companies must rethink how they hire, train and reward and retain their employees" (Ready, D. A., & Conger, J. A. 2007). Employee engagement is considered to be the most powerful and effective HR practice that facilitates to make greatest contribution by the people who are capable of creating added value in product and service which is rare and

* Associate Professor & Head Department of Accounting and Finance, M.O.P. Vaishnav College for Women, Chennai, Tamil Nadu, India. Email: pravid1284@yahoo.com

inimitable. It enhances job involvement, job satisfaction, career satisfaction, and organizational commitment (Noorliza, K. and Hansom. 2006). Employee engagement is considered a prime priority for senior executives. In this current and challenging globalized economy, business leaders need high-performing workforce for growth and survival. They recognize that a highly engaged workforce can increase innovation, productivity, and bottom-line performance and thus enables employees to be committed to the organization (Harward Business school publishing, 2013). There is clear evidence to suggest what employees look for in their work is a mixture of both tangible and intangible elements that creates a stimulating environment where their contribution is recognized and suitably rewarded (Kaul, V. M., 2011). Organizational commitment is said to be the force that binds an employee's course of action to achieve one or more targets (Cohen, 2003 Liu, Y; Cohen, 2010). According to Athey R., (2004), "individuals need greater elasticity in their career paths, and organizations need greater elasticity from employees". Finally, as a current trend, individuals are less-focused on short-term rewards and taking a long-term view of their career development. Hence, they are choosy about their organization. Thus, organizations have started converging on talent-management issues and focus on providing a successful career path to its employees. Thus, key to the realization of the banks generally rests on the efficacious application of talent-retention strategies.

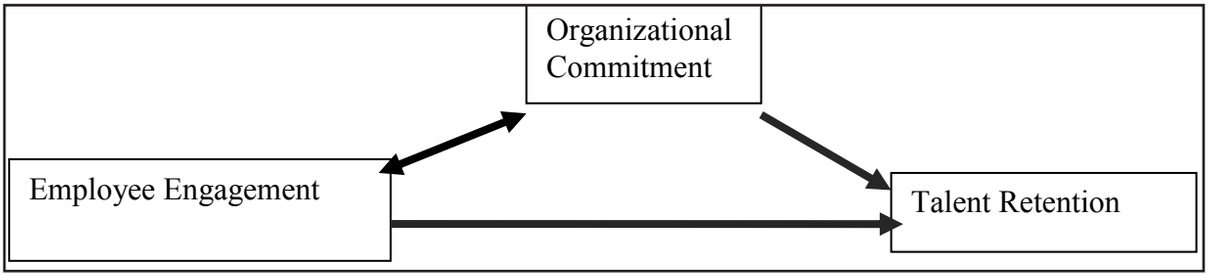


Fig. 1: Conceptual Framework

Conceptual Model and Hypotheses Development

The conceptual model of the research is developed based on the literature survey in which two main constructs and their components are integrated together as shown in Fig. 1. Each path between the constructs and the components represents the hypothetical relationship to be verified using structural equation model.

Research Methodology

Primary data were collected for the purpose of the research needs and secondary data are also used.

Research Instrument

The survey method of data was implemented to gather the primary data. The structured questionnaire was adopted to accumulate the primary data from the sample of 500 employees working in private sector banks located at Chennai city.

Sampling Details

The sample for this study includes 500 managers working with private sector banks from the city of Chennai. Convenience method of sampling was employed for selecting the respondents for the study. The population of the study included assistant managers, managers, and senior managers working in private sector bank branches in Chennai.

Statistical Tools

t-Test for independent samples and structural equation model were used. The independent samples' *t*-test was used to analyse if there is any significant difference between the gender of the employees towards the perception of engagement, organizational commitment, and talent retention; whereas, structural equation modeling (SEM) was used to explore the relationship between the components of engagement, organizational commitment, and talent retention. Three types of analyses for scale development were used namely, exploratory factor

analysis (EFA), confirmatory factor analysis (CFA), and structural equation modeling (SEM). Firstly, exploratory factor analysis was used to identify the number of factors. Secondly, confirmatory factor analysis was used to validate the emergent factor structure through EFA. CFA deals with measurement models on the relationship between latent variable and observed measure. And lastly, structural equation modeling technique was used, a statistical methodology that requires a confirmatory approach (hypothesis testing) to analyse a structural theory. In EFA, KMO, and Bartlett's test, communalities, total variance explained, rotation component matrix was used for talent retention strategies, organizational commitment, and employee retention construct. In CFA, measurement model was used as a confirmatory tool for testing the measurement theory. CFA resulted in three-factor model for employee engagement, four-factor model for organizational commitment, and three-factor model for talent retention were used. Second-order structural equation modeling was used to test the hypothesis. In SEM, three constructs and 10 factors were used to run the model.

Findings

The results provide valuable insights into the engagement, organizational commitment, and employee retention. It highlights the methods managers use to handle talent issues.

Table 1: Demographic Details

Personal and Occupational profile Variables	*Respondents Details*	*Number of Respondents*	*Percentage of Respondents*
Age	21-30	190	38
	31-40	130	26
	41-50	95	19
	Above 50	85	17
Total		500	100
Gender	Male	280	56
	Female	220	44
Total		500	100
Educational Background	Post graduates	175	35
	Graduates	200	40
	Professional qualified	125	25
Total		500	100
Designation	Senior Managers	150	30
	Managers	140	28
	Assistant Managers	210	42
Total		500	100
No. of years of experience	0-5 years	100	20
	6-10 years	110	22
	11-15 years	125	25
	16-20 years	75	15
	Above 20 years	90	18
Total		500	100

- Age of the employees: 38% of the respondents were in the age group of 21-30 followed by 26% of respondents in the age group of 31-40, followed by 19% in the age group of 41-50 and 17% above 50 years of age.
- Educational background: 35% of the respondents were postgraduates, 40% of the respondents were professionally qualified, and 25% of the respondents were graduates.

- Gender: majority of respondents were male accounting for about 56% and female respondents accounted for 44%.
- Designation: 42% of the respondents were assistant managers, 28% of the respondents were managers, and 30% of the respondents were senior managers.
- Number of years of experience of the respondents: 18% of the respondents have above 20 years of experience, 15% of them have 16-20 years of experience, 25% of them have 11-15 years of experience, 22% of them have 6-10 years of experience, and 20% of them have less than five years of experience.

Factor Analysis

Factor analysis is a tool of multivariate analysis that is based on the interrelationship between a set of variables. By applying factor analysis, numerous variables are analysed such that it can be explained in a single factor. Degroot, et al., (1982) states that factor analysis is used to reduce a number of variables into overall groups.

Employee Engagement, Organizational Commitment, and Talent Retention Factors

Exploratory principal components analysis using a Varimax rotation was used to summarize the items into an underlying set of employee engagement, organizational commitment, and talent retention factors. All the factor loadings of 0.5 or above were identified in the factor matrix; (EFA) has been used to identify the various factors. Principal Component Analysis method is used and the following results are obtained:

Table 2: Kaiser-Meyer-Olkin and Bartlett's Test

		Employee engagement	*Organisational commitment*	*Talent retention*
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.844	.901	.751
Bartlett's Test of Sphericity	Approx.Chi-Square	2351.390	2979.966	1322.75
	Df	.105	105	45
	Significance	.000	.000	.000

Source: Computed Data

(Kaiser H F.1970) Meyer-Olkin and Bartlett's Test of Sphericity provides information about the factorability of the data. As a measure of sampling adequacy, KMO is a test of the amount of variance within the data that could be explained by factors. The Kaiser-Meyer-Olkin Measure of sampling adequacy values are 0.844, 0.901, 0.751 as in Table 2 and Bartlett's Test of Sphericity with approximate Chi-Square values are 2351.390, 2979.966 and 1322.75, respectively. These values are statistically significant at 5% level. Therefore, it can be concluded that the sample size of the research is adequate for the factors and all the variables considered for the research.

Employee Engagement Factors

It is found that 15 variables pertaining to employee engagement are reduced into three predominant factors with total variance of 51.262. These factors also possess individual variances, 22.230%, 15.138% and 13.893%. The Eigen values above 1 are noticed for the three factors. The variable loadings for each factor are measured using Rotated Component Matrix. The Rotated Component Matrix shows the 15 engagement practices used in banks operating in Chennai is extracted into three principal factors. These factors are explained below with the respective variables.

Table 3: Factor Analysis of Employee Engagement: Variables and Factor Loadings

F.no	*Variable*	*Factor loading*	*Name given to the factor*
F1	Open to New Ideas	.718	Team drivers
	Makes employee development a priority	.712	
	Cares about employees	.645	
	Mutual trust amidst teams	.641	
	Resolves conflict immediately	.617	
	Members reach out and help each other	.557	
	Provides job specific training	.525	

F.no	Variable	Factor loading	Name given to the factor
F2	Linking projects to personal development	.682	Opportunity drivers
	Linking projects to career development	.660	
	Prospect to work in new functional areas	.630	
	Opportunity to work in new business units	.562	
F3	Provides future orientation	.768	Organizational drivers
	Creates a culture of innovation	.566	
	Motivates risk taking	.533	
	Harnesses integrity and diversity	.501	

Source: Computed data

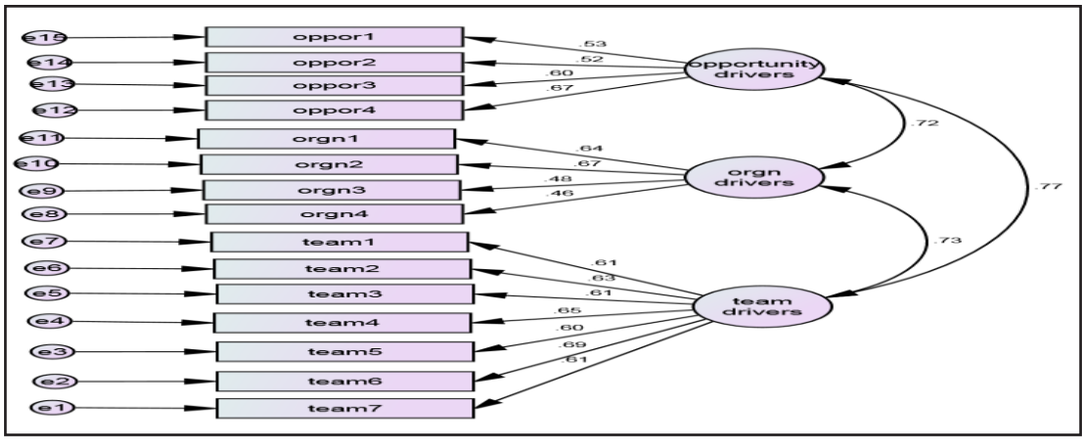


Fig. 2: Employee Engagement CFA

Confirmatory Factor Analysis - The Driver of Employee Engagement

To test the validity of the scales, AMOS was used. The data were selected for assumptions of CFA. For the employee engagement scale, CFA results revealed that the three-factor model. Single-headed arrows represent linear dependents. Double-headed arrows reveal that opportunity has significant effect on organization, organization on team drivers. Thus, it can be concluded that employee engagement had three drivers namely, team drivers, opportunity drivers, and organizational drivers. The CFA provided a satisfactory fit to the data as indicated in Table 4. All estimated loadings like, GFI, AGFI, CFI, NFI, RMA, and RMSEA were significant.

Table 4: Employee Engagement - Model Fit

Measure	Threshold
Chi-square/df (CMIN/DF)	2.87
P-value for the model	.000
Goodness-of-Fit Statistic (GFI)	.908
Adjusted Goodness-of-Fit Statistic (AGFI)	.953
Comparative Fit Index (CFI)	.901
Normed-Fit Index (NFI)	.961
Tucker-Lewis index (TLI)	.942
Incremental Fit Index (IFI)	.971
Root Mean Square Residual (RMR)	.899
Root Mean Square Error of Approximation (RMSEA	.907

Organizational Commitment Factors

It is found that 15 variables pertaining to organizational commitment are reduced into four predominant factors with total variance of 63.264%. These factors also possess individual variances, 20.865%, 15.648%, 13.415%, and 13.337%. The Eigen values above 1 are noticed for the four factors. The variable loadings for each factor are measured using Rotated Component Matrix.

Table 5: Factor Analysis of Organizational Commitment - Variables and Factor Loadings

F.no	*Variable*	*Factor loading*	*Name given to the factor*
F1	I am extremely glad that I chose this organization There is much to be gained by sticking with this organization I find it easy to agree with this organization's policies I really care for this organization. For me this is the best of all organizations to work. My values and the organization's values are very similar	.675 .669 .644 .625 .605 .587	Trustworthy commitment
F2	I would accept any type of job to work for this organization This organization really inspires the very best in me I will not leave this organization	.830 .703 .660	Passionate commitment
F3	I project this organization positively to others I am loyal to this organization.	.776 .715	Contented Commitment
F4	Working for this organization is a definite boon I am proud to say that I belong to this organization	.781 .656	Affirmative commitment

Source: Computed data

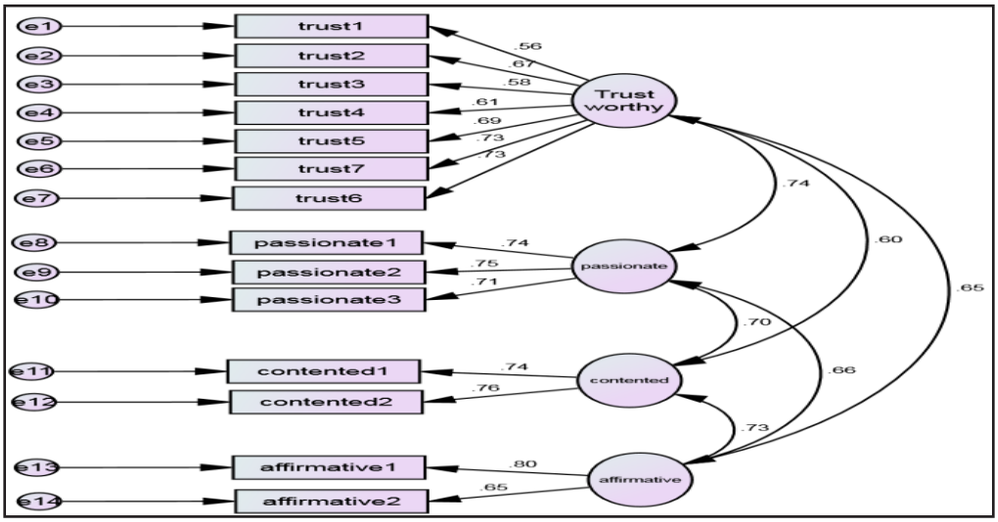


Fig. 3: Organizational Commitment CFA

Confirmatory Factor Analysis - Organization Commitment Factors

The data were selected for assumptions of CFA. For the organizational commitment scale, CFA results revealed a four-factor model. Single-headed arrows represent linear dependents. Double-headed arrows reveal that trust has significant effect on passionate, passionate on contented, and contented on affirmativeness. The CFA provided a satisfactory fit to the data as indicated in Table 6.

Table 6: Organization Commitment - Model Fit

Measure	*Threshold*
Chi-square/df (CMIN/DF)	1.993
P-value for the model	.000
Goodness-of-Fit Statistic (GFI)	.956
Adjusted Goodness-of-Fit Statistic (AGFI)	.934
Comparative Fit Index (CFI)	.948
Normed-Fit Index (NFI)	.986
Tucker-Lewis index (TLI)	.978
Incremental Fit Index (IFI)	.948
Root Mean Square Residual (RMR)	.58
Root Mean Square Error of Approximation (RMSEA	.65

Employee Retention

It is found that 10 variables pertaining to employee retention are reduced into three predominant factors with total variance of 61.977. These factors also possess individual variances, 22.209%, 21.309%, and 18.459%. The Eigen values above 1 are noticed for the three factors. The variable loadings for each factor are measured using RCM. The RCM shows the 10 employee retention in banks operating in Chennai is extracted into three principal factors, namely intent to stay, career path, and work-life balance. These factors are explained below with the respective variables.

Table 7: Factor Analysis of Employee Retention - Variables and Factor Loadings

F.no	*Variable*	*Factor loading*	*Name given to the factor*
F1	I am happy with the work environment	.809	Intent to stay
	Links performance to pay	.782	
	Compensation in par with similar industries	.744	
F2	I have career progression at workplace	.760	Provide career path
	Opportunity to revive a struggling business	.760	
	Opportunity to work in a different country	.758	
	Opportunity to work in new projects	.742	
F3	I am able to devote adequate time for personal life	.804	Work life balance
	I am satisfied with work life balance.	.738	

Source: Computed data

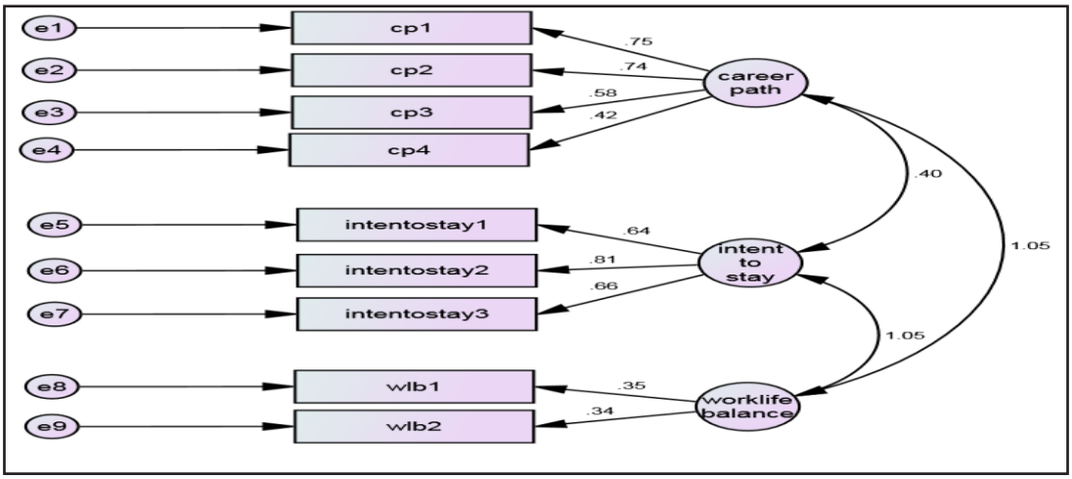


Fig. 4: Talent Retention CFA

Confirmatory Factor Analysis - Talent Retention Factors

To test the validity of the scales AMOS was used. The data were selected for assumptions of CFA. The results revealed a three-factor model namely providing a career path, intent to stay, and work-life balance. Single-headed arrows represent linear dependents. The double-headed arrows connected in the path diagram reveal that career path has a significant effect on intent to stay and

intent to stay on work-life balance. The CFA provided a satisfactory fit to the data as indicated in Table 8. All estimated loadings like, GFI, AGFI, CFI, NFI, RMA, and RMSEA were significant

Table 8: Talent Retention - Model Fit

Measure	*Threshold*
Chi-square/df (CMIN/DF)	1.802
P-value for the model	.000
Goodness-of-Fit Statistic (GFI)	.962
Adjusted Goodness-of-Fit Statistic (AGFI)	.950
Comparative Fit Index (CFI)	.987
Normed-Fit Index (NFI)	.925
Tucker-Lewis index (TLI)	.947
Incremental Fit Index (IFI)	.987
Root Mean Square Residual (RMR)	.89
Root Mean Square Error of Approximation (RMSEA	.78

Hypotheses

H_1 Organizational commitment and employee engagement are positively related to talent retention.

Structural Equation Modeling for Ascertaining the Impact of Organizational Commitment and Employee Engagement on Talent Retention

The effect of organizational commitment and engagement on talent retention among bank managers working in Chennai was tested using structural equation modeling approach. Structural equation modeling is an adept method of assessing the measurement error where it can be incorporated commonly in observed and latent variables. Therefore, the association among measured variables team work, organizational drivers, and opportunity drivers trustworthy passionate, contented, affirmative, and the latent variable namely talent retention were assimilated in structural equation modeling. Fig. 5 illustrates the SEM model based on the standardized regression coefficients.

The current research hypotheses have been delineated on the source of the model fit summary which sketched underneath and by means of research conducted on the effect of organizational commitment & employee engagement on talent retention, the subsequent hypotheses is projected:

H_1 Organizational commitment and employee engagement are positively related to talent retention.

Thus, it can be inferred from the above the coefficients of organizational commitment are 0.03 and that of employee engagement are 0.87 which signifies that with every increase in organizational commitment and employee engagement, the talent can be retained in banks.

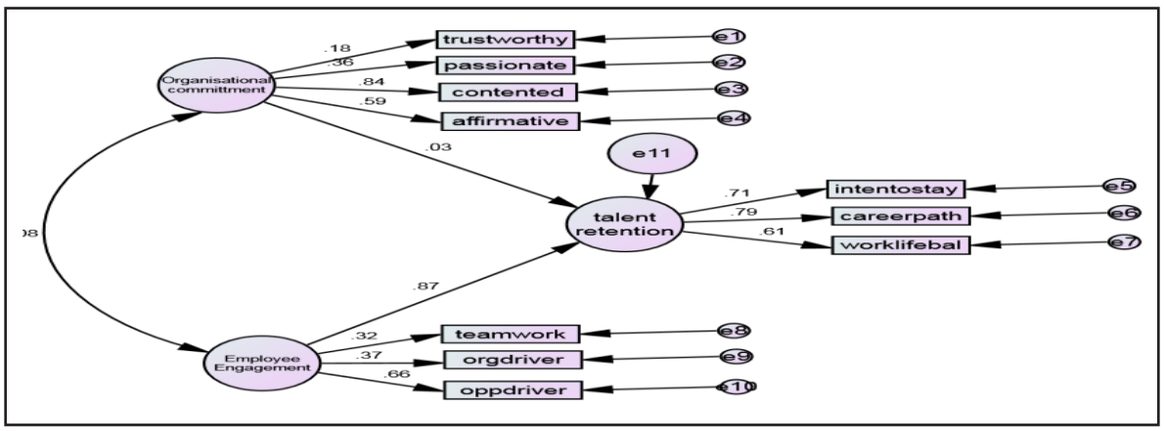


Fig. 5: Conceptual Model

Table 9: Conceptual Model Fit

Measure	*Threshold*
Chi-square/df (CMIN/DF)	1.91
P-value for the model	.000
Goodness-of-Fit Statistic (GFI)	.906
Adjusted Goodness-of-Fit Statistic (AGFI)	.930
Comparative Fit Index (CFI)	.901
Normed-Fit Index (NFI)	.851
Tucker-Lewis index (TLI)	.815
Incremental Fit Index (IFI)	.901
Root Mean Square Residual (RMR)	.58
Root Mean Square Error of Approximation (RMSEA	.51

Table 9 shows the model fit summary of the research model. It is understood that the significance value of *p* is 1.91 which is superior to 0.05 which is a perfect fit. The goodness fit indeed and adjusted goodness fit index values are more than 0.90, which indicates it is an acceptable model fit. The value of comparative fit index is 0.90, which also represents a worthy fit to the model and the values of RMR and RMSEA are 0.58 and 0.51, respectively, which specific that it is also an acceptable model. Thus, the hypotheses, organizational commitment, and employee engagement is positively related to talent retention.

Table 10: Structural Equation Model for Testing the Framework

Constructs and measures	*Standardized*	*Unstandardized*	*P- sig value*
Employee engagement - Employee retention	.87	2.36	<0.001
Organizational commitment - Employee retention	.03	.12	<0.001
Organisational Commitment			
Organizational commitment - Trust	.18	1.00	
Organizational commitment - Passionate	.36	2.00	<0.001
Organizational commitment - contented	.84	4.11	<0.001
Organizational commitment - affirmative	.59	3.46	<0.001
Employee engagement			
Employee engagement - team drivers	.32	1.00	
Employee engagement - organizational driver	.37	1.10	<0.001
Employee engagement - opportunity drivers	.66	2.60	<0.001
Talent retention			
Talent retention - intent to stay	.71	1.00	
Talent retention - Provide career path	.79	1.12	<0.001
Talent retention -work life balance	.61	.91	<0.001

Table 10 summarizes the effect of organizational commitment and engagement on talent retention with standardized and unstandardized estimates. It is observed that the unstandardized regression coefficient of employee engagement is 0.87 and organizational commitment is 0.03 which signifies the partial effect over talent retention by considering that the other variables are not having an influence over talent retention. The estimate denotes that talent retention will increase by 0.87 and 0.03 for every unit rise in employee engagement strategies and organizational commitment at the given level of significance.

The unstandardized coefficient value of contented is 4.11 which represents the effect of contented employees on organizational commitment. Similarly, people with a positive attitude towards the organization are said to exhibit more commitment towards their workplace.

The unstandardized coefficient value of opportunity driver is 2.60, which signifies the effect of opportunity drivers on

employee engagement. Likewise, organizational drivers also have a positive effect on employee engagement.

Hypotheses Testing

The independent samples' *t*-test was used to analyse if there is any significant difference between gender of the employees towards the perception of talent retention, organizational commitment, and employee engagement.

Hypotheses

There is no significant difference among male and female employees with respect to their perception towards talent retention, organizational commitment, and employee engagement.

Table 11 presents the results of the independent samples' *t*-test, which reveals that the significance values of team work, opportunity drivers, organizational drivers, career path, intent to stay, work-life balance, passionate, affirmative, contented, and trustworthy are less than 0.05%. Therefore, it is recognized that there is a significant difference between male and female employees with respect to their perception towards employee engagement, organizational commitment, and talent retention. It also exposes that female employees are more passionate, trustworthy, and have a greater intention to stay in comparison to their male counterparts.

Table 11: *t*-test

S.no	*Name of the factor*	*Gender*	*Mean*	*t-value*	*p-value*
1	Team drivers	Male	9.81	2.935	0.003*
		Female	8.49		
2	Organizational drivers	Male	9.13	2.353	0.019*
		Female	8.15		
3	Opportunity drivers	Male	9.83	3.402	0.001*
		Female	8.33		
4	Trustworthy	Male	7.67	2.844	0.010*
		Female	8.62		
5	Passionate	Male	8.76	2.457	0.016*
		Female	9.16		
6	Affirmative	Male	7.48	3.173	0.002*
		Female	6.56		
7	Contented	Male	8.01	3.201	0.001*
		Female	7.99		
8	Intent to stay	Male	8.47	2.844	0.009*
		Female	9.41		
9	Providing a career path	Male	9.22	2.444	0.015*
		Female	8.19		
10	Work life balance	Male	7.58	3.168	0.002*
		Female	6.82		

Managerial Implications

Banks need to create an environment that talented people will want to stay in the organization. There are a number of "stay" factors, which appeal to an individual's wants and needs that provoke them to stay in an organization. Sound values, compelling vision, create exciting jobs that will stimulate, challenge, and stretch capable people. Moreover, banks need to ensure that the recruited talent is effectively coached, mentored, given feedback, and appropriately rewarded so that they are retained in the organization. Banks need to identify the gaps, and train and guide employees in order to make effective use of talent. Managers agreed that engaging employees, creating development roles, and providing a career path will enable the talent to grow and blossom. Brundage H, Koziel M (2010) highlighted the importance of talent

retention by saying that effective talent retention is a continuous process and it must be part of organizational culture. Thus, providing a career path will enable retaining the right talent.

Conclusion

Successful organizations have one factor in common-getting the right talent nurtured, retained, and valued. Thus, it becomes evident for banks today to get to know that people are increasingly the prime assets of any organization. Private sector banks must not only focus in terms of having an excellent talent pool, but also in their ability to retain their best talent. Effectively engaging is the beginning of effective retention.

Matching tasks and talents is a great challenge. Talent management ensures that the existing employees in the organization are properly utilized (Cheloha, R., & Swain, J., 2005). Ensuring right person is assigned for right job at the right time is equally important (Kesler, G. C., 2002). Identifying the strategic positions and the turnover risks associated with these positions play a vital role on talent retention. Deliberated and integrated set of initiatives allow the employees to align themselves with the organizational goals and objectives. Banks have to strategize how best to utilize the people's talents by identifying areas for internal development that are necessary for ensuring future success. Thus, banks have to move away from simple administration to some objective-oriented approach for retaining key employees (Hussein M 2009). Thus, today organizations need to steer their energies to get a hold on to their talent. This can be done by devising excellent retention strategies such as:

- Providing career opportunities & assignments with diverse set of job responsibilities and
- Redesigning development programmes using latest technology and providing instant feedback.
- This will pave the way for engaging them experientially and providing them with a valued pay off.

References

Athey, R. (2004). Why Acquisition and Retention Strategies Don't Work. Retrieved December 14, 2015, from http://www.barrettrose.com/acquisition-retention-strategies-work

Brousseau, K. R., Driver, M. J., Hourihan, G., & Larsson, R. (2006). The seasoned executive's decision making style. *Harvard Business Review, 84*(*2*), 110-21,

Brundage, H., & koziel, M (2010). Retaining top talent. *J. Account., 209*(5).

Cheloha, R., & Swain, J. (2005). Talent management system key to effective Succession planning. *Canadian HR Reporter, 18*(17), 5-7.

Cohen, A. (2003). *Multiple commitments in the workplace: An integrative approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Liu, Y., & Cohen, A. (2010). Values, commitment, and OCB among Chinese employees. *International Journal of Intercultural Relations, 34,* 493-506.

Degroot M. H., Ferber, R., Frankel, M. R., Seneta, E., Watson, G. S., Kotz, S., et al. (Eds.). (1982). *Encyclopedia of Statistical Sciences. Faa' Di Bruno's Formula to Hypothesis Testing.* New York: John Wiley and Sons.

Eichinger, R. W., & Lombardo, M. M. (2004). Learning agility as a prime indicator of potential. *Human Resource Planning, 27*(4), 12-16.

Grossman, R. J. (2007). IBM's HR takes a risk. *HR Magazine, 52*(4), 54-59.

Hand field J. H., & Axelrod, B. (1997). *The war for talent survey.* McKinseyQuarterly.

Hussein, M. (2009). Hiring and Firing with ethics. *Hum. Resource. Manage. Digest, 17*(4), 37-40.

Hymowitz, C. (2008). In the lead: They ponder layoffs, but executives still face gaps in talent. *The Wall Street Journal*.

Kaiser, H. F. (1970). A Second generation little jiffy. *Psychometrika, 35,* 401-415.

Kaul, V. M. (2011). *NHRD Network Journal, 4*(3), 8-15.

Kesler, G. C. (2002). Why the leadership bench never gets deeper: Ten insights about executive talent development. *Human Resource Planning, 25*(1), 32-44.

Noorliza, K., & Hasni, M. (2006). The effects of total quality management practices on employees work related attitudes. *The TQM Magazine, 18*(1), 30-33.

Ready, D. A., & Conger, J. A. (2007). Make your company a talent factory. *Harvard Business Review, 85*(6), 68-77.

The Impact of Employee Engagement on Performance - Harvard Business School Publishing - review report 2013.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 25-30
http://publishingindia.com/ijbai/

QFD-Based Smart Cane Design: A Technology to Assist Visually Impaired

S. M. Akash*, Mustafizur Rahman*

Abstract

This paper reports on a study that helps visually impaired people to steer uncountable confidently. The study hypothesizes that a Wise Cane that alerts visually impaired people about obstacles before might facilitate them in walking with less accident. The white cane, because of its primitive vogue, is unable to produce the blind and visually impaired level of independence that is possible with stylish technology. The aim of the paper is to handle the event work of a cane, which is able to communicate with the users through sound alert, and vibration, that is mentioned as 'Wise/Smart Cane'. The event work involves writing and physical installation. A series of test unit is distributed on the wise cane and the results are mentioned. This study found that the Wise Cane functions well as meaning, in alerting users about the obstacles before.

Keywords: Arduino, Assistive Technology, Ergonomic design, House of Quality, RFID, Ultrasonic Sensor HC-SR04, Visually Impaired

Introduction

The survey of United Nations Agency (World Health Organization) allotted in 2011 tells that in the world regarding one hundred and twenty-fifth of the human population is visually impaired and amongst them regarding 100 percent is blind.

The main drawback with blind folks is quality. This paper proposes a tool for visually impaired people that can offer them navigation. Long white Cane may be an ancient quality tool that will not observe obstacles within the path of a blind man. By modifying this cane with some physical parts and sensors, this can become good cane. Good white cane is specially designed to observe obstacles, which can facilitate the blind to navigate care freely. The vibration feedback and therefore the sound alert can keep the user alert and significantly cut back accidents.

Consumer and medical technology have created vital advancements over the past 60 years. However, the practicality of canes for the visually impaired remains restricted, counting on the user's ability to physically observe objects and forcing the user to be entirely accountable for their safety. This burden may be alleviated with the more security of associate object detector. Additionally, the quality of the white cane has no vary of physical choices. It places an extra burden on the user by forcing an amendment in handle grip betting on however jam-packed the environment is. The white cane is needed by the user to adapt to the cane instead of having a cane, which will adapt to the user.

To address these shortcomings, the good Cane project examines, however, Canes may be technologically equipped to boost their practicality in an exceeding method. This Cane will be conjointly economically accessible. The goal for the Smart Cane project is to eliminate this drawback by planning, building and testing a cane for the blind that utilizes laptop and sensory technology to produce object detection capabilities and freedom of physical variables. Once the project is completed, the cane style is quantitatively and qualitatively examined to see its success as a product.

Bibliography

This section describes applicable connected works on the event of Smart Canes meant for visually impaired people.

The white cane originated in Europe in 1921. Once James Biggs, a creative person of UN agency had lost his vision, began to color his walking Cane white to alert others to his presence. Once veterans of warfare II came to America with vision impairment and sightlessness and they wished to own an equivalent level of independence as that they had before the war. Owing to this, the white walking Cane

* Department of Industrial and Production Engineering, Jessore University of Science and Technology, Jessore, Bangladesh.
Email: akashsm322@gmail.com, mostafizkuet09@gmail.com

was altered into the long cane type that is still current these days. At present, eighty-two of the world's blind population is at the age of fifty and higher than. More or less ninetieth of the world has visually impaired board-developing nations because of the dearth of care and medical treatments. These figures square measure vital once considering the population that the good Cane is addressing.

According to A. Nurulnadwan, M.R. Nur-Hazwani, and A.M. Ariffin (A. Nurulnadwan, M.R. Nur-Hazwani, and A.M. Ariffin, 2009), technology will facilitate in reducing several barriers that individuals with disabilities face. These sorts of technologies square measure stated as helpful technology (AT).

There square measure many sorts of disabilities, as well as physical disabilities, deaf and visually impaired. AT has been used in helping them. However, developing Associate in Nursing AT is pricey (Herman N. J., 1999) and creating their price high.

According to Mazo and Rodriguez the blind Cane is one amongst the helping tools for the visually-impaired and it's extremely vital. Per (Herman N. J., 1999), one amongst the most issues of the visually impaired, is that almost all of those folks have lost their physical integrity. In addition, they are doing not place confidence in themselves. This statement has been tried by Bouvrie (Bouvrie J. V. 2007), within which Associate in Nursing experiment name "Project Prakash" has been applied. It absolutely was meant at testing the visually impaired to utilize their brain to spot the set of objects. Per Yangtze and Song (Chang C. C. and Song K. T., 2000), this may even be applied to completely different scenario. Once the visually impaired walk into replacement surroundings, they are going to notice it troublesome to con the locations of the thing or obstacles. These examples demonstrate the difficulties of visually impaired folks.

The Guide Cane is intended to assist the visually impaired users navigate safely and quickly among obstacles and different hazards (Borenstein J. and Ulrich I., 2001). Guide Cane is employed just like the wide used white cane, wherever the user holds the Guide Cane ahead of the user whereas walking. The Guide Cane is significantly heavier than the white cane, because of it uses a servomotor. The wheels square measure equipped with encoders to see the relative motion. The servomotor, controlled by the intrinsic pc, will steer the wheels left and right relative to the cane. To find obstacles, the Guide Cane is supplied with 10 unhearable sensors. A mini joystick situated at the handle permits the user to specify a desired direction of motion. Guide Cane is much heavier than the normal white Cane and additionally it is arduous to stay as a result of it cannot be doubled.

Smart Cane is one invention that was originally the creation of a standard blind cane however; it is equipped with a device system. This invention resembles Guide Cane wherever this invention contains a range of unhearable sensors and servomotors. This invention is intended with the aim at serving to the blind in navigating. Inaudible sensors have to compel to sight and avoid obstacles or objects set ahead of the user. Meantime the fuzzy controller is needed to work out the directions, which will be dead as an example to show right, left or stop. Like Guide Cane, this invention additionally encompasses a management button on the handle, and the button has four completely different directions. This invention has equivalent weaknesses because the Guide Cane wherever there will be a retardant to avoid wasting house or to position the sensible cane. Besides that, value is additionally a weakness during this project because it uses inaudible sensors and variety of servo motors. If the price is just too high, users do not seem to be ready to afford for it because of the typical financial gain of the visually impaired folks is comparatively little.

Smart Cane has been designed by students from Central Michigan University wherever this invention uses oftenest Identification (RFID). RFID is used to sight objects or obstacles ahead of the user and detects the RFID tag that has been placed in many areas to navigate the users. This invention is simply sort of a traditional stick however is provided with a bag, worn by the user. The bag provides electricity power to the invention and informs the user through speakers within the bag. For users of United Nations agency does not have the flexibility to listen to, there square measure special gloves, which will vibrate at each finger, during which vibrations in every finger have different meanings. However, this invention has many weaknesses and is just appropriate for tiny areas. This is often as a result of it solely detects the realm with RFID tag otherwise this invention solely works as an everyday blind cane. Additionally, this invention needs a high value if it's utilized in the external atmosphere as a result of the larger space that require to be labeled, the upper value is required.

Mechatronic Blind Stick may be a guiding system, designed to facilitate the daily work among the visually impaired folks. This invention has several similarities with the sensible Blind Cane. During which this invention

uses inaudible sensors and sound vibrations. However, this invention additionally has many weaknesses; it cannot be closed and troublesome to stay. Additionally, this invention is not equipped with sensors to sight the water areas.

METHODOLOGY

Materials and Methods

There were two main elements that were the main target of the good Cane style process: one is changing the fundamental mechanics of the standard white cane and the other one is group action technology as to create it "smart". It had been all over that each would be addressed, initially by creating the cane adjustable by adding a device that may extend the very of the user may observe. The detection of the potential obstacles would then be transmitted to the user through vibrations within the handle.

The Arduino

The sensible Cane's sensors and motors are powered by an Arduino microcontroller. The Arduino could be a programmable electronic platform that permits users to produce prototypes. Together with a board and alternative items of instrumentality equipment, the Arduino is often wont to create varied electronic input, output, and sensory systems. Other than basic electronic hardware, a large variety of advanced devices, as well as sensors, is created to be compatible with the Arduino system. The Arduino artificial language is C based mostly and maybe wont to produce a large type of programs. The Arduino is additionally created a lot of accessible by its low price. Most boards (including the Uno, which the sensible Cane uses) price is $30.

Pulse Width Modulation

The Arduino permits for input and output by plugging wires into 'pins'. Input pins scan information (such as info from a sensor) and square measure capable of taking in a very continuous varies of values. Thus, through sensors, the Arduino may be incessantly updated with info regarding the surroundings around it. The output pins send a current to any device connected to them, like a motor or a lightweight bulb. Not like the input pins, there square measure solely two prospects for the voltage: 5V or 0V. Bit by bit dynamical the speed of the motor needs never-ending amendment in voltage that is not potential with the output pins. However, the Arduino will leave (and has special output pins dedicated to) Pulse Width Modulation (abbreviated as PWM). Rather than go over several voltages, the voltage apace changes from 0V to 5V. PWM simulates a gradual amendment from one voltage to a different, providing something connected to the pin to vary at a time. As an example, if 5V square measure being outputted one-fifth of the time, this can be called a two hundredth duty cycle, and the simulated voltage is one-fifth of 5V, i.e. 1V.

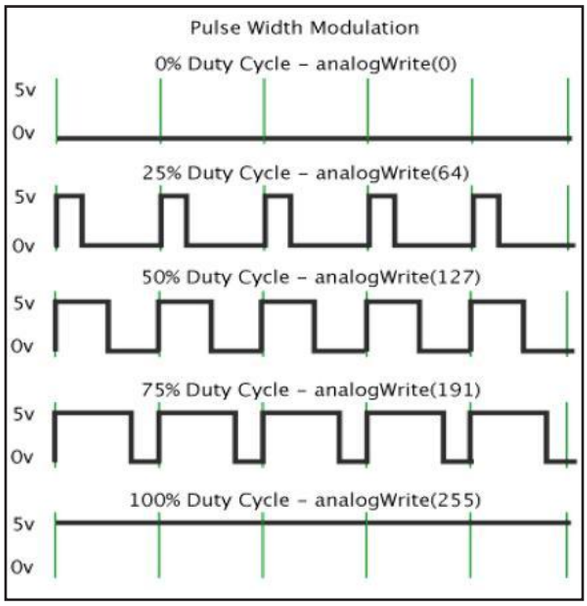


Fig. 1: This Diagram Shows how the Number of Pulses Sent Changes Depending on the Analog Write Value

A operate designed into the Arduino, analog Write (), permits a program to create use of the PWM operate just by plugging in a very price starting from zero to 255, with the latter being the utmost double voltage (a continuous output of 5V).

The Ultrasonic Sensor

The detector employed in the sensible Cane is that the HC-SR04. It emits associate ultrasound at 40,000 cycles that travels through the air associated. If there is an object or obstacle on its path, it will recover to the module. Considering the period and therefore the speed of the sound one will calculate the gap. The HC-SR04 inaudible Module has four pins- Ground, VCC, Trig, and Echo. The bottom and therefore the VCC pins of the module must be

connected to the bottom and therefore the five volts pins on the Arduino board severally and therefore the trig and echo pins to any Digital I/O pin on the Arduino board. So as to come up with the ultrasound one must set the Trig on a High State for 10 µs. which will send associate eight cycles sonic burst which is able to travel at the speed sound and it'll be received within the Echo pin. The Echo pin can output the time in microseconds the acoustic wave traveled. For instance, if the article is 10 cm aloof from the detector, and therefore the speed of the sound is 340 m/s or 0.034 cm/µs the acoustic wave can get to travel concerning 294 ? However, what we may get from the Echo pin are double that variety because of the acoustic wave must travel forward and bounce backward. thus so as to induce the gap in cm we'd like to multiply the received period of time price from the echo pin by 0.034 and divide it by 2.

Creating the Ergonomic Design

Besides the sensory system itself, the most innovative and important aspect of the Smart Cane is the ergonomic design. The design began with preliminary measurements and sketches. Then, the cane was designed and modeled in 3-D using Autodesk AutoCAD software.

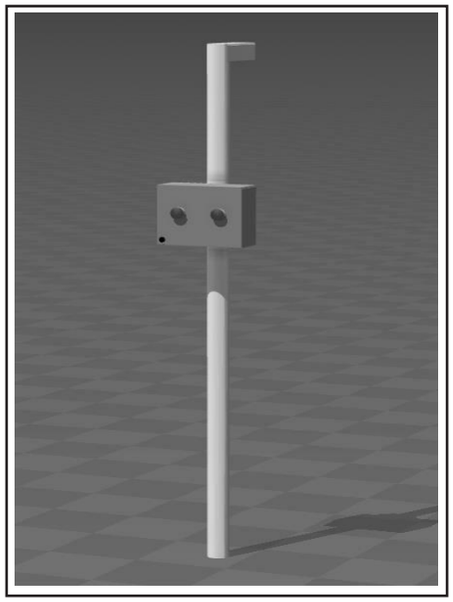


Fig. 2: Design of the Smart Cane

This allowed for the Arduino, ultrasonic sensor, breadboard, vibration motor, battery and the on-off switch to be placed inside a box attached to the cane.

Result Analysis

Few details had been obtained when analyzing the customer review as Quality Function Deployment (QFD) listed below:

- Increase of the working range of the cane
- Durability
- Recharge ability of the power source
- Water proof covering for the electronic parts
- Light weight
- Low price

The research work was done focusing on those criteria obtained from the customer reviews and the smart cane is able to detect the obstacle in path such as chair, pole, human, wall, and door up to 2-meter efficiently.

Discussion

Key Findings of the research work is listed below:

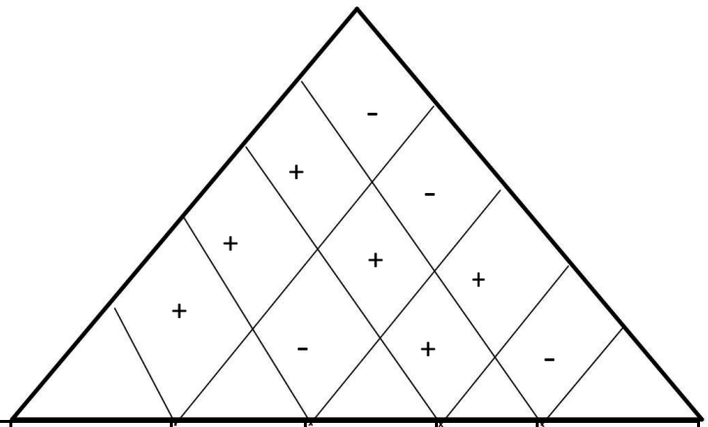
- The device increases the confidence of the users due to which travel time for majority of the users is reduced.
- During the study, users suggested various modifications such as reduced weight, reliability in detection of obstacles and improved path finding abilities.
- Training is very critical for the acceptance and use of this device.

Every customer had his or her own unique reviews about the Smart Cane. However, some criteria were common in every case and by following the process of QFD, the final specifications for the smart cane were found and focusing on those specifications, the cane has produced. Moreover, the findings were discovered from the reviews of the final product known as the Smart Cane.

Conclusion

The main purpose of this study is to supply an image, which will notice objects or obstacles in front of users and feeds warning back, within the styles of sound and vibration, to users. From the tests dole out on its functions reveal that the developed image that is known as Sensible Cane has achieved its objectives.

Table 1: House of Quality of Smart Cane



Customer Requirements	Customer Importance	Rechargeable battery	Ultrasonic Sensor	Water Proof Covering	Steel Parts	Aluminium Parts
Increase the working range of the cane	40	7	10	10	7	10
Reliable	20	7	7	3	10	7
Rechargeable power source	20	10	7	7	3	3
Light weight	10	3	7	3	7	10
Low price	10	7	7	7	7	10
Target values		720	820	700	680	800
Ranking		3	1	4	5	2

Relation, Symbol & Rating (0 – 10):
Strong Relationship = 10,
Moderate Relationship = 7,
Weak Relationship = 3,
Positive Relation= +,
Negative Relation= –

This study would suggest that an influence provide meter reading will be put in to watch its power standing. The inaudible sensor's operating vary may be variable by the user. GSM hunter may be wont to acknowledge the user regarding his/her position. This Smart and Sensible Cane technology will be used widely among visually impaired personals and create them a little more well-off in their life and also the Sensible Cane in-built this project is additionally low cost and approachable for pretty much each user.

References

Aziz, N., Roseli, N. H., & Mutalib, A. A. (2009). Assistive Courseware for Visually-Impaired. *Lecture Notes in Computer Science Visual Informatics: Bridging Research and Practice,*905-915. doi:10.1007/978-3-642-05036-7_86

Bouvrie, J. V., & Sinha, P. (2007). Visual object concept discovery: Observations in congenitally blind children, and a computational approach. *Neurocomputing, 70*(13-15), 2218-2233. doi:10.1016/j.neucom.2006.01.035

Chang, C. C., & Song, K. (1996). Ultrasonic sensor data integration and its application to environment perception. *Journal of Robotic Systems, 13*(10), 663-677. doi:10.1002/(sici)1097-4563(199610)13:10<663::aid-rob3>3.0.co;2-t

Fathauher, G. H. (1999). Ultrasonic Sensor. *IEEE transaction on Ultrasonic, 3*, 57.

Hapsari, G. I., Mutiara, G. A., & Kusumah, D. T. (2017). Smart cane location guide for blind using GPS. *2017 5th International Conference on Information and Communication Technology (ICoIC7)*. doi:10.1109/icoict.2017.8074697

Herman-Kinney, N. J. (1995). *Deviance a symbolic interactionist approach* (Vol. 24, 236-241). Dix Hills, NY: General Hall.

Huang, W., McNamara, H., Molodan, D., Pasarkar, A., & Rizzo, R. (n.d.). Smart Cane. Retrieved December, 2017, from http://soe.rutgers.edu/sites/default/files/imce/pdfs/gset-2014/Smart Cane Final.pdf

Magori, V., & Walker, H. (1999). Ultrasonic Presense Sensor. *IEEE transaction on Ultrasonic, Ferroelctrics and Frequency control, 34* (2), 2-3.

Mutalib, A. A., & Maarof, F. (2010). Guidelines of Assistive Courseware (AC) for hearing impaired students. *Knowledge Management International Conference 2010 (KMICe2010), Organized by: UUM College of Art & Sciences, Universiti Utara Malaysia,*186-191.

Mutiara, G. A., Hapsari, G. I., & Rijalul, R. (2016). Smart guide extension for blind cane. *2016 4th International Conference on Information and Communication Technology (ICoICT)*. doi:10.1109/icoict.2016.7571896

Nurulnadwan, A. (2011). Assistive Courseware for the Visually Impaired based on Theory of Multiple Intelligence and SECI Model. *American Journal of Economics and Business Administration, 3*(1), 150-156. doi:10.3844/ajebasp.2011.150.156

Roseli, N. H., Aziz, N., & Mutalib, A. A. (2010). The enhancement of assistive courseware for visually impaired learners. *2010 International Symposium on Information Technology*. doi:10.1109/itsim.2010.5561368

Upase, S. U. (2016). Speech recognition based robotic system of wheelchair for disable people. *International Conference on Communication and Electronics Systems (ICCES)*. doi:10.1109/cesys.2016.7889851

Ulrich, I., & Borenstein, J. (2001). The GuideCane-applying mobile robot technologies to assist the visually impaired. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,31*(2), 131-136. doi:10.1109/3468.911370

International Journal of Business Analytics and Intelligence
6 (1) 2018, 31-34
http://publishingindia.com/ijbai/

Optimal Release Time of Software under Fuzzy Environment with Testing Effort

Asit Dwivedi*, Deepak Kumar**

Abstract

Software Release Time Problem (SRTD) focuses on the timely release of software keeping the constraints of Quality and Cost. Several SRTD models have been proposed in literature since 1970s. Software Reliability Growth Models (SRGM) give the mean number of failures during the process of debugging. In perfect debugging, fault is removed completely; whereas in imperfect debugging, faults are not removed perfectly during debugging process. A probability is always attached with the imperfect debugging. This paper focuses on the optimal release time of an imperfect debugging model with constants as cost and reliability in fuzzy environment with testing effort. Results are illustrated numerically, the problem of timely release of software in imperfect debugging is addressed, and optimal release time with given constraints is given.

Keywords: Fuzzy Membership Function, Fuzzy Optimization, Imperfect Debugging, Release Time, Software Release Time Problem (SRTD), Software Reliability Growth Model (SRGM)

Introduction

The main aim of software engineering is to develop software that are reliable in nature. As software plays a crucial role in today's world not only in business and human application as well as in daily lifestyle of an individual. Various works have contributed towards developing the software that is error free in nature; however, developing a reliable software is a major challenge faced by the software-developing industry, as developing such a software consumes lots of resources such as cost and power, with various processes to check its reliability. While developing reliable software, the project manager must be aware of the accurate information regarding how to grow software reliability to effectively maintain the project and the budget sanctioned to it (Khatri et al., 2012).

Various methods were described to produce the software reliabilty growth models using the techinques of soft computing like fuzzy logic and nueral networks (Bector et al., 2005). Various articles also studied the prediction of faults that get accumulated during the software-testing process using various parametric models as well as non-parametric models (Huang et al., 1997). In this paper authors have tried to provide solution for estimation of the defect fix effort using neural networks; for project management, fuzzy logic and neural network were utilized.

This paper presents a software reliability growth model which is used to solve the problem under fuzzy environment, which is given in literature. The efficiency of the given model is already predicted.

Software Reliability Growth Model

Assumptions

The model given in this paper follows the following assumptions:

1. Elimination of faults in the software and surveillance of malfunction follow-up NHPP.
2. During the running of software, the error present in software leads to software failure.
3. Instantaneous efforts takes place after the interpretation of failure to segregate the cause of malfunction; to separate it, various efforts are initiated.
4. Failure rate of software is straightforwardly related to the residual faults in the software.

* Research Scholar, Mewar University, Chittorgarh, Rajasthan, India. Email: asitdwivedi@rediffmail.com
** Mewar University, Chittorgarh, Rajasthan, India.

Notations

$m_r(t)$: expected faults removed in t time

a : opening no. of hidden errors of software

b : rate of error removal

w(t) : effort of testing at a given time t

W(t) : collective effort of testing in a given duration, i.e., $\frac{d}{dt}W(t) = w(t)$

Ts: duration of warranty.

S-shaped SRGM with Testing Effort

The following model is given by Kapur et al. (1991). The equation given below illustrates fault-removal phenomenon with testing effort:

$$\frac{dm_r(t)}{dt} \times \frac{1}{w(t)} = b(t)(a - m_r(t)) \quad \frac{dm_r(t)}{dt} \times \frac{1}{w(t)} = b(t)(a - m_r(t)) \tag{1}$$

where, $b(t) = \left(\frac{b^2 W(t)}{1 + bW(t)} \right)$

On solving equation (1) given above with primary conditions $m_r(t=0)=0$ and $W(t=0)= 0$, the following mean value function can be found:

$$m_r(t) = a\left[1 - \left((1 + bW(t))e^{-bW(t)}\right)\right] \tag{2}$$

Fuzzy Problem Formulation

Total minimum estimated cost of testing with intended constrictions on reliability is given by the optimal software release time.

Function of the Cost

The estimate cost reliant on cost of the testing per unit time (Co_1), cost sustain in the debugging before releasing the software (Co_2), and the cost sustain in the debugging after releasing the software (Co_3) (Kapur et al., 1999). The function of cost can be given by,

$$Co\,(W(T)) = Co_1 W(T) + Co_2 m(T) + Co_3(m(T+T_S) - m(T)) \tag{3}$$

Function of Reliability

The function of reliability for software is defined by,

$$R((T+Ts)|T) = e^{-(mf(T+Ts)-mf(T))} \tag{4}$$

Minimize Co (W(T))

Subject to $R\,(Ts|T) \gtrsim R_0$

$W\,(T) \lesssim W_0$

$T \geq 0$ (5)

The following symbol $\gtrsim$ or $\lesssim$ represents the fuzzy greater or equal to and fuzzy less or equal to optimization techniques of fuzzy logic such as fuzzy mathematical programming is used instead of the crisp optimization techniques to solve the problems regarding the optimization of fuzzy.

Release Time Problem Soultion: Optimization Using Fuzzy

Under this part of paper, using Zimmermann's (2001) approach we discuss various steps pertaining to solve the problem related to software release time under the influence of fuzzy environment as given in section of fuzzy problem formulation. Using the fuzzy environment, the problem (5) is reaffirmed as:

Find T

Such that $Co(W(T)) \leq Co_0$

$R(Ts|T) \geq R_0$

$W(T) \leq W_0$

$T \geq 0$ (6)

For every fuzzy inequality, we define membership functions $\nu_i(T)$, i=1, 2, 3 given in the problem as:

$$\nu_1(T) = \begin{cases} 1 & ; Co(T) \leq Co_0 \\ \frac{Co^* - Co(T)}{Co^* - Co_0} & ; Co_0 < Co(T) \leq Co^* \\ 0 & ; Co(T) > Co^* \end{cases}$$

$$\nu_2(T) = \begin{cases} 1 & ; R(Ts|T) \geq R_0 \\ \frac{R(Ts|T) - R^*}{R_0 - R^*} & ; R^* \leq R(Ts|T) < R_0 \\ 0 & ; R(Ts|T) < R^* \end{cases}$$

$$\nu_3(T) = \begin{cases} 1 & ; W(T) \leq W_0 \\ \frac{W^* - W(T)}{W^* - W_0} & ; W_0 < W(T) \leq W^* \\ 0 & ; W(T) > W^* \end{cases} \tag{7}$$

Here, cost tolerance is given by Co*, R* defines reliability level tolerance, and W* gives testing efforts. In order solve inequalities with in fuzzy system corresponding to problem given in section (6), fuzzy decisions are made using Bellman and Zadeh's (1973) principle. The crisp optimization problem is given by,

Maximize α

Subject $v_i(T) \geq \alpha, \quad \varphi = 1, 2, 3; \quad \beta \geq 0, \quad T \geq 0$ (8)

Optimal solution for problem (8) can be found if feasible. Problem (8) can also be solved with help of mathematical programming prior to incorporation of parameter value.

Mathematical Example

It's already presumed that parameters a and b of the SRGM are already predicted and tested on the dataset composed by Obha (1984). Predicted values of constraints over the dataset provided by Obha (1984) are a=354.78 and b=0.0889. Further, it is assumed that values of Co_1, Co_2, Co_3, and T_s are already known. We have taken here Co_1=12, Co_2=30, Co_3=50, and T_s=5. Its presumed total budget cost for management is Co_0=11000, and required reliability for software at delivery time R_0=0.89 and testing effort required is W_0=70 and tolerance level of required cost, reliability, and testing effort are given as Co*=13000, R*=0.999, and W*= 150. The fuzzy optimization technique has been used to find the solution with help of above given values of constants and parameters. The cost, reliability, and testing effort membership functions plotted on reliability are shown in figures.

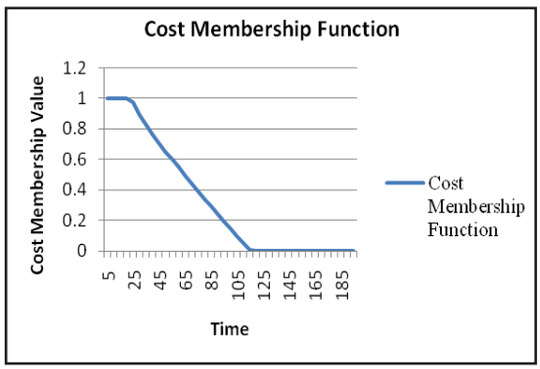


Fig. 1: Cost Membership Function

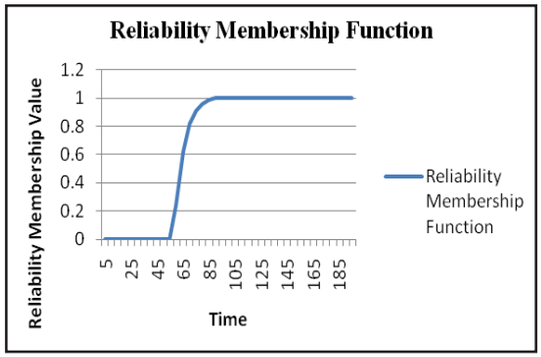


Fig. 2: Reliability Membership Function

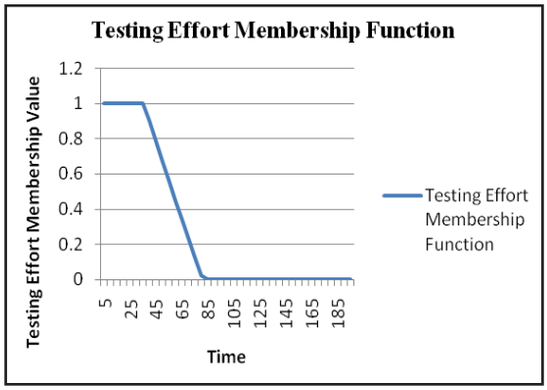


Fig. 3: Testing Membership Function

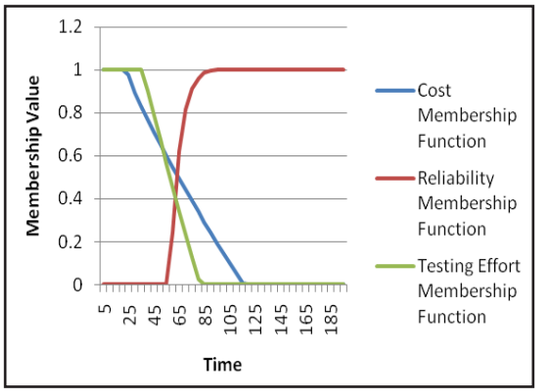


Fig. 4

From Fig. 4, on solving the problem (8) for the optimum software release time, we get T=50 and α=0.6731.

Conclusion

In the given paper fuzzy optimization has been discussed to solve problem pertaining to software release time with

the help of constraints to lower the software cost in account to testing effort and software reliability. A mathematical example along with the result is discussed for solving the fuzzy logic optimization problem.

Authors of this paper have considered cost as the objective function. Due to increasing competition in the growing software industry, software developers try to achieve and consider various things simultaneously, resulting in the optimization of software release time. In future, bi-criterion software release time problem with the help of fuzzy optimization techniques can be discussed.

References

Bector, C. R., & Chandra, S. (2005). *Fuzzy mathematical programming and fuzzy matrix games*. Berlin, Heidelberg: Springer-Verlag.

Bellman, R. E., & Zadeh, L. A. (1973). Decision making in a fuzzy environment. *Management Science*, *17*(4), 141-164.

Huang, C. Y., Kuo, S. Y., & Chen, J. Y. (1997). Analysis of a software reliability growth model with logistic testing effort function. *Proceeding of 8th International Symposium on software reliability engineering*, 378-388.

Kapur, P. K., & Garg, R. B. (1991). Optimal release policies for software systems with testing effort. *International Journal of Systems Science, 22*(9), 1563-1571.

Kapur, P. K., Garg, R. B., & Kumar, S. (1999). *Contributions to hardware and software reliability.* Singapore: World Scientific Publishing Co. Ltd.

Khatri, S. K., Kumar, D., Dwivedi, A., & Mrinal, N. (2012). Software Reliability Growth Model with Testing Efforts using Learning Function. *IEEE explore digital library, Proceedings of International Conference on Software Engineering (Conseg)*, 1-5.

Obha, M. (1984). Inflection S-shaped Software Reliability Growth Model. In Osaki S., Hatoyama Y. (eds) *Stochastic Models in Reliability Theory.* Lecture notes in Economics and Mathematical Systems, *235,* 144-162. Berlin, Heidelberg: Springer.

Zimmermann, H. J. (2001). *Fuzzy set theory and its applications*. New York: Springer.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 35-46
http://publishingindia.com/ijbai/

Estimation of Hedging Effectiveness Using Variance Reduction and Risk-Return Approaches: Evidence from National Stock Exchange of India

Mandeep Kaur*, Kapil Gupta**

Abstract

Present study estimates the hedging effectiveness by applying variance-reduction framework and risk-return framework using near month contracts of three benchmark indices (NIFTY50, NIFTYIT, and BANKNIFTY) traded at National Stock Exchange of India (NSE) for the sample period from June 2000 to March 31, 2017 by using nine optimal hedge ratio models. Out of these nine models, six are constant hedging models and three are time-varying hedging models. The study finds that using variance-reduction framework, highest hedging effectiveness is achieved using Ordinary Least Square model; whereas, 1:1 naïve hedge ratio gives lowest hedging effectiveness. On the other hand, when hedging effectiveness is estimated in a risk-return framework, naïve hedge ratio gives highest hedging effectiveness; whereas, OLS gives the least estimate. Secondly, the coefficients of both optimal hedge ratio as well as hedging effectiveness have increased during post-crisis period implying an increase in the cost of hedging. These findings suggests that conventional hedging models are more efficient than highly complicated time-varying hedging models for estimating optimal hedge ratio, these findings are consistent with the findings of Lien (2005), Bhaduri and Durai (2007), Bhargava (2007), Mandal (2011), Wang et al. (2015).

Keywords: Hedging Effectiveness, Optimal Hedge Ratio, Equity Futures Market, Generalized Auto-regressive Conditional Heteroscedasticity (GARCH), Constant Hedge Ratio, Time-Varying Hedge Ratio

JEL: C13, C22, C32, D81, D82, G12, G14, N25, and O16

Introduction

The globalization of financial markets as well as political and economic disturbances around the world have increased the exposure to financial risk. Therefore, as a need to hedge the financial risk, derivative contracts have been introduced which includes futures contracts, options contracts, swaps, swaptions, and so on. Literature observes that although futures market plays a significant role in hedging price risk, price discovery and increasing cash market efficiency, yet hedging is considered the primary function of futures market.

The co-movement and long-term equilibrium relationship between spot and futures market enables hedger to offset price fluctuations in underlying asset prices by taking opposite position in both spot and futures market. However, numerous studies[1] document the fact that such a relationship gets disturbed in the shortrun due to the presence of market frictions such as: noise trading, infrequent trading of component stocks of underlying index, difference in the trading cost in both the markets, violation of assumptions of cost of carry model, etc. Such disturbances lead to basis risk, which mandates a hedger to estimate the required number of futures contracts to achieve superior hedging effectiveness (according to specific objective function to be optimized).

While designing an efficient hedge strategy, the objective of investors to hedge is of prime consideration. There are three different views on hedging based upon investor's objective to hedge. The traditional theory assumes

[1] Castelino (1992); Figlewski (1984); Stoll and Whaley (1990)

* Research Scholar, Department of Management, I. K. Gujral Punjab Technical University, Kapurthala, Punjab, India.
Email: kaur_mandeep13@ymail.com
** Assistant Professor, Department of Management, I. K. Gujral Punjab Technical University, Kapurthala, Punjab, India.
Email: kapilfutures@gmail.com

investor as a pure-risk avoider, whereas Working (1953) views hedger as a pure risk-taker speculating on the spread between futures and cash prices. The third theory adopts a hybrid approach and claims that a hedger neither purely avoids risk nor does he increases his risk to highest levels. Instead, a hedger prefers a portfolio that optimizes his level of risk and return. This theory, known as Portfolio Hedging Theory, became the most widely accepted framework for designing hedge strategy.

The literature on estimation of optimal hedge ratio initiated with the proposal of Minimum-Variance Hedge Ratio (MVHR) framework suggested by Ederington (1979), Johnson (1960), and Stein (1961). Johnson (1960) and Stein (1961) prepared a theoretical background for estimating MVHR, known as Portfolio theory, based upon which, Ederington (1979) suggested that MVHR could be estimated as the ratio of covariance of spot-futures returns and variance of futures returns. In this view, Ederington (1979) suggested single regression equation (Ordinary Least Square (OLS)) that regresses cash returns upon futures return for estimating optimal hedge ratio. Ederington's OLS is the most simplest of all the models, and thus highly appreciated by a large body of literature (Bhargava & Malhotra (2007); Bonga & Umoetek (2016); Deaves (1994); Lien (2005); Lien et al. (2002); Lee & Chien (2010); Malliaris & Urrutia (1991); Mandal (2011); Moon et al. (2009)).

Further, despite huge popularity of Ederington's model, voluminous literature[2] observes that Ederington's OLS hedge ratio does not account for heteroscedasticity, i.e., it ignores the fact that financial time-series exhibits time-varying volatility; therefore, OLS technique results in biased estimate of optimal hedge ratio. In order to overcome these limitations, ARCH model proposed by Engle (1982) and generalized by Bollerslev (1986) has been extensively used in literature to improve the effectiveness of hedge. For instance, Choudhary (2004) compared the hedging performance of Ederington's OLS and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model and concluded that hedging performance is improved by using time-varying GARCH relative to OLS technique. As such, voluminous literature (Bekkerman, 2011; Giha & Zuppiroli, 2014; Kroner & Sultan, 1993; Myers, 1991; Park & Switzer, 1995; Srinivasan, 2011; Yang & Allen, 2004), and found improved hedging effectiveness using time-varying hedging models.

As discussed previously, numerous studies claim superior performance of the time-varying hedge ratios. Alternatively, a strand of literature favours the use of constant hedging models. For instance, Lee and Chien (2010) analysed hedging performance of TAIFEX index futures contracts by comparing standard OLS model with GARCH and found that OLS provided superior hedging effectiveness over GARCH model. Similarly, Awang et al, (2014), Bonga and Umoetek (2016) and Mandal (2011)compared OLS technique with other time-varying models and concluded that OLS performed better than time-varying hedging models.

Furthermore, Ederington (1979) suggested a measure of hedging effectiveness, which is measured as proportionate reduction in standard deviation of returns from hedged portfolio. The hedge ratio that gives highest hedging effectiveness is popularly known as MVHR. Ederington's measure of hedging effectiveness has been widely appreciated in the literature (Bhargava and Malhotra (2007), Bhaduri and Durai (2007); Chen et al. (2002); Floros and Vougas (2004, 2006); Gupta and Singh (2009); Holmes (1995); Hou and Li (2013); Lypny and Powella (1998); Men and Men (2008); Park and Switzer (1995); Pradhan (2011); Yang and Allen (2005)) mainly due to its simplicity to compute and understand.

Furthermore, despite huge popularity of Ederington's measure of hedging effectiveness, a strand of literature criticizes it on the ground that it focuses solely on variance reduction and ignores any changes in portfolio returns. In other words, hedging is viewed as comprising of minimization of risk only; whereas, on the contrary, Brailsford et al. (2001) and Penning and Meulenberg (1997) suggest that hedging should comprise of both risk reduction as well as return maximization. Therefore, in order to overcome this limitation, few models have been proposed in the literature (see, Chang and Shanker (1987); Hsin et al. (1994); Howard and D'Antonio (1984); Lindahl (1991), etc.), which take into consideration changes in expected return on hedged as well as unhedged portfolio

[2] Basher and Sadorsky (2016); Bekkerman (2011); Choudhary (2003); Choudary (2004); Floros and Vougas (2004); Floros and Vougas (2006); Lee and Yoder (2007); Lypny and Powalla (1998); Moschini and Myers (2002); Park and Switzer (1995); Srinivasan (2011); Yang and Allen (2005).

in addition to risk minimization. For, instance, Howard and D'Antonio (1984) suggested a risk-return measure of hedging effectiveness based upon Sharpe's measure, which is further elaborated in Equation (11).

Apart from the previously discussed issues on optimal hedge ratio and hedging effectiveness, it is observed that futures trading is not only popular in developed markets of the world, but is also equally popular in emerging markets like India, which is evident from the fact that Indian equity futures market consistently rank amongst the top five markets of the world for the last decade. However, in India, to the best of our knowledge, only few attempts have been made to examine the hedging effectiveness[3] and these studies have primarily focused on examining a superior methodology for determining optimal hedge ratio in variance-reduction framework. To the best of our knowledge, none of these studies has attempted to examine hedging effectiveness in a risk-return framework. Therefore, present study is an attempt to examine the hedging effectiveness in a risk-return framework as suggested by Howard and D'Antonio (1984), in addition to estimating hedging effectiveness based upon measure proposed by Ederington (1979) and optimal hedge ratios using nine econometric models. In addition, an attempt has been made to study the impact of financial crisis on optimal hedge ratio and hedging effectiveness.

[3] Bhaduri and Durai, 2007; Rao and Thakur, 2008; Gupta and Singh, 2009, Pradhan, 2011, Haq and Rao, 2013, Kumar and Pandey, 2013, Malhotra, 2015

Database and Research Methodology

In India, L.C. Gupta committee recommended the introduction of equity derivative market, which initiated with the launch of equity futures contracts at NSE and BSE from June 2000. Since inception of derivative market, equity futures market has shown a phenomenal growth[4] both in terms of volume of contracts and number of products, which include stock and index futures and options contracts, interest rate futures contracts, currency futures and options, etc. As far as the present study is concerned, only equity futures contracts have been considered and the sample size of the study comprises of near month futures contracts of NIFTY, NIFTYIT, and BANKNIFTY. These indices have been selected based on their uninterrupted trading history and high liquidity and the data has been collected from official website of the National Stock Exchange of India (NSE), i.e., www.nseindia.com. The period of the study is from inception date of respective indices until March 31, 2017 as presented in Table 1 below:

[4] During 2015–16, the total number of index futures traded on NSE is 140,538,674 and stock futures is 234,243,967, source; www.nseindia.com

Table 1: Sample Size and Sample Period of Study

Symbol	*Period of study*	*Number of Observations*		*Total*
		Pre-Crisis	*Post-Crisis*	
NIFTY50	June 12, 2000 – March 31, 2017	1898	2290	4188
NIFTYIT	August 29, 2003 – March 31, 2017	1092	2290	3382
BANKNIFTY	June 13, 2005 – March 31, 2017	638	2290	2928

Unit-root Test

The first step to analyze the time-series data is to investigate if the series under study is stationary or non-stationary. This is because estimation of optimal hedge ratio involves regression analysis where cash market returns are regressed upon futures market returns. Hence, to avoid the spurious statistical results, stationarity of series is a pre-requisite. Augmented Dickey Fuller (ADF) test has been used to investigate the presence of unit-roots. It has been observed that the both spot and futures prices are non-stationery at levels. Therefore, the price series have been transformed into return series by taking the natural log of first difference of prices, which is found to be stationery[5]. Thus, cash and futures returns have been used for estimating optimal hedge ratio.

[5] The estimated results are not reported in the paper, butcan be provided on demand.

Estimation of Optimal Hedge Ratio

In the present study, optimal hedge ratio has been estimated using nine econometrical procedures andan efficient hedge ratio would be the one that provides the highest reduction in the portfolio variance. The models are explained as follows:

Model 1: Naive one-to-one Model

Naïve hedge ratio is a model free estimation procedure, whichassumes that futures and cash market observes perfect correlation. Therefore, optimal hedge ratio suggested by this model is one, which implies equal investment in both futures and spot market.

Model 2: Ordinary Least Square (OLS)

The second is OLS Method also known as single-equation method in which cash market returns are regressed upon futures returns to estimate optimal hedge ratio as given in Equation (1). Suggested by Ederington (1979), this method is the most widely used for estimating OHR as discussed in Section 2 and is specified as follows:

$$R_{s,t} = \alpha_0 + \beta_1 R_{f,t} + \mu_t \quad (1)$$

In the given regression Equation (1), $R_{s,t}$ represent returns from cash market, $R_{f,t}$ represent returns from futures market, α_0 is the intercept term, β_1 is the optimal hedge ratio and μ_t is the error term.

Model 3: Autoregressive Moving Average OLS

The standard OLS model mentioned in Equation (1) does not take into account serial correlation of stock returns, i.e., the present stock prices are dependent upon its past values;therefore, the estimated coefficient of optimal hedge ratio may be biased. In other words, stock prices are not random and any information set continues to affect stock prices for some time. Autocorrelation in stock returns has become stylized in the financial literature Therefore, autoregressive terms are incorporated in Equation (1) and the modified estimation procedure is presented in Equation (2) as follows:

$$R_{s,t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{s,t-i} + \beta_1 R_{f,t} + \mu_t \quad (2)$$

The autoregressive terms in Equation (2) is represented by $\left(\sum_{i=1}^{p} \alpha_i R_{s,t-i}\right)$. The order of the autoregressive terms is determined according to Schwartz Information Criteria.

Model 4: Modified OLS

It is observed that futures prices are an unbiased predictor of cash prices and basis, as an error correction term, corrects the deviation between current spot price and its equilibrium price. Therefore, Equation (2) has been further improved by Gupta and Singh (2009) by including first lag of both futures return and basis as presented in Equation (3), where $R_{f,t-1}$ represents lagged futures return and $(R_{f,t-1} - R_{s,t-1})$ represents lagged basis.

$$R_{s,t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{s,t-i} + \beta_1 R_{f,t} + \beta_2 R_{f,t-1} + \beta_3 (R_{f,t-1} - R_{s,t-1}) + \varepsilon_t \quad (3)$$

Model 5: Vector Autoregression (VAR)

VAR overcomes the limitation of OLS regression equation (Equation 1) by modelling the serial correlation of residual series, which OLS fails to capture. VAR model can be specified as under:

$$R_{s,t} = \sum_{i=1}^{M} \alpha_i R_{s,t-i} + \sum_{j=1}^{N} \beta_j R_{f,t-j} + \mu_{st} \quad (4)$$

$$R_{f,t} = \sum_{k=1}^{O} \alpha_k R_{s,t-k} + \sum_{l=1}^{P} \beta_1 R_{s,t-1} + \mu_{ft} \quad (5)$$

After running the given regression equations, optimal hedge ratio can be estimated as ratio of covariance of $\mu_{s,t}$ and variance of μ_{ft}. However, this model fails to capture the long-run cointegration between spot and futures prices.

Model 6: Vector Error Correction (VEC) Model

Ghosh (1993) and Lien (2004) argue that when spot-future prices are cointegrated in the longrun, the OLS equation gives an underestimated value of the optimal hedge ratio. Therefore, VAR model with an error correction term (known as VECM) is used to account for long-run co-integrating relationship in addition to capturing short-run lead-lag relationship. The VECM model can be specified as below:

$$R_{f,t} = \alpha_{0f} + \sum_{i=1}^{p} \alpha_{if}(F_{t-i} - S_{t-i}) + \sum_{j=1}^{q} \beta_f R_{f,t-j} + \sum_{k=1}^{m} \beta_f R_{s,t-k} + \mu_{ft} \quad (6)$$

$$R_{s,t} = \alpha_{0s} + \sum_{i=1}^{p} \alpha_{is}(F_{t-i} - S_{t-i}) + \sum_{l=1}^{n} \beta_s R_{s,t-l} + \sum_{h=1}^{o} \beta_s R_{f,t-h} + \mu_{st} \quad (7)$$

The optimal hedge ratio using VECM can be estimated as ratio of covariance of ($\mu_{s,t}$) and variance of (μ_{ft}), as computed in case of VAR model above.

Model 7: Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

The estimation procedures discussed earlier (Equations (1) though (7)) assume that volatility of asset returns remain constant; however, literature[6] argues that covariance and variance of returns are timevarying. Therefore, in order to capture the time-varying volatility of stock returns, GARCH model, proposed by Bollerslev (1986), has been used to estimate optimal hedge ratio. GARCH models the time-varying volatility by using a variance equation (Equation (8)) along with the mean equation as given in Equation (2). The variance equation is as follows:

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + \upsilon_t \quad (8)$$

In Equation (8), h_t is the conditional volatility, α_i is the coefficient of autoregressive term and β_j is the coefficient of GARCH term.

Model 8: Exponential GARCH (EGARCH)

Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model, proposed by Nelson (1991), captures such asymmetric relationship between conditional volatility and conditional mean. The specification of EGARCH model is as follows:

$$h_t = {}^{3}{}_{1} + {}^{3}{}_{2} \left| \frac{\mu^2_{t-1}}{h_{t-1}} \right| + {}^{3}{}_{3} \frac{\mu^2_{t-1}}{h_{t-1}} + {}^{3}{}_{4} h_{t-1} \quad (9)$$

Model 9: Threshold ARCH (TARCH)

Numerous studies (like, Karpoff (1987), Veronesi (1999), etc.) find that the reaction of investors vary with the type of information received in the market which generate different levels of volatility. For instance, Veronesi (1999) finds that investors tend to overreact to bad news in good times and under-react to good news in bad times. Therefore, in order to capture the asymmetric investor's reactions, TARCH model is used, whose variance equation is represented as follows:

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \alpha_k \varepsilon_{t-1}^2 \xi_{t-i} + \sum_{j=1}^{p} \beta_j h_{t-j} + \upsilon_t \quad (10)$$

In Equation (10), $\varepsilon_{t-1}^2 \xi_{t-i}$ represents the dummy variable having value one if the news is negative and zero for non-negative news.

Estimation of Hedging Effectiveness

The previous statistical procedure suggests the optimal hedge ratio(s); however, the effectiveness of these estimated optimal hedge ratio(s) shall have to be computed based upon two criteria: first, variance-reduction criterion suggested by Ederington (1979) and second, risk-return criterion suggested by Howard and D'Antonio (1984). The hedge ratio that gives the highest hedging effectiveness in each of the two methods would be proposed as efficient hedge ratio.

Framework 1: Variance-Reduction Framework

After estimating the optimal hedge ratio(s) using the given econometric procedures, their effectiveness has been tested by using a measure suggested by Ederington (1979). The method suggested by Ederington measures hedging effectiveness as proportionate decline in portfolio variance and optimal hedge ratio that declines the portfolio variance to the maximum extent is considered as an efficient hedge ratio. The Ederington's hedging effectiveness is based upon Sharpe's measure and is calculated as follows:

$$\text{Hedge effectiveness} = \frac{\text{Var(U)} - \text{Var(H)}}{\text{Var(U)}} \quad (11)$$

In Equation (11), variance of unhedged portfolio is the same as variance of cash returns, whereas the variance of hedged portfolio is measured as $\sigma_s^2 + h^{*2}\sigma_f^2 - 2h^*\sigma_{s,f.}$

Framework 2: Risk-Return Framework

As already mentioned in Section I, the variance reduction measure of hedging effectiveness gained huge popularity

[6] Engle(1982), Bollerslev (1986), Lypny and Powalla (1998) and Floros and Vougas (2004)

in the academic literature. However, this method does not take into account the return on hedged and unhedged portfolio(s). Therefore, Howard and D'Antonio (1984) suggested a measure of hedging effectiveness (λ) that incorporates the return component.Equation (12) specifies the estimation of hedging effectiveness which is measured as ratio of slope of risk-return relative from hedged portfolio and risk-return relative from unhedged portfolio.

$$HE = \theta \Big/ \frac{\bar{r}_s - i}{\sigma_s} \qquad (12)$$

Where, $\theta = \frac{\bar{R}_p - i}{\sigma_p}$

$\bar{R}_p$ = expected return from hedged portfolio

σ_p = standard deviation of returns from hedged portfolio

i = risk-free rate of return

$\bar{r}_s$ = expected return from unhedged portfolio

σ_s = standard deviation of returns from unhedged portfolio

Results and Analysis

The descriptive statistics of cash and futures returns for all the indices under study is reported in Table 2. All the stocks show excess kurtosis and their coefficient of skewness is negative implying that the return series are leptokurtic in nature. These statistics indicate that the returns are not normal which is further supported by Jarque-Bera test that rejects the null hypothesis that cash and futures returns are normal.

Table 2: Descriptive Statistics of Returns

Contract	variables	Count	Mean	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
NIFTY50	Futures Return	4187	0.000425	0.015731	-0.493672	12.83305	16029.05(0.00)
	Cash Return	4187	0.000427	0.015061	-0.296783	11.83627	12872.65(0.00)
	Basis	4188	25.02533	43.84938	1.049673	3.455654	757.6089(0.00)
NIFTYIT	Futures Return	3381	-9.20E-05	0.045839	-43.78679	2257.988	6.65E+08(0.00)
	Cash Return	3381	-9.13E-05	0.045663	-43.98293	2271.964	6.73E+08(0.00)
	Basis	3382	8.270341	41.38112	-0.231291	59.98556	424078.3 (0.00)
BANKNIFTY	Futures Return	2927	0.000559	0.020685	0.038455	7.632446	2396.083(0.00)
	Cash Return	2927	0.000556	0.020167	0.075152	7.537706	2300.967(0.00)
	Basis	2728	17.07549	38.06139	0.636343	4.800262	542.7752(0.00)

Further, optimal hedge ratio(s) have been estimated using nine econometrical procedures: Naive, Ederington's Model, ARMA (p,q), Modified OLS, VAR, VECM, GARCH (p,q), EGARCH (p,q), and TARCH (p,q) and the results are reported in Table 3. Two important observations can be seen. Firstly, in case of all these indices, Ederington's OLS gives the lowest coefficient of hedge ratio. Secondly, coefficient of optimal hedge ratio(s) estimated through constant hedge ratio models is relatively smaller than the hedge ratio estimated through time varying models, i.e., GARCH, EGARCH, and TARCH. These findings imply that the constant hedge ratio models offer economical hedging as compared to time-varying models because lower coefficient of hedge ratio means lower investment in the futures contracts.[7] Furthermore, these results remain consistent during pre-crisis as well as post-crisis period (Table 4), which implies that the state of market does not affect the hedging model to be used to estimate hedging effectiveness. However, a slight increase in the coefficients of hedge ratio for NIFTY and BANKNIFTY during post-crisis period has been observed; whereas in case of NIFTYIT, coefficients of only time-varying hedge ratios have been increased.

[7] Lower hedge ratio implies lower investment in futures contracts.

Table 3: Estimation of Optimal Hedge Ratio

Contract	Naïve	OLS	ARMA OLS	Modified OLS	VAR	VECM	GARCH (1,1)	EGARCH (1,1)	TARCH (1,1)
NIFTY50	1	0.936	0.953	0.941	0.939	0.940	0.955	0.967	0.962
NIFTYIT	1	0.993	0.998	0.996	0.996	0.996	0.998	0.997	1.002
BANKNIFTY	1	0.967	0.979	0.970	0.968	0.968	0.981	0.987	0.984

Furthermore, Table 5 reports the hedging effectiveness in the form of variance reduction, proposed by Ederington (1979), after taking hedging position with the estimated optimal hedge ratio(s). It is observed that constant hedging models (OLS, Modified OLS, VAR , and VECM) give highest hedging effectiveness[8]. On the other hand, traditional 1:1 naïve gives poorest hedging effectiveness. Moreover, there is very insignificant difference between the coefficients of hedging effectiveness estimated from different hedge ratio models understudy. These findings are consistent with the findings of Lien et al. (2002), Lien (2005) and Maharaj et al. (2008) who find no significant improvement in hedging effectiveness using sophisticated econometric methods. Moreover, the impact of financial crisis of 2008 on hedging effectiveness has been studied and results have been reported in Table 6 and it is interesting to note that OLS model dominates over other hedging models in obtaining highest hedging effectiveness, while remaining unaffected by the impact of financial crisis 2008. Another observable fact is that hedging effectiveness increases after the global financial crisis, for all indices understudy, except NIFTYIT.

[8] These findings are consistent with the findings of Lien et al (2002), Moosa (2003), Lien (2005), Maharaj et al. (2008), Bhargava and Malhotra (2008), Rao and Thakur (2008)).

Table 4: Optimal Hedge Ratio Over Pre- and Post-Crisis Period

Contract	*Period*	*Naïve*	*OLS*	*ARMA OLS*	*Modified OLS*	*VAR*	*VECM*	*GARCH (1,1)*	*EGARCH (1,1)*	*TARCH (1,1)*
NIFTY50	Pre-crisis	1	0.919	0.941	0.921	0.923	0.925	0.936	0.931	0.935
	Post-crisis	1	0.950	0.964	0.958	0.949	0.950	0.967	0.976	0.976
NIFTYIT	Pre-crisis	1	0.995	0.999	0.998	0.997	0.998	0.999	0.997	0.996
	Post-crisis	1	0.982	0.988	0.983	0.993	0.994	1.004	1.007	1.005
BANKNIFTY	Pre-crisis	1	0.953	0.976	0.959	0.957	0.962	0.982	0.985	0.982
	Post-crisis	1	0.970	0.979	0.969	0.971	0.971	0.982	0.988	0.984

Table 5: Hedging Effectiveness in Variance-Reduction Framework

Contract	*Naïve*	*OLS*	*ARMA OLS*	*Modified OLS*	*VAR*	*VECM*	*GARCH (1,1)*	*EGARCH (1,1)*	*TARCH (1,1)*
NIFTY50	96.083	96.539	96.507	96.536	96.538	96.537	96.490	96.431	96.458
NIFTYIT	99.359	99.365	99.362	99.364	99.364	99.364	99.362	99.361	99.357
BANKNIFTY	97.969	98.088	98.069	98.088	98.088	98.088	98.066	98.044	98.054

Table 6: Hedging Effectiveness in Variance-Reduction Framework Over Pre- and Post-Crisis Period

Contract	*Period*	*Naïve*	*OLS*	*ARMA OLS*	*Modified OLS*	*VAR*	*VECM*	*GARCH (1,1)*	*EGARCH (1,1)*	*TARCH (1,1)*
NIFTY50	Pre-crisis	93.805	94.547	94.489	94.547	94.545	94.543	94.512	94.531	94.519
	Post-crisis	97.894	98.168	98.147	98.160	98.168	98.168	98.136	98.091	98.091
NIFTYIT (previously CNXIT)	Pre-crisis	99.553	99.556	99.555	99.555	99.556	99.555	99.555	99.556	99.557
	Post-crisis	96.437	96.473	96.468	96.472	96.459	96.457	96.422	96.408	96.415
BANKNIFTY	Pre-crisis	95.930	96.182	96.115	96.179	96.178	96.171	96.084	96.064	96.085
	Post-crisis	98.418	98.512	98.502	98.512	98.512	98.512	98.498	98.479	98.492

Furthermore, Table 7 reports the hedging effectiveness estimated using risk-return criteria proposed by Howard and D'Antonio (1984) that incorporates both risk and return components on hedged portfolio. It is observed

that naïve hedge ratio gives highest hedging effectiveness for NIFTY, NIFTYIT, and BANKNIFTY; whereas, Ederington's OLS hedge ratio gives lowest hedging effectiveness. Furthermore, the impact of financial crisis on hedging effectiveness has been examined and results have been reported in Table 8. It is observed that Naïve hedge ratio gives highest hedging effectiveness (except NIFTYIT post-crisis); whereas, Ederington's OLS gives lowest hedging effectiveness. Moreover, it is found that there has been an increase in hedging effectiveness during post crisis (except NIFTYIT post-crisis).

Table 7: Hedging Effectiveness in Risk-Return Framework

Contract	*Naïve*	*OLS*	*ARMA OLS*	*Modified OLS*	*VAR*	*VECM*	*GARCH (1,1)*	*EGARCH (1,1)*	*TARCH (1,1)*
NIFTY50	1.2452	1.2374	1.2395	1.2379	1.2377	1.2379	1.2399	1.2412	1.2407
NIFTYIT	1.3245	1.3234	1.3241	1.3238	1.3239	1.3238	1.3242	1.3239	1.3247
BANKNIFTY	1.16203	1.1594	1.1604	1.1594	1.1594	1.1595	1.1605	1.1609	1.1608

Table 8: Hedging Effectiveness in Risk-Return Framework over Pre- and Post-Crisis Period

Contract	*Period*	*Naïve*	*OLS*	*ARMA OLS*	*Modified OLS*	*VAR*	*VECM*	*GARCH (1,1)*	*EGARCH (1,1)*	*TARCH (1,1)*
NIFTY50	Pre-crisis	1.1053	1.1013	1.1024	1.1014	1.1015	1.1016	1.1022	1.1019	1.1021
	Post-crisis	1.2810	1.2741	1.2760	1.2753	1.2740	1.2742	1.2765	1.2778	1.2778
NIFTYIT	Pre-crisis	1.2899	1.2891	1.2897	1.2896	1.2894	1.2896	1.2897	1.2894	1.2893
	Post-crisis	1.3857	1.3822	1.3834	1.3825	1.3843	1.3845	1.3864	1.3869	1.3867
,BANKNIFTY	Pre-crisis	1.0570	1.0557	1.0563	1.0559	1.0558	1.0559	1.0565	1.0566	1.0565
	Post-crisis	1.1983	1.1954	1.1963	1.1953	1.1954	1.1955	1.1965	1.1971	1.1967

Further, it is observed that NIFTYIT has been an exception to the previous findings related to optimal hedge ratio and hedging effectiveness. Such exception may be due to severe impact of global financial crisis on Indian IT industry and investor sentiment. This is quite evident from the correlation between spot and futures prices as well as trading of IT stocks, both of which declined after the global financial crisis (see Tables 9 and 10).

Table 9: Descriptive Statistics of Volume of Futures Contract

Symbol	*Period*	*Count*	*Mean*	*Minimum*	*Maximum*	*Std. Dev.*
NIFTY50	Pre	1898	135556.2	19	1338598	183077.1
	Post	2290	415608.1	14371	1343511	207469.5
NIFTYIT	Pre	1092	471.1612	0	3683	480.5841
	Post	2290	315.1019	1	3395	289.932
BANKNIFTY	Pre	638	2011.188	27	10453	1.409485
	Post	2290	73973.49	557	343417	46689.565

Table 10: Correlation Coefficient between Cash and Futures Returns

Symbol	*Period*	*Count*	*Correlation Coefficient*
NIFTY50	Pre-Crisis	1897	0.973
	Post-Crisis	2290	0.991
NIFTYIT	Pre-Crisis	1091	0.999
	Post-Crisis	2290	0.983
BANKNIFTY	Pre-Crisis	637	0.982
	Post-Crisis	2290	0.994

Conclusion

Indian equity futures market has recorded voluminous growth since its inception in year 2000 and to the best of our knowledge, there have been only a few attempts to study it (see, Bhaduri and Durai, 2007; Gupta and Singh, 2009; Haq and Rao, 2013; Malhotra, 2015; Pradhan, 2011; Rao and Thakur, 2008) all of which have restricted their scope to examine the hedging effectiveness in a traditional risk minimization framework that ignores the changes in return on hedged portfolio. Therefore, present study is an attempt to examine hedging effectiveness by two methods: variance reduction criteria proposed by Ederington (1979) and risk-return criteria proposed by Howard and D' Antonio (1984) by using three benchmark indices of NSE (NIFTY, NIFTYIT, and BANKNIFTY) from their respective date of inception till March 31, 2016. Also, an attempt has been made to study the impact of financial crisis of 2008 by segregating the return series into pre-crisis period (inception date–December 31, 2007) and post-crisis period (January 1, 2008–March 16, 2016).

In the present study, optimal hedge ratio has been estimated using constant hedging models (Ederington's OLS model, Modified OLS, VAR, VECM, and ARMA (p,q))and time-varying hedging models (GARCH, EGARCH, and TARCH). It is observed that coefficient of optimal hedge ratio estimated through constant hedging models is comparatively smaller than the hedge ratios estimated through time-varying models. Secondly, after segregating the data series into pre- and post-crisis period, it is observed that hedge ratios during the post-crisis period are relatively higher than pre-crisis period, which implies that the cost of hedging has been increased after the financial crisis.

Furthermore, it is found that in a variance-reduction framework, Ederington's OLS hedge ratio gives highest hedging effectiveness (except for NIFTYIT where VECM gives highest hedging effectiveness); whereas, Naïve hedge ratio gives the lowest hedging effectiveness. These findings are consistent with the findings of Collins (2000), Lien et al. (2002), Moosa (2003), Lien (2005), Bhargava and Malhotra (2007), Rao and Thakur (2008) and remain consistent when the data series is segregated into pre- and post-crisis period. However, when hedging effectiveness is computed in a risk-return framework, Naïve hedge ratio gives highest hedging effectiveness (except for CNXIT post-crisis), whereas OLS gives lowest hedging effectiveness. Once again, the results obtained remains consistent when series is segregated into pre and post crisis period.

From the given findings few important implications can be drawn. Firstly, constant hedging models give highest hedging effectiveness whether estimated based on variance-reduction criteria or risk-return criteria proposed by Howard and D'Antonio (1984). These findings are consistent with the findings of Maharaj et al. (2008) and Wang et al. (2015) who question if sophisticated econometrical procedures can really help in achieving highest hedging effectiveness. However, on the contrary, these findings are inconsistent with numerous studies[9], which suggest that time-varying hedging models dominate constant hedging models. The reason for such anomaly may be attributed to the fact that hedging model to be used may be country specific (Hou and Li, 2013). Secondly, since both the measures of hedging effectiveness suggest different optimal hedging models, selection of right hedging model becomes vital for investor, which depends upon his objective to hedge. Thirdly, there has been increase in estimates of both optimal hedge ratio and hedging effectiveness (except CNXIT) during post-crisis period, which implies an increase in the cost of hedging. The reason for increase in both these estimates can be due to increase in correlation coefficient computed over the post-crisis period.

References

Awang, N., Azizan, N. A., Ibrahim, I., & Said, R. M. (2014). Hedging effectiveness stock index futures market: An analysis on malaysia and singapore futures markets. *Recent Advances in Economics, Management and Development,* pp. 24-34.

Basher, S. B., & Sadorsky, P. (2016). Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH. *Energy Economics, 54,* 235-247.

[9] Myers (1991), Park and Switzer (1995), Lypny and Powella (1998), Yang (2001), Kavussanos and Nomikos (2000), Moschini and Myers (2002), Floros and Vougas (2004), Yang and Allen (2004), Choudhry (2004), Kofman and McGlenchy (2005), Floros and Vougas (2006), Bhaduri and Durai (2007), Lee and Yoder (2007), Kumar et al. (2008), Srinivasan (2011), Hou and Li (2013).

Bekkerman, A. (2011). Time-varying hedge ratios in linked agricultural markets. *Agricultural Finance Review, 71*(2), 179-200.

Bhaduri, S. N., & Durai, S. R. S. (2008). Optimal hedge ratio and hedging effectiveness of stock index futures: Evidence from India. *Macroeconomics and Finance in Emerging Market Economies, 1*(1), 121-134.

Bhargava, V., & Malhotra, D. K. (2007). Determining the optimal hedge ratio: Evidence from cotton and soybean markets. *The Journal of Business and Economic Studies, 13*(1), 38.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*(3), 307-327.

Bonga-Bonga, L., & Umoetok, E. (2016). The effectiveness of index futures hedging in emerging markets during the crisis period of 2008-2010: Evidence from South Africa. *Applied Economics, 48*(42), 3999-4018.

Brailsford, T. J., & Cusack, A. J. (1997). A comparison of futures pricing models in a new market: The case of individual share futures. *Journal of Futures Markets, 17*(5), 515-541.

Brooks, C., & Chong, J. (2001). The cross-currency hedging performance of implied versus statistical forecasting models. *Journal of Futures Markets, 21*(11), 1043-1069.

Butterworth, D., & Holmes, P. (2001). The hedging effectiveness of stock index futures: evidence for the FTSE-100 and FTSE-Mid250 indexes traded in the UK. *Applied Financial Economics, 11*(1), 57-68.

Castelino, M. G. (1992). Hedge effectiveness: Basis risk and minimum-variance hedging. *Journal of Futures Markets, 12*(2), 187-201.

Chan, L. H., & Lien, D. D. (2001). Cash settlement and price discovery in futures markets. *40*(3/4), 65-77.

Chang, J. S. K., & Shanker, L. (1987). A risk-return measure of hedging effectiveness: a comment. *Journal of Financial and Quantitative Analysis, 22*(3), 373-376.

Chang, C. –Y., Lai, J. –Y., & Chuang, I. –Y. (2010). Futures hedging effectiveness under the segmentation of bear/bull energy markets. *Energy Economics, 32*(2), 442-449.

Chen, S. –Y., Lin, C. –C., Chou, P. –H., & Hwang, D. –Y. (2002). A comparison of hedge effectiveness and price discovery between TAIFEX TAIEX index futures and SGX MSCI Taiwan index futures. *Review of Pacific Basin Financial Markets and Policies, 5*(2), 277-300.

Chen, C. –C., & Tsay, W. –J. (2011). A Markov regime-switching ARMA approach for hedging stock indices. *Journal of Futures Markets, 31*(2), 165-191.

Chen, S. –S., Lee, C. –F., & Shrestha, K. (2004). An empirical analysis of the relationship between the hedge ratio and hedging horizon: A simultaneous estimation of the short-and long-run hedge ratios. *Journal of Futures Markets, 24*(4), 359-386.

Chen, S. –S., Lee, C. –F., & Shrestha, K. (2001). On a Mean—Generalized Semivariance Approach to Determining the Hedge Ratio. *Journal of Futures Markets, 21*(6), 581-598.

Chen, P., Zirong, Z., & Liu, J. (2016). Estimation and Comparative of dynamic optimal hedge ratios of China gold futures based on ECM-GARCH. *International Journal of Economics and Finance, 8*(3), 236.

Choudhry, T. (2004). The hedging effectiveness of constant and time-varying hedge ratios using three Pacific Basin stock futures. *International Review of Economics and Finance, 13,* 371–385.

Chou, W. L., Denis, K. K. F., & Lee, C. F. (1996). Hedging with the Nikkei index futures: The convential model versus the error correction model. *The Quarterly Review of Economics and Finance, 36*(4), 495-505.

Collins, R. A (2000). The risk management effectiveness of multivariate hedging models in the US soy complex. *Journal of Futures Markets, 20*(2), 189-204.

Das, J. K., & Chakraborty, G. (2015). The hedging performance of commodity futures in India: An empirical study on some agricultural commodities. *International Journal of Information, Business and Management, 7*(3), 162.

Deaves, R. (1994). Naive versus conditional hedging strategies: The case of Canadian stock index futures. *Canadian Journal of Administrative Sciences, 11*(3), 264-270.

Dimson, E., & Mussavian, M. (1998). A brief history of market efficiency. *European Financial Management, 4*(1), 91-103.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *The Journal of Finance, 34*(1), 157-170.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United

Kingdom inflation. *Econometrica: Journal of the Econometric Society,* 987-1007.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society,* 251-276.

Figlewski, S. (1984). Hedging performance and basis risk in stock index futures. *The Journal of Finance, 39*(3), 657-669.

Floros, C., & Vougas, D. V. (2004). Hedge ratios in Greek stock index futures market. *Applied Financial Economics, 14*(15) 1125-1136.

Floros, C., & Vougas, D. (2006). Hedging effectiveness in Greek stock index futures market, 1999-2001. *International Research Journal of Finance and Economics, 5*, 7-18.

Ghosh, A. (1993). Cointegration and error correction models: Intertemporal causality

between index and futures prices. *Journal of Futures Markets, 13,* 193-198.

Gupta, K., & Singh, B. (2009). Estimating the optimal hedge ratio in the Indian Equity Futures market. *IUP Journal of Financial Risk Management, 6*(3/4), 38.

Herbst, A. F., McCormack, J. P., & West, E. N. (1987). Investigation of a lead-lag relationship between spot stock indices and their futures contracts. *Journal of Futures Markets, 7*(4), 373-381.

Holmes, P. (1995). Ex ante hedge ratios and the hedging effectiveness of the FTSE-100 stock index futures contract. *Applied Economics Letters, 2*(3), 56-59.

Hou, Y., & Li, S. (2013). Hedging performance of Chinese stock index futures: An empirical analysis using wavelet analysis and flexible bivariate GARCH approaches. *Pacific-Basin Finance Journal, 24,* 109-131.

Howard, C. T., & D'Antonio, L. J. (1984). A risk-return measure of hedging effectiveness. *Journal of Financial and Quantitative Analysis, 1*9(1), 101-112.

Hsln, C. –W., Kuo, J., & Lee, C. –F. (1994). A new measure to compare the hedging effectiveness of foreign currency futures versus options. *Journal of Futures Markets, 14*(6), 685-707.

Johnson, L. L. (1960). The theory of hedging and speculation in commodity futures. *The Review of Economic Studies, 27*(3), 139-151.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis, 22*(1), 109-126.

Kavussanos, M. G., & Nomikos, N. K. (2000). Constant vs. time-varying hedge ratios and hedging efficiency in the BIFFEX market. Transportation Research Part E: Logistics and Transportation Review, Vol. 36, No. 4, pp. 229-248.

Kenourgios, D., Samitas, D., & Drosos, P. (2008). Hedge ratio estimation and hedging effectiveness: the case of the S&P 500 stock index futures contract. *International Journal of Risk assessment and management, 9,* 1-2, 121-134.

El-Khatib, Y., & Hatemi-J, A. (2013). Asymmetric optimal hedge ratio with an application. *In Electrical Engineering and Intelligent Systems*, pp. 231-237.

Kofman, P., & McGlenchy, P. (2005). Structurally sound dynamic index futures hedging. *Journal of Futures Markets, 25*(12), 1173-1202.

Koutmos, G., & Pericli, A. (1998). Dynamic hedging of paper with T bill futures. *Journal of Futures Markets, 18*(8), 925-938.

Kroner, K. F., & Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis, 28*(4), 535-551.

Kumar, B., Singh, P., & Pandey, A. (2008). Hedging effectiveness of constant and time varying hedge ratio in Indian stock and commodity futures markets. Available at SSRN 1206555.

Kumar, B., & Pandey, A. (2013). Market efficiency in Indian commodity futures markets. *Journal of Indian Business Research, 5*(2), 101-121.

Lee, H. –T., & Yoder, J. K. (2007). A bivariate Markov regime switching GARCH approach to estimate time varying MVHRs. *Applied Economics, 39*(10), 1253-1265.

Lee, H. –C., & Chien, C. –Y. (2010). Hedging performance and stock market liquidity: Evidence from the taiwan futures market. *Asia-Pacific Journal of Financial Studies, 39*(3), 396-415.

Lien, D. (2005). A note on the superiority of the OLS hedge ratio. *Journal of Futures Markets, 25*(11), 1121-1126.

Lien, D., Tse, Y. K., & Tsui, A. K. C. (2002). Evaluating the hedging performance of the constant-correlation GARCH model. *Applied Financial Economics, 12*(11), 791-798.

Lindahl, M. (1991). Risk-return hedging effectiveness measures for stock index futures. *Journal of Futures markets, 11*(4), 399-409.

Lypny, G., & Powalla, M. (1998). The hedging effectiveness of DAX futures. *The European Journal of Finance, 4*(4), 345-355.

Maharaj, E. A., Moosa, I., Dark, J., & Silvapulle, P. (2008). Wavelet estimation of asymmetric hedge ratios: Does econometric sophistication boost hedging effectiveness? *International Journal of Business and Economics, 7*(3), 213.

Malliaris, A. G., & Urrutia, J. L. (1991). The impact of the lengths of estimation periods and hedging horizons on the effectiveness of a hedge: Evidence from foreign currency futures. *Journal of Futures Markets, 11*(3), 271-289.

Malhotra, M. (2015). Evaluating the Hedging Performance of Oil and Oilseeds Futures in India. *Paradigm, 19*(2), 184-196.

Mandal, A. (2011). Hedging effectiveness of stock index futures contracts in the Indian derivative markets. *International Journal of Financial Management, 1*(2), 1-20.

Men, X., & Men, X. (2009). Hedging effectiveness of Hong Kong stock index futures contracts. *International Journal of Business and Management, 3*(8), 98.

Moon, G. –H., Yu, W. –C., & Hong, C. –H. (2009). Dynamic hedging performance with the evaluation of multivariate GARCH models: Evidence from KOSTAR index futures. *Applied Economics Letters, 16*(9), 913-919.

Moosa, I. (2003). The sensitivity of the optimal hedge ratio to model specification", Finance Letters, Vol. 1, No. 1, pp. 15-20.

Moschini, G. –C., & Myers, R. J. (2002). Testing for constant hedge ratios in commodity markets: a multivariate GARCH approach. *Journal of Empirical Finance, 9*(5), 589-603.

Myers, R. J. (1991). Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets, 11*(1), 39-53.

Nelson, D. B. (1991), Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society,* 347-370.

Park, T. H., & Switzer, L. N. (1995). Time-varying distributions and the optimal hedge ratios for stock index futures. *Applied Financial Economics, 5*(3), 131-137.

Pattarin, F., & Ferretti, R. (2004). The Mib30 index and futures relationship: Econometric analysis and implications for hedging. *Applied Financial Economics, 14*(18), 1281-1289.

Pennings, J. M. E., & Meulenberg, M. T. G. (1997). Hedging efficiency: A futures exchange management approach. *Journal of Futures Markets, 17*(5), 599-615.

Pradhan, K. (2011). The hedging effectiveness of stock index futures: Evidence for the S&P CNX nifty index traded in India. *South East European Journal of Economics and Business, 6*(1), 111-123.

Rao, S. V. D. N., & Thakur, S. K. (2008). *Optimal hedge ratio and hedge efficiency: An empirical investigation of hedging in Indian derivatives market.* Society of Actuaries, Monograph Publication, pp. 1-27.

De Salles, A. A. (2013). An investigation of some hedging strategies for crude oil market. *International Journal of Energy Economics and Policy, 3*(1), 51.

Srinivasan, P. (2011). Hedging effectiveness of constant and time-varying hedge ratio in Indian commodity futures markets: Evidence from The Multi-Commodity Exchange. *IUP Journal of Financial Economics, 9*(3), 7-27.

Stein, J. L. (1961). The simultaneous determination of spot and futures prices. *The American Economic Review, 51*(5), 1012-1025.

Stoll, H. R., & Whaley, R. E. (1990). The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis, 25*(4), 441-468.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *The Review of Financial Studies,* 12(5), 975-1007.

Wang, Y., Wu, C., & Yang, L. (2015). Hedging with futures: Does anything beat the naïve hedging strategy?. *Management Science, 61*(12), 2870-2889.

Working, H. (1953). Futures trading and hedging. *American Economic Review, 43,* 314-343.

Yang, M. J., & Lai, Y. C. (2009). An out-of-sample comparative analysis of hedging performance of stock index futures: Dynamic versus static hedging. *Applied Financial Economics, 19*(13), 1059-1072.

Yang, W., & Allen, D. E. (2005). Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets. *Accounting & Finance, 45*(2), 301-321.

Revoredo-Giha, C., & Zuppiroli, M. (2014). Exploring the hedging effectiveness of European wheat futures markets during the 2007-2012 period. *Procedia Economics and Finance, 14,* 90-99.

International Journal of Business Analytics and Intelligence
6 (1) 2018, 47-57
http://publishingindia.com/ijbai/

Applicability of Random Forests Forecasting to International Currency Trade: An Investigation Through Language

Kamakshaiah Musunuru*, S.S. Prasada Rao**

Abstract

The goal of this research is to study the performance of foreign exchange trade in both India and China. India and China raised rapidly in recent times and there is abundant of speculation that these countries might reach to the level of few other developed nations as far as international trade is concerned. Whereas there isn't any doubt that these countries emerging as economic powers in the Asia-Pacific region, a lot of effort is required at international platform with respect to trade and commerce. One of such areas of competition is international currency trade. The aim of this study is to understand trends of currency trade in order to predict how likely these countries are going to emerge as best in the region. The study used certain secondary datasets from very reliable and authenticated sources. As far as statistical techniques are concerned, random walk forecasting methods were employed to test the study hypothesis. The study gathered certain evidence that though there are similarities in present and past performance, it is not likely to be the same in the future. However, the study concludes that random forests forecasting as a methodology is highly useful in studying trends in the data.

Keywords: Asia-Pacific Region, Currency Trade, International Trade, Random Walk Forecasting, Time Series Analysis

Introduction

By September 1, 2017 foreign exchange reserves in India increased to USD $398,120 million. The averaged foreign exchange reserves were USD $206,149.58 million between 1998 and 2017. By September of 2017, the same had reached to an all-time high level of USD $398,120 million and USD $29,048 million which is a record low in September of 1998 (TE, 2017).

There were certain funny but unfavourable speculations done against American stock markets in 2007. The stock market performance was expected to be low by the very first quarter of 2007; this was the speculation at that time. However, the actual story was different. Many companies surpassed the market speculations and did performed well. Experts show the evidence in support of the fact that about the influence of ever aggrandizing economics of India and China (Bloomberg, 2017). Perhaps, this is one of the reasons for these countries to be viewed as major competitors in the region (Dar and Ahmad, 2014). The other reasons could be-ever increasing consumerism, getting redefined through empowerment in education and affluence supported by credit disbursement from financial institutions. These lead to a plausible reasoning that India's and China's growth invincibly affects the US's growth and remain a robust driver for their economies. The other reason that supports the aforementioned idea is the amount of innovation that is happening in these countries. Markets have affordable products driven by innovation that might affect consumerism grow further and economies evolve more robust.

Some of the neoclassical researchers like Jagadish Sheth coined a new strategy known as Chindia. Some of the world's renowned MNCs like Nokia and GE successfully implemented a Chindia strategy. These researchers say that the shift in focus on these two emerging Asian economies

* Assistant Professor, GITAM School of International Business, GITAM University, Visakhapatnam, Andhra Pradesh, India. Email: kamakshaiah.m@gmail.com

** Director, Academic Affairs & Professor, GITAM Institute of Management, GITAM University, Visakhapatnam, Andhra Pradesh, India. Email: director_academicaffairs@gitam.edu

is necessary not only for the Western countries, but also Japan and South Korea, who will find themselves in a similar position as their population starts to age (Sangani, P., 2008).

One of the best anecdotes of rising importance of these countries can be illustrated by Coca-Cola. Coca-Cola is one of the world's giant entities that believed in Chindia strategy. While the cola biggie's growth has almost flattened out in the United States, in 2006, 70% of its growth and 80% of its profits came from international business, a significant part of this from China. Similarly, German auto major Volkswagen was among the first to enter the Chinese market and had a market share of 50% in the early 1990's (Tang, R., 2009). A few years ago, this came down to 18%, but the company figures out that it's still a healthy market share in what is soon going to be the world's largest auto market with 60 brands, compared with 37 in the United States.

Albeit the aforementioned facts, the economies of India and China have grown rapidly over the past couple of decades, and it is widely accepted that these two emerging giants will transform the global economy in numerous ways over the coming decades. Despite the importance of these countries, their strengths and weaknesses, the sources of their growth, and the missing ingredients to sustain high growth rates are not widely known.

Indian Economy

Since liberalization began in the 1980s, GDP growth has surged in India. The Elephant metaphor did not reflect the recent speed of India's transformation, which has been more like a tiger. Since 2003–2007, GDP growth has averaged 8.6% (NRC, 2010). India's growth would continue and increase in the coming decade if economic reforms continue, and are expanded, and large-scale structural changes are undertaken to support growth (IBEF, 2017). Exports have doubled in three years, and software exports doubled in the last two years. The exports-to-GDP ratio is "extremely low," he said, even though huge increases in foreign investment—over $21 billion—is comparable to that seen in China. India can adapt quickly, as evidenced by India's telecommunications revolution. From 5 million telephone lines in 1991, India now has over 200 million lines. India's demography will very likely help sustain this growth. India's population is younger than China's and exhibiting a rising rate of personal savings.

Chinese Economy

There are also counter views that China and India are not comparably sized global giants. China's trade is six times larger than India's (Gathani, B., 2004). Even more striking is the fact that the increase in China's trade level in 2007 ($433 billion, valued using MER) was greater than India's total trade. India's share of the global economy today is still less than half of what it was at independence in 1948. India's economy is expanding rapidly; but, its trade is still less than 1% of the global total; whereas, China's trade is the second or third largest (TNAP, 2004). A similar disparity exists in foreign investment. For these reasons, Lardy expressed more optimism about China's growth than about India's. The competitive environment in China is more favourable and intense than it is in India, where certain sectors are protected from import competition. In China, with reduced tariffs, domestic firms face competition not just from foreign imports but from foreign firms operating in China. China spends three times as much on infrastructure as India.

China's main challenge is to rebalance its growth strategy, moving towards one that relies more on domestic demand and less on exports. Currently, household consumption is only 36% of GDP; whereas, in India, that figure is 50–60% (CRISIL, 2011). For sustained economic development, India needs more manufacturing, a more liberalized trade environment, and more flexible labour markets.

The conventional wisdom is: "India does software; China does hardware. Those are their paths to expansion." But China's hardware exports are growing much faster than India's software exports, which make up less than 5% of India's GDP. India will need to take advantage of relatively low wage rates to build up its labour-intensive manufacturing sectors.

Comparing the Two Countries

Total Factor Productivity (TFP) growth rates are important. Capital deepening—that is, an increase in capital intensity, usually measured as capital stock per labour hour—also plays a dramatic role in growth,

especially in China, and is the "major explanatory factor" in the differences between the two countries' per capita annual growth. India averaged 4.8% between 2000 and 2005, about half of China's 8.1% annual per capita GDP growth rate. This difference is also seen in the R&D expenditure differences: R&D intensity in India is less than 1%; in China, it is 1.4% (Angang, H., et al., 2003).

India has competitive costs and wage levels, but it needs large-scale firms to compete successfully. Labour market restrictions in India are that country's greatest challenge. At the state level, though, India is deregulating and making labour markets more flexible. In China, where private firms are more productive than public firms, there is a great need to extend privatization (The Economist, 2017). China is restructuring rapidly and deepening regional specializations. India's financial markets are more developed than China's but India has a greater need to reduce regulatory restrictions in financial product markets (Merrill, S., Taylor, D., & Poole, R., 2010).

Review of Literature

Rossi, B. (2005) applied newly developed tests for nested models that are robust to the presence of parameter instability. The empirical evidence shows that for some countries we can reject the hypothesis that exchange rates are random walks. This raises the possibility that economic models were previously rejected not because the fundamentals are completely unrelated to exchange rate fluctuations, but because the relationship is unstable over time and, thus, difficult to capture by Granger Causality tests or by forecast comparisons. The authors also analysed forecasts that exploit the time variation in the parameters and find that, in some cases, they can improve over the random walk.

Rossi, B. (2005) did certain study on foreign exchange markets and its speculation. The study finds that the fluctuations in oil/commodity prices have been considered; however, they do not fully explain the volatility. The paper considers various economic indicators, selects the most impactful, and finalizes a model. The variables with the most impact come down to unemployment rate, oil price, consumer confidence index, and wheat prices. Ultimately, through in and out-of-sample forecasting, the authors could find that the model is not fully able to or accurately forecast short-term fluctuations, and yet does well in expressing long-term forecasts.

Zhang, S., et al. (2007) examines the monetary model of exchange rate determination for the US dollar exchange rates against the currencies of Canada, Japan, and the United Kingdom. The study utilized the co-integration technique for testing long-run relationship, and vector error correction model for short-run dynamics and out-of-sample forecasting. The existence of co-integration supports the long-run relationship among nominal exchange rate and a number of fundamental variables. The out-of-sample forecasting indicates that the nominal exchange rate forecasts from the VEC monetary model can be superior to random-walk based forecasts in a projection period of less than one year. This conclusion implies that the monetary model of exchange rate determination is a reliable tool for policy makers to evaluate their currency and the monetary authority should expect a much shortened response time to the monetary policy impulse in the surging trend of international economic integration.

Lam, L., et al. (2008) compares the forecast performance of the Purchasing Power Parity model, Uncovered Interest Rate Parity model, Sticky Price Monetary model, the model based on the Bayesian Model Averaging technique, and a combined forecast of all the above models with benchmarks given by the random-walk model and the historical average return. Empirical results suggest that the combined forecast outperforms the benchmarks and generally yields better results than relying on a single model.

Irena, M., Andrius, B., (2013) presents the main fundamental exchange rate forecasting models and discusses the advantages and drawbacks of the mentioned models. The research should help to explain why the forecasts can be not accurate.

Research Methods

The research is basically an exploratory study. The aim of the research is to perform prediction on foreign exchange rates for both India and China. So, besides from being exploratory, the research also falls in the ambit of comparative analysis as in methodology. India and China thought to be racing in the global economy and much of the Asian wealth depends on these two giants. There is more description given in the Introduction vide comparison.

Objectives

The aim of the study is to understand the current status of foreign exchange markets for both India and China. However, since this is a scholarly article and gives more emphasis on certain vigorous research methods. The study uses certain valid forecasting techniques like mean forecasting method in comparison with Random (Walk) Forests. These methods serve as mechanisms for forecasting foreign exchange rates for study sample, i.e., India and China. So, coming to the objectives, the following statements serve as objectives to the study:

- To study and evaluate if there exists any random behaviour of foreign exchange rates of India.
- To study and evaluate if there exists any random behaviour of foreign exchange rates of China.
- To forecast foreign exchange rates and understand the foreign exchange markets of India and China.
- To evaluate linear forecasting models such as random walk and its applicability on foreign exchange rates.

Hypotheses

As the aim of the study is to understand the performance of these two countries with respect to foreign exchange, the following statements deserve to be hypotheses for the study:

H_1: There exist significant differences in performance of currency trade between India and China in past and at present.

H_2: There will be significant differences in performance of foreign exchange trade between India and China.

Data

The data used for this study are secondary and they are retrieved from World Bank repository. World Bank serves abundant of economic yet financial data for variegated needs of academic and research. The data for this study were retrieved from http://data.worldbank.org/indicator/PA.NUS.FCRF?locations=IN. Table 1 shows the first few records of study data.

Table 1: First Few Records of Study Data

Country Name	*China*	*India*	*Country Name*	*China*	*India*
Country Code	*CHN*	*IND*	*Country Code*	*CHN*	*IND*
Indicator Name	*Official exchange rate (LCU per US$, period average)*	*Official exchange rate (LCU per US$, period average)*	*Indicator Name*	*Official exchange rate (LCU per US$, period average)*	*Official exchange rate (LCU per US$, period average)*
1960	2.461809455	4.761900004	2011	6.461461327	46.67046667
1961	2.461809455	4.761900004	2012	6.312332827	53.43723333
1962	2.461809455	4.761900004	2013	6.195758346	58.59784542
1963	2.461809455	4.761900004	2014	6.143434094	61.02951446
1964	2.461809455	4.761900004	2015	6.227488673	64.15194446
1965	2.461809455	4.761900004	2016	6.644477829	67.19531281

The above data show the foreign exchange value for both India and China. The dataset is 61 × 3 order data matrix. All rows represent years and columns represent year-wise foreign exchange rates for both India and China. The base currency for both exchange rates is USD. For instance, the foreign exchange rate for India for 2016 is 67.19 which means one USD is equal to approximately 67 Indian Rupees.

Research Methods

This study used two important theoretical models of forecasting. The first is forecasting by means method (Means Forecast) and the second is forecasting by random walk (RW) method. The first method, i.e., means forecasting assumes that the underlying data distribution

is identically independent distribution. As such, the means method is defined as follows:

$$Y_t = \mu + Z_t$$

where Y_t is the actual data and μ is the mean of the distribution Z_t is the error term, which is believed to be normally distributed. The forecasting function is given by,

$$Y_{n+h} = \mu$$

where is sample estimate of distribution mean.

The concept of random walk belongs to a machine learning methodologies known as random forest. Random forests or random decision forests are an ensemble learning method for classification, regression, and other tasks, which operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set. So, there will be two sets while fitting the actual data through random forests or walk. The first one is the training (set) data and the other is test (set) data. Usually, it is first the training which is performed over a subspace and later the prediction will be verified with the other subspace of the same set.[1] The random walk with drift model is,

$$Y_t = C + Y_{t-1} + Z_t$$

where Z_t is a normal *iid* error. Forecasts are given by,

$$Y_{n+h} = Ch + Y_n$$

If there is no drift, the drift parameter c=0. Forecast standard errors allow for uncertainty in estimating the drift parameter. Interestingly, there exists a close analogy between ARIMA and Random Walk. A general model can be given as,

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{p-t+1} + \epsilon_t$$

where α is constant and rest of the parameters are coefficients. If the model is brought under constraints that α, β are zero, then this general model turned into random walk and the second constraint, i.e., β =0 makes the model random walk with a drift.

Statistical Tools

MS Excel together with R is used for this study. Excel was used only to prune or cure the data *ala* editing cases and fields, merging rows and columns etc. R is one of the programming language and also known as *lingua franca* of statistics. R has wonderful mechanisms to perform all sorts of analyses and its capabilities can be extended to publishing quality visuals. There are approximately 600 packages for all variegated needs of statistical analyses. R is used to perform forecasting on sample data. All the analysis was done using one of the renowned and widely used R package known as forecast. This package has many mechanisms to perform time series analysis on survey or sampled data. Both means forecasting and random walk methods were performed by using forecast package in R.

Analysis

The analysis was done by three main tasks. *First*, importing the datasets to R as in the form of CSV files. *Second*, performing modeling, as in both training and testing, on the sampled data. *Third*, forecasting and plotting. The data are time-bound data which means for R it is time series data. The very first column of the dataset is time variable. R has automatic conversions to deal with time-bound data. The very first method in the analysis is visual investigation. The following is the time series graph which provides certain insights into the dataset.

Fig. 1: (a) is time series plot for China's foreign exchange rate. Figure 1 (b) is time series plot for India's foreign exchange rate. The plots also have graphs for both autocorrelation function (ACF) and partial autocorrelation functions (PACF). It is not very obvious that there is any visible speculative trend. However, there are certain ups and troughs; as such, it might be premature to infer that there exist seasonal changes. The data need to be tested for stationarity and seasonality. However, a careful observation on ACF plots gives us a valuable insights that both datasets are roughly random. The following is the R Output for Augmented Dickey Fuller Tests on sample datasets.

[1] In machine learning, the random subspace method, also called attribute bagging or feature bagging, is an ensemble learning method that attempts to reduce the correlation between estimators in an ensemble by training them on random samples of features instead of the entire feature set.

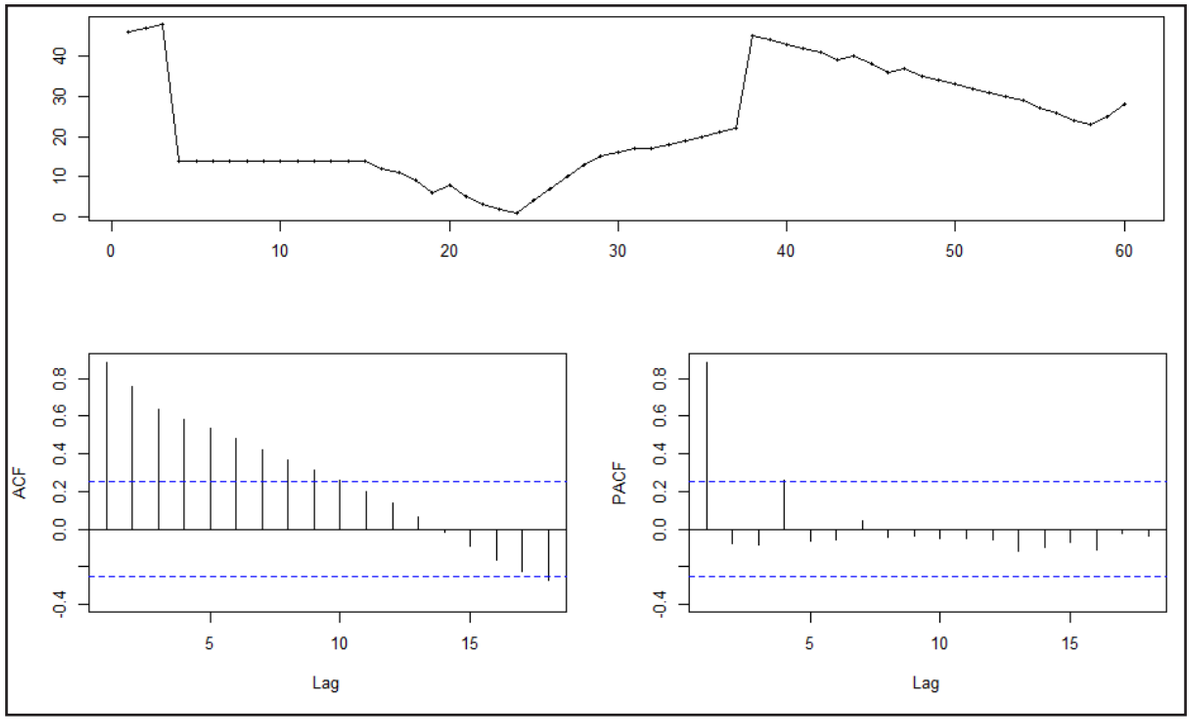


Fig. 1: (a) Time Series Plot for China Foreign Exchange Rate

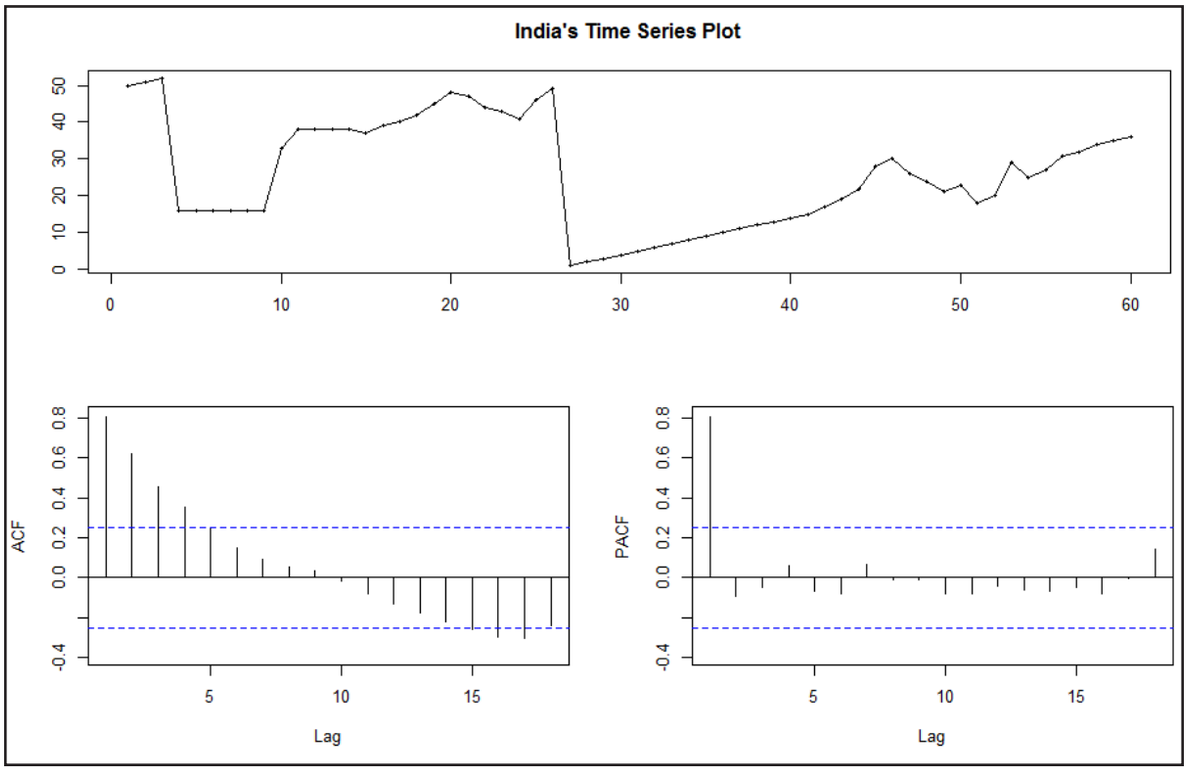


Fig. 1: (b) Time Series Plot for India Foreign Exchange Rate

R Output 1: Tests for Stochasticity.

```
> adf.test(feds[,2])

        Augmented Dickey-Fuller Test

data:  feds[, 2]
Dickey-Fuller = -1.5061, Lag order = 3,
p-value = 0.7744
alternative hypothesis: stationary

> adf.test(feds[,3])

        Augmented Dickey-Fuller Test

data:  feds[, 3]
Dickey-Fuller = -2.0542, Lag order = 3,
p-value = 0.5529

alternative hypothesis: stationary
```

The above output was taken from R for ADF test on time series data of China and India. In both cases the *p* value observed as greater than 0.05. So, at 5% significance level, it is not possible to reject the null hypothesis that the data are not stationary. So, it is clear that the data are stochastic. This gives a notion that the random walk method might be suitable for prediction. The following is the test for the seasonality.

R Output 2: Tests for Seasonality.

```
> PP.test(as.numeric(feds[,2]))

      Phillips-Perron Unit Root Test

data:  as.numeric(feds[, 2])

Dickey-Fuller = -3.0176, Truncation lag
parameter = 3, p-value = 0.1635

> PP.test(as.numeric(feds[,3]))

      Phillips-Perron Unit Root Test

data:  as.numeric(feds[, 3])

Dickey-Fuller = -2.4712, Truncation lag
parameter = 3, p-value = 0.3843
```

There is a wide variety of tests to check the significance of seasonality of time series data.[2] The above R Output shows the Phillips-Perron Test for autocorrelations. The *p* value is found to be sufficiently greater than 0.05; so, it is not possible to reject null hypothesis that the data under study are sufficiently random. Phillips-Perron Test tests only autocorrelations. So, any decision regarding seasonality is merely based on correlation but not on error variable. It might be better to recheck or reevaluate this decision through any other alternative mechanisms. The forecast package of R has different means to test seasonality of time series data (Hyndman R.J., 2017 & Hyndman R. J., and Khandakar Y., 2008). [3]

R Output 3: Seasonality check for foreign exchange rates for China and India.

```
> library(forecast)

> fit <- tbats(as.numeric(feds[,2]))

> seasonal <- !is.null(fit$seasonal)

> seasonal

[1] FALSE

> fit <- tbats(as.numeric(feds[,3]))

> seasonal <- !is.null(fit$seasonal)

> seasonal
> fit <- ets(feds[,3])
> ll <- logLik(fit)
> fit$loglik
```

[2] One of the other ways in testing significance of seasonality is through log-likelihood. The log-likelihood as a ratio follows chi-square distribution. Time series data when fitted give log-likelihood of

[3] The package `forecast` has sufficient methods to take care of seasonality. Hyndman RJ, *et al.* used exponential smoothing state space model with Box-Cox transformation along with ARMA errors. One way could be through log-likelihood. Both methods `ETS` and `TBATS` in `fma` and `forecast` yield log-likelihood. It might be possible to compute chi-square statistic through log-likelihood value based on which the decision can be taken that whether the underlying distribution has seasonality or not. For instance, for Indian time series data can be as follows:
The *p* value is 1 which exactly means there isn't any whit of seasonality in the Chinese time series data. However, a careful observation could reveal that the Pillips-Perron Test looks to be better than Chi-square test. One of the important weaknesses in Chi-square test is that actually, it is omnibus test, which means the statistic will be testing the sample size instead of sample attribute.

```
[1] -225.2308
> ll
'log Lik.' -225.2308 (df=3)
> 1-pchisq(ll, 3)
'log Lik.' 1 (df=3)
```

[1] FALSE

The seasonal component is absent in both the data variables, i.e., China and India. So, now it seems plausible to conclude that the data are sufficiently random for prediction. Now that it is known that the data are sufficiently random, it might be possible to predict through random walk method.

Prediction Through Random Walk Method

By the definition, random walk is machine learning method which needs two inputs - one, training set and the other, test set. As it was mentioned in the research methods, the entire dataset will be divided into two subsets, one for the training and the other for testing. Such mechanisms are not so naïve in the area of machine learning. The study dataset has 60 rows (records) and 3 columns (fields).

```
> dim(feds)
[1] 60  3
```

So, it is quite rational to make two subsets each with 30 records keeping fields as it is in the dataset. The random walk method is implemented on the first 30 records of the study data. The following is the procedure to perform random walk method in R.

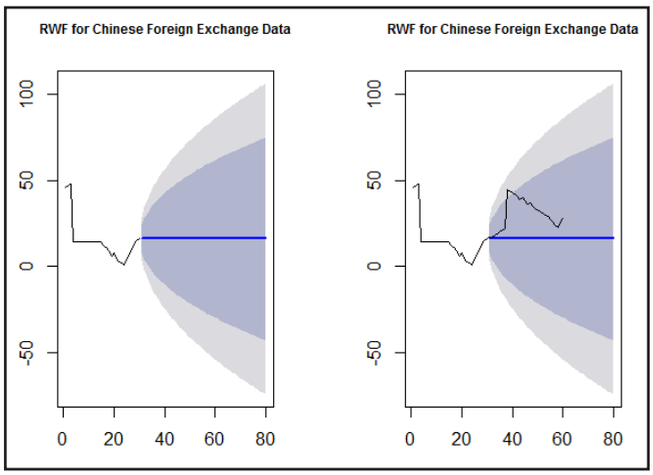


Fig. 2 (a): Random Walk Forecasting Plot for Chinese Foreign Exchange Data.

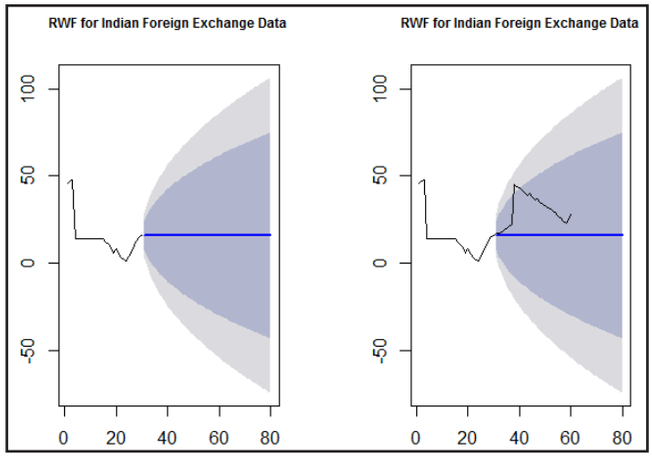


Fig. 2: (b) Random Walk Forecasting Plot for Indian Foreign Exchange Data

The following are the fitted and residuals of the RWF model in R.

Table 2: Fitted Values and Residuals for Random Walk Forecasting for Chinese Foreign Exchange Rate Data

```
> cbind(fit$fitted, fit$residuals, fit2$fitted, fit2$residuals)
Time Series:
Start = 1
End = 30
Frequency = 1
     fit$fitted  fit$residuals  fit2$fitted  fit2$residuals
1    NA          NA             NA           NA
2    46          1              50           1
3    47          1              51           1
4    48          -34            52           -36
5    14          0              16           0
6    14          0              16           0
7    14          0              16           0
8    14          0              16           0
9    14          0              16           0
```

10	14	0	16	17
11	14	0	33	5
12	14	0	38	0
13	14	0	38	0
14	14	0	38	0
15	14	0	38	-1
16	14	-2	37	2
17	12	-1	39	1
18	11	-2	40	2
19	9	-3	42	3
20	6	2	45	3
21	8	-3	48	-1
22	5	-2	47	-3
23	3	-1	44	-1
24	2	-1	43	-2
25	1	3	41	5
26	4	3	46	3
27	7	3	49	-48
28	10	3	1	1
29	13	2	2	1
30	15	1	3	1

Figure 2 (a) shows both training and actual series for random walk forecasting on Chinese foreign exchange data. Figure 2 (b) shows both training and actual series for random walk forecasting on Indian foreign exchange data. From both the figures, it seems that the forecasting is same for both countries. It makes little confusion that whether it is due to the fault of the model or method. Table 1 makes it clear about fitted values and residuals for the both countries. From the table, it very clear that the values are not same. For instance, Table 3 shows the actual data; from the data, it is clear that the actual values for China and India are 1.49 and 7.86, respectively, for the year 1980. And from Table 2 it is clear that the fitted value is 2 (1.49) and 43 (7.86) with their respective residuals, i.e. –1 and –2, respectively. So, the model values are different.

Table 3: Tail Part of the Study Dataset

```
> feds[24:30, ]
Country.Name       China       India
24         1980 1.498399999 7.862944701
25         1981 1.704533333 8.658522817
26         1982 1.892541666 9.455131933
27         1983 1.975674999 10.09889824
28         1984 2.320041666 11.36258333
29         1985 2.936658333    12.36875
30         1986 3.452791667 12.61083333
```

The following are the accuracy measures for both China and Indian foreign exchange rate data.

Table 4: Accuracy Measures for Both China and India RWF Method

```
> accuracy(fit1)
                    ME      RMSE       MAE       MPE     MAPE      MASE        ACF1
Training set 0.2241903 0.6555713 0.4936788 -95.06225 143.5371 0.7063767 0.003487063
> accuracy(fit2)
                     ME     RMSE      MAE       MPE     MAPE MASE       ACF1
Training set -1.586207 11.73853 4.758621 -166.3677 181.4842    1 -0.0488535
```

It shows that the RWF method is better for Chinese foreign exchange rate compared with that of India. RMSE for China is lesser than that of India, and it is true for rest of the other measures. The accuracy measures for China show that the RWF prediction works better for Chinese foreign exchange trade compared with that of India.

CONCLUSION

Indian and China are two robust economies in Asia-Pacific region. There is lot of literature that these two markets have abundant of influence on trade and economy of the region. The aim of this paper is to study the foreign exchange trade of India in contrast to China. The study assumes that the foreign exchange trade in India and China is not significantly different. The analysis was performed on secondary datasets of foreign exchange rates for both countries using random walk forecasting methodology. The analysis shows that though there isn't much difference in foreign exchange status for present and past for both India and China, the analysis shows that the predictions through methodology, i.e., random walk forecasting are not same. While random walk method could do well for China, is failed for India due to unknown reasons. Hence, though there isn't evidence in support of first hypothesis, i.e., the differences are not significant with respect to present and past performance, the same observation can't be applied for the other hypothesis. The study could find that the differences in performance with respect to foreign exchange could be significant in future. So, the study concludes that though the foreign exchange rates could appear to be governed by certain random processes, but there exist differences for predictions. More research is required to know the reasons as to why different methods are needed for prediction in spite of the similarities in the data.

Acknowledgements and Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

The Economist. (2017). India Foreign Exchange Reserves 1998-2017. Retrieved from https://tradingeconomics.com/india/foreign-exchange-reserves

Bloomberg, (April 07, 2017). Why India could be the winner of a US-China trade war. Retrieved from https://economictimes.indiatimes.com/news/economy/foreign-trade/why-india-could-be-the-winner-of-a-us-china-trade-war/articleshow/58059316.cms

Dar, B. A., & Ahmad, S. (2014). Major bilateral issues between China and India. *Arts Social Sci J 5*(64). doi: 10.4172/2151-6200.1000064

Sangani, P. (Feb 08, 2008). India, China to impact global economy. Retrieved from http://economictimes.indiatimes.com/articleshow/2765861.cms

Tang, R., (2009). The Rise of China's Auto Industry and Its Impact on the U.S. Motor Vehicle Industry. CRS Report for Congress. Retrieved from https://fas.org/sgp/crs/row/R40924.pdf

NRC. (2010). *The Dragon and the Elephant: Understanding the development of innovation.* National Academy of Sciences. ISBN: 978-0-309-15160-3. Pg. No. 6.

IBEF, (2017). About Indian Economy Growth Rate & Statistics. Retrieved from https://www.ibef.org/economy/indian-economy-overview

Gathani, B. (2004). People of Indian origin to play larger role in development. Retrieved from http://www.thehindubusinessline.com/2004/10/08/stories/2004100800540900.htm

TNAP. (2004). The Dragon and the Elephant: Understanding the Development of Innovation Capacity in China and India. Retrieved from https://www.nap.edu/catalog/12873/the-dragon-and-the-elephant-understanding-the-development-of-innovation

CRISIL. (2011). Raising the growth bar. Retrieved from https://www.crisil.com/pdf/corporate/india-raising-the-bar.pdf

Baldacci, E., Callegari, G., Coady, D., Ding, D., Kumar, M., Tommasino, P., & Woo, J. (2010). Public Expenditures on Social Programs and Household Consumption in China. International Monetary Fund Working Paper. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.167.6981&rep=rep1&type=pdf

Angang, H., Linlin, H., & Zhixiao, C. (2003). China's economic growth and poverty reduction (1978-2002). Retrieved from https://www.imf.org/external/np/apd/seminars/2003/newdelhi/angang.pdf

The Economist. (2017). Reform of China's ailing state-owned firms is emboldening them. Retrieved from https://www.economist.com/news/finance-and-

economics/21725293-outperformed-private-firms-they-are-no-longer-shrinking-share-overall

Merrill, S., Taylor, D., & Poole, R., (2010). *The Dragon and the Elephant: Understanding the development of innovation capacity in China and India.* The National Academies Press.

Rossi, B. (2005). Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.497.1697&rep=rep1&type=pdf

Palombizio, E., & Morris, I. (2012). Forecasting Exchange Rates using Leading Economic Indicators. Palombizio and Morris, 1:8. Retrieved from http://dx.doi.org/10.4172/scientificreports.402

Zhang, S., & Lowinger, T. C. (2007). The monetary exchange rate model: Long-run, short-run, and forecasting performance. *Journal of Economic Integration 22*(2), 397-406

Lam, L., Fung, L., & Yu, I.-W. (2008). Comparing Forecast Performance of Exchange Rate Models. Working Paper 08/2008. June 2008. Retrieved from http://www.hkma.gov.hk/media/eng/publication-and-research/research/working-papers/HKMAWP08_08_full.pdf

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The elements of statistical learning* (2nd ed.). Springer. ISBN 0-387-95284-5.

Hyndman, R. J. (2017). Forecast: Forecasting functions for time series and linear models. R package version 8.1. Retrieved from http://github.com/robjhyndman/forecast.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 26*(3), 1-22. Retrieved from http://www.jstatsoft.org/article/view/v027i03

Guidelines for Authors

International Journal of Business Analytics and Intelligence welcomes original manuscripts from academic researchers and business practitioners on the topics related to descriptive, predictive and prescriptive analytics in business. The authors are also encouraged to submit perspectives and commentaries on business analytics, cases on managerial applications of analytics, book reviews, published-research paper reviews and analytics software reviews based on below mentioned guidelines:

Journal follows online submission for peer review process. Authors are required to submit manuscript online at http://manuscript.publishingindia.com

Title: Title should not exceed more than 12 Words

Abstract: The abstract should be limited to 150 to 250 words. It should state research objective(s), research methods used, findings, managerial implications and original contribution to the existing body of knowledge

Keywords: Includes 4–8 primary keywords which represent the topic of the manuscript

Main Text: Text should be within 4000-7000 words Authors' identifying information should not appear anywhere within the main document file. Please do not add any headers/footers on each page except page number. Headings should be text only (not numbered).

Primary Heading: Centered, capitalized, and italicized.

Seconday Heading: Left justified with title-style capitalization (first letter of each word) and italicized.

Tertiary Heading: Left justified and indented with sentence-style capitalization (first word only) in italics.

Equations: Equations should be centered on the page. If equations are numbered, type the number in parentheses flush with the left margin. Please avoid using Equation Editor for simple in-line mathematical copy, symbols, and equations. Type these in Word instead, using the "Symbol" function when necessary.

References: References begin on a separate page at the end of paper and arranged alphabetically by the first author's last name. Only references cited within the text are included. The list should include only work the author/s has cited. The authors should strictly follow APA style developed by American Psychological Association available at American Psychological Association. (2009). Publication manual of the American Psychological Association (6th Ed.). Washington, DC.

Style Check

To make the copyediting process more efficient, we ask that you please make sure your manuscript conforms to the following style points:

Make sure the text throughout the paper is 12-point font, double-spaced. This also applies to references.

Do not italicize equations, Greek characters, R-square, and so forth. Italics are only used on p-values.

Do not use Equation Editor for simple math functions, Greek characters, etc. Instead, use the Symbol font for special characters.

Place tables and figures within the text with titles above the tables and figures. Do not place them sequentially at the end of the text. Tables and figures must also be provided in their original format.

Use of footnotes is not allowed; please include all information in the body of the text.

Permissions

Prior to article submission, authors should obtain all permissions to use any content if it is not originally created by them.

When reproducing tables, figures or excerpts from another source, it is expected to obtain the necessary written permission in advance from any third party owners of copyright for the use in print and electronic formats. Authors should not assume that any content which is freely available on the web is free to use. Website should be checked for details of copyright holder(s) to seek permission for resuing the web content

Review Process

Each submitted manuscript is reviewed first by the chief editor and, if it is found relevant to the scope of the journal, editor sends it two independent referees for double blind peer review process. After review, the manuscript will be sent back to authors for minor or major revisions. The final decision about publication of manuscript will be a collective decision based on the recommendations of reviewers and editorial board members

Online Submission Process

Journal follows online submission for peer review process. Authors are required to register themselves at http://manuscript. publishingindia.com prior to submitting the manuscript. This will help authors in keeping track of their submitted research work. Steps for submission is as follows:

1. Log-on to above mentioned URL and register yourself with "International Journal of Business Analytics & Information"
2. Do remember to select yourself as "Author" at the bottom of registration page before submitting.
3. Once registered, log on with your selected Username and Password.
4. Click "New submission" from your account and follow the 5 step submission process.
5. Main document will be uploaded at step 2. Author and Co-author(s) names and affiliation can be mentioned at step 3. Any other file can be uploaded at step 4 of submission process.

International Journal of Business Analytics and Intelligence

SUBSCRIPTION DETAILS

Dispatch Address:-
The Manager,
International Journal of Business Analytics and Intelligence
Plot No-56, 1st Floor
Deepali Enclave, Pitampura
New Delhi -110034
Ph - 9899775880

Subscription Amount for Year 2018

	Print	Print + Online
Indian Region	Rs 2700	Rs 3500
International	USD 150	USD 180
Price mentioned is for Academic Institutions & Individual. Pricing for Corporate available on request. Price is Subject to change without prior notice.		

Payment can be made through D.D./at par cheque in favour of "Publishing India Group" payable at New Delhi and send to above mentioned address.

Disclaimer

The views expressed in the Journal are of authors. Publisher, Editor or Editorial Team cannot be held responsible for errors or any consequences arising from the use of Information contained herein. While care has been taken to ensure the authenticity of the published material, still publisher accepts no responsibility for their inaccuracy.

Journal Printed at Anvi Composers, Paschim Vihar.

Copyright

Printed and published by Publishing India Group, New Delhi. Any views, comments or suggestions can be addressed to – Coordinator, IJBAI, info@publishingindia.com

www.manuscript.publishingindia.com

PublishingIndia.com

Publishing India Group

Plot No. 56, 1st Floor, Deepali Enclave
Pitampura, New Delhi-110034, India
Tel.: 011-47044510, 011-28082485
Email: info@publishingindia.com
Website: www.publishingindia.com

ISSN 2321-1857
01
9 772321 185001