

Indira Gandhi national Open University

Analysis of product development process in a machine learning firm

166545289 – Mathews Baby

TABLE OF CONTENTS

1. Introduction	4
1.1. Introduction to the project	5
1.2. Objectives	10
1.3. Scope of the study	10
1.4. Research methodology	11
1.5. Limitations of the study	13
2. Literature review	15
2.1. Profile of IT industry	17
2.2. Profile of AI and ML industry	21
2.2.1. Introduction for ML	21
2.2.2. Machine learning applications	25
2.2.3. Management aspect of ML product development	27
2.2.4. Technical aspect of ML product development	29
2.3. Profile of the organization	39
2.3.1. Introduction - organization	39
2.3.2. Applications	42
2.3.3. Domains	47
2.3.4. Product development division	49
2.3.5. Higher management team	56
2.3.6. Future endeavors	57
2.3.7. Sister companies	58

3. Data collection	63
3.1. Simple percentage analysis	64
3.2. Correlation matrixes	108
3.3. Gap analysis	123
4. Findings, Suggestions and conclusion	135
4.1. Findings	136
4.2. Suggestions	138
4.3. Conclusion	139
5. Bibliography	140
6. Appendix	142
6.1. Questionnaire	143
6.2. Glossary	150

CHAPTER 1

INTRODUCTION

World around us is moving so fast and I always wondered to see how things work so smoothly. I always had a thought like everything happens with the basic means of some rule based algorithms written in any of the 100s of programming languages like Cobol, Java, dot net and so on. But eventually I started realizing that none of this would have happened with only rule based systems being in place. Something extra intelligent is there around us which are effectively used for a lot of things in life to happen so easily.

Gradually my interest began to grow to know more on these intelligent systems which is left unknown for me. Then, I started to see and think the pieces I come across to relate and understand the astonishing power of these technologies. That is when my mind got stuck in some of the writings I read. One year back, in North India, when a destructive flood happened, it was the tremendous and accurate weather forecast that helped things to be moved back and protect the lives from the massive calamity. Weather forecast is something we always blame about, but I always wondered the tedious effort of evaluating several measures and dimensions, comparing to the geographical conditions, previous effects on the place, which all finally turns into the forecast we see. This was not something merely using the base technologies; it somewhere had an extra intelligence used.

Similarly, tech news on a new agenda by Amazon was another piece which attracted me. They were launching forecast delivery of items to our nearby courier point without our knowledge. They mine our search, all the data available on us to intelligently conclude on our coming decision of buying something in future. They forecast it accurately and deliver to the nearby point of our location,

so that when we plan to search on this to buy, it attracts us easily with a short delivery time. What an intelligent use of technology to grow the business!

Slowly, my interest became a necessity for me. Consider the smart phones we have with us, they have become the most indispensable tech product that we own today and we use it almost all the time for meeting our each needs. When using a smart assistant application, whether it's Google Assistant, Alexa, Siri, or Bixby, you more or less know that these are not working on top of some simple technology. All these day to day devices we use, own an intelligent system which learn from its experience and give us back the better the next time. Apart from the photograph taken from a high spec camera, the same picture shot on a single lens mobile camera is getting more clarity and each time with more dimensions used. Moreover of using high specific lens as in a camera, the mobile devices make use of the intelligent technology to learn and adapt each time on the pictures taken. From the precise product recommendations that we see on Amazon, Flipkar, Paytm or Facebook pages to the self-intelligent cars; there are wide variety of products and applications which kept me wondering on the ability to effectively use the data within and mine it to be useful for each ones lives.

After doing a thorough study on these topics I realized that the system in place solving all of the above problems is Artificial intelligence - the biggest game changing technology we have now. Majority of people are not clear on what artificial intelligence or machine learning is and what their difference is. I have asked several of my friends or colleagues and found that most of them still

consider AI as a simple robotics which is not true. Out of my curiosity I put some effort to learn the basics of machine learning and AI technology.

Artificial Intelligence is the broader concept of machines being able to carry out tasks in a way that we would consider “smart”. In Generalized concept AI is systems or devices which can in theory handle any task – are less common, but this is where some of the most exciting advancements are happening today. It is also the area that has led to the development of Machine Learning. Often referred to as a subset of AI, it’s really more accurate to think of it as the current state-of-the-art.

Machine Learning is a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves. Machine learning (ML) is a category of an algorithm that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic premise of machine learning is to build algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available.

It is so clear that, all of our lives are impacted by artificial intelligence on a daily basis. Whether we are using our smartphones, surfing the internet, buying products online, using navigation, spending time on social media or listening to songs on our favorite music streaming service, AI is impacting our choices in one way or another. For most of us, it’s our mobile phones which help us do our jobs more efficiently and which generally making our lives easier. And the best

part – this is only the beginning. In fact, AI has the potential to increase business productivity for every business across every industry.

After learning the technical side, I started thinking about the management aspect of machine learning industry. To cope up with the customer demand in this field without losing the quality of service and product, it is very much compulsory to follow some different but efficient product development system which effectively uses the ability of machine learning.

The business sector has a critical role in meeting the aspirations of the society and increasing the satisfaction. These can be enabled by developing and successfully launching new products, which also results as an additional source of satisfaction. However, product development can be a costly and a time-consuming activity. It, therefore, carries risk as well as reward. To succeed, businesses must follow a structured development and launch process.

Monopoly is almost extinct in the prevailing market scenario and survival and growth depends on quality and value creation. This study is an attempt to study the development process of a product in a machine learning firm. The study critically analyses the importance of structured process and recommend methods to bring down the time and cost involved and enhancing the quality of the product and there by customer satisfaction.

It is only few companies that purely work in the machine learning domain in India. To do a study that conveniently I can finish up within this short span of time and resources, I choose Curvelogics advanced technology solutions Pvt. Ltd. Since the company has office at Trivandrum where I stay and already have previous contacts with the senior management team there, I thought it would be easier to conduct study at Curvelogics.

Curvelogics is a techno analytical firm focuses on Artificial Intelligence and Advanced Analytics related products and services through a professionally organized product development System. The company has released multiple products in areas ranging from Robotics, Technology Infrastructure for Sales, Marketing, Pricing, Campaign Management, Social Media so on and all the way up to Genomics.

The study intends to have an overall analysis of the standard of product development management system in Curvelogics advanced technology solutions Pvt. Ltd, so as to find out the significance of time and quality standards in the modern industry and its contribution to the overall growth of society and help to implement further improvisation and monitor its effectiveness in the overall context.

Objectives

The objectives of the study are as follow

- To evaluate the process development
- To study the process of generating ideas for the product
- To suggest improvements in processes
- To assess how employees view the processes

Each organization is started with an aim of growing and succeeding in their respective field. Apart from the vision and spark to start with an organization, proper thinking along with several other factors only would lead to the success. An organization grows effectively when the employees and employer starts focusing together on the management and functioning of each department. Instead of focusing on the end product alone, the management on the path to reach the same would have a great impact on the future of the organization and the employees. Thus, one and most important factor which leads to the success is the proper management and functioning of the product development department. This is the same reason being the spotlight of the project focusing on the functioning of the product development department of the organization giving emphasize to their functions.

The scope of the study includes analyzing the exiting functioning of the product management department, understanding the flaws, improvising the process and focus on areas to improve the efficiency of the same. With this, we expect the study to provide a proper guideline to raise the product development process

efficiency for the organization and thereby to the machine learning industry. The study coming up with an improvised proper process structure to be followed would help to improve the product quality as well, which in turn will benefit the organization. This therefore can help creating more products with advanced machine learning applications such as speech recognition, complex analytics and computer vision. The study would also focus on the employee perspective, rather than the traditional process following of focusing on the customer or employer perspective. To benefit to this, a questionnaire is circulated among the employees and their feedback have been taken into account for creating the guideline on the processing of product development department. Thus, the overall scope of the study would include all the above perspectives to build a guideline for improvising the product development to develop a more efficient product squeezing all the advanced applications in machine learning.

Data collection for the study done through a survey conducted among the employees. The questionnaire for the same prepared after interviewing the staff from the company and collecting secondary data. Thus prepared structured questionnaire distributed to selected employees constituting from different department, designation, age and gender. Tools like Psychologist Rensis Likert's five level likert scale have used in this method. Data thus collected is used as the primary data for analysis.

Secondary data are mainly collected from the reports and documents provided by the company as well as the literature content available in website and text books described in the reference part. Annual status report, effort analysis and feedback

available from earlier studies from the organization and related firms in the same industry also catered as a course for much of the secondary data.

Sampling method used for selecting the employees for the survey is simple random. Simple random is a subset of a statistical population in which each member of the subset has an equal probability of being chosen. Since the number of employees is very few, we choose hundred from them for our task constituting equal ratio of employees from different department, designation, age, gender and industry experience.

The raw data thus collected is presented mainly using simple table, bar diagram and pie chart. And later to interpret the data in a meaningful manner we have used tools like Simple Percentage Analysis, Correlation matrix and Gap analysis.

Simple percentage analysis is the method to represent raw streams of data as a percentage (a part in 100 per cent) for better understanding of collected data.

A correlation matrix is a table showing correlation coefficients between variables. Each cell in the table shows the correlation between two variables. Each random variable (X_i) in the table is correlated with each of the other values in the table (X_j). This allows you to see which pairs have the highest correlation.

A gap analysis is the strategic tool used for analyzing the gap between the target and anticipated results, by assessing the extent of the task and the ways, in which gap might be bridged. It involves making a comparison of the present

performance level of the entity or business unit with that of standard established previously.

There are many factors, which leads to the success of an organization. One and the most important among them are the proper management and functioning of the product development department. Hence the present study mainly focuses on the functioning of product development department giving emphasize to their functions. We expect the study can provide proper guidelines to raise the product development process efficiency for the company and machine learning industry itself. Since knowingly or unknowingly all of us are impacted with the advancement of this field, this can help creating more products with advanced machine learning applications such as speech recognition, complex analytics and computer vision and thus raise the value for the product by raising the comfort of the life of the end user. Also the study can give an idea on the view of the processes from the perspective of employees with the help of the questionnaire.

This industry is purely technology oriented, so to have a detailed study on their process, in-depth knowledge about technologies like machine learning, artificial intelligence and Internet of Things is very required. As we are not expert in these fields we have faced difficulties understanding the complex technical aspect of prototype development. This field is driven with the availability of data, a small data breach can impact severely the company's business. Hence all the internal data could not be analyzed as there are access issues related to confidentiality policies. Majority of the product and services offered by the company are of B2B nature. Because of the confidentiality agreement with the client companies, the

Analysis of product development process in a machine learning firm

management is reluctant much of the time revealing the details of financial deals and exact product and service details. This industry is very vast, comprised of big companies like Google , Facebook to small startups. To have a proper study on all of these it would require more time. Since the study has to be conducted in a short span of time, it was not possible to cover all areas.

CHAPTER 2

LITRATURE REVIEW

Information technology (IT) is an example of a general purpose technology that has the potential to play an important role in economic growth, as well as other dimensions of economic and social development. IT may have a special role to play in growth and development simply because of empirical characteristics that apply at the current time. In particular, the recent and continuing rapid innovation in IT make it a dynamic sector that is an attractive candidate as a contributor to growth for that reason alone, much as the automobile industry was targeted by the Japanese after World War II. On the other hand, there may be features of IT that make it attractive from a theoretical perspective on economic growth. For example, IT may be one of the sectors in which countries such as India have, or can develop, a comparative advantage. Even if this is so, IT is likely to share this characteristic with several other sectors.

A somewhat more special characteristic of IT may be that it is a 'general purpose technology' (GPT, Bresnahan and Trajtenberg, 1995), distinguished by pervasiveness, technological dynamism and innovational complementarities. In this case, IT is one of a special few technologies: other examples of GPTs include steam and electricity (both advances in power delivery systems) and synthetic materials. Finally, IT may be unique in its impact on growth. In this view, IT has a special role in the process of innovation, because it affects the rate at which potential new ideas are converted into additions to the usable stock of knowledge in ways that nothing else can. The formalization of this special role is based on the model of recombinant growth.

PROFILE OF IT INDUSTRY

Information Technology (IT) Industry in India has played a key role in putting India on the global map. IT industry in India has been one of the most significant growth contributors for the Indian economy. The industry has played a significant role in transforming India's image from a slow moving bureaucratic economy to a land of innovative entrepreneurs and a global players in providing world class technology solution and business services. The industry has helped India transform from a rural and agriculture-based economy to a knowledge based economy. Information Technology has made possible information access at gigabit speeds. It has made tremendous impact on the lives of millions of people who are poor, marginalized and living in rural and far flung topographies. Internet has made revolutionary changes with possibility of e-government measures like e-health, e-education, etc. Today , whether its filing Income tax returns or applying for passports online or railway e-ticketing, it just need few clicks of the mouse. India's IT potential is on a steady march towards global competitiveness, improving defense capabilities and meeting up energy and environment challenges amongst others. IT- ITeS and BPO sector in India, with the main focus on increasing technology adoption and developing platforms, has aggregated revenue of USD 175 billion in FY2019, while generating direct and indirect employment for over 17 million people. Out of 175 billion, exports revenue (including Hardware) has reached USD 110 billion in FY2019 while domestic revenues (including Hardware) of about USD 55 billion. The information Technology industry has gained a brand image knowledge economy due to its development from software exporters to providing IT –BPO sectors. IT industry has registered a notable growth because of rich and varied expansion into

verticals, well differentials service offering and increasing growth penetration. The phenomenal success of this industry is attributable to favorable government policies, rich and burgeoning demand conditions, healthy growth of the related industries and competitive environment prevalent in the industry. The interplay of these forces has put the industry on the global map.

India is the topmost offshoring destination for IT companies across the world. Having proven its capabilities in delivering both on-shore and off-shore services to global clients, emerging technologies now offer an entire new gaunt of opportunities for top IT firms in India. Social, Mobility, Analytics and Cloud (SMAC) are collectively expected to offer a US\$ 1 trillion opportunity. Cloud represents the largest opportunity under SMAC, increasing at a CAGR of approximately 30% to around US\$ 650-700 billion by 2020. The social media is the second most lucrative segment for IT firms, offering a US\$ 250 billion market opportunity by 2020. The Indian e-commerce segment is US\$ 12 billion in size and is witnessing strong growth and thereby offers another attractive avenue for IT companies to develop products and services to cater to the high growth consumer segment.

Future of IT industry in India IT sector alone has the potential to take Indian economy to a higher level. IT has the ability to increase efficiency and productivity of every sector of the country. The demand for IT sector is likely to increase due to the following reasons:

- Reduction in the price of semiconductor that in turn resulted in the reduction in price in electronic equipment enabled more and more people to buy computers and use it wherever possible.
- There has been significant rise in the number of internet user and the number is likely to increase in future. This happened due to the lowering of broadband services and availability of large number of portable services. The use of internet has definitely increased the number of online users of various services. And this will surely increase in future.
- Rapid urbanization of developing countries like India has made computer household equipment. As a result computer literacy has increased. Since a large number of people still live in rural areas, there's a high probability that the increase in the number of computer and internet user will increase exponentially.

But there are also some risks associated with this sector:

- There is uncertainty in global economy due to the slow recovery in many major markets. Moreover there is instability in major European nations. This could lead to cut in IT budget which will ultimately result in declining demand.
- Rules and regulations that are being implemented in major markets like US will surely lead to a demand compression. The VISA restrictions in US are in fact targeted at the IT industry. This step was taken to curb outsourcing and will directly affect the demand in future.
- Increased competition could result in price competition. Indian IT majors like TCS, Infosys, Wipro, CTS etc. are close substitutes of each other. So

there is increasing pressure on the IT majors to make their service different from others. This, however, will not affect the overall demand.

- Wage inflation and overall increase in cost could reduce the profit margin. Appreciating rupee will further add to this problem. To deal with this companies have to focus on increasing productivity.

Having said this, Artificial intelligence is strongly disrupting the whole IT industry in the last decade. Even though majority of the nature is similar to normal IT industry, there are things which are specific for AI. Most importantly the nature of product services are deterministic in normal IT product where as it is mostly probabilistic in the case of machine learning. The process is also thus different in both.

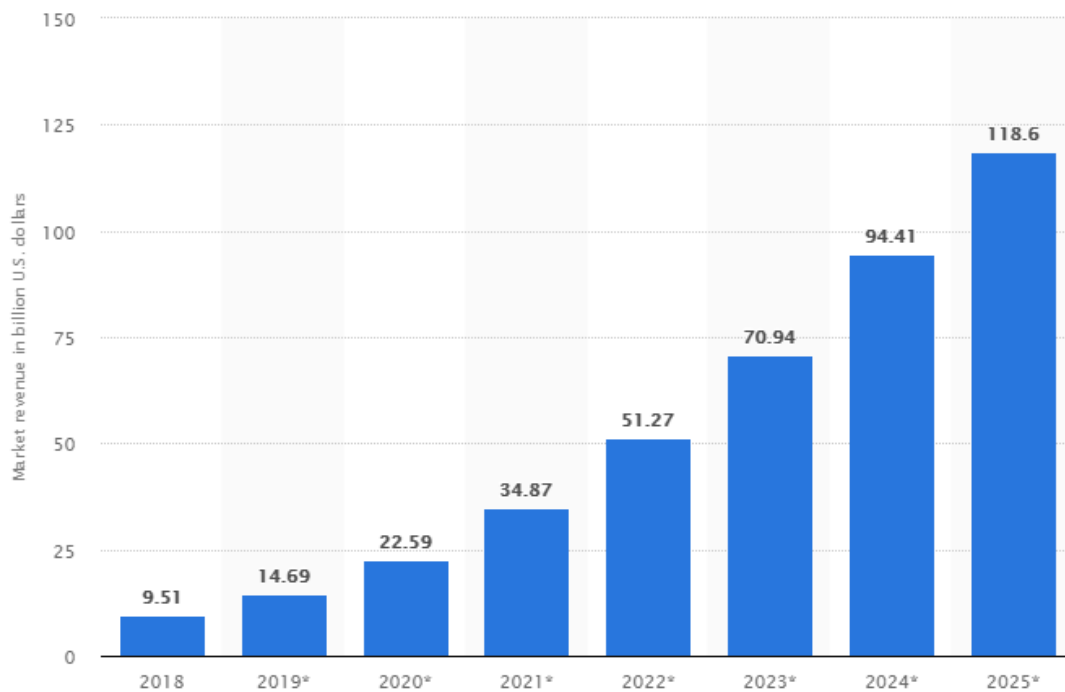
MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE INDUSTRY

Machine learning can be considered a part of AI, as most of what we imagine when we think about AI is machine-learning based. Machine learning is, at its core, the process of granting a machine or model access to data and letting it learn for itself.

This idea is relatively new. In the past, we believed robots would need to learn everything from us. But the human brain is sophisticated; not all of the actions and activities it coordinates can be easily described. In 1959, Arthur Samuel came up with the brilliant idea that we shouldn't have to teach computers, but rather, we could let them learn on their own. He coined the term "machine learning" to describe his theory, which is now a standard definition for the ability of computers to learn autonomously.

This industry is so vast with regards to its functionalities and application. Companies across different domains use machine learning to raise the efficiency in their functionality. According to the market research firm Tractica, the global artificial intelligence software market is expected to experience massive growth in the coming years, with revenues increasing from around 9.5 billion U.S. dollars in 2018 to an expected 118.6 billion by 2025. The overall AI market includes a wide array of applications such as natural language processing, robotic process automation, and machine learning.

Analysis of product development process in a machine learning firm

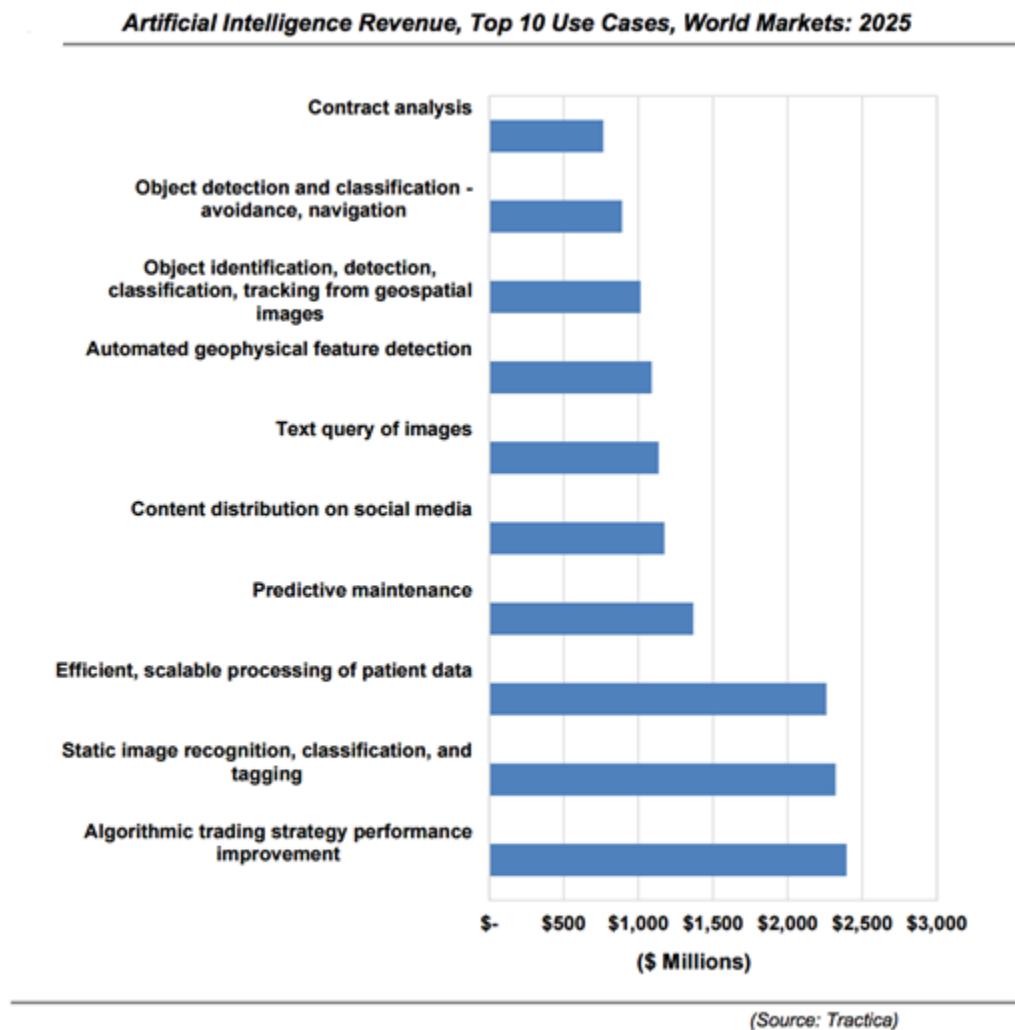


© Statista 2019

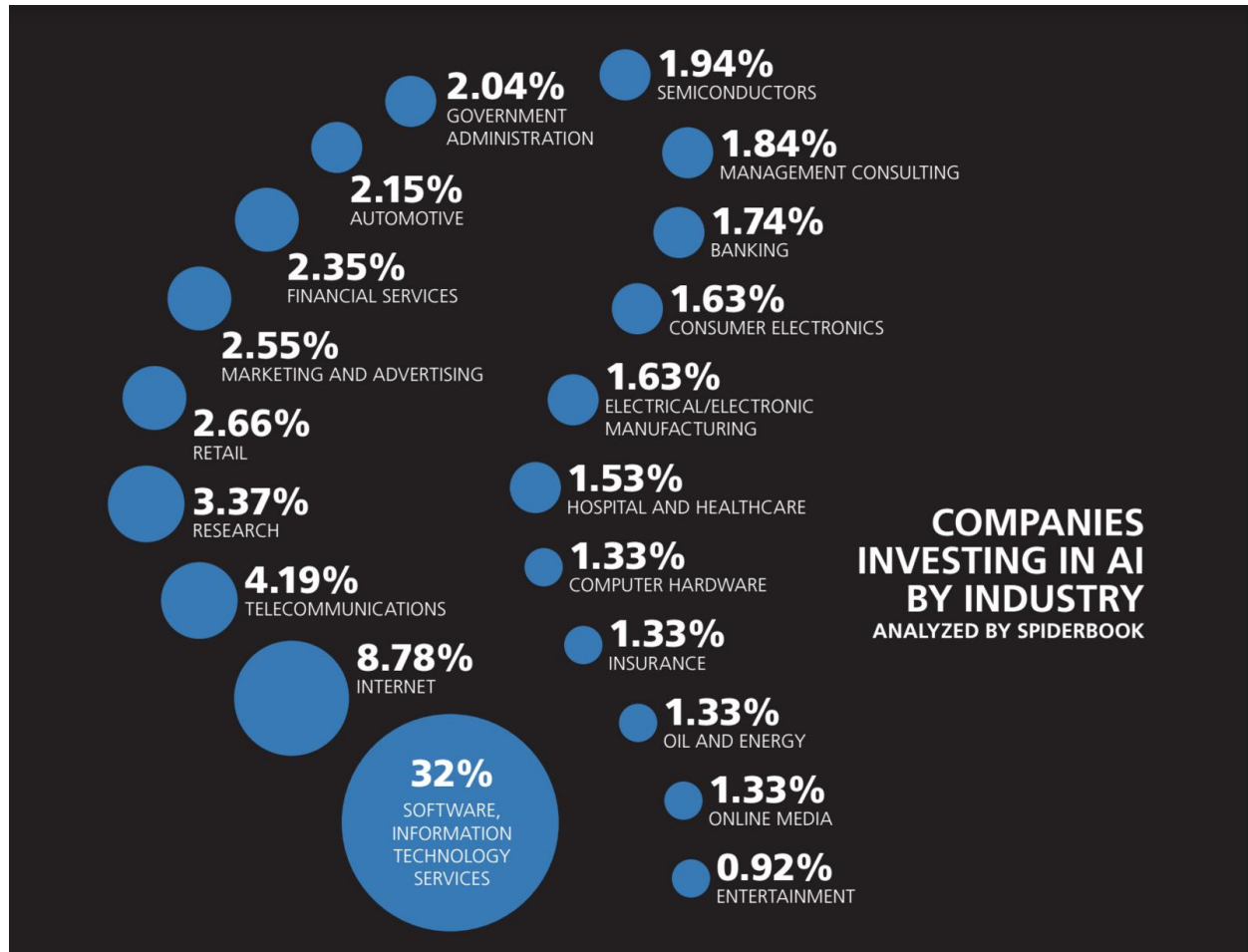
Revenues

from the artificial intelligence software market worldwide from 2018 to 2025

According to Tractica's study out of the top 10 use cases related to AI industry in 2015, 60% of the use cases are related to big data and 40% are related to image or object recognition, which is a good high-level characterization of how the market is expected to shape up.



There are only 1,500 companies in North America that are doing anything related to AI today, even using its narrow, task-based definition. That means less than one percent of all medium-to-large companies across all industries are adopting AI. A breakdown of AI adoption does not provide a fair picture of the current level of AI maturity in the market. The following figure by Spiderbook details how many companies are using AI technologies beyond lab experiments that is those developing applications based on it or deploying it across the company.



From the above study it is clear that how much diversified the industry is and what all domains start heavily invest in machine learning. Describing about any of the individual company is difficult as the application and processes vary tremendously across them. So following given details are from generalized across the industry.

Common Machine Learning Applications

- **Natural language processing:**

Google Translate is not just a dictionary, actually it created from a set of machine learning algorithms that updates the service over time based on input from users, like new words and syntax. Siri, Alexa, Cortana, and, most recently, Google Assistant all rely on natural language processing to recognize speech and synthesis, allowing them to understand or pronounce words they have never encountered before.



Source : [Oracle+datascience.com/ machine-learning-overview](https://oracle+datascience.com/machine-learning-overview)

- **Recommendation systems:**

On Netflix, Amazon, and Facebook, everything that is recommends to user depends on search activity, likes, and previous behavior. These websites deliver recommendations across platforms, devices, and apps.

Machines match sellers with buyers, movies with prospective viewers, photos with people who want to see them — all of which improves our lives and online experiences significantly. Amazon has such amazing machine learning algorithms in place that it can predict with high certainty what you'll buy and when you'll buy it. The company even owns a patent for “anticipatory shipping,” a system that ships a product to the nearest warehouse so you can order and receive your item on the same day.

- **Algorithmic trading:**

Algorithmic trading is a process that involves random behavior, ever-changing data, and a variety of factors — from political to judicial — that are far away from traditional finance. While financiers cannot predict much of that behavior, machine learning algorithms can — and they respond to changes in the market much faster than a human.

There are plenty of other business implementations of machine learning. You can predict if an employee will stay with your company or leave. You can decide if a customer is worth your time, if they'll likely buy from a competitor, or not buy at all. You can optimize processes, predict sales, and discover hidden opportunities.

Then we have autonomous vehicles. What was once merely a vision of science fiction is now a reality; millions of miles have already been driven by cars that don't require a human operator. This, once again, originated

from a set of machine learning algorithms that enabled cars to learn how to drive safely and effectively.

Management aspect of machine learning product development

Machine learning management requires a more experimental approach than traditional software engineering. It is more experimental because it is an approach that involves learning from data instead of programmatically following a set of human rules. Because the nature and approach of ML projects is more experimental, industry people and their companies won't know what will happen until they try it (i.e., more probabilistic rather than deterministic). This may present challenges for product management as many product managers (PM) have been trained on shipping projects with a deterministic approach.

PMs can leverage that intuition to calibrate the tradeoffs of various approaches given their company's data "and how it can be used to solve customer problems.

Majority of the whole product development process in a machine learning company is same as the software development life cycle (SDLC) practiced in all IT companies for the software development process.

SDLC consist of the following activities

1. Planning: The most important parts of software development, requirement gathering or requirement analysis are usually done by the most skilled and experienced software engineers in the organization. After the requirements

are gathered from the client, a scope document is created in which the scope of the project is determined and documented.

2. Implementation: The software engineers start writing the code according to the client's requirements.
3. Testing: This is the process of finding defects or bugs in the created software.
4. Documentation: Every step in the project is documented for future reference and for the improvement of the software in the development process. The design documentation may include writing the application programming interface (API).
5. Deployment and maintenance: The software is deployed after it has been approved for release.
6. Maintaining: Software maintenance is done for future reference. Software improvement and new requirements (change requests) can take longer than the time needed to create the initial development of the software.

In this project we are mainly studying about the Planning, Implementation and testing activities of a machine learning product. Planning face is pretty much same as the normal SDLC procedure whereas prototype development differs technically which is explained in the next section 'Technical aspect of machine learning'. Testing differs in the context that unlike normal software testing where we could segregate which are passed and failed test cases, here to make a decision whether a model works fine or not is difficult. It is due to the fact that machine learning products are not deterministic, but they are probabilistic.

Technical aspect of machine learning prototype development

Creating a machine learning algorithm ultimately means building a model that outputs correct information given that we've provided input data. Consider model as a black box: inputs go in and outputs come out — but the processes in between are fairly complex. For instance, if we want to create a model that predicts the weather tomorrow based on meteorological information from the past few days, we would feed the model metrics such as temperature, humidity, and precipitation. The output would be the weather forecast for tomorrow. But we can't just assume the model is accurate. First, we must train the model. Training is a key concept in machine learning; it's the process through which a model learns how to make sense of input data.

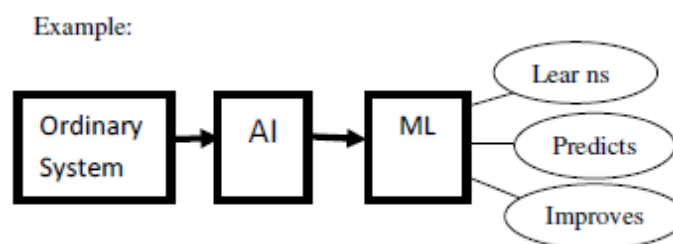
Machine Learning is an idea to learn from examples and experience, without being explicitly programmed. Instead of writing code, you feed data to the generic algorithm, and it builds logic based on the data.

- **Life without machine Learning:** when we search information on any topics on google, it will collect all the information what we want and present it to us accordingly but if there was no google, then we have to go through different books, articles and even we could find the relevant answer.
- **Life with Machine Learning:** we are getting easily connects with our old friends by associating one friend with others is like facebook and shopping product from different vendors online like from amazon, flipkart ...etc which has made shopping easy.

Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans and animals: learn from experience.

Machine learning algorithms use computational methods to learn information directly from data without relying on a predetermined equation as a model.

The algorithms adaptively improve their performance as the number of samples available for learning increases. Deep learning is a specialized form of machine learning. Machine Learning is a field which is raised out of Artificial Intelligence (AI). Applying AI, we wanted to build better and intelligent machines. But except for few mere tasks such as finding the shortest path between point A and B, we were unable to program more complex and constantly evolving challenges. There was a realization that the only way to be able to achieve this task was to let machine learn from itself. This sounds similar to a child learning from its self. So machine learning was developed as a new capability for computers. And now machine learning is present in so many segments of technology, that we don't even realize it while using it.



Application of machine learning

Suppose we provide a system with input data that contains the photo of the students of a university. Then we do:

- First analyze data. Then it tries to find patterns such as dress color, height, size etc.
- Based on these patterns, the system tries to predict different types of student belong to particular course, fittest student...etc. and partition them.
- Finally, it keeps all tracks of the decisions; it took to make sure that it is learning. Then next time when we ask the machine to predict and segregate the different types of student. Then it does not go through the entire processes again. That's the machine learning works.

Generally, the field of machine learning is divided into three sub domains:

1. Supervised learning.
2. Unsupervised learning.
3. Reinforcement learning.

Supervised learning:

The machine learns from the training data that is labeled. So we have to supervise machine learning while training it to work by its own. Supervised learning requires training with labeled data which has inputs and desired outputs. Supervised learning uses classification and regression techniques to develop predictive models.

Classification techniques predict discrete responses—for example, whether an email is genuine or spam, or whether a tumor is cancerous or benign.

Analysis of product development process in a machine learning firm

Classification models classify input data into categories. Typical applications include medical imaging, speech recognition, and credit scoring.

Regression techniques predict continuous responses—for example, changes in temperature or fluctuations in power demand. Typical applications include electricity load forecasting and algorithmic trading.

Unsupervised learning:

The machine learns from the training data but without labeled. Unsupervised learning finds hidden patterns or intrinsic structures in data. It is used to draw inferences from datasets consisting of input data without labeled responses.

Clustering is the most common unsupervised learning technique. It is used for exploratory data analysis to find hidden patterns or groupings in data.

Applications for cluster analysis include gene sequence analysis, market research, and object recognition.

For example, if a cell phone company wants optimize the locations where they build cell phone towers, they can use machine learning to estimate the number of clusters of people relying on their towers. A phone can only talk to one tower at a time, so the team uses clustering algorithms to design the best placement of cell towers to optimize signal reception for groups, or clusters, of their customers.

Reinforcement Learning:

The machine learns on its own i.e. by its mistake and experiences.

Suppose a new born baby put the finger to the burning candle flame which hurts, so next time when the baby sees the candle burning, then it recalls what has

happen last time and would repeat again. This is how reinforcement learning works.

Learning Methods	Data processing tasks	Learning algorithm
Supervised	Classification/ Regression/Estimation	Support Vector Machine
		Naive Bayes
		Hidden Markov Model
		Bayesian network
Unsupervised	Clustering/ Prediction	Neural Networks
		K-means
		Gaussian Mixture Model
Reinforcement	Decision making	Q-learning
		R-learning
		TD learning

The right kind of machine learning solution depends on:

- 1) The problem statement: If the problem is to predict the future stock market price, the supervised learning would work best.
- 2) The size, quality & nature of the data: If the data is clotted, then we can go for unsupervised learning. If the data is categorized manner, then we go for supervised learning.
- 3) Complexity of the algorithm: If we are going for predicting stock market price ,then we can go for reinforcement learning which would be vary time consuming then go for supervised learning.

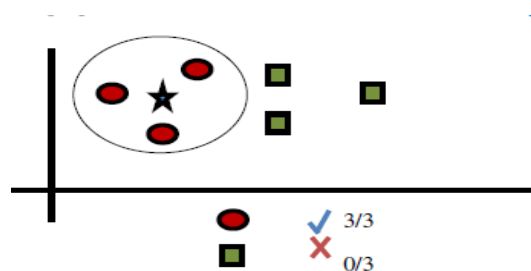
Common machine learning algorithms

There are various algorithms in machine learning but key algorithms are k-Nearest Neighbors, linear regression, decision tree and naive bayes.

1. k- nearest Neighbors

k-Nearest Neighbors (kNN) algorithm works in a way that a new data point is assigned to a neighboring group to which it is most similar.

It can be used for both classification and regression problems. However, it is more widely used in classification problems in the industry. K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases by a majority vote of its k neighbors. The case being assigned to the class is most common amongst its K nearest neighbors measured by a distance function. In k-Nearest Neighbors, 'k' can be an integer greater than 1. So, for every new data point we want to classify, we compute to which neighboring group it is closest to it or similar to it.



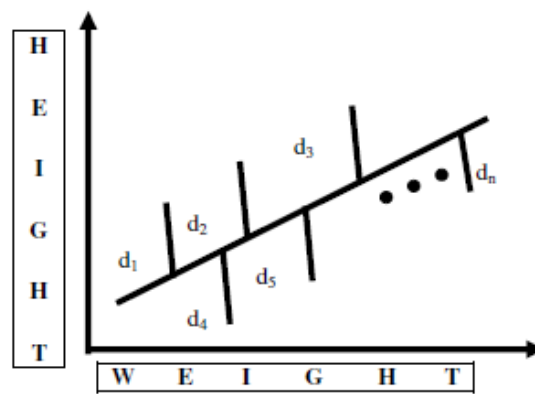
2. Linear Regression

Linear regression is a process used for estimating the relationships among variables. Here, one of the variables is dependent on one or more variables.

Analysis of product development process in a machine learning firm

It is used to estimate real values (cost of houses, number of calls, total sales etc.) based on continuous variable(s). Here, we establish relationship between independent and dependent variables by fitting a best line. This best fit line is known as regression line and represented by a linear equation:

$$Y = a * X + b$$



$$D = d_1^2 + d_2^2 + \dots + d_n^2$$

The regression line has the least value of D.

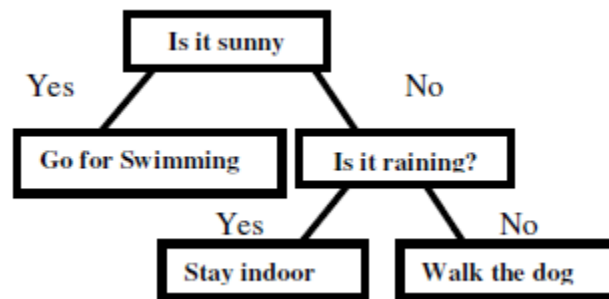
3. Decision Tree

This is one of my favorite algorithms. It is a type of supervised learning algorithm that is mostly used for classification problems. Surprisingly, it works for both categorical and continuous dependent variables.

A decision tree is a graph that uses a branching method to illustrate every possible outcomes of a decision. By using branching method, it realizes the problem and makes the decision based on the conditions.

Analysis of product development process in a machine learning firm

Suppose, somebody sitting at home and thinking for swimming, then so



Therefore we follow the decision tree every day and realize the problem and take the decision accordingly.

4. Naive Bayes

Naive Bayes is a classification technique based on Bayes' theorem with an assumption of independence between predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

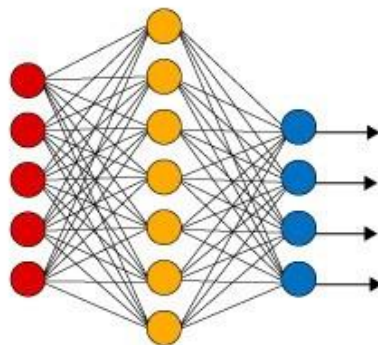
Naive Bayes classification is done when there is larger dataset. The naive Bayes classifier technique is based on conditional probability and it particularly suited when the complexity of the inputs is high.

$$P(C|A) = \frac{P(C|A) P(C)}{P(C)}$$

5. Neural network & Deep neural network

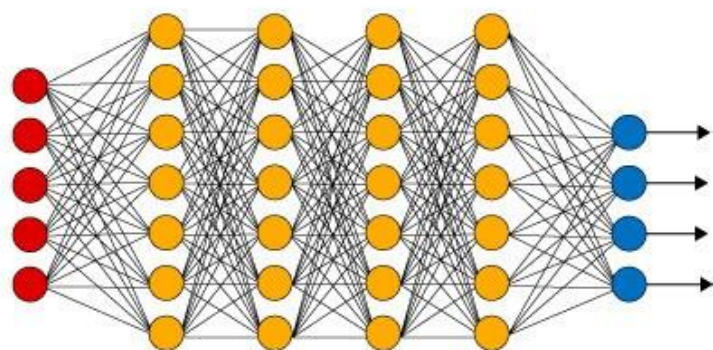
A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria.

Simple Neural Network



● Input Layer

Deep Learning Neural Network



● Hidden Layer

● Output Layer

A neural network contains layers of interconnected nodes. Each node is a perceptron and is similar to a multiple linear regression. The perceptron feeds the signal produced by a multiple linear regression into an activation function that may be nonlinear.

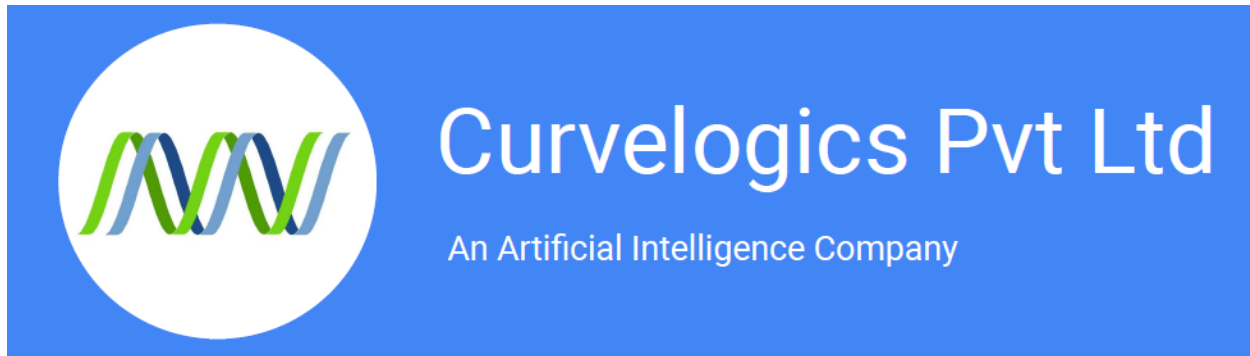
In a multi-layered perceptron (MLP), Perceptrons are arranged in interconnected layers. The input layer collects input patterns. The output layer has classifications or output signals to which input patterns may map. For instance, the patterns may comprise a list of quantities for technical indicators about a security; potential outputs could be “buy,” “hold” or “sell.”

Hidden layers fine-tune the input weightings until the neural network’s margin of error is minimal. It is hypothesized that hidden layers extrapolate salient features

Analysis of product development process in a machine learning firm

in the input data that have predictive power regarding the outputs. This describes feature extraction, which accomplishes a utility similar to statistical techniques such as principal component analysis.

PROFILE OF THE ORGANIZATION



Curvelogics Pvt. Ltd is an Indian provider of machine learning services and solutions, headquartered in Bangalore, Karnataka, India. Resmi Sreekanta Panicker founded Curvelogics in 2012 in Bangalore. The company has offices in Bangalore and Trivandrum.

Curvelogics provides machine learning and artificial intelligence IT Services and Solutions for Globally acclaimed companies. Established in 2012 and headquartered in Bangalore, Curvelogics has actively started operations in Trivandrum during 2016. Curvelogics is a technology startup with profound domain expertise across the following industries: Healthcare & Insurance, Retail, Financial Services, Transportation & Logistics, Manufacturing & Automotive, Telecommunication and Media & Entertainment. By focusing on the business model of fewer CLIENTS, more ATTENTION', Curvelogics strives for excellence in providing their clients with the best service and commitment to long-term client success. With more than 100 employees,

Curvelogics's growth and clientele have been impressive. Curvelogics use a client-centric engagement model that combines local, senior with the cost, scale, and quality advantages of off-shore operations. This client-centric focus forms the

basis for how Curvelogics, as a company, operate and serve their clients. Commitment to long-term client success empowers every employee to provide value and flexibility beyond the contract.

The expertise found within Curvelogics 's plays a key role in company's success with clients. The company delivers pragmatic machine learning solutions that allow clients to consistently achieve their most critical business objectives. Our Partner Program complements the Centers of Excellence. The program aims to develop strategic relationships with best-of-breed organizations to provide Curvelogics developers with advance access to new technology and educational resources.

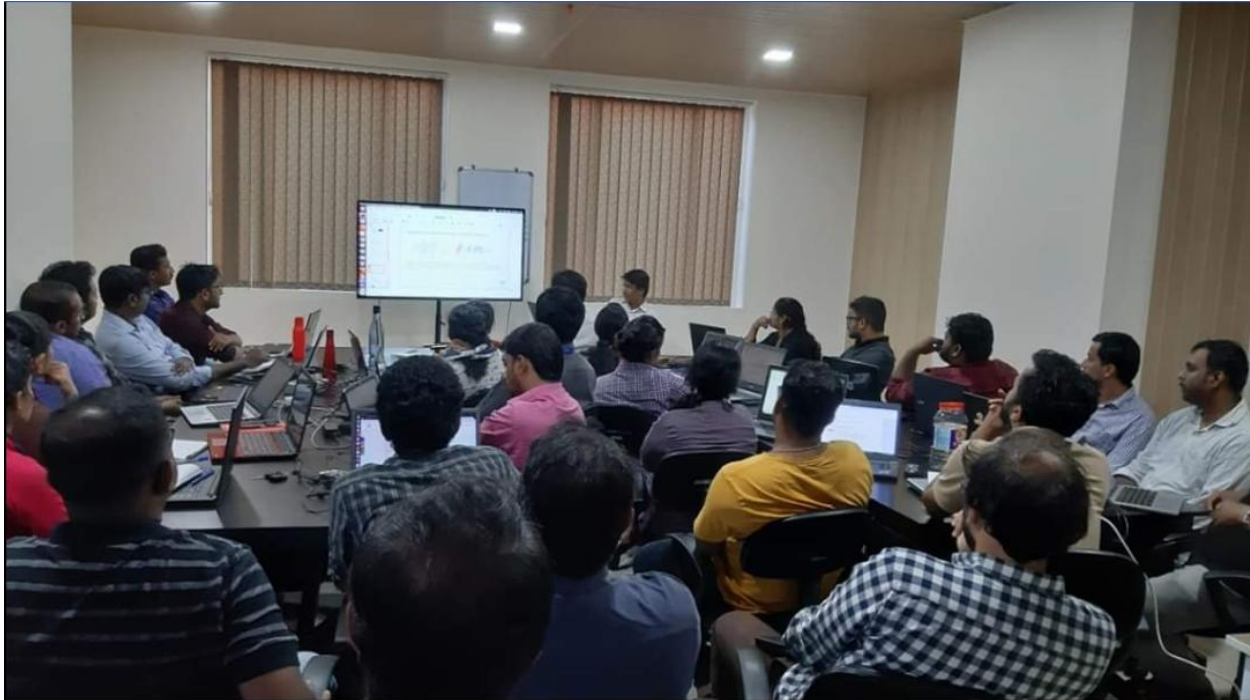
Every year Curvelogics has been attracting multiple visits from global companies. Senior executives visit Curvelogics and it has created visibility for Kerala. Every fifth startup being set up in India is from Kerala, the company also has a series of plans to provide training programs for new entrepreneurs. The recruitment process of Curvelogics is pushing the academics and urging them to get them ready for the job. Curvelogics is planning to work with the universities and schools here to facilitate the training in latest disrupting technologies like machine learning, robotics and Internet of Things. So when they graduate, they can be employed without the long cycle of waiting and waste of time and money. This approach is beginning to yield very good results. The company believes that this can create a lot more employment in Kerala. Curvelogics along with data science academy frequently host seminars and paper presentations for college students and IT professionals on topics like artificial intelligence and deep neural network. The firm has joined hands with APJ Abdul Kalam Technological University (KTU)

Analysis of product development process in a machine learning firm

and Cochin university of science and technology students in the state for their internship program —Bridging the Digital Divide.



Computer vision and genomics were company's original application fields to start with. Now they are extended to a wide variety of list. All fields are getting heavily disrupted by the digital changes. The world over print media is getting challenged by other forms of media like social media and digital media and thus associated data.



Applications :

1. Virtual Personal Assistants

Siri, Alexa, Google Now are some of the popular examples of virtual personal assistants. As the name suggests, they assist in finding information, when asked over voice. All you need to do is activate them and ask “What is my schedule for today?”, “What are the flights from Germany to London”, or similar questions. For answering, your personal assistant looks out for the information, recalls your related queries, or send a command to other resources (like phone apps) to collect info. You can even instruct assistants for certain tasks like “Set an alarm for 6 AM next morning”, “Remind me to visit Visa Office day after tomorrow”.

Machine learning is an important part of these personal assistants as they collect and refine the information on the basis of your previous involvement with them. Later, this set of data is utilized to render results that are tailored to your preferences.

2. Predictions while Commuting

Traffic Predictions: We all have been using GPS navigation services. While we do that, our current locations and velocities are being saved at a central server for managing traffic. This data is then used to build a map of current traffic. While this helps in preventing the traffic and does congestion analysis, the underlying problem is that there are less number of cars that are equipped with GPS.

Machine learning in such scenarios helps to estimate the regions where congestion can be found on the basis of daily experiences.

Online Transportation Networks: When booking a cab, the app estimates the price of the ride. When sharing these services, how do they minimize the detours? The answer is machine learning. Jeff Schneider, the engineering lead at Uber ATC reveals in a an interview that they use ML to define price surge hours by predicting the rider demand. In the entire cycle of the services, ML is playing a major role.

3. Videos Surveillance

Imagine a single person monitoring multiple video cameras! Certainly, a difficult job to do and boring as well. This is why the idea of training computers to do this job makes sense.

The video surveillance system nowadays are powered by AI that makes it possible to detect crime before they happen. They track unusual behaviour of people like standing motionless for a long time, stumbling, or napping on benches etc. The system can thus give an alert to human attendants, which can ultimately help to avoid mishaps. And when such activities are reported and counted to be true, they help to improve the surveillance services. This happens with machine learning doing its job at the backend.

4. Social Media Services

From personalizing your news feed to better ads targeting, social media platforms are utilizing machine learning for their own and user benefits. Here are a few examples that you must be noticing, using, and loving in your social media accounts, without realizing that these wonderful features are nothing but the applications of ML.

People You May Know: Machine learning works on a simple concept: understanding with experiences. Facebook continuously notices the friends that you connect with, the profiles that you visit very often, your interests, workplace, or a group that you share with someone etc. On the basis of continuous learning, a list of Facebook users are suggested that you can become friends with.

Face Recognition: You upload a picture of you with a friend and Facebook instantly recognizes that friend. Facebook checks the poses and projections in the picture, notice the unique features, and then match them with the people in your friend list. The entire process at the backend is complicated and takes care of the precision factor but seems to be a simple application of ML at the front end.

Similar Pins: Machine learning is the core element of Computer Vision, which is a technique to extract useful information from images and videos. Pinterest uses computer vision to identify the objects (or pins) in the images and recommend similar pins accordingly.

5. Email Spam and Malware Filtering

There are a number of spam filtering approaches that email clients use. To ascertain that these spam filters are continuously updated, they are powered by machine learning. When rule-based spam filtering is done, it fails to track the latest tricks adopted by spammers. Multi Layer Perceptron, C 4.5 Decision Tree Induction are some of the spam filtering techniques that are powered by ML.

Over 325, 000 malwares are detected everyday and each piece of code is 90–98% similar to its previous versions. The system security programs that are powered by machine learning understand the coding pattern. Therefore, they detect new malware with 2–10% variation easily and offer protection against them.

6. Online Customer Support

A number of websites nowadays offer the option to chat with customer support representative while they are navigating within the site. However, not every website has a live executive to answer your queries. In most of the cases, you talk to a chatbot. These bots tend to extract information from the website and present it to the customers. Meanwhile, the chatbots advance with time. They tend to understand the user queries better and serve them with better answers, which is possible due to its machine learning algorithms.

7. Search Engine Result Refining

Google and other search engines use machine learning to improve the search results for you. Every time you execute a search, the algorithms at the backend keep a watch at how you respond to the results. If you open the top results and stay on the web page for long, the search engine assumes that the results it displayed were in accordance to the query. Similarly, if you reach the second or third page of the search results but do not open any of the results, the search engine estimates that the results served did not match requirement. This way, the algorithms working at the backend improve the search results.

8. Product Recommendations

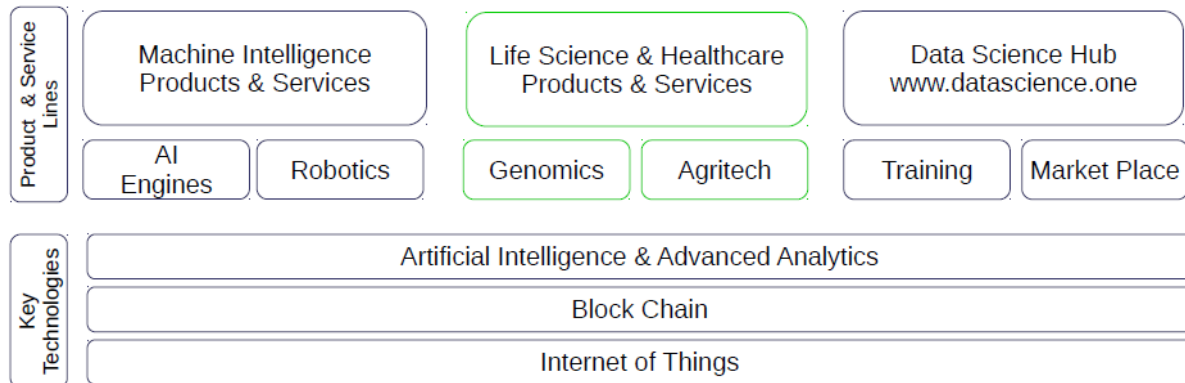
You shopped for a product online few days back and then you keep receiving emails for shopping suggestions. If not this, then you might have noticed that the shopping website or the app recommends you some items that somehow matches with your taste. Certainly, this refines the shopping experience but did you know that it's machine learning doing the magic for you? On the basis of your behaviour with the website/app, past purchases, items liked or added to cart, brand preferences etc., the product recommendations are made.

9. Online Fraud Detection

Machine learning is proving its potential to make cyberspace a secure place and tracking monetary frauds online is one of its examples. For example: Paypal is using ML for protection against money laundering. The company uses a set of tools that helps them to compare millions of transactions taking place and

Analysis of product development process in a machine learning firm

distinguish between legitimate or illegitimate transactions taking place between the buyers and sellers.



Applications in different domains

The value of machine learning technology has been recognized by companies across several industries that deal with huge volumes of data. By leveraging insights obtained from this data, Curelogics works in an efficient manner to control costs as well as get an edge over other competitors for the clients. These are some domains Curvelogics is implementing machine learning –

- **Financial Services**

Companies in the financial sector are able to identify key insights in financial data as well as prevent any occurrences of financial fraud, with the help of machine learning technology. The technology is also used to identify opportunities for investments and trade. Usage of cyber surveillance helps in identifying those individuals or institutions which are prone to financial risk, and take necessary actions in time to prevent fraud.

- **Marketing and Sales**

Companies are using machine learning technology to analyze the purchase history of their customers and make personalized product recommendations for their next purchase. This ability to capture, analyze, and use customer data to provide a personalized shopping experience is the future of sales and marketing.

- **Government**

Government agencies like utilities and public safety have a specific need FOR ML, as they have multiple data sources, which can be mined for identifying useful patterns and insights. For example sensor data can be analyzed to identify ways to minimize costs and increase efficiency. Furthermore, ML can also be used to minimize identity thefts and detect fraud.

- **Healthcare**

With the advent of wearable sensors and devices that use data to access health of a patient in real time, ML is becoming a fast-growing trend in healthcare. Sensors in wearable provide real-time patient information, such as overall health condition, heartbeat, blood pressure and other vital parameters. Doctors and medical experts can use this information to analyze the health condition of an individual, draw a pattern from the patient history, and predict the occurrence of any ailments in the future. The technology also empowers medical experts to analyze data to identify trends that facilitate better diagnoses and treatment.

- **Transportation**

Based on the travel history and pattern of traveling across various routes, machine learning can help transportation companies predict potential problems that could arise on certain routes, and accordingly advise their customers to opt for a different route. Transportation firms and delivery organizations are increasingly using machine learning technology to carry out data analysis and data modeling to make informed decisions and help their customers make smart decisions when they travel.

- **Oil and Gas**

This is perhaps the industry that needs the application of machine learning the most. Right from analyzing underground minerals and finding new energy sources to streaming oil distribution, ML applications for this industry are vast and are still expanding.

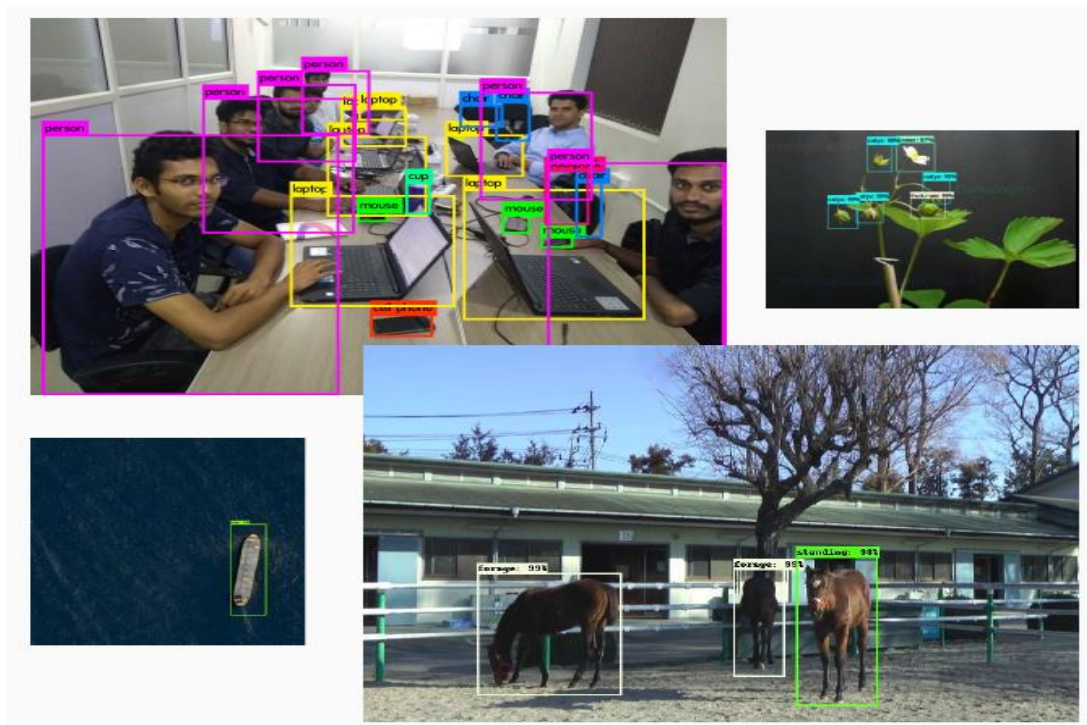
PRODUCT DEVELOPMENT DIVISION OF CURVELOLOGICS

(Exact details cannot be published because of client purchase agreements)

Along with the service based work for the clients, Curvelogics has a strong research and development center for their own products. They have released multiple products in the last couple of years which then made aggressive demand among a number of other business companies. Since majority of the products are meant for business entities, they are following a B2B business model.

- **Object detection**

Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance.



Every object class has its own special features that helps in classifying the class – for example all circles are round. Object class detection uses these special features. For example, when looking for circles, objects that are at a particular distance from a point (i.e. the center) are sought. Similarly, when looking for squares, objects that are perpendicular at corners and have equal side lengths are needed. A similar approach is used for face identification where eyes, nose, and

lips can be found and features like skin color and distance between eyes can be found.

- **Video interview**

HR today needs to be empowered with Artificial Intelligence for businesses to thrive without losing the personal approach to communication. AI-powered video interview bots can help HR heads in saving manpower engaged in setting up interviews and evaluate candidates without human biases, even if only subconsciously.

With automated interviews, recruiters no more need to worry about scheduling interviews, time restrictions, and interviewer biases. Advanced AI along with technology are an enabler for human resources. Aspiring Minds, a company that uses Artificial Intelligence to enable meritocracy in the labour market, offers a virtual interview platform that not only mimics actual face-to-face interview but also automatically evaluates a candidate just like a human interviewer would have.

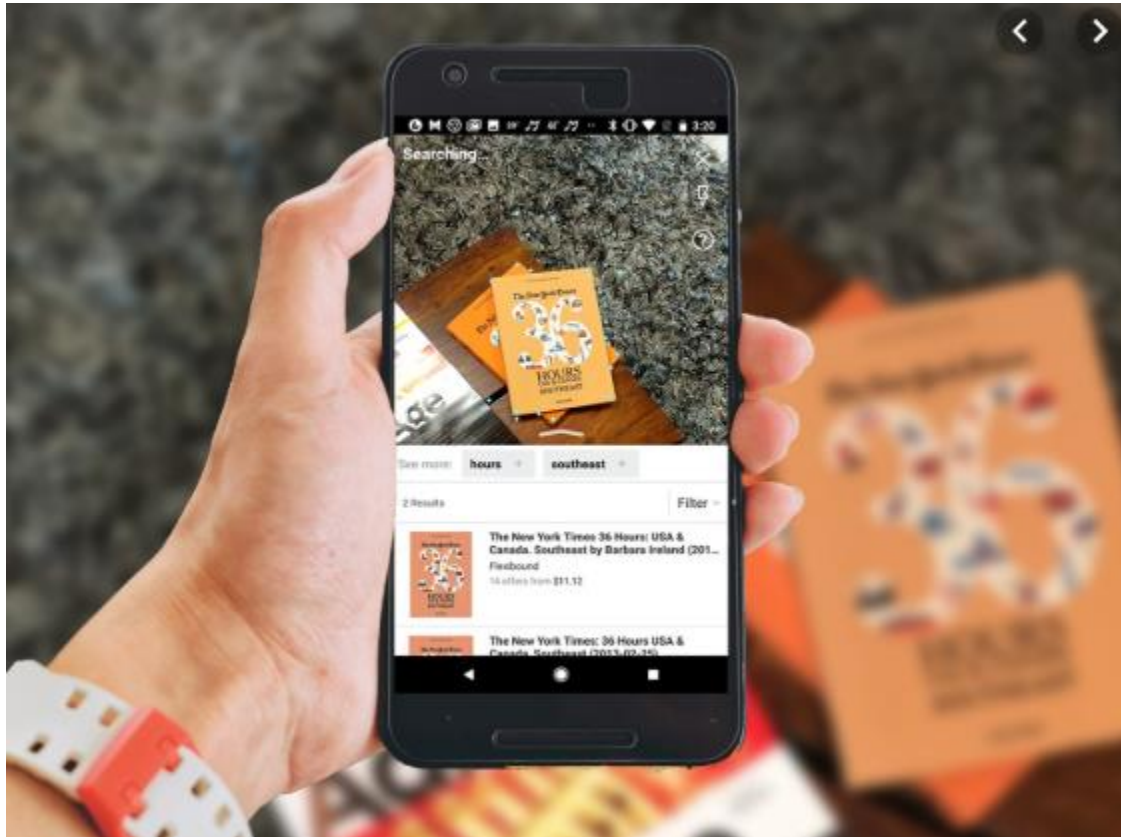
- **Facial recognition system**

A facial recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems work, but in general, they work by comparing selected facial features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analysing patterns based on the person's facial textures and shape.



- **Visual search engine**

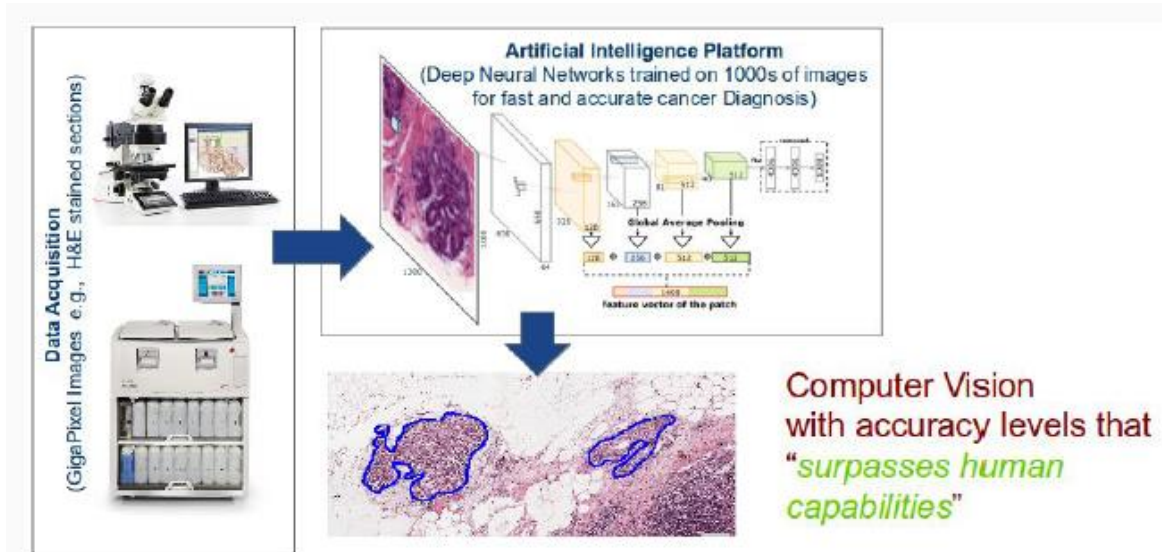
A Visual Search Engine is a search engine designed to search for information on the World Wide Web through the input of an image or a search engine with a visual display of the search results. Information may consist of web pages, locations, other images and other types of documents. This type of search engines is mostly used to search on the mobile Internet through an image of an unknown object (unknown search query). Examples are buildings in a foreign city. These search engines often use techniques for Content Based Image Retrieval.



A visual search engine searches images, patterns based on an algorithm which it could recognize and gives relative information based on the selective or apply pattern match technique

- **Computer Aided Diagnosis**

In medical imaging field, computer-aided detection (CADe) or computer-aided diagnosis (CADx) is the computer-based system that helps doctors to take decisions swiftly . Medical imaging deals with information in image that the medical practitioner and doctors has to evaluate and analyze abnormality in short time. Analysis of imaging in medical field is very crucial task because imaging is basic modality to diagnose any diseases at the earliest but acquisition of image is not to harm the human body.



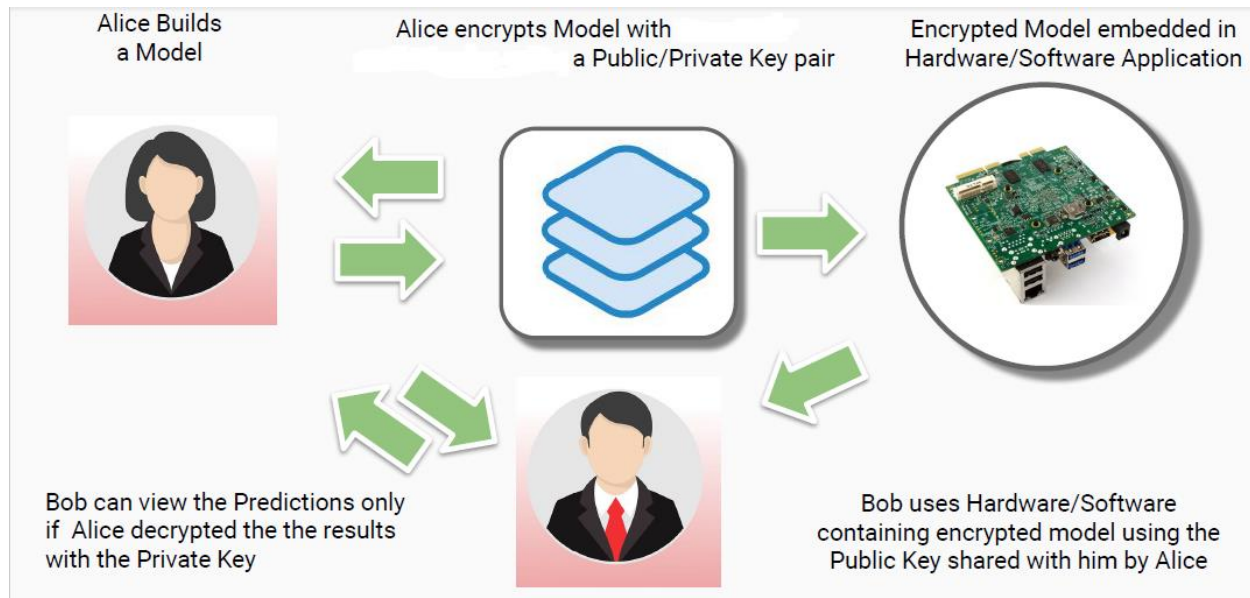
Imaging techniques like MRI, X-ray, endoscopy, ultrasound, etc. if acquired with high energy will provide good quality image but they will harm the human body; hence, images are taken in less energy and therefore, the images will be bad in quality and low contrast. CAD systems are used to improve the quality of the image, which helps to interpret the medical images correctly and process the images for highlighting the conspicuous parts.

- **Tool to protect ML models from piracy risk**

Machine learning models are key intellectual property assets in the current age of AI and ML, but reverse engineering and piracy poses serious risks to those asset's value and importance. Following is the working of the tool.

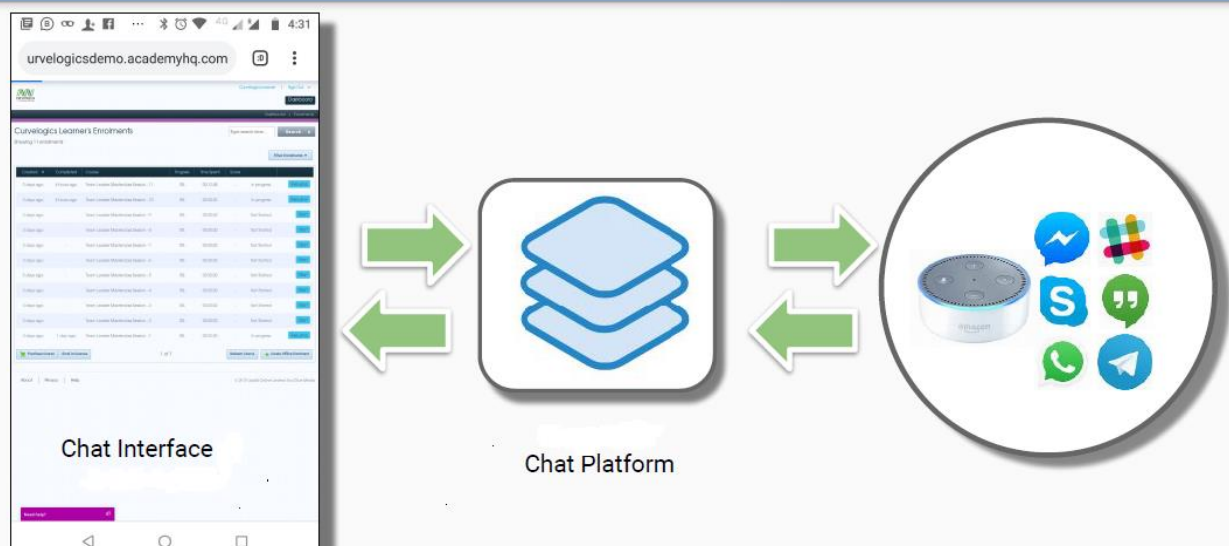
Traditional e-learning platforms do not have hyper personalization capabilities. Time spent on personal mobile devices by users has been increasing at a rapid pace compared to any other form factor.

Analysis of product development process in a machine learning firm



- **AI chatbot platform for enhancing existing learning and design applications**

How it works



Analysis of product development process in a machine learning firm

This tool enable enhancement of existing e- learning platform by providing chat interface with hyper specialization capability and behavior tracking. The conversational media enhances learner engagement.

HIGHER MANAGEMENT TEAM

Brijesh Madhavan



Brijesh Madhavan, PhD

CEO, Curvelogics
and co-founder Gnobiosys

20+ years in Data Science, Machine
Learning and Advanced Analytics

Profile: Brijesh is the CEO of Curvelogics Pvt Ltd and a co-founder of Gnobiosys Computational Genomic Labs. His focus is to grow Curvelogics into a successful Artificial Intelligence (AI) company through cutting edge R&D and developing AI product lines centered around Robotics, IoT, Block Chain, Genomics, Big Data and Advanced Analytics. He has a keen interest in creating a stronger Artificial Intelligence ecosystem in India and contributes to this by mentoring youngsters in building a strong career in Data Science and AI engineering.

Experience: He has about 20 year of experience in advanced Statistical Modeling and Machine learning. He has held senior leadership roles heading Advanced analytics, Big Data and Artificial Intelligence teams in MNCs such as HP, IHS and EY. He has also been a mentor to several budding entrepreneurs and startups.

Yuvaraj Ranganathan**Yuvaraj Ranganathan, PhD**

Chief Data Scientist, Curvelogics
and co-founder Gnobiosys

10+ years in Data Science,
Computational Genomics

Profile: Yuvaraj is the Principal Data Scientist at Curvelogics and is also the cofounder of Gnobiosys Computational Genomic Labs. He is focused on developing Artificial Intelligence products that cater to the life-science and healthcare industry.

Besides this he is also the Curriculum Architect of our data science training initiative. Experience: He has over 10 years of fundamental research experience

from premiere institutes such as Indian Institute of Science (IISc) and Agrovzykum, Czech Republic. During his stint at IISc, his research focus was on decoding communication in systems involving chemical signalling channels. During this research he extensively used statistical and machine learning algorithms to analyse massive amounts of machine generated data. He is also the person who standardized analytical methods for some of best research institutes in Czech Republic.

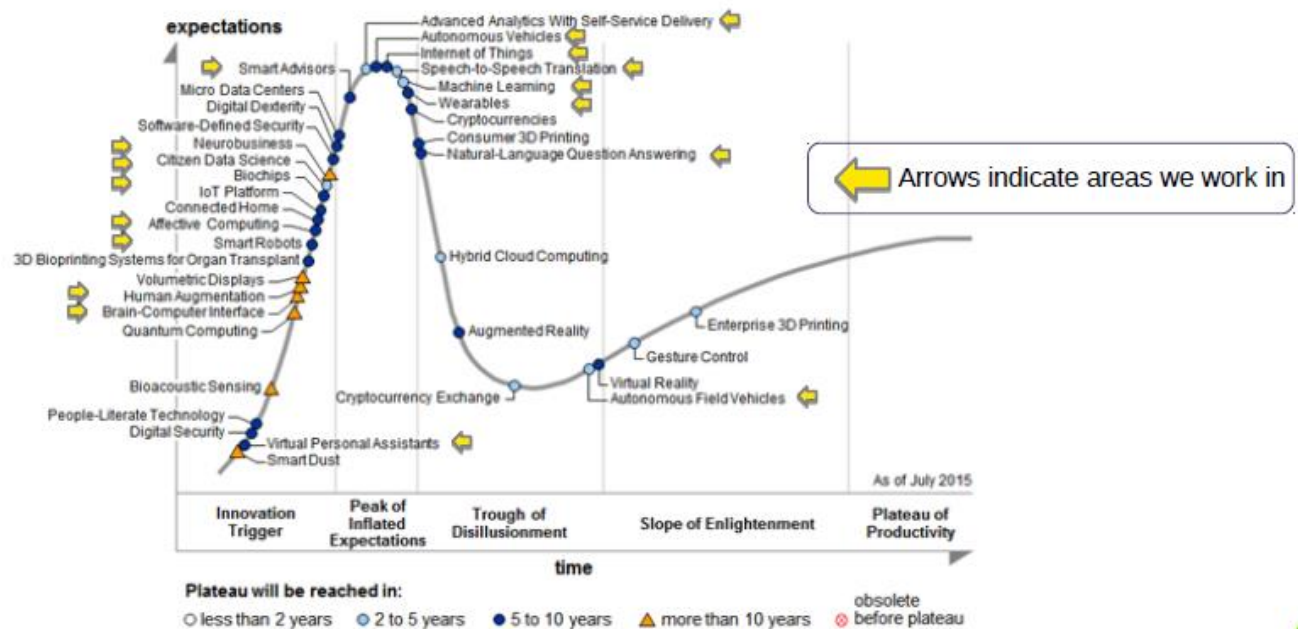
FUTURE ENDEAVORS

The company is expected to grow rapidly in different machine learning domains in the next couple of years by hiring more employees expertsed in different fields.

Analysis of product development process in a machine learning firm

Following is a graph depicted on hype curve by Gartner with the expected expansion.

Gartner's Technology Hype Curve 2015



SISTER COMPANIES

The same management is having 2 other companies working closely with Curvelogics in their tasks.

1. Data science academy
2. Gnobiosys

DATA SCIENCE ACADEMY



Data Science Academy consists of expert trainers with several years of industry experience besides Human resource expertise. The team consists of people with PhDs from premier institutes who have applied AI and advanced analytics in a variety of areas, and these areas range from Robotics, Technology Infrastructure Sales and Marketing, Pricing, campaign Management, Social

Media etc., all the way up to Genomics.

As leaders of Data Science teams in the corporate world we have always found it very challenging to get talented machine learning and big data practitioners over the past decade and always relied on building teams ground up by coaching engineers, statisticians and other quants graduates. With the current boom in Artificial Intelligence and Machine Learning, and highly competitive nature of the industry, organizations do not have the luxury of building teams using in-house training approach which is at the cost of precious time of experts within their organization. As a solution to this challenge, we set up Data Science Academy both to up skill enthusiasts who are currently working and to train fresh graduates by helping in bridging the academia-industry gap.

Over the years we have been able to build a large network of data science and analytics professionals which give us access to a large pool of talented Data

Analysis of product development process in a machine learning firm

Science Professionals and this makes our talent acquisition services that we offer highly effective for our clients.

Skill Development Services - Data Science Academy

1. Data Science, Big Data and AI Trainings: At Data Science Academy, our skill development and corporate training division, we have devised innovative training techniques that makes learning Data science a seamless effort. We provide instructor led classroom based trainings.

2. Expert Trainers: Trainings are delivered by industry leaders who have applied Advanced Data Science Techniques both in solving real world business problems and academic research.

3. Focus on Emerging Technologies: We provide quality training in Artificial Intelligence, Big Data, Machine learning , Neural networks , Advanced Analytics and similar technological areas.

4. Internships: We also facilitate Internship for deserving candidates having good Academics records with our partner/client's organisations.

Talent Acquisition Services

1. Talent Acquisition: We provide Talent acquisition services and this includes recruitment, executive selection, and bulk recruitment services.

2. Relevant Shortlisting: Unlike most recruitment consulting firms, the shortlisted talent that we bring to you as potential hires are already screened for technical capabilities, this reduces the effort from your side and ensures speedy closure of vacant positions.

Analysis of product development process in a machine learning firm

3. Expert Skill Evaluation: Technical screening is done by our leadership team who have served as heads of large Advanced Analytics, Big Data and AI practices in MNCs like HP, EY and IHS Inc.

4. Online/Offline Aptitude Testing: We provide services for conducting online and offline aptitude test for Data Science, Machine Learning, Advanced Analytics, Artificial Intelligence, Big Data and other associated technologies.

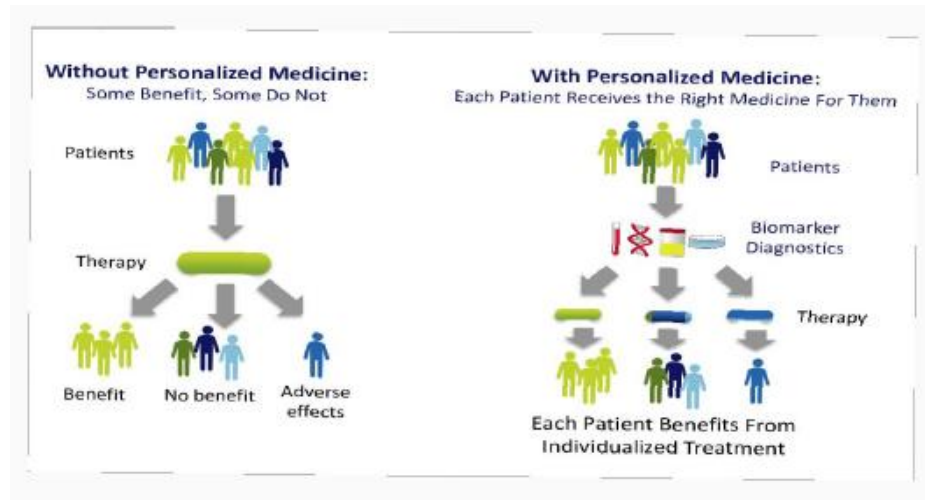
GNOBIOSYS



Gnobiosys consists of expert in the field of machine learning and genetic science. The team consists of people with PhDs from premier institutes who have applied AI and advanced analytics specialized in genetic related domains.

Using this expertize the company is trying to make some disruptive breakthroughs in the genetic field. It works in more of an R&D nature. Major works in the company revolves around the following problem.

Computational genomics – disease diagnosis using gene sequence



- Identify genetic factors that predispose
- Predict the severity of the disease
- Personalize your medication and use
- Discover carrier status for a particular disease
- Find genetic factors that your children could acquire
- Other traits such as Irises color, digit ratio, Male pattern baldness
- Genomics based dietary recommendations

CHAPTER 3

DATA COLLECTION

Data collection was the lengthier process in this project. Since the study was about a completely new and technically complex industry, the time taken to collect the secondary data and understanding the technology is comparatively slower. The source for secondary data includes referring plethora of literatures and journals available in internet and books given in the Bibliography of this report. After understanding the necessary information about the technologies like machine learning and robotics, I have collected data about the product development process in IT and machine learning industry.

Primary data for the study collected mainly through conducting a survey among the employees of the company. The necessary questions for the survey prepared after having a talk with couple of senior managers in the company about the details of their works, also have done a thorough research on literatures and earlier studies on the same from internet. Later this questionnaire distributed using Google forms to hundred employees picked using simple random sampling method, with each member of the subsets (gender, designation, age and industry experience) has an equal probability of being chosen. The human resource department in the company helped in collecting the data from employees at the earliest and the response thus received later interpreted after converting into the presentable formats given below.

Simple Percentage Analysis

The questionnaire consists of 35 questions. In this section we are simply representing the responses received directly from employees after converting it into percentage figure. Percentage analysis is the method to represent raw streams of data as a percentage (a part in 100 per cent) for better understanding of collected data.

$$\text{Percentage of respondents} = \frac{\text{No of Respondents}}{\text{Total Respondents}} \times 100$$

Sample table :

Si. No.	Designation	No. of Respondents	Percentage (%)
1	A	B	C

A – Type of designation

B – Number of response for the option A

C- Percentage figure of response A out of total response.

There are some questions which asked employees to give rank for the choices as per their preference. To represent it in percentage figure we gave weightage points for the response and converted each choices points to percentage out of total points for all choices. For example the question,

? Rank your preference for the idea generation methods in the company

- ☐ Idea from employees
- ☐ Client
- ☐ Market study reports/external agencies

Here Rank 1 choice will be given 3 points, 2 points for Rank 2 and 1 point for Rank 3 choice. After calculating the sum of points for each choice, will find the percentage on the basis of the calculated points for each option.

Analysis of product development process in a machine learning firm

Sample table format for rank questions will be as follows

Si. No.	Idea generation source	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	A	B	C	D	E	F

A – Option for type of Idea generation source

B – Number of response with 'A' as their 1st preference

C- Number of response with 'A' as their 2nd preference

D- Number of response with 'A' as their 3rd preference

E- Total points after giving weightage to responses B, C and D,

I.e., $B*3 + C*2 + D*1$

F – Percentage figure of E out of the total points of all idea generation methods.

Table: 1
Designation

Si. No.	Designation	No. of Respondents	Percentage (%)
1	Senior Manager	5	5
2	Technical Manager	10	10
3	Non-Technical Manager	3	3
4	Technical staff	70	70
5	Non-technical staff	12	12

Source: Primary data

Data analysis

- 70% of the Employees are technical staff, whereas 12% of them are non-technical staff. 10% of them are having the designation Technical Manager and 3% of the employees are working as non-technical manager. Only 5% of the employees are senior managers.
- The most of the employees working in this company are from technical domain.

Table: 2**Age**

Si. No.	Age	No. of Respondents	Percentage (%)
1	Below 30 years	57	57
2	30 – 39 years	41	41
3	40 – 49 years	2	2
4	Above 49 years	0	0

Source: Primary data

Data analysis

- 57% of the Employees are below 30 years of age, whereas 41% of them are in between 30 and 39 years of age. Rest of the 2% employees is between 40 and 49 years of age.
- It is clear that majority of the people are having age below 39 in this firm.

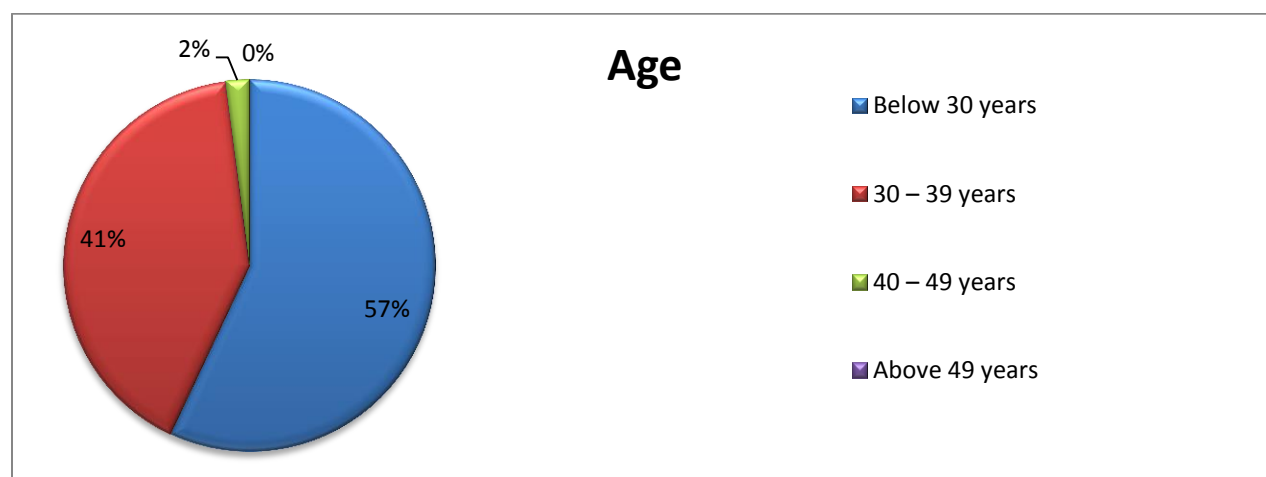


Table: 3**Gender**

Si. No.	Gender	No. of Respondents	Percentage (%)
1	Male	51	51
2	Female	49	49

Source: Primary data

Data analysis

- 51% of the Employees are male and 49% are female. That shows around same percentage of male and female in the firm.

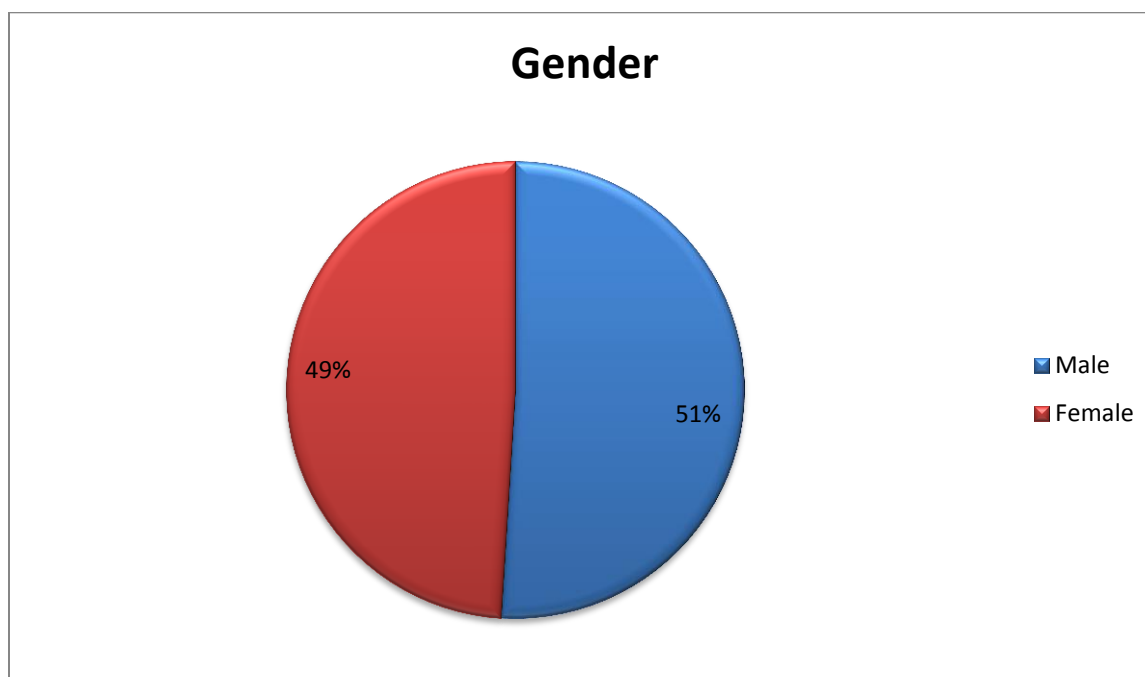


Table: 4**Total industry experience**

Si. No.	Total industry experience	No. of Respondents	Percentage (%)
1	Below 2 years	26	26
2	3 – 5 years	31	31
3	5 – 10 years	25	25
	10 – 20 years	18	18
4	Above 20 years	0	0

Source: Primary data

Data analysis

- 26% of the Employees are having only below 2 years of experience while 31% of the employees are having 3 to 5 years of experience. 25% of people are having 5-10 years of industry experience and only 18% with 10 to 20 years of experience.
- It is clear that majority of the people are having experience less than 10 years in this company.

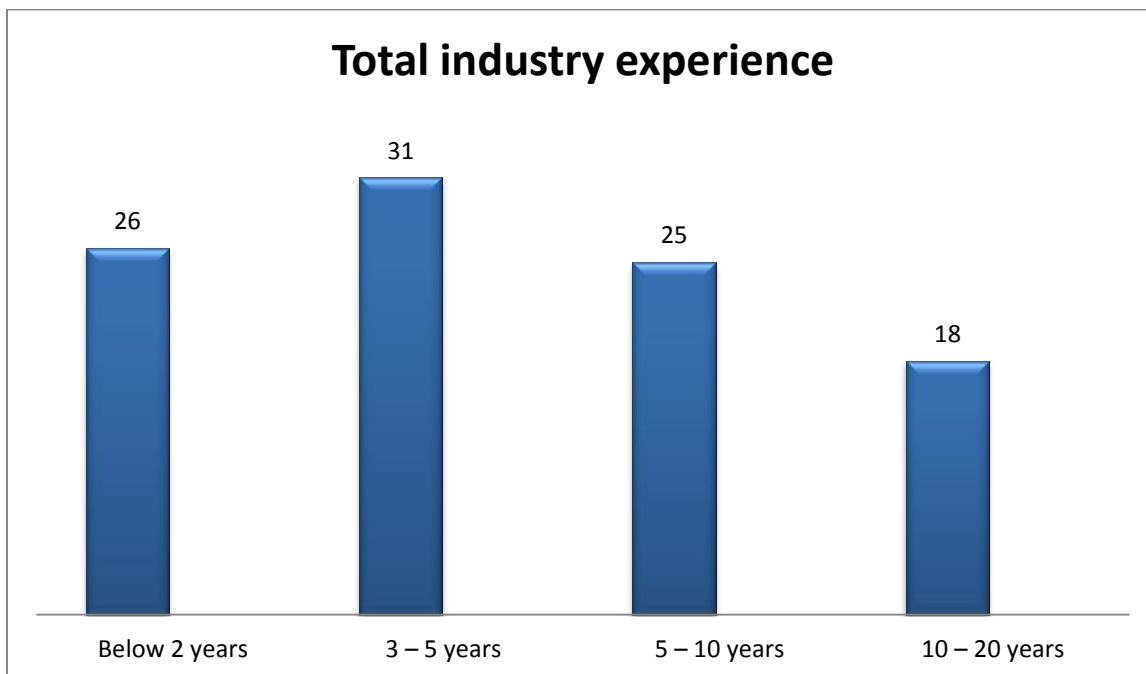


Table: 5**Kind of machine learning domain generally works**

Si. No.	Machine learning domain	No. of Response	Percentage (%)
1	Computer vision	82	82
2	Genomics	83	83
3	Chatbots	86	86
4	Natural language processing	80	80
5	Trend analysis	0	0

Source: Primary data

Computer vision : Computer vision is a field of computer science that works on enabling computers to see, identify and process images in the same way that human vision does, and then provide appropriate output. It is like imparting human intelligence and instincts to a computer.

Genomics : Genomics is the study of whole genomes of organisms, and incorporates elements from genetics. Genomics uses a combination of recombinant DNA, DNA sequencing methods, and bioinformatics to sequence, assemble, and analyse the structure and function of genomes.

Chatbot : A chatbot is a piece of software that conducts a conversation via auditory or textual methods. Such programs are often designed to convincingly simulate how a human would behave as a conversational partner.

Natural language processing : Natural language processing is a subfield of linguistics, computer science, information engineering, and artificial intelligence

concerned with the interactions between computers and human languages, in particular how to program computers to process and analyze large amounts of natural language data.

Trend analysis: A market trend analysis is an analysis of past and current market behavior and dominant patterns of the market and consumers. An important aspect of conducting a trend analysis for an organization is to obtain insights on the market scenario, consumer preferences, and the macroeconomic environment.

Data analysis

- 82% of the Employee's response is Computer vision while 83% think Genomics is also part of their work. 86% people think they are working in chatbot related works and 80% of people believe they are mainly working in Natural language processing. None of the people responded with the option Trend analysis.
- All four domains computer vision, genomics, chatbots and natural language processing have equal role in this firm.

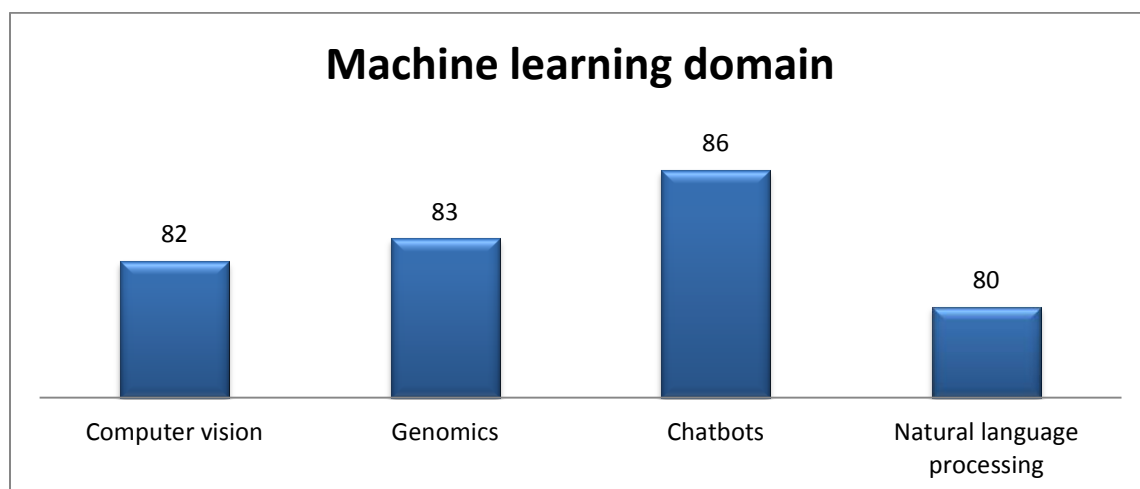


Table: 6**Kind of problems attended generally**

Si. No.	Kind of problems attended	No. of Respondents	Percentage (%)
1	Existing problems	84	84
2	New problem	16	16

Source: Primary data

Data analysis

- 84% of the Employees think the company deals with existing problems, whereas remaining 16% think the company generally deals with new problems.
- Majority of the people think company mainly deals with existing problems.

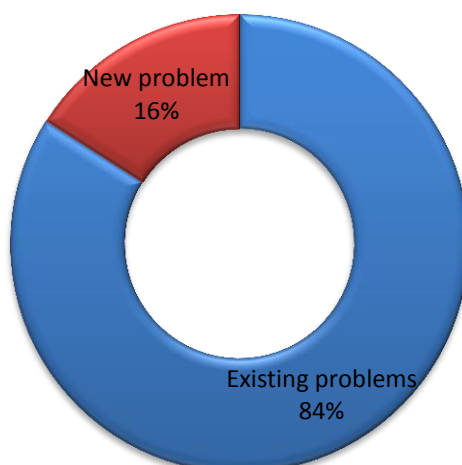
Kind of problems attended

Table: 7**Idea generation source**

Si. No.	Idea generation source	No. of Respondents	Percentage (%)
1	From employees	3	3
2	Client/Customer	17	17
3	Market study reports/external agencies	80	80

Source: Primary data

Data analysis

- 80% of the Employees think the main source for idea is from market study by internal and external agency. 17% of the people think the main idea source is from clients. Only 3% of the people think idea generation from employees actually plays a vital role in the company.
- The majority of the idea comes from either through market study or client requirement.

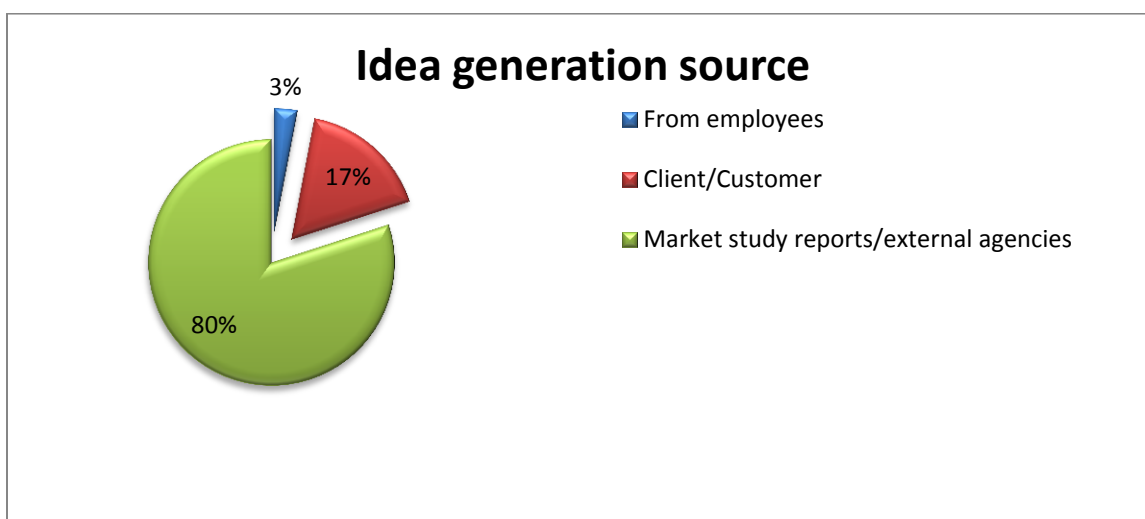


Table: 8**Preference for the idea generation methods**

Si. No.	Idea generation source	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	Idea from employees	0	3	85	91	15
2	Client	9	76	15	194	33
3	Market study reports/external agencies	91	21	0	315	52

Source: Primary data

Data analysis

- Here we are ranking the preference for idea generation according to employee's perspective. After going points to each rank we then converted it to percentage format.
- 52% of the score is for market study by internal and external agency. 33% of the point is for the ideas from clients where as 15% of the score is for ideation workshop among employees.

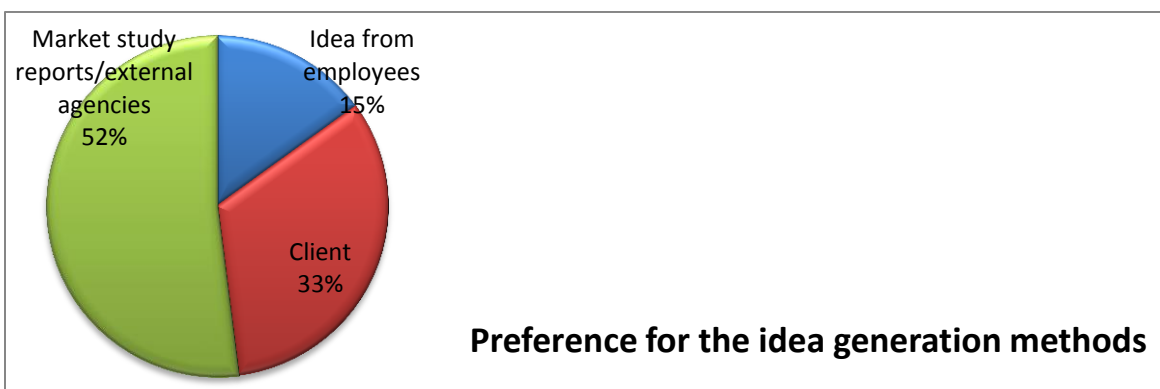


Table: 9**Employee's participation in idea generation**

Si. No.	Employee's participation in idea generation	No. of Respondents	Percentage (%)
1	Always advisable	65	65
2	Usually advisable	25	25
3	Sometimes advisable	10	10
4	Rarely advisable	0	0
5	Never advisable	0	0

Source: Primary data

Data analysis

- 65% of the Employees think employees participation in idea generation always welcomes where as 25% of the mare are of the view that usually it is advisable. Rest 10% thinks it is sometimes advisable.
- All of the employees welcome the concept of ideation hackathon among employees is fruitful.

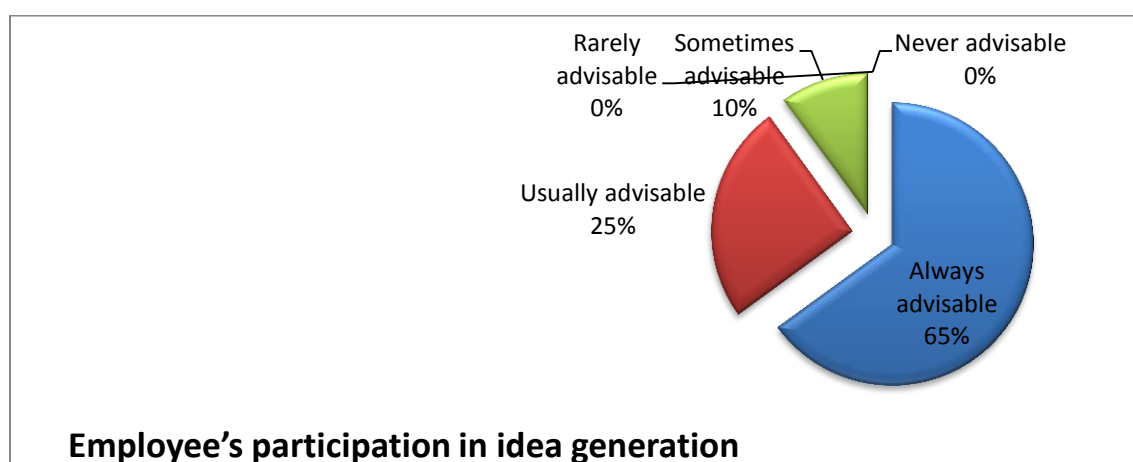


Table: 10**Rating the idea generation process**

Si. No.	Rating the idea generation process	No. of Respondents	Percentage (%)
1	Very high quality	68	68
2	High quality	23	23
3	Neither high nor low quality	9	9
4	Low quality	0	0
5	Very low quality	0	0

Source: Primary data

Data analysis

- 68% of the Employees think the idea generation process in the company is having very high quality where as 23% consider the process as high quality. Only the rest 9% consider the process as neither high nor low quality.
- Majority of the people consider the idea generation process in the company great.

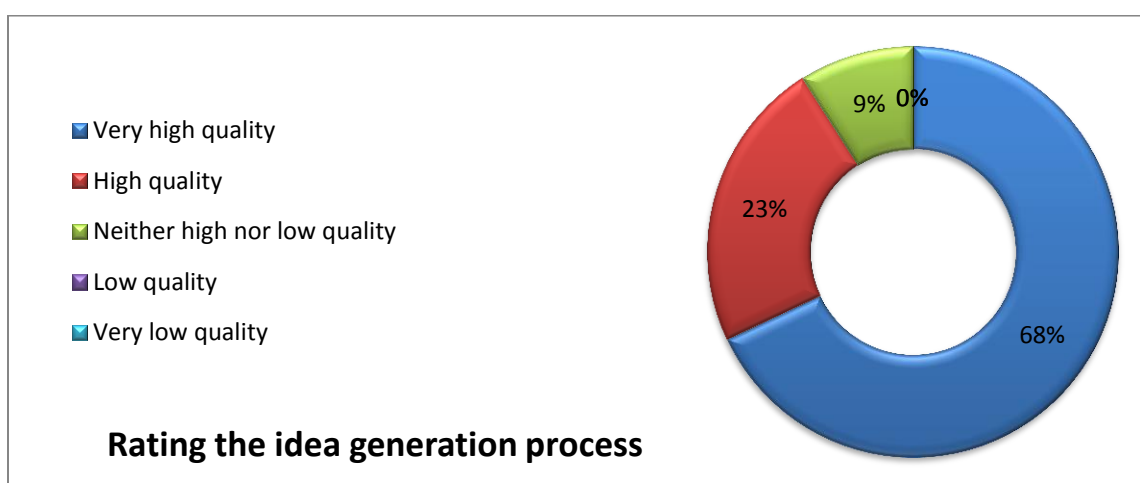


Table: 11**Team assignment for the feasibility study**

Si. No.	Team assignment for the feasibility study	No. of Respondents	Percentage (%)
1	Senior level management team	83	83
2	Pre designated team for all feasibility study	17	17
3	Random selection of members from each departments	0	0

Source: Primary data

Data analysis

- 83% of the Employees think senior level management team do the feasibility study whereas the rest 17% of the people think there is a pre designated team for all feasibility study.
- None of them think that the random selection of team members happens in the firm.

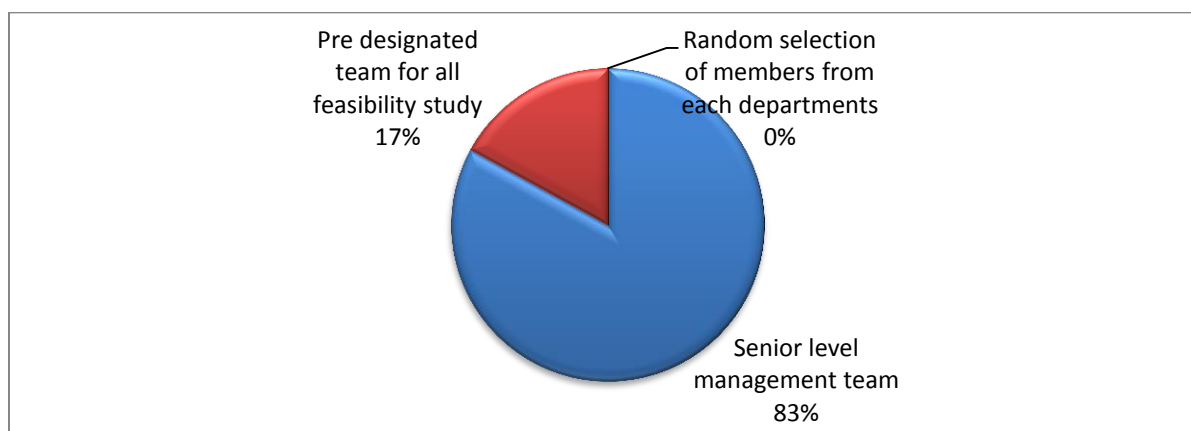


Table: 12**Preference for assigning teams for feasibility study**

Si. No.	Preference for assigning teams for feasibility study	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	Random selection of members from each departments	0	0	100	100	16.7
2	Pre designated team for all feasibility study	21	79	0	221	36.8
3	Senior level management team	79	21	0	279	46.5

Source: Primary data

- Here we are ranking the preference for the assignment of teams for feasibility study according to employee's perspective. After giving points to each rank we then converted it to percentage format.
- 46.5% of the score is for assigning senior level management for the feasibility study where as 36.8% of employees consider a pre designated team for all feasibility will be more useful. 16.7% of the point is for random selection of members from different departments.

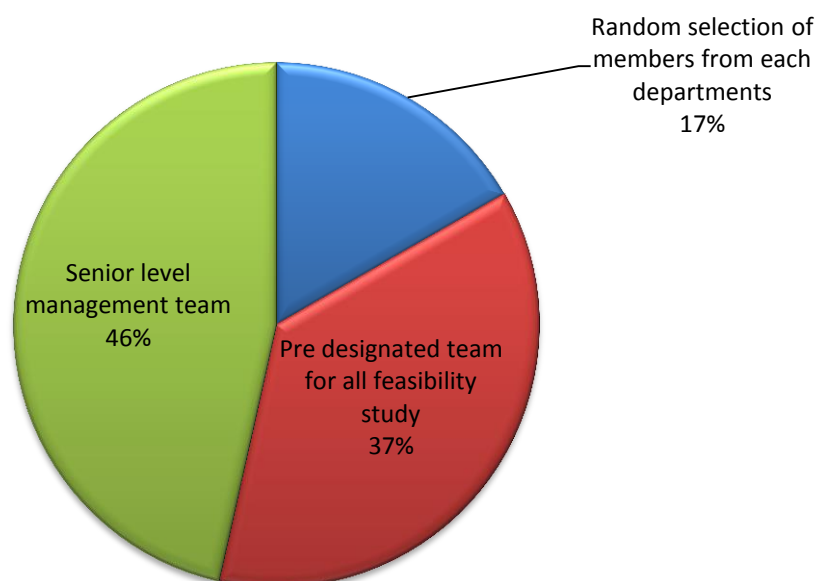
Preference for assigning teams for feasibility study

Table: 13**Leadership of team for studying the feasibility of selected ideas**

Si. No.	Team assignment for the feasibility study	No. of Respondents	Percentage (%)
1	Member from marketing	2	2
2	Member from technical	98	98
3	Member from finance	0	0
4	Member from HR	0	0

Source: Primary data

Data analysis

- 98% of the Employees think technical department leads the team for feasibility study where as 2% people think it is marketing team. Even though we had options such as finance and HR, only 0% of the people voted for them.
- Approximately Zen percentage people believe it is from technical department.

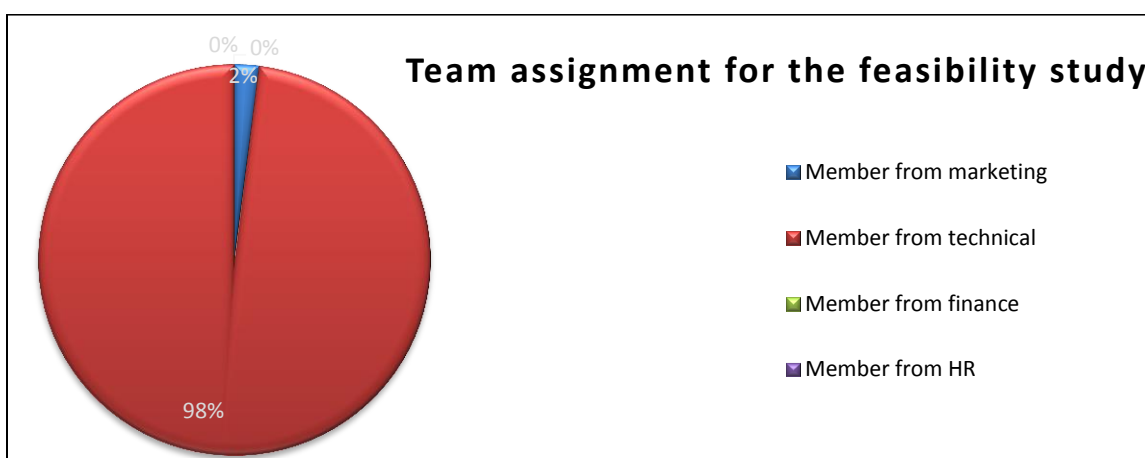


Table: 14**Idea selection process in the company**

Si. No.	Idea selection process	No. of Respondents	Percentage (%)
1	By voting majority	9	9
2	Consensus after discussion	91	91
3	Analytical	0	0

Source: Primary data

Data analysis

- An idea will be selected in the company mainly using consensus after discussion.
- 91% employees says that the idea selection process is always through consensus after discussion and only 9 % says that's through the majority voting but no one support Analytical selection.

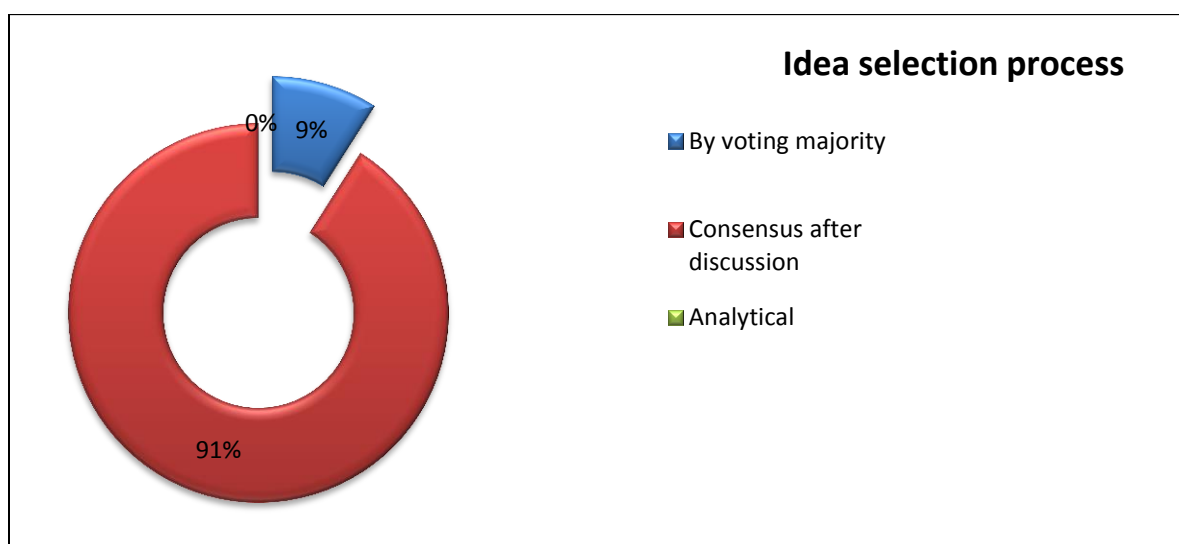


Table: 15**Preference for Idea Selection Process**

Si. No.	Preference for Idea Selection Process	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	By voting majority	8	67	25	183	31
2	Consensus after discussion	81	15	4	277	47
3	Analytical	8	17	75	133	22

Source: Primary data

Data analysis

- After giving points to each rank we then converted it to percentage format.
- 47% of the score says that Consensus after discussion is the preferred idea selection method, whereas 31% of employees consider majority voting would be good. Only 22% opt for the Analytical Process.

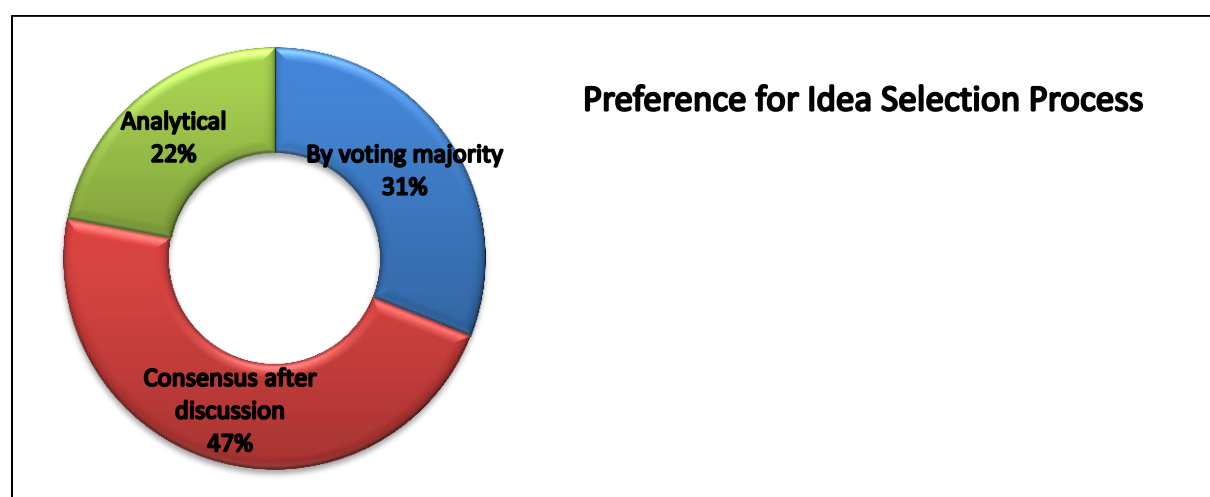


Table: 16**Rank the influencing factors to evaluate the feasibility of an idea**

Si. No.	Factors to evaluate the feasibility of an idea	Ranks					Points	Percentage
		Rank 1	Rank 2	Rank 3	Rank 4	Rank 5		
1	Market acceptance	72	13	13	2	0	442	30
2	Cost of production	4	7	18	67	4	240	17
3	Technology	24	75	1	0	0	423	28
4	Competitors	0	4	61	24	11	258	17
5	Profit	0	1	7	7	85	124	8

Source: Primary data

- Here we are ranking the preference for influencing factors to evaluate the feasibility of an idea according to employee's perspective. After giving points to each rank we then converted it to percentage format.
- 30% of the score says that Market acceptance is the important influencing factor where 28% believe Technology should be the influencing factor.
- Cost of production and Competitors share the same support of 17%. Profit came with least point 8%.

Influencing factors to evaluate the feasibility of an idea

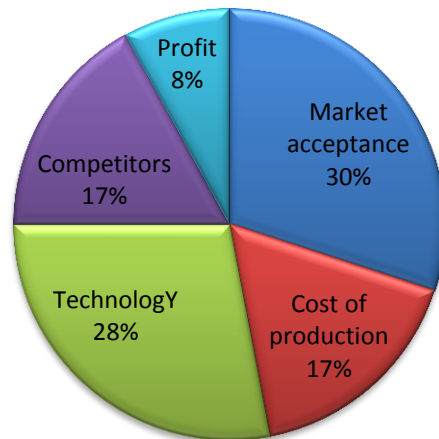


Table: 17**Estimate the approximate time and cost for a product**

Si. No.	Estimate the approximate time and cost for a product	No. of Respondents	Percentage (%)
1	Experience from similar projects carried out by us	79	79
2	Competitors cost/quote	5	5
3	Own estimate	16	16

Source: Primary data

- 79% of employees say that the experience from similar projects carried out by the company itself helps them to find the approximate time and cost estimate for a product.
- Only 16 % says that it is by their own new estimate, only around 5 % says that it is according to competitors quote.

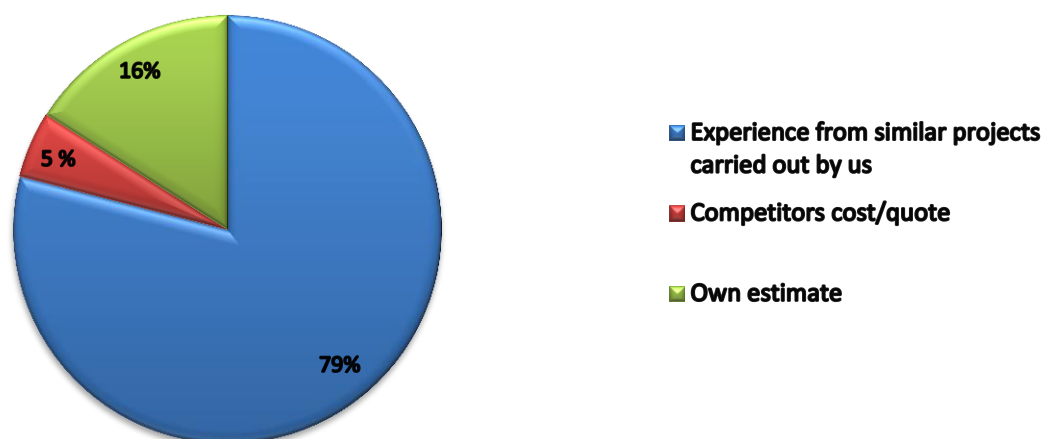
Estimate the approximate time and cost for a product

Table: 18**Methods for the detailed study on the selected ideas**

Si. No.	Methods for the detailed study on the selected ideas	No. of Respondents	Percentage (%)
1	Assigning different teams with different objectives	85	85
2	Studying reports of similar products within the company	15	15
3	Competitors reports	0	0

Source: Primary data

- Majority of employees think the method for detailed study on selected idea is assigning different teams with different objectives with a support of 85% response.
- 15% employees consider studying reports of similar products within the company itself is happening for the detailed study.
- None of them think the competitors report plays any role in this matter in the company.

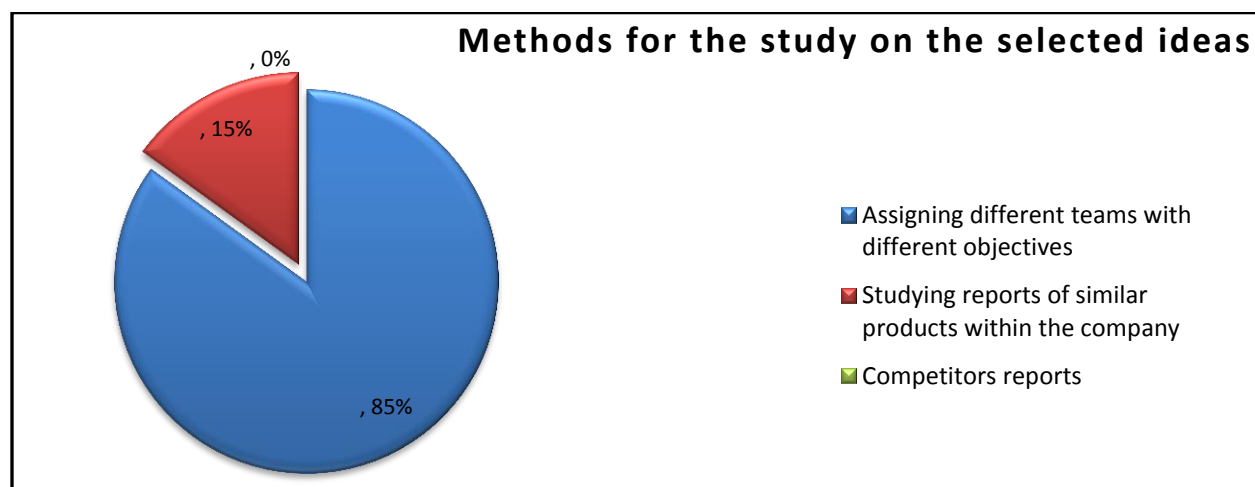
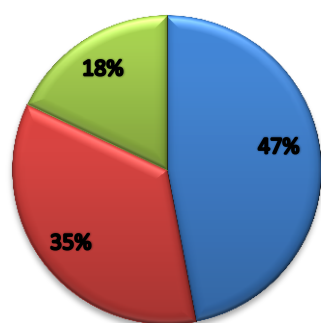


Table: 19**Preference for the process of detailed study of selected idea**

Si. No.	Preference for the process of detailed study of selected idea	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	Assigning different teams with different objectives	82	15	3	279	47
2	Studying reports of similar products within the company	13	75	12	201	35
3	Competitors reports	5	10	85	120	18

Source: Primary data

- After giving points to each rank we then converted it to percentage format.
- 47% of the score for the preferred process of detailed study of selected idea is given to assigning different team with different objective, whereas 35% of employees consider Studying reports of similar products within the company will be good. Also 22% points have given for Competitors reports.

Preference for the process of detailed study of selected idea

- Assigning different teams with different objectives
- Studying reports of similar products within the company
- Competitors Reports

Table: 20**Team leader for studying the details of selected ideas**

Si. No.	Team leader for studying the details of selected ideas	No. of Respondents	Percentage (%)
1	Member from marketing	2	2
2	Member from technical	98	98
3	Member from finance	0	0
4	Member from HR	0	0

Source: Primary data

- 98% employees think member from the technical team as the leader for studying the details of selected ideas in the company.
- Only 2% people consider a member from the marketing team lead the team. No one chose member from finance or HR.

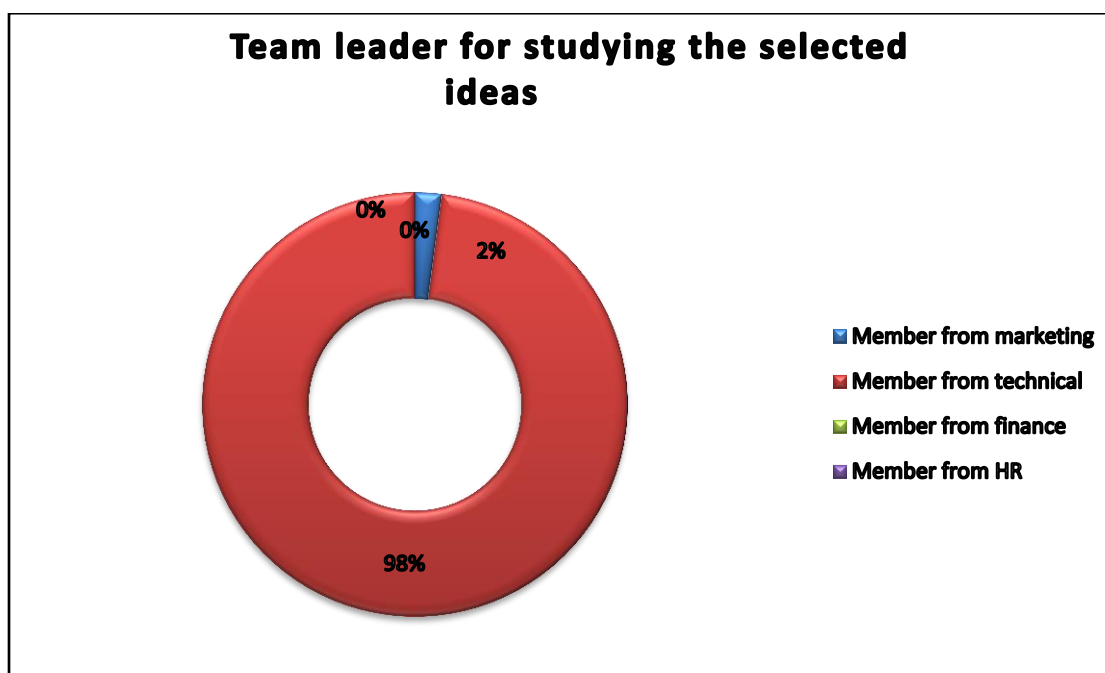


Table: 21**Preference for the important fields covered in the detailed study of a product**

Si. No.	Important fields covered in the detailed study of a product	Ranks				Points	Percentage (%)
		Rank 1	Rank 2	Rank 3	Rank 4		
1	Marketing	20	67	13	0	307	31
2	Technical	81	15	4	0	377	38
3	Finance	2	16	78	4	216	21
4	HR	0	0	5	95	105	10

Source: Primary data

- After giving points to each rank we then converted it to percentage format.
- Employees prefer both technical and marketing should be the important fields in the detailed study. They got 38 and 31% scores respectively. Finance and HR having 21 % and 10 % scores respectively.

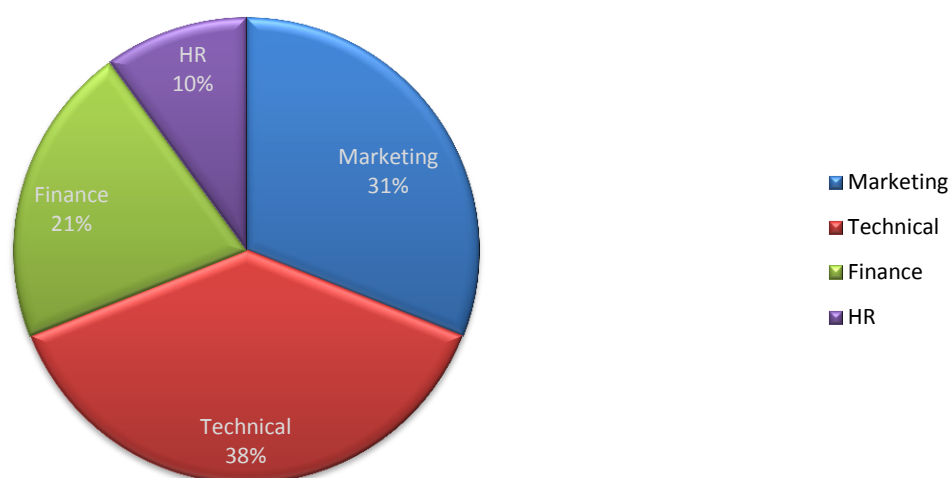
Important fields covered in the detailed study

Table: 22**Rate the process for the detailed study of an idea in the company**

Si. No.	Rate the process for the detailed study of an idea in the company	No. of Respondents	Percentage (%)
1	Very high quality	2	2
2	High quality	23	23
3	Neither high nor low quality	73	73
4	Low quality	2	2
5	Very low quality	0	0

Source: Primary data

- Majority are with the feedback that their process for the detailed study of an idea is neither very high quality nor very low quality, they believe that, it is 'neither high or nor low' quality, 73% response for that option. Also some believes it's of high quality 23% while 2% each believe the detailed study is having very high quality and having low quality.

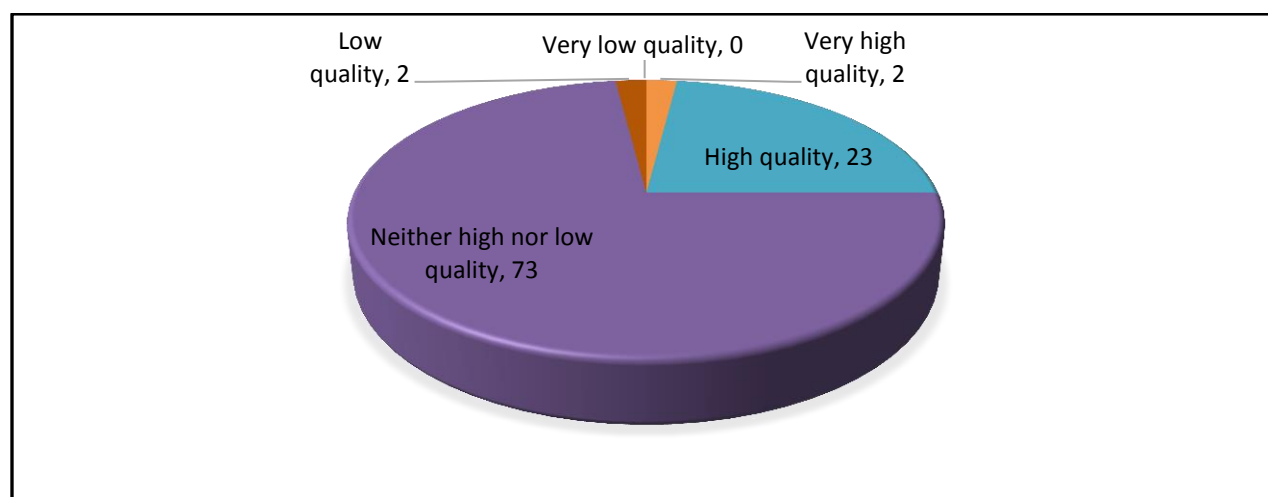


Table: 23**Teams assigned for product development in the company**

Si. No.	Teams assigned for product development	No. of Respondents	Percentage (%)
1	New team with sufficient expertise in each functionality	87	87
2	Reassigning work to a team had similar project	9	9
3	Inviting application from employees	4	4

Source: Primary data

- The team assigned for product development is by the new team with sufficient expertise in each functionality so that 87 % of employees opted this one.
- 9% employees prefer to be reassigning the work to a team had similar project, only 4 % thinks Inviting application from employees will be helpful.

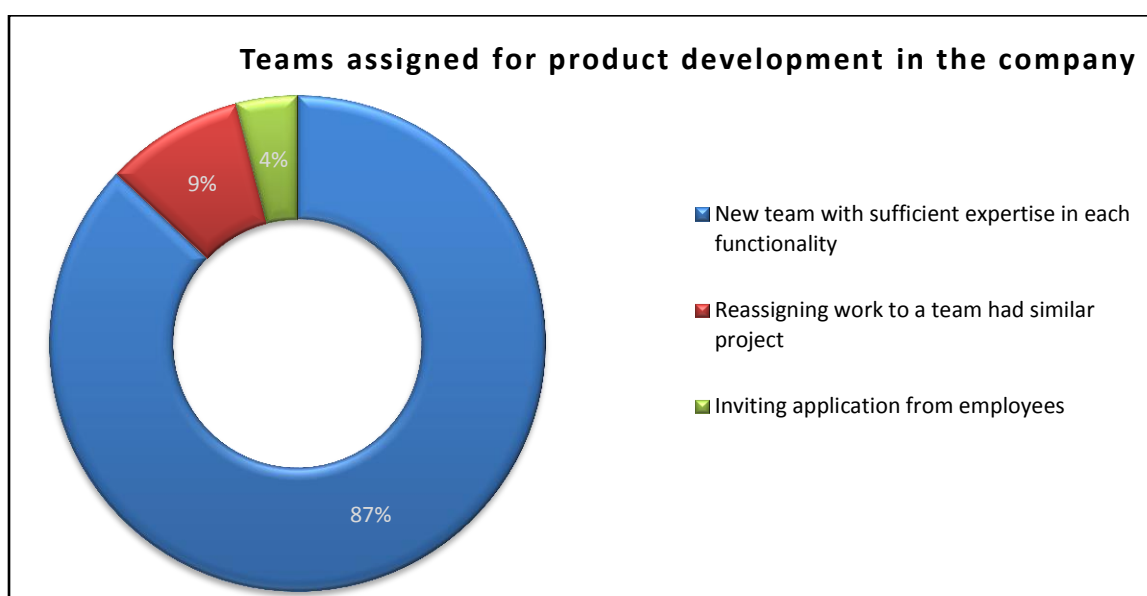


Table: 24**Rank the preference for the assignment of teams for product development**

Si. No.	Preferences	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	New team with sufficient expertise in each functionality	2	71	26	174	29
2	Reassigning work to a team had similar project	86	10	3	281	47
3	Inviting application from employees	12	19	71	145	24

Source: Primary data

- Here we are ranking the preference for the assignment of teams for product development according to employee's perspective.
- After giving points to each rank we then converted it to percentage format.
- According to this table, 47% of the points were assigned for the reassignment of work to a team had similar experience whereas 29% and 24% employees prefer new team and inviting application among employees respectively.

Analysis of product development process in a machine learning firm

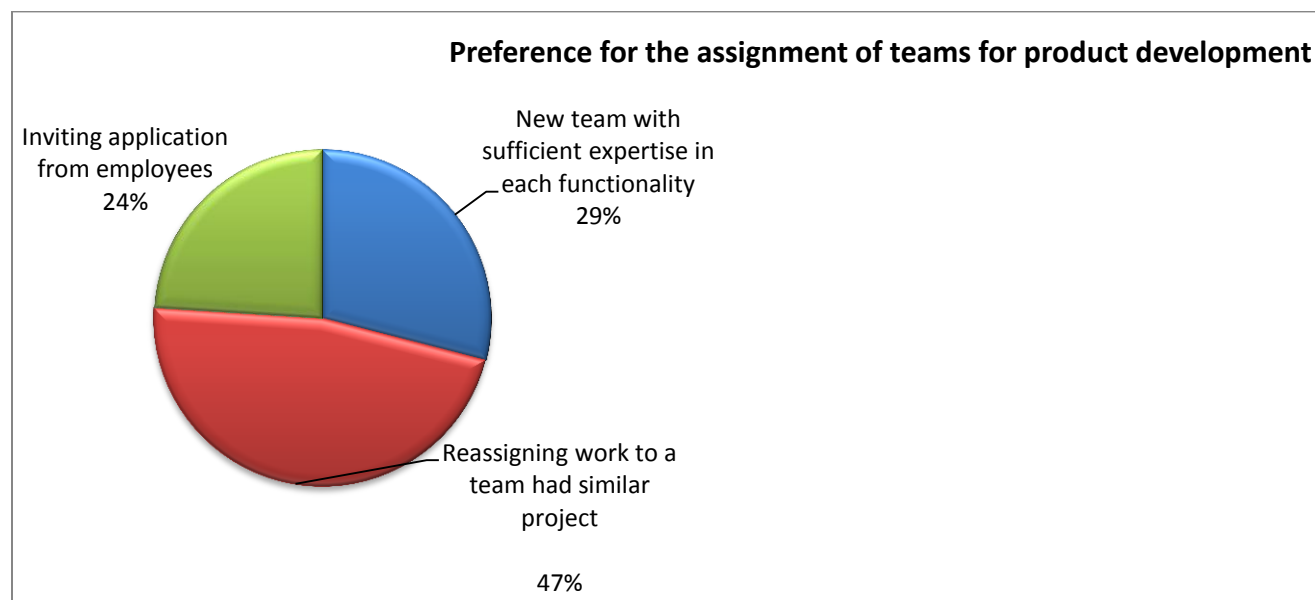


Table: 25**Serious issues in the last 6 months**

Si. No.	Issues	No. of Response	Percentage (%)
1	Once	5	5
2	Twice	10	10
3	Thrice	12	12
4	More than thrice	73	73

Source: Primary data

- 73% of the employees responded that they had more than 3 serious issues in the last 6 months.
- 12% told three times whereas 10% of the employees think that only two serious issue they had in the last 6 months. Rest 5 responded that only one issue was there.

Serious issues in the last 6 months

■ Once
 ■ Twice
 ■ Thrice
 ■ More than thrice

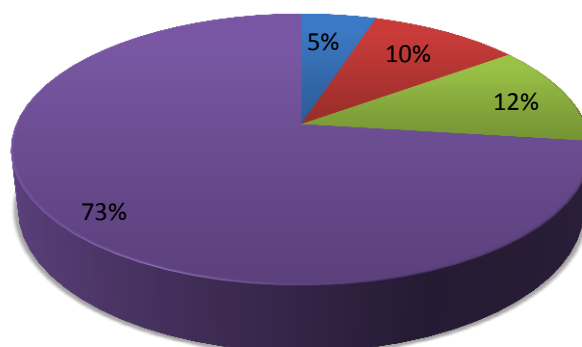


Table:26**Code review methods**

Si. No.	Support statements	No. of Response	Percentage (%)
1	Self	16	12
2	Peer	13	10
3	Lead	97	78
4	Architect	0	0

Source: Primary data

- 78% of the response is in favor of lead review whereas self and peer review is chosen by 12% and 10% responses respectively.
- It is clear from the survey that majority of respondents choose lead review since it is the most evident one in the company.

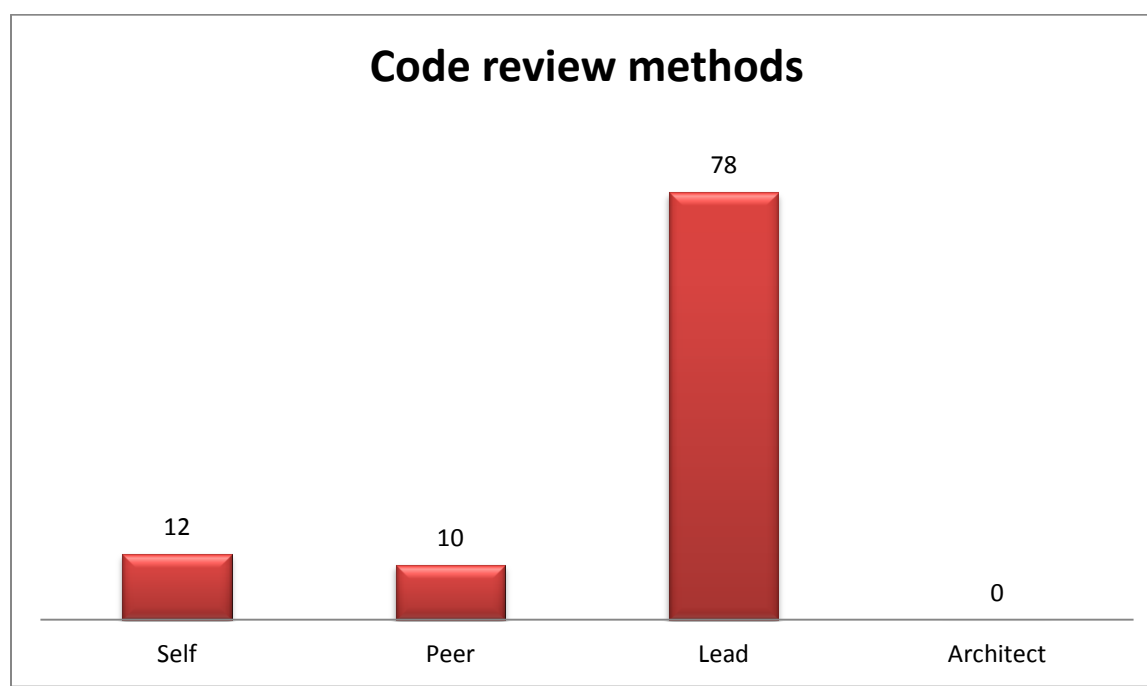


Table: 27**Capability of workforce in working in new technology**

Si. No.	Support statements	No. of Respondents	Percentage (%)
1	Definitely would	15	15
2	Probably would	85	85
3	Probably would not	0	0
4	Definitely would not	0	0

Source: Primary data

- 85% of the employees think the workforce probably would be able to work in any new technology whereas the rest 15% think that the employees definitely would be able to take any challenging new technology.

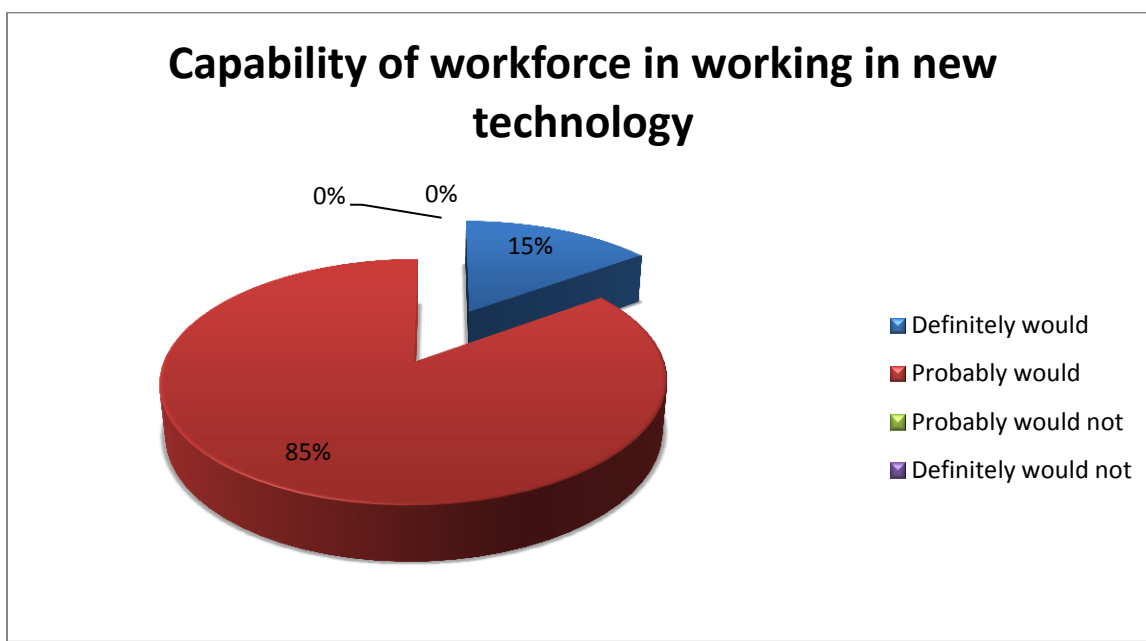


Table: 28**Positives of prototype development process**

Si. No.	Methods	No. of Respondents	Percentage (%)
1	Approach for development	80	31
2	The way in which selecting members for the development team	94	35
3	Availability of subject matter experts	88	34

Source: Primary data

- 31% of the employees think approach for development is a positive in prototype development where as 35% think team structure play a positive role and 34% of response was in support of the availability of subject matter expert.
- It is clear from the survey that around same response received for all 3 options given.

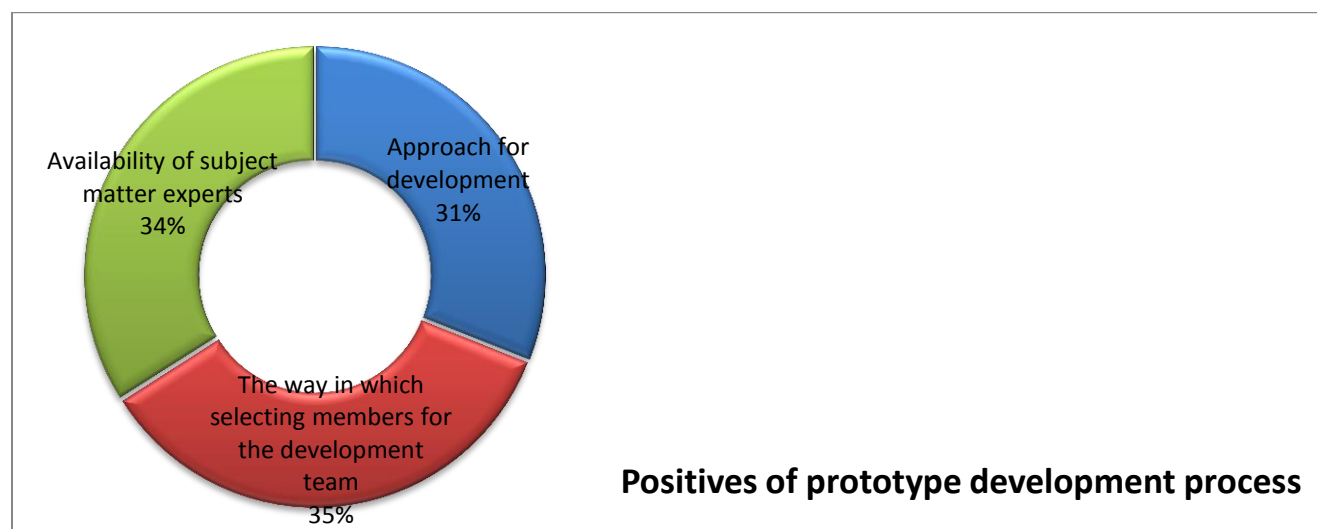


Table: 29**Rank the positives of prototype development**

Si. No.	Areas	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	Approach for development	64	14	20	240	40
2	The way in which selecting members for the development team	10	79	10	198	33
3	Availability of subject matter experts	26	7	70	162	27

Source: Primary data

- After giving points to each rank we then converted it to percentage format. According to this table, 40% of the points were assigned for the approach for development whereas nearly 33% think the tea structure for development play a positive role.
- The other 27% believe the availability of subject matter experts is the important positive in the process.



Table: 30**Rank the areas need improvement for the prototype development**

Si. No.	Areas	Ranks			Points	Percentage
		Rank 1	Rank 2	Rank 3		
1	Approach for development	21	16	63	158	26.3
2	The way in which selecting members for the development team	39	23	28	191	31.8
3	Availability of subject matter experts	40	61	9	251	41.8

Source: Primary data

- After giving points to each rank we then converted it to percentage format. According to this table, 20% of the points were assigned for the approach for development whereas nearly 24% think the team structure for development play a positive role.
- The other 29% believe the availability of subject matter experts is the important positive in the process.

**Areas need improvement
for the prototype development**

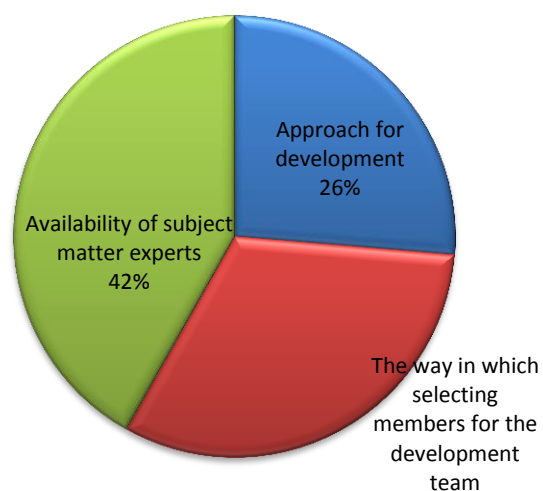


Table: 31**Assigning the task of unit testing**

Si. No.	Method	No. of Respondents	Percentage (%)
1	Developer along with the development	90	90
2	Assigning to a separate team member	10	10

Source: Primary data

- According to this table, 90% of the employees think developer itself do the unit testing along with the development where as 10% think a separate team member do the needful unit testing.

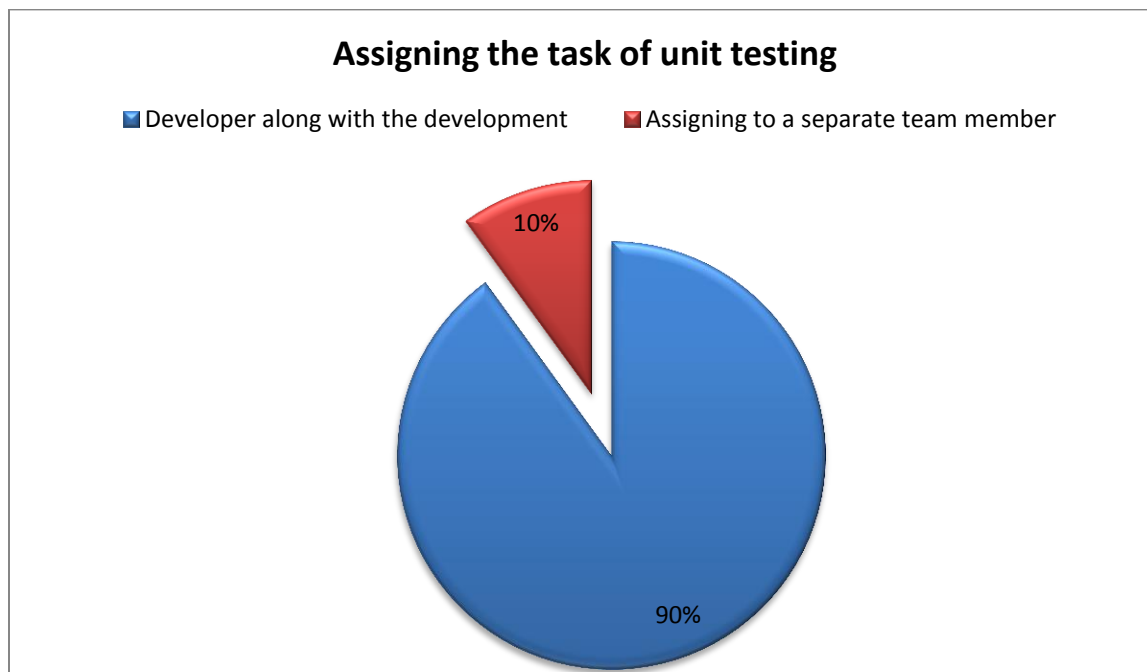


Table: 32**Process of requirement gathering for testers**

Si. No.	Support statement	No. of Respondents	Percentage (%)
1	Developer	31	31
2	Self-learning from design and requirement documents	10	10
3	Business Analysts and Developers	59	59

Source: Primary data

- 59% of the employees think testers gather information about requirement from BA and developer while 31% believe only the source is developer.
- Rest 10% think they are gathering details from design and requirement documents.

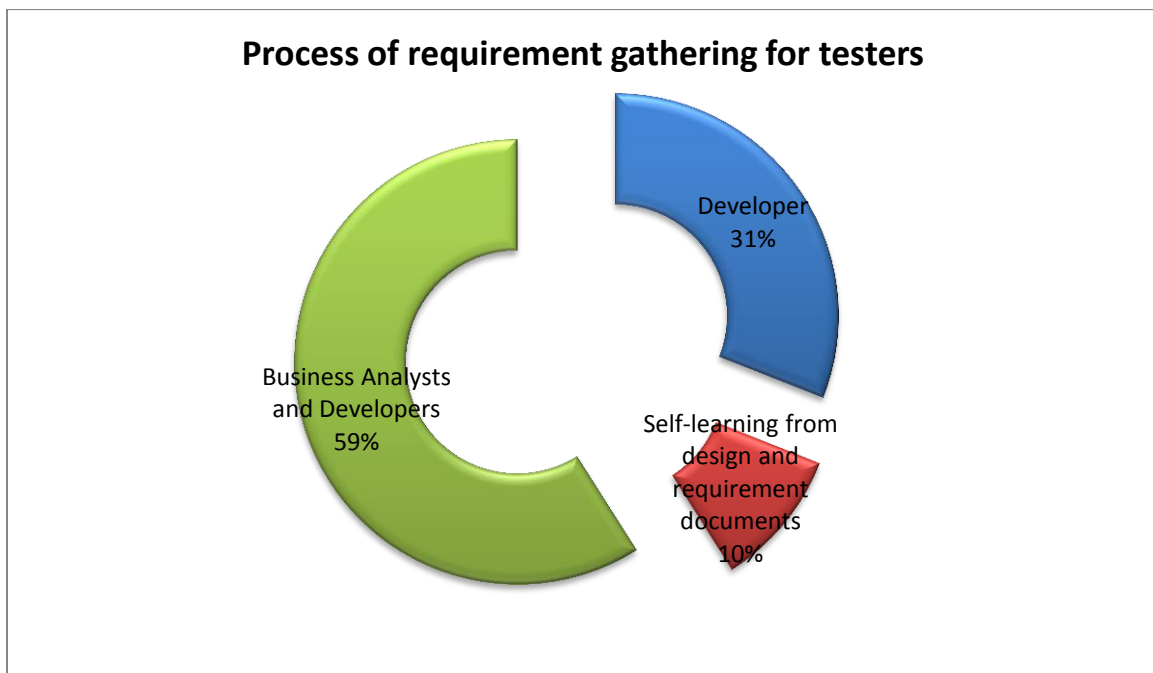


Table: 33**Effectiveness of assigning testing to external companies**

Si. No.	Support statement	No. of Respondents	Percentage (%)
1	Strongly Agree	8	8
2	Somewhat Agree	80	80
3	Neither Agree nor disagree	12	12
4	Somewhat disagree	0	0
5	Strongly Disagree	0	0

Source: Primary data

- 88% of the Employees either strongly or somewhat agree the idea of giving testing task to any external companies.
- Other 12% neither agree nor disagree the same idea.

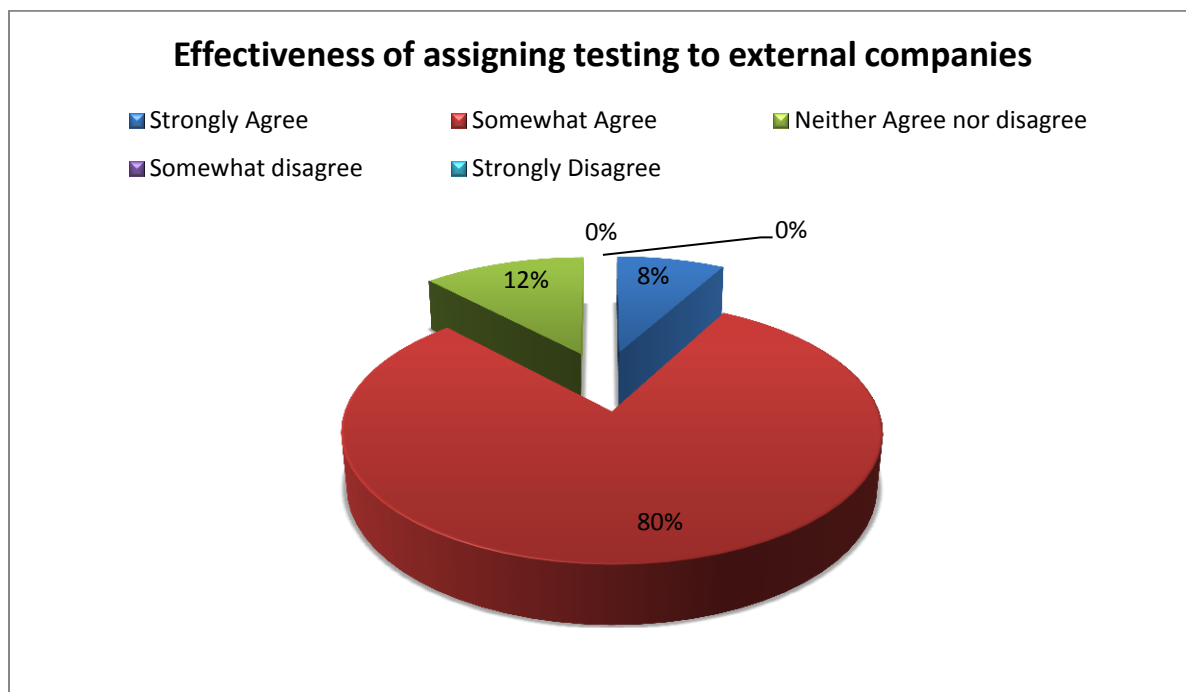


Table: 34**Positives of overall testing process in your company**

Si. No.	Areas	No. of Respondents	Percentage (%)
1	Enough allocation of time for testing	86	77
2	Detailed review for Unit testing, SIT and UAT results	22	19
3	Testing is done by other company	0	0
4	Very active user base to check detailed user acceptance	3	4

Source: Primary data

- 77% of the Employees think enough allocation of time for the testing is the best thing about the testing process in the company where as other 19% support the detailed review process off all unit testing, SIT and UAT results.
- Other 4% believe the active user base of the company helps in the proper testing of products from the company.

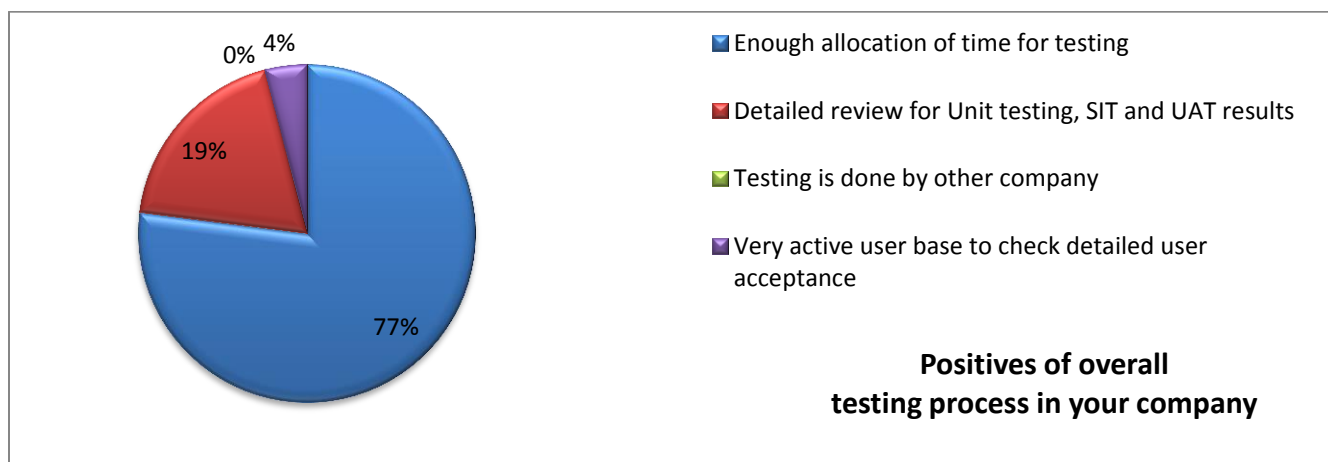
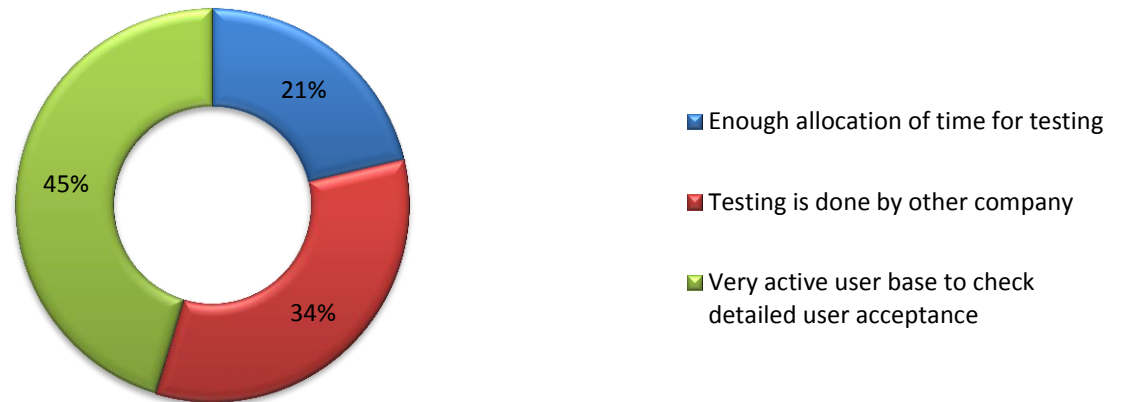


Table: 35**Rank the areas those need improvements for the overall testing process**

Si. No.	Areas	Ranks				Points	Percentage
		Rank 1	Rank 2	Rank 3	Rank 4		
1	Enough allocation of time for testing	10	5	25	60	165	17
2	Detailed review for Unit testing, SIT and UAT results	20	40	25	15	265	27
3	Testing is done by other company	20	15	40	25	230	23
4	Very active user base to check detailed user acceptance	50	40	10	0	340	33

Source: Primary data

- After giving points to each rank we then converted it to percentage format.
- According to this table, 33% of the points were assigned for the need of more active user base to check the user acceptance testing where as 23% are for the improvements in assigning external companies for the testing and among the rest, 27% is for detailed review and 17% is for allocating more time for testing.

Rank the areas those need improvements for the overall testing process

Correlation matrixes

A correlation matrix is a table showing correlation coefficients between variables. Each cell in the table shows the correlation between two variables. Each random variable (X_i) in the table is correlated with each of the other values in the table (X_j). This allows you to see which pairs have the highest correlation and possible to make interpretation connecting variables from different tables.

Sample table: Machine learning domain Vs Designation

Si. No.	Machine learning domain	C (D)		
		#	%	%
1	A(B)	E	F	G

A – Type of machine learning domain from table 6

B – Total number of response for 'A' from table 6

C – Type of designation from table 1

D – Total number of response for 'C' from table 1

E – Number of response for option 'A' by employees with designation 'C'

F – Percentage figure of 'E' out of 'B'

G – Percentage figure of 'E' out of 'D'

Table: 36**Machine learning domain Vs Designation**

Si. No.	Machine learning domain	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Computer vision (82)	5	6	100	8	10	80	2	2	66	58	71	83	9	11	75
2	Genomics (83)	3	4	60	9	11	90	2	2	66	59	71	84	10	12	83
3	Chatbots (86)	4	5	80	9	10	90	0	0	0	66	77	94	7	8	58
4	Natural language processing (80)	2	2	40	7	9	70	0	0	0	63	79	90	8	10	66
5	Trend analysis (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- Out of 5 senior managers all have chosen computer vision.
- Trend analysis is not opted by anyone with any designation.
- The distribution of technical staff across all domains except trend analysis is almost same.
- Around 70% to 80% of all domains except trend analysis is chosen by technical staff.

Table: 37**Machine learning domain Vs Total industry experience**

Si. No.	Machine learning domain	Below 2 years(26)			3 – 5 years(31)			5 – 10 years(25)			10 – 20 years(18)			Above 20 years(0)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Computer vision (82)	22	28	85	25	30	81	20	24	80	15	18	83	0	0	0
2	Genomics (83)	21	25	81	26	32	84	21	25	84	15	18	83	0	0	0
3	Chatbots (86)	23	27	88	28	33	90	17	20	68	18	20	100	0	0	0
4	Natural language processing (80)	19	24	73	26	33	84	16	20	64	17	23	94	0	0	0
5	Trend analysis (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- Trend analysis is not chosen by any of the employees.
- All of the senior employees with 10 to 20 years of experience have opted the option Chatbot.

Table: 38**Kind of problems attended generally Vs Designation**

Si. No.	Machine learning domain	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Existing problems (84)	5	6	100	9	10	90	2	2	67	58	69	83	10	12	83
2	New problem (16)	0	0	0	1	6	10	1	6	33	12	75	17	2	13	17

Data analysis

- All of the senior managers with broader knowledge across all domains prefer existing problem
- Minority of both technical and non-technical staff responded with the option 'New problem'

Table: 39**Idea generation source Vs Total industry experience**

Si. No.	Idea Generation Source	Below 2 years (26)			3 – 5 years (31)			5 – 10 years (25)			10 – 20 years(18)		
		#	%	%	#	%	%	#	%	%	#	%	%
1	From employees (3)	2	67	8	1	33	3	0	0	0	0	0	0
2	Client/Customer (17)	8	47	30	7	41	23	1	6	4	1	6	6
3	Market study reports/external agencies (80)	16	20	62	23	29	74	24	30	96	17	21	94

Data analysis

- 80% of the employees irrespective of experience consider the market study report as the main source of ideas for the company.
- Less experienced employees seems to believe that the company considers their opinions and suggestions for the idea generation while senior employees think opposite.
- None of the employees above 5 years of experience actually believe their suggestions are taken to higher lever of idea pooling.
- Irrespective of the experience, employees believe client requirement has considerable role in fixing the ideas for product.

Table: 40**Employee's participation in idea generation Vs Designation**

Si. No.	Employee's participation in idea generation	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff (12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Always advisable (65)	4	6	80	8	12	80	3	5	100	42	65	60	8	12	67
2	Usually advisable (25)	1	4	20	2	8	20	0	0	0	20	80	29	2	8	17
3	Sometimes advisable (10)	0	0	0	0	0	0	0	0	0	8	80	11	2	12	16
4	Rarely advisable (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Never advisable (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- Irrespective of the designation majority of the employees support the concept of receiving idea from employees.
- Even though majority is in favor of this concept, 11% and 16% of technical and non-technical staff respectively responded the concept as only sometimes advisable.

Table: 41**Rating the idea generation process Vs Designation**

Si. No.	Rating the idea generation process	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Very high quality (68)	5	7	100	8	12	80	3	4	100	43	63	62	9	13	75
2	High quality (23)	0	0	0	2	9	20	0	0	0	19	83	27	2	9	17
3	Neither high nor low quality (9)	0	0	0	0	0	0	0	0	0	8	89	11	1	11	8
4	Low quality (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Very low quality (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- Majority of the employees irrespective of designation consider the idea generation process as either 'high' or 'very high quality'.
- Out of 5 senior managers all of them rated idea generation process as very high quality
- 89% of the employees rated idea generation as neither high nor low is from technical staff.

Table: 42**Team assignment for the feasibility study Vs Total industry experience**

Si. No.	Team assignment for the feasibility study	Below 2 years(26)			3 – 5 years(31)			5 – 10 years(25)			10 – 20 years(18)			Above 20 years(0)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Member from marketing (2)	2	100	100	0	0	0	0	0	0	0	0	0	0	0	0
2	Member from technical (98)	24	24	92	31	32	100	25	26	100	18	18	100	0	0	0
3	Member from finance (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Member from HR (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- Irrespective of the experience, employees believe technical team has greater voice in the feasibility study.
- Only two of the very junior employees responded with the option marketing

Table: 43**Idea selection process Vs Designation**

Si. No.	Idea selection process	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	By voting majority (9)	0	0	0	0	0	0	0	0	0	8	89	11	1	11	8
2	Consensus after discussion (91)	5	5	100	10	11	100	3	3	100	62	68	89	11	12	92
3	Analytical (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- 91% of the employees are in prefer making decision by consensus after discussion
- Only few from technical(11 %) and non-technical (8%) responded with the voting option.
- All of the managers including senior, technical and non-technical support consensus only.
- None of the employees think company follow analytical method in idea selection

Table: 44**Methods for the detailed study on the selected ideas Vs Designation**

Si. No.	Methods for the detailed study on the selected ideas	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Assigning different teams with different objectives (85)	5	6	100	10	12	100	3	6	100	56	66	80	11	13	92
2	Studying reports of similar products within the company (15)	0	0	0	0	0	0	0	0	0	14	93	20	1	7	8
3	Competitors reports (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- All of the managers including senior, technical and non-technical fully supported the option assigning teams with objectives.
- 20% and 8% technical and non-technical employees respectively think the company does detailed study of the idea by studying reports from similar project.
- None think competitors report play any role in the process.

Table: 45**Team leader for studying the details of selected ideas Vs Designation**

Si. No.	Team leader for studying the details of selected ideas	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Member from marketing (2)	0	0	0	0	0	0	0	0	0	2	100	3	0	0	0
2	Member from technical (98)	5	5	100	10	10	100	3	3	100	68	69	97	12	12	100
3	Member from finance (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Member from HR (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- 98% of people think that the technological teams influence greatly in the detailed study of ideas.
- Irrespective of designation people support the option technical team

Table: 46**Teams assigned for product development in the company Vs Total industry experience**

Si. No.	Teams assigned for product development in the company	Below 2 years(26)			3 – 5 years(31)			5 – 10 years(25)			10 – 20 years(18)			Above 20 years(0)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	New team with sufficient expertise in each functionality (87)	19	22	73	25	29	81	25	28	96	18	21	100	0	0	0
2	Reassigning work to a team had similar project (9)	4	44	15	5	56	16	0	0	0	0	0	0	0	0	0
3	Inviting application from employees (4)	3	75	12	1	25	3	0	0	0	0	0	0	0	0	0

Data analysis

- Across all experience, especially employees more than 3 years of think that the company always prefers for constituting new team with expertise in each functionality.
- Very few of the employees with experience less than 5 years chosen reassigning to similar project and inviting employees preferences

Table: 47**Serious issues in the last 6 months Vs Designation**

Si. No.	Serious issues in the last 6 months	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Once (5)	0	0	0	0	0	0	0	0	0	4	80	6	1	20	8
2	Twice (10)	0	0	0	0	0	0	0	0	0	9	90	13	1	10	8
3	Thrice (12)	0	0	0	0	0	0	0	0	0	12	100	17	0	0	0
4	More than thrice (73)	5	7	100	10	14	100	3	4	100	45	62	64	10	14	84

Data analysis

- All managers fully admitted that they have encountered more than 3 mistakes for a prototype in the last 6 months.
- But junior designated people have response ranging from 1 to more than 3.

Table: 48**Capability of workforce in working in new technology Vs Designation**

Si. No.	Capability of workforce in working in new technology	Senior Manager (5)			Technical Manager (10)			Non-Technical Manager (3)			Technical staff (70)			Non-technical staff(12)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Definitely would (15)	0	0	0	0	0	0	0	0	0	13	87	19	2	13	17
2	Probably would (85)	5	6	100	10	12	100	3	4	100	57	67	81	10	12	83
3	Probably would not (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Definitely would not (0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Data analysis

- About 85% of the employees believe that workforce probably would capable of working in any new technology.
- Majority of the employees (87%) who responded definitely are from technical staff.

Table: 49**Process of requirement gathering for testers Vs Total industry experience**

Si. No.	Process of requirement gathering for testers	Below 2 years(26)			3 – 5 years(31)			5 – 10 years(25)			10 – 20 years(18)			Above 20 years(0)		
		#	%	%	#	%	%	#	%	%	#	%	%	#	%	%
1	Developer (31)	13	42	50	12	39	39	6	19	24	0	0	0	0	0	0
2	Self-learning from design and requirement documents (10)	5	50	19	3	30	10	2	20	8	0	0	0	0	0	0
3	Business Analysts and Developers (59)	8	14	31	16	27	52	17	29	68	18	31	100	0	0	0

Data analysis

- Majority of the employees especially those who are above 5 years of experience think that the testing team gather information about the requirement from either BA or developer.
- Employees below 5 years of experience chose developer alone
- It seems a clear consensus is absent for this question among employees.

Gap analysis

Strategic tool used for analyzing the gap between the target and anticipated results, by assessing the extent of the task and the ways, in which gap might be bridged. It involves making a comparison of the present performance level of the entity or business unit with that of standard established previously.

Sample table

Si. No.	Current idea generation source	Preference for the idea generation methods →	B
		%	D
1	A	C	E

A – Type of idea generation (Current)

B – Type of idea generation (Employees preference)

C – Percentage of response for A

D – Percentage of response for B

E – Gap between the current and preference

$$\text{i.e. } E = C - D$$

Table: 50**Current idea generation source Vs Preference for the idea generation methods**

Si. No.	Current idea generation source ↓	Preference for the idea generation methods →	From employees	Client/Customer	Market study reports/external agencies
		%	0	9	91
1	From employees	3	-3	6	88
2	Client/Customer	17	-17	-8	74
3	Market study reports/external agencies	80	-80	-71	11

Data analysis

- As per the table it is clear that the employees have an interest in raising the dependency on market study/external agency report for the idea generation with a raise of 11% from the current dependency of 80%.

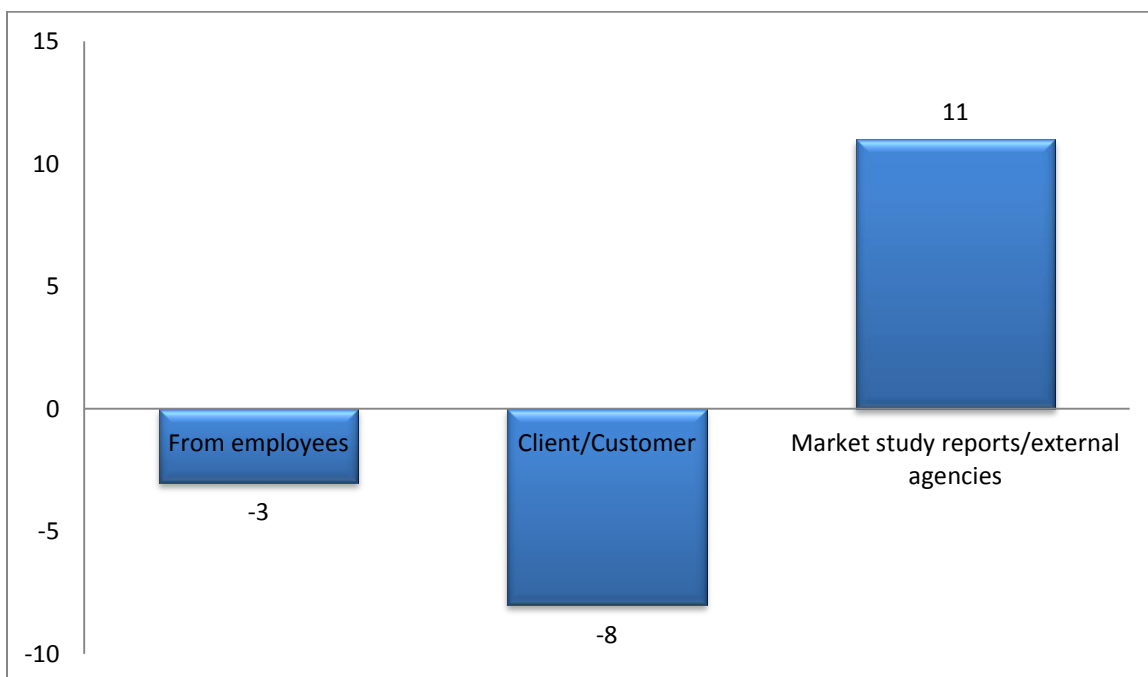


Table: 51

**Current team assignment for the feasibility study vs. preferred team assignment
for the feasibility study methods**

Si. No.	Current idea team assignment for the feasibility ↓	Preferred team assignment for the feasibility study methods →	Senior level management team	Pre designated team for all feasibility study	Random selection of members from each departments
		%	79	21	0
1	Senior level management team	83	-4	-62	-83
2	Pre designated team for all feasibility study	17	62	4	-17
3	Random selection of members from each departments	0	79	21	0

Data analysis

- Employees still prefer senior level management for the feasibility study.
- There is raise of 4% in the interest for pre designated team for the feasibility study.
- The concept of Random selection of team members not prefers by anyone

Analysis of product development process in a machine learning firm

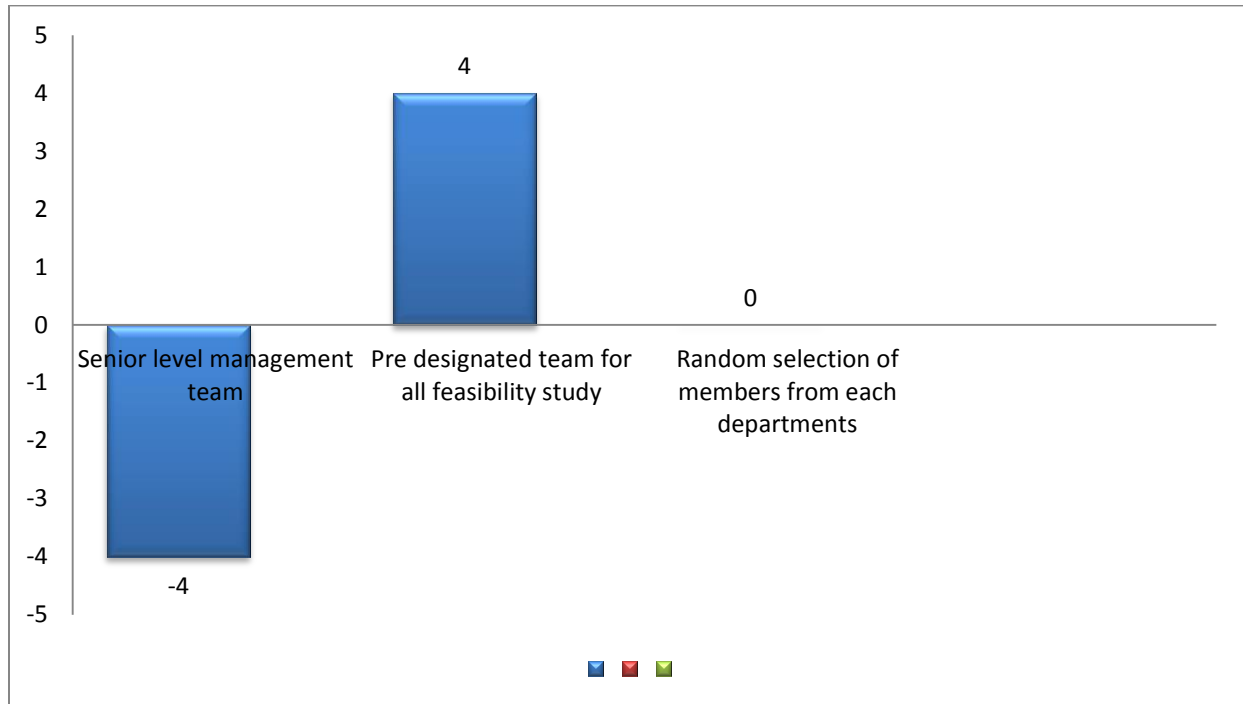


Table: 52**Current Idea selection process vs. Preferred Idea selection process**

Si. No.	Current Idea selection process ↓	Preferred Idea selection process →	By voting majority	Consensus after discussion	Analytical
		%	8	81	8
1	By voting majority	9	-1	72	-1
2	Consensus after discussion	91	-83	-10	-83
3	Analytical	0	8	81	8

Data analysis

- Current idea selection processes such as consensus and voting majority remains what employee's preference.
- But it is clear that, about 8% employees think it is time to start experimenting some analytical method to do the idea selection

Analysis of product development process in a machine learning firm

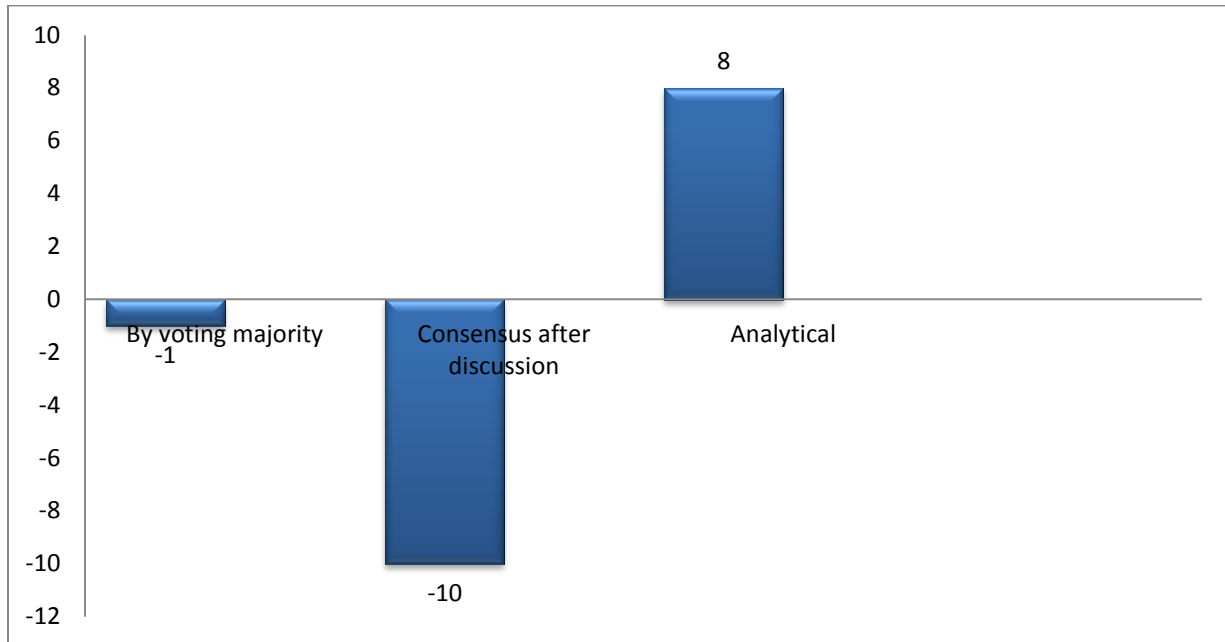


Table: 53

**Current process of team assignment for product development vs. preferred
process of team assignment for product development**

Si. No.	Current process of team assignment for product development ↓	Preferred process of team assignment for product development →	New team with sufficient expertise in each functionalit y	Reassigning work to a team had similar project	Inviting application from employees
		%	2	85	12
1	New team with sufficient expertise in each functionality	87	-85	-2	-75
2	Reassigning work to a team had similar project	9	-7	76	3
3	Inviting application from employees	4	-2	81	8

Data analysis

- Primary preference for the team assignment for the product development is very different from how they do it now. Currently they are assigning the work to a new team with expert from different specialties. The interest for the same idea reduced to 2 from the current response of 87.
- The concept of assigning work to same team raised from 9% to 85.

Analysis of product development process in a machine learning firm

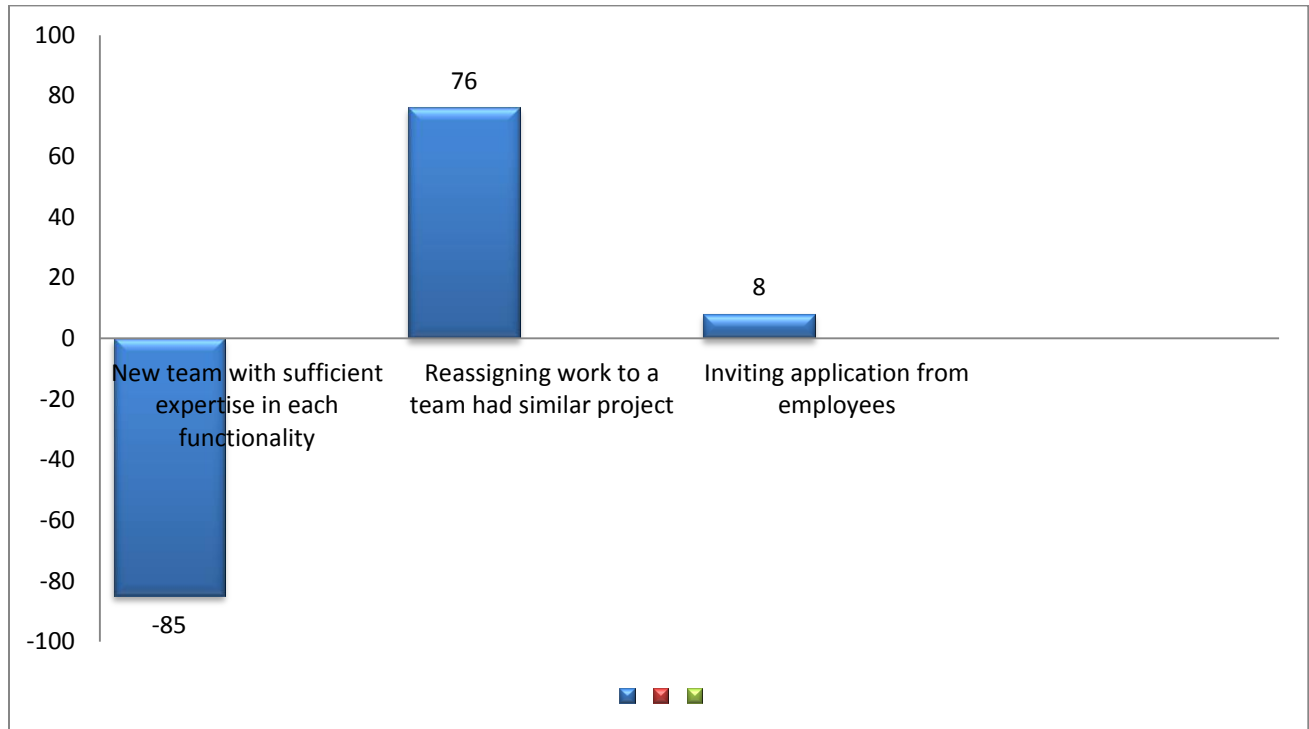


Table: 55**Positives of prototype development vs. necessary improvements for the prototype development**

Si. No.	Positives of prototype development ↓	Necessary improvements for the prototype development →	Approach for development	The way in which selecting members for the development team	Availability of subject matter experts
		%	26	32	42
1	Approach for development	40	-14	-8	2
2	The way in which selecting members for the development team	33	-7	-1	9
3	Availability of subject matter experts	27	-1	5	15

Data analysis

- Almost 15% employees prefer to improve the availability of subject matter experts in the prototype development.
- The employees are fine with the way in which team selection happens now.
- Since approach for the development is fine now, employees don't see any major need for its improvements.

Analysis of product development process in a machine learning firm

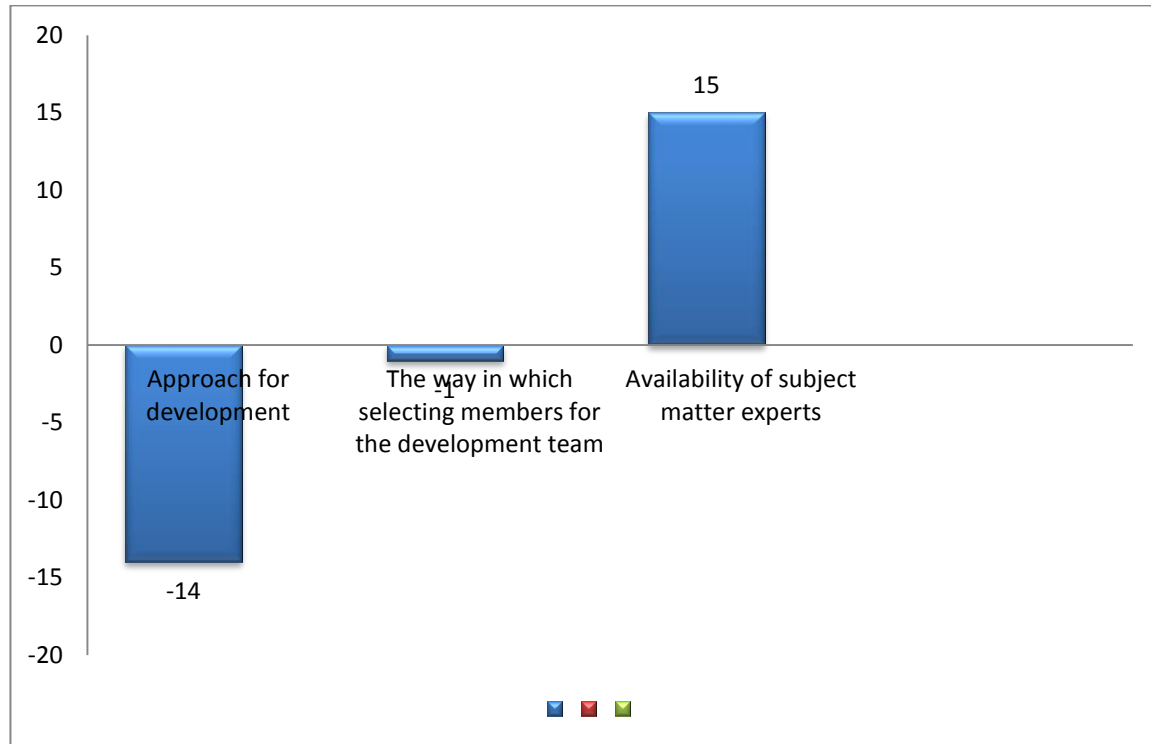


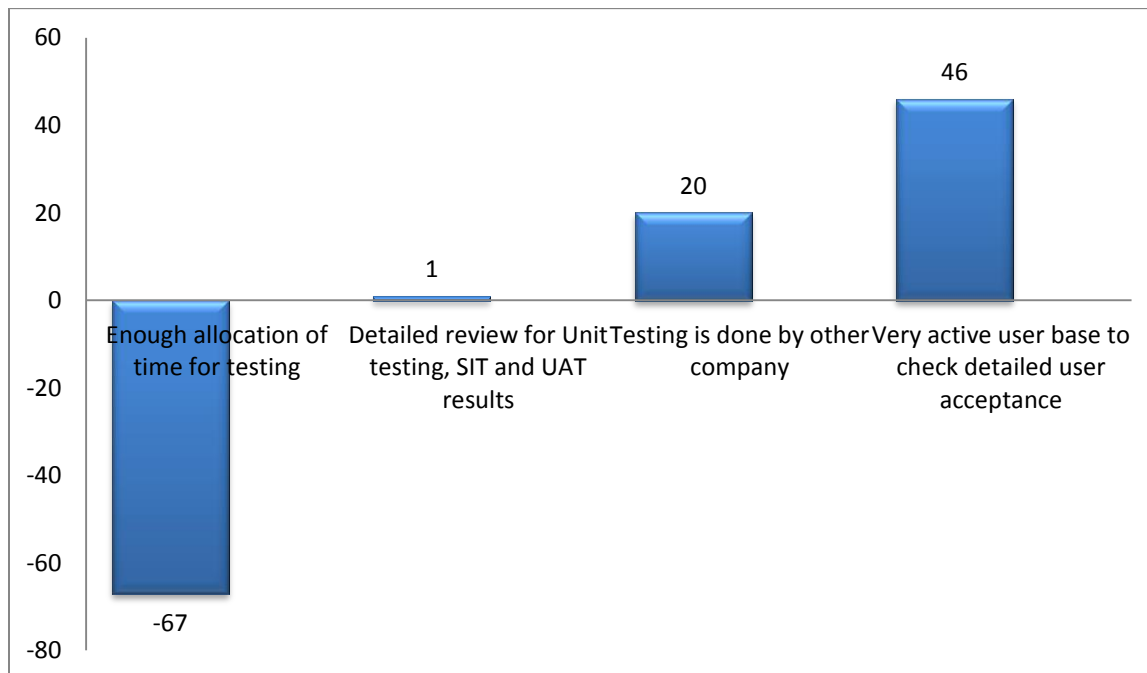
Table: 56

**Positives of overall testing process in your company
vs. necessary improvements for the overall testing**

Si. No.	Positives of overall testing process in your company ↓	Necessary improvements for the overall testing →	Enough allocation of time for testing	Detailed review for Unit testing, SIT and UAT results	Testing is done by other company	Very active user base to check detailed user acceptance
		%	10	20	20	50
1	Enough allocation of time for testing	77	-67	-57	-57	-27
2	Detailed review for Unit testing, SIT and UAT results	19	-9	1	1	31
3	Testing is done by other company	0	10	20	20	50
	Very active user base to check detailed user acceptance	4	-71	16	16	46

Data analysis

- Employees prefer to have strong user base to properly test their product. It is a very weak section of the firm. As a product based company it is always important to have strong user base to test the product.
- It seems Time and review process is sufficiently good, so that nobody wish it to be an improvement.
- Assigning testing to other companies also need to consider as there is a 20% raise for its improvement interest



CHAPTER 4

FINDINGS, SUGGESTIONS AND CONCLUSION

FINDINGS

- Organization should start work in domain like Trend analysis.
- Organization is giving high weightage to technology than factors like marketing and profit while making decisions.
- The non-technical staff do not have a proper idea of the process
- There is always a difference in the patter of response based on industry experience. The response from employees with less experience is heavily scattered and they have a tendency to opt some exceptional choices none other from the organization have chosen.
- The organization concentrates mainly on solving existing problem available, however they are not reluctant to deal with new problems
- Less experienced employees believe that the company seriously considers their opinions and suggestions for the idea generation while senior employees think opposite.
- The organization values ideas from the employees for new product and a service which is not seemed to happen often in company. Especially majority of the senior managers also supported the concept of idea pooling from employees.

- Decisions are taken with consensus
- The organization does not rely on competitor's technology or estimates to handle issues
- The responders were reluctant to share the faults in the process of solving problems
- The employees believe that workforce is capable of working in any new technology.
- Employees prefer to trust more on external agency report rather than other sources for idea generation.
- The preferred team assignment for the product development is very different from what they have currently. Employees are more comfortable with staying in the same team rather than the current process of moving to new team for each project.
- The organization doesn't have strong user base to properly test their product.

SUGGESTIONS

- Company need to expand their domains to experiment with machine learning technology.
- The focus of the business should include marketing, finance and profit.
- More interaction between technical and non-technical employees is desirable.
- The organization should take analytical methods to take decisions which can result in more radical and rational decisions.
- It is very important to choose the right evaluation method to test the models. Organization should make the testing process more structured.
- Company should seriously consider the possibility of starting a pre designated team for the feasibility studies
- Upgrading the tech experience for the existing employees and hiring more experienced employees is desirable.
- To strengthen the feedback mechanism from end user, rectify the problem with user base issue

CONCLUSION

Data has become a valuable resource, and it's cheaper than ever to capture and store. Through the use of artificial intelligence, specifically Process-Based Machine learning, manufacturers can use data to significantly impact their bottom line by greatly improving production efficiency, product quality, and employee safety. The opportunities for companies working on machine learning product and services are immense. In this competitive world providing proper satisfaction for the customer is the key to success. Having an efficient development system is necessary for constantly producing such quality products.

This research work undertaken in Trivandrum for Curvelogics advanced technology solutions Pvt. Ltd intends to have an overall analysis of the product development at Curvelogics advanced technology solutions Pvt. Ltd in order to understand the process and recommend methods to increase the overall efficiency of the process and enhancing customer satisfaction.

From the study it is clear that there are fields in company need improvement for a better product development process. With an excellent workforce that is adaptable to the changes and ready to learn new technologies, Curvelogics has a high scope in achieving better customer satisfaction with implementing necessary changes in the system.

BIBLIOGRAPHY

1. Beckley, J. H., Paredes, D., & Lopetcharat, K. (2012). Product innovation toolbox: a field guide to consumer understanding and research. John Wiley & Sons.
2. Bergson, H. (1911). Creative evolution. Cosmino Publications
3. Brown, T. (2008). Design thinking. Harvard Business Review
4. Chan Kim, W., & Mauborgne, R. (2005). Value innovation: a leap into the blue ocean. Journal of Business Strategy
5. Cooper, R. G. (2000). Product innovation and technology strategy. Research Technology Management
6. Cooper, R. G., & Edgett, S. J. (2010). Developing a product innovation and technology strategy for your business. Research-Technology Management
7. Ulrich, K. T., & Eppinger, S. D. (1995). Product design and development (5th Ed.)
8. Al-Jarrah, Omar Y. et al. "Efficient Machine Learning for Big Data: A Review." Big Data Research 2 (2015): 87-93.
9. Russel, S. and Norvig, P. (2003). Artificial Intelligence: A Modern Approach. 2nd Edition. New York: Prentice-Hall.
10. Gallaughier, J. (2012). Understanding Software: A Primer For managers. Information Systems: A Manager's Guide to Harnessing Technology (1.4 ed.). Irvington, NY, United States: Flatworld Knowledge, Inc.
11. Leau, Y. B., Loo, W. K., Tham, W. Y., & Tan, S. F. (2012). Software Development Life Cycle AGILE vs Traditional Approaches. In 2012 International Conference on Information and Network Technology (ICINT 2012) IPCSIT (Vol. 37).
12. McMurtrey, M. (2013). A Case Study of the Application of the Systems Development Life Cycle (SDLC) in 21st Century Health Care: Something Old, Something New?. Journal of the Southern Association for Information Systems,1.
13. Agogino, A.M., A.S. Almgren. 1987. Techniques for integrating qualitative reasoning and symbolic computation in engineering optimization. Engrg. Optim. 12 117-135.

14. Ali, A., M.U. Kalwani, D. Kovenock. 1993. Selecting product development projects: Pioneering versus incremental innovation strategies. *Management Sci.* 39(March) 255-274. Allen, T.J. 1977. *Managing the Flow of Technology*. M.I.T. Press, Cambridge, MA
15. Bayus, B.L., S. Jain, A.G. Rao. 1997. Too little, too early: Introduction timing and new product performance in the personal digital assistant industry. *J. Marketing Res.* 34(Feb.) 50-63.
16. Bhattacharya, Shantany, V. Krishnan, Vijay Mahajan. 1998a. Capturing the effect of technology improvements on product positioning and introduction sequence decision-making. Working paper, The University of Texas at Austin Department of Management, Austin, TX.
17. Website: <https://towardsdatascience.com/introduction-to-machine-learning-db7c668822c4>.
18. Henderson, R.M., K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Admin. Sci. Quart.* 35(Mar.) 9-30.
19. Kalish, S., G.L. Lilien. 1986. A market entry timing model for new technologies. *Management Sci.* 32(Feb.) 194-205. , V. Mahajan, E. Muller. 1995. Waterfall and sprinkler new product strategies in competitive global markets. *Internat. J. Res in Marketing* 12(July) 105-119.
20. Kalyanaram, G., V. Krishnan. 1997. Deliberate product definition: Customizing the product definition process. *J. Marketing Res.* 34(May) 276-285.
21. Website: <https://in.mathworks.com/discovery/machinelearning.html>
22. Website: *The Systems Development Life Cycle Basics*. (n.d.). McGraw Hill Connect. Retrieved on July 12, 2013 from <http://highered.mcgraw-hill.com/sites/dl/free/0073376736/596593/AppendixD.pdf>
23. Website : <http://www.curvelogics.com>
24. Annual report from Curvelogics advanced technology solutions Pvt. Ltd
25. Client purchase agreements from Curvelogics advanced technology solutions Pvt. Ltd

CHAPTER 5

APPENDIX

QUESTIONNAIRE

This study is an attempt to analyze the development process of a machine learning product at Curvelogics advanced technology solutions Pvt. Ltd, Trivandrum. The study critically analyses the importance of structured process and recommend methods to bring down the time and cost involved and to enhance the quality of the product and thereby customer satisfaction. I kindly request you to fill this questionnaire which forms a part of my comprehensive project report. The information provided by you will be kept confidential and used for the purpose of this project.

1) Name :

2) Designation :

3) Age:

☐ Below 30 years

☐ 40 – 49 years

☐ 30 – 39 years

☐ Above 49 years

4) Gender

☐ Male

☐ Female

5) Total industry experience:

☐ Above 20 years

☐ 3-5 years

☐ 10-20 years

☐ 0-2 years

☐ 5-10 years

6) Which machine learning domain do you work generally?

☐ Computer vision

☐ Natural language
processing

☐ Genomics

☐ Trend analysis

☐ Chat boats

Analysis of product development process in a machine learning firm

7) What kind of problems is attended generally?

☐ Existing problems

☐ New problem

8) Where do you get most of the ideas for new products?

☐ From employees

☐ Client/Customer

☐ Market study reports/external agencies

☐ Other, specify _____

9) Rank your preference for the idea generation methods given above (giving 4 to the most preferred and there on)

☐ Idea from employees

☐ Client

☐ Market study reports/external agencies

☐ Other, specify

10) Employee's participation in idea generation is advisable or not?

☐ Always

☐ Rarely

☐ Usually

☐ Never

☐ Sometimes

11) How do you rate the idea generation process in your company?

☐ Very good

☐ Bad

☐ Good

☐ Very bad

☐ Need improvements

12) How are teams assigned for the feasibility study

☐ Random selection of members from each department

☐ Pre designated team for all feasibility study

☐ Senior level management team

☐ Other, specify _____

13) Rank your preference for the methods of assigning teams for feasibility study (giving 4 to the most preferred and there on)

- ☐ Random selection of members from each departments
- ☐ Pre designated team for all feasibility study
- ☐ Senior level management team
- ☐ Other, specify _____

14) Who leads the team for studying the feasibility of selected ideas

- ☐ Member from marketing
- ☐ Member from HR
- ☐ Member from technical
- ☐ Other, specify _____
- ☐ Member from finance

15) How the idea selection process works in the company?

- ☐ By voting majority
- ☐ Other, specify _____
- ☐ Consensus after discussion
- ☐ Analytical

16) Rank your preference for the idea selection process (giving 4 to the most preferred and there on)

- ☐ By voting majority
- ☐ Other, specify _____
- ☐ Consensus after discussion
- ☐ Analytical

17) How do you evaluate the feasibility of an idea? Rank the influencing factors.

- ☐ Market acceptance
- ☐ Competitors
- ☐ Cost of production
- ☐ Profit
- ☐ Technology

18) What is the means of finding the approximate time and cost estimate for a product

- ☐ Experience from similar projects carried out by us
- ☐ Own estimate
- ☐ Competitors cost/quote
- ☐ Other, specify _____

19) How teams do the detailed study on each selected ideas?

- ☐ Assigning different teams with different objectives
- ☐ Studying reports of similar products within the company
- ☐ Competitors reports
- ☐ Other, specify _____

20) Rank your preference for the process of detailed study of selected idea (giving 4 to the most preferred and there on)

- ☐ Assigning different teams with different objectives
- ☐ Studying reports of similar products within the company
- ☐ Competitors reports
- ☐ Other, specify _____

21) Who leads the team for detailed study of selected idea

- ☐ Member from marketing
- ☐ Member from technical
- ☐ Member from finance
- ☐ Member from HR
- ☐ Other, specify _____

22) What are the important fields covered in the detailed study of a product, Rank them in preference

- ☐ Marketing
- ☐ Technical
- ☐ Finance
- ☐ HR
- ☐ Other, specify _____

23) How do you rate the process for the detailed study of an idea in your company?

- ☐ Very good
- ☐ Good
- ☐ Need improvements
- ☐ Bad
- ☐ Very bad

24) How are teams assigned for product development in the company?

- ☐ New team with sufficient expertise in each functionality
- ☐ Reassigning work to a team had similar project
- ☐ Inviting application from employees
- ☐ Other, specify _____

25) Rank your preference for the assignment of teams for product development (giving 4 to the most preferred and there on)

- ☐ New team with sufficient expertise in each functionality
- ☐ Reassigning work to a team had similar project
- ☐ Inviting application from employees
- ☐ Other, specify _____

26) How many times a prototype model have faced serious issues which then required serious rework from beginning in the last 6 months

- ☐ Once
- ☐ Twice
- ☐ Thrice
- ☐ More than thrice

27) What all kinds of code review methods are there in your company?

- ☐ Self
- ☐ Peer
- ☐ Lead
- ☐ Architect

28) In your opinion, whether the workforce in the company are capable of handling works in any new technology or domain

- ☐ Definitely would
- ☐ Probably would
- ☐ Probably would not
- ☐ Definitely would not

29) What are the positives of prototype development process in your company?

- ☐ Approach for development
- ☐ The way in which selecting members for the development team
- ☐ Availability of subject matter experts
- ☐ Other, specify _____

30) Rank the positives of prototype development process in your company?

- ☐ Approach for development
- ☐ The way in which selecting members for the development team
- ☐ Availability of subject matter experts
- ☐ Other, specify _____

31) Rank the areas that need improvements for the prototype development process?

- ☐ Approach for development
- ☐ The way in which selecting members for the development team
- ☐ Availability of subject matter experts
- ☐ Other, specify _____

32) How are unit testing done in your company?

- ☐ Developer along with the development
- ☐ Assigning to a separate team member
- ☐ Other, specify _____

33) Who explains the requirements and functionalities of the product to testers

- ☐ Developer
- ☐ Self-learning from design and requirement documents
- ☐ Developer and Business analyst

34) Do you think if testing is done by other company, will it be more effective?

- ☐ Strongly Agree
- ☐ Somewhat Agree
- ☐ Neither Agree nor disagree

Analysis of product development process in a machine learning firm

- ☐ Somewhat disagree
- ☐ Strongly Disagree

35) What are the positives of overall testing process in your company? Rank them (giving 5 to the most preferred and there on)

- ☐ Enough allocation of time for testing
- ☐ Detailed review for Unit testing, SIT and UAT results
- ☐ SIT is done by other company
- ☐ Very active user base to check detailed user acceptance
- ☐ Other, specify _____

36) Rank the areas those need improvements for the overall testing process in your company?

(Giving 5 to the most preferred and there on)

- ☐ Enough allocation of time for testing
- ☐ Detailed review for Unite testing, SIT and UAT results
- ☐ SIT is done by other company
- ☐ Very active user base to check detailed user acceptance
- ☐ Other, specify _____

GLOSSARY

Word	Explanation
AI	Artificial Intelligence
B2B	Business to Business
Big data	Extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.
Bixbi	Bixby is a virtual assistant developed by Samsung Electronics
BPO	Business process outsourcing (BPO)
Chatbots	A chatbot is merely a computer program that fundamentally simulates human conversations.
Cobol	A computer programming language. It is an acronym for Common Business Oriented Language
Complex analytics	It refers to the use of advanced algorithmic methods to process large unstructured data sets effectively
Computer vision	It is an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos.
Continuous variable	Continuous variables are numeric variables that have an infinite number of values between any two values
Credit scoring	A credit score is a numerical expression based on a level analysis of a person's credit files, to represent the creditworthiness of an individual.
Data Science	Data science is a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data
Dependent & independent variables	Dependent variables rely on other variables to find their value, and independent variables do not rely on other variables to find their value.
Deterministic	Algorithm, model, procedure, process, etc., whose resulting behavior is entirely determined by its initial state and inputs, and which is not random or stochastic
DNA	Deoxyribonucleic acid, or DNA, is a molecule that contains the instructions an organism needs to develop, live and reproduce.
Dot net	It is a framework that provides a programming guidelines that can be used to develop a wide range of applications

Analysis of product development process in a machine learning firm

Genomics	Genomics is the study of whole genomes of organisms, and incorporates elements from genetics
GPS	Global Positioning System
Healthcare	Healthcare is the maintenance or improvement of health via the prevention, diagnosis, and treatment of disease, illness, injury, and other physical and mental impairments in people
IoT	Internet of things
IT	Information Technology
ITeS	IT enabled Services
Java	A computer programming language.
Lucrative	producing a great deal of profit
Macroeconomic	Macroeconomics is the branch of economics that studies the behavior and performance of an economy as a whole
MBA	Master of Business Administration
Meritocracy	government or the holding of power by people selected according to merit.
ML	Machine Learning
Netflix	Netflix is the world's leading internet entertainment service
Paypal	PayPal is one of the largest online payment processors in the world
Pervasiveness	The quality of spreading widely or being present throughout an area or a group of people.
Predictive Maintenance	It is the maintenance that monitors the performance and condition of equipment during normal operation to reduce the likelihood of failures
R&D	Research and development work
Siri	Siri is an intelligent assistant that offers a faster, easier way to get things done on your Apple devices. Even before you ask
Spiderbook	Spiderbook is an account-based sales and marketing platform that uses data science to give you qualified accounts with the same knowledge and intuition as a strategic account manager.
Technological dynamism	This is a concept which was defined by Albu (2009) as the rate of exchange in the level of predictability of new technologies
Uber ATC	Uber advanced technology center - leading online cab facility