



# Leveraging Analytics to Optimize Call Center Operations

**CORECOMPETE**  
DELIVERING VALUE FROM BIG DATA

## Table of Contents

1.0	Introduction.....	<a href="#">3</a>
2.0	Background of EasyCredit.....	<a href="#">3</a>
3.0	Overview of Analytical Techniques.....	<a href="#">4</a>
4.0	Implementation of machine learning and optimization.....	<a href="#">6</a>
4.1	Easy Credit Exploratory Analysis.....	<a href="#">7</a>
4.2	Caller Profiling using K-Means clustering.....	<a href="#">7</a>
4.3	Agent Profiling using K-Means clustering.....	<a href="#">8</a>
4.4	Optimization – Skill based agent assignment.....	<a href="#">8</a>
4.5	Designing Queue Simulation to support Resource Planning.....	<a href="#">10</a>
4.6	Technology Stack.....	<a href="#">10</a>
5.0	Conclusion.....	<a href="#">11</a>
6.0	About Core Compete.....	<a href="#">11</a>
7.0	References.....	<a href="#">12</a>

## 1 INTRODUCTION

Today companies that run call centers are on the lookout for ways to improve customer service and reduce costs. To achieve these goals, the focus has historically been on outsourcing strategies or on improving their digital presence, to reduce inbound calls. The area often overlooked by these companies is the growth in computing power and cloud computing. Technologies that allow analysis and processing of huge volumes of data, in many cases, in real-time. Given these developments, companies looking to optimize call center operations need to use analytics as a strategic lever.

This paper discusses a hypothetical lending organization, EasyCredit. Its is struggling with addressing high volume calls to their customer care call center. EasyCredit is planning to use data science to optimize their call handling. This paper aims to describe the conceptual design, to utilize the power of Machine learning for optimization and simulation, to streamline call handling. Machine learning algorithms have been used to profile the complexity of callers and profile customer service executives/agents receiving the calls. Simulation has been used to design the abstract of a real system.

## 2 BACKGROUND OF EASYCREDIT

EasyCredit is a lending organization that issues credit card to customers. They run a call center to resolve customers complaints and grievances. They receive roughly 50,000 calls a day and 80,000 applications a day during promotional periods.

- EasyCredit offers three types of cards: Classic, Advance and Signature
- 210 Agents working in 3 different 8-hour shifts during the day
- Average call handle time is a measure of operational efficiency. It also directly impacts customers' satisfaction level
- Operational efficiency is measured as = (Wait Time + Call time + Post Call time) / Total number of calls received

Daily Call Time Required	Time Available of Agent (Daily)
Number of calls in a day: 30,000	Number of agents: 400
Avg call time: 10 Minutes	Working hours per agent: (8X60 Minutes)
Total Call Time: 300,000 Minutes	Total Agent time: 192,000 Minutes

EasyCredit is facing challenges in effectively servicing all the calls, as volume of calls is very high with varying complexity of applications. Historical average call handle time is 15 minutes, including 5 minutes of waiting time to connect to the agent, leading to higher customer dissatisfaction.

A analysis of demand vs supply is in the table above.

- The discrepancy between required agent time and available agent time is high
- Agents are busy most of the time. Wait time for a caller to connect to agent is 5-minute, leading to high abandonment rate.
- As complaints/queries of customer are resolved, there is yet high level of customer dissatisfaction.

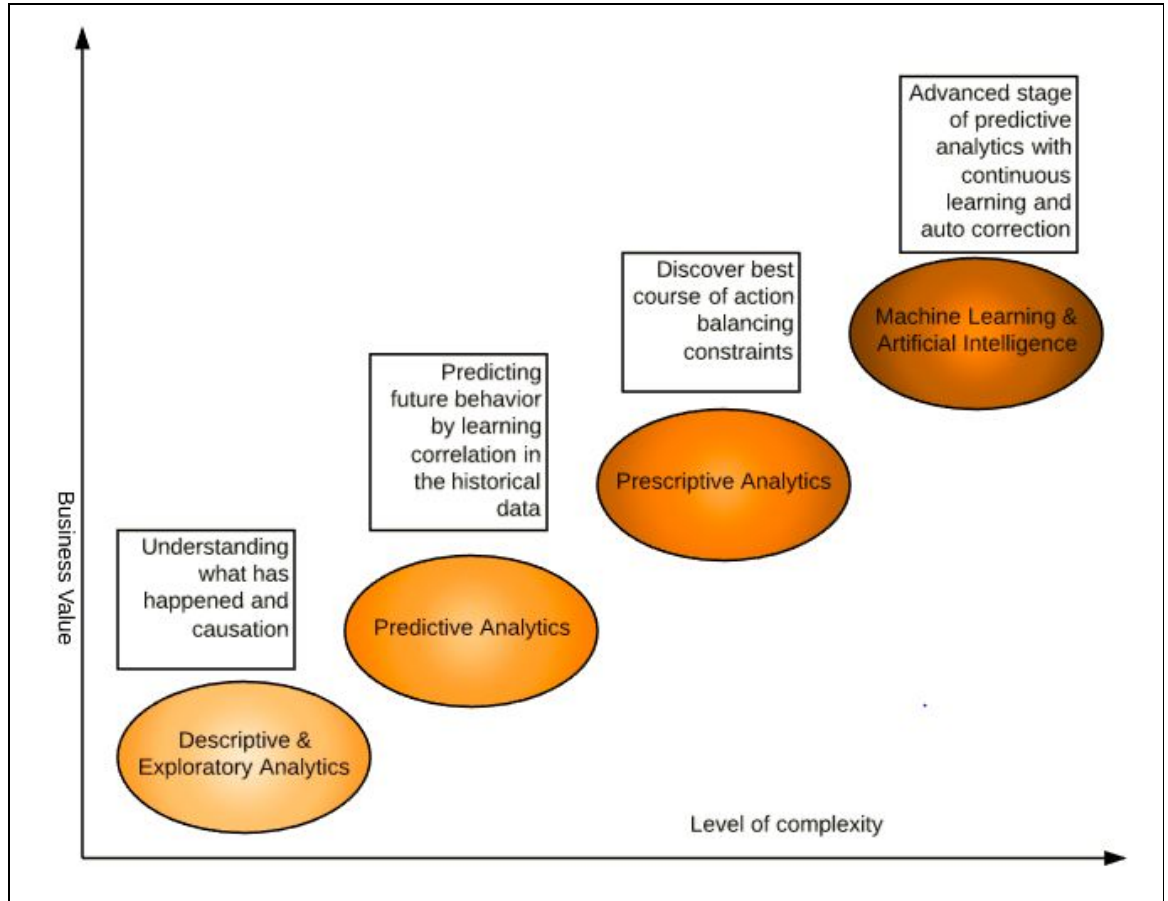


EasyCredit has decided to use Machine learning and optimization techniques to reduce the call handling time.

### 3 OVERVIEW OF ANALYTICAL TECHNIQUES

The application of analytical techniques to solve the business problem has evolved significantly in last 30 years. These methods allow users to form a relationship between variables. The relationship can be as simple as linear relation or complex ones which are usually not interpretable. Analytics can be classified into four categories

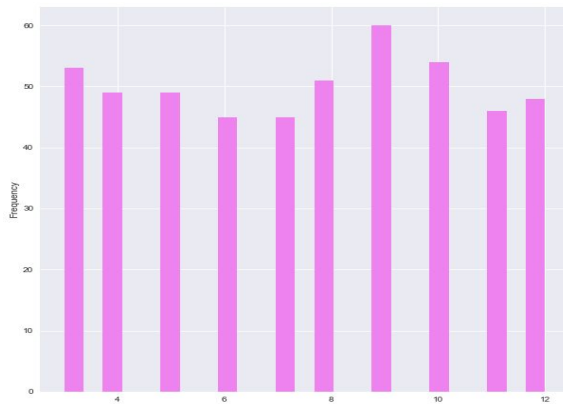
1. Descriptive and exploratory analytics: Also called Business Intelligence, where the interest is in what has already taken place. It is used for strategic performance assessment. For example, total revenue by market in a quarter, new customers acquired in a month, to check the effectiveness of a campaign, etc.
2. Predictive Analytics: Historical data is used to predict probability of an event in future. Predicted event is a binary or categorical field in most cases, but there are techniques used to predict the continuous variable as well. The advanced algorithms such as Logistic Regression, Decision Tree, Random Forest, Naïve Bayes classifier are used to predict the event. In customer analytics, these algorithms are used to predict customer purchasing propensity in future or probability to default, etc.
3. Prescriptive Analytics: These analytics methods prescribe specific action to be taken. Advance mathematical algorithms are used to optimize a specific objective.
4. Machine Learning and Artificial Intelligence: Within the field of analytics, machine learning is a method used to devise complex models and algorithms, that makes machines learn without being programmed. It uses neural network models to understand the complex relationship between variables.



**4 IMPLEMENTATION OF MACHINE LEARNING AND OPTIMIZATION**

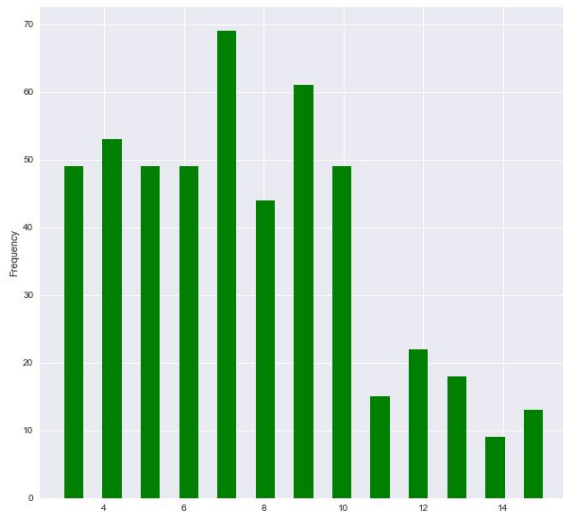
**4.1 Easy Credit Exploratory Analysis**

Initial data analysis shows that call handle time of an agents varies based on the caller’s attitude and behavior.



Histogram of Call Handle Time taken by an example Agent on a day

It has been also observed that for a given caller profile, call handle time varies by an agents skill sets and experience.



Histogram of Call Handle Time taken by an example Agent on a day

These observations led to following hypotheses

- a. Callers can be grouped by those having similar profile and historical behavior
- b. Agents can be grouped based on their skill sets and experience to handle callers
- c. Certain agents are more suitable for certain kind of customers.

With the above hypotheses as basis, Machine learning algorithm was developed to cluster callers and agents into homogeneous groups.

A chart was created to map the average call handle time of a caller profile to agent profile. This information can be used to patch the callers to the right agents, to minimize overall call duration using optimization techniques.

Process of optimization of call handle time by using skill-based agent allocation, can be developed in three steps. Finally, a simulation model was developed to construct the attributes of EasyCredit call center, which was then used to analyze call center operations in different scenarios.

## 4.2 Caller Profiling using K-Means clustering

Grouping of callers and agents can be done using heuristic approaches, but best results are not received when number of features are more than two. Clustering, is an unsupervised machine learning algorithm which is widely used to create homogenous groups of the observations, using huge set of features, when the data is not labeled. There are several clustering algorithms available. K-Means clustering technique, was used for profiling callers and agents as it provides control on the number of clusters created. The idea is to define K centers, called centroids, one for each cluster. The next step is to associate, each point in the data with nearest centroids. When all the points are assigned centroids, centroids are recalculated. After we have these K new centroids, points are reassigned to new centroids. This process is repeated until the centroids do not change within a threshold value.

Customers of EasyCredit were segmented, using a variety of variables using K-Means clustering technique. Cluster analysis of customer base is a creative and iterative process which requires a good understanding of statistics and customer demographics. The customer base is segmented using following categories of variables.

- Demographic profile of caller's variables such as age, sex, income, education, residential districts, language etc.
- Channel usage based on day and time
- Complaints log
- Transaction and shopping behavior
- Online vs POS transactions
- Type of card holding etc.

Three segments were developed which were profiled and identified as following

1. High Complexity
2. Medium Complexity
3. Low Complexity

The average call handling time by segment is clearly different. High complex customers taking more time from agents, whereas low complex customer take lesser time on calls.

### 4.3 Agent Profiling using K-Means clustering

Customer service executives or agents of EasyCredit were categorized using K-Means clustering technique. The agent base is segmented using following types of variables.

- Demographic profile of agents, variables such as age, sex, education, language etc.
- Work experience
- Number of calls handled
- Average time of call
- Average time taken post call

Three segments were developed which were profiled and identified as following

4. High Skilled
5. Medium Skilled
6. Low Skilled

When time taken is plotted by agent segment and caller segment, it was observed that time taken by high skilled agent group is significantly lower than time taken by low skilled agent group.

### 4.4 Optimization – Skill based agent assignment

Linear programming is a method to achieve the best outcome in a mathematical model whose requirements are represented by linear relationships. Linear programming is a special case of mathematical programming. (Source: <https://en.wikipedia.org>). Linear programming is an optimization technique for a linear objective function, constrained by linear equality and linear inequality. This technique has been in use for resource planning problems, commodity flow planning problem etc. Agent allocation to caller can be optimized to reduce call handle time using linear programming technique.

A linear programming optimization problem was formulated to solve agent allocation problem to reduce the average call handle time. Optimization will connect the callers to appropriate agents based on their profiles, to have a harmonious conversation and to resolve a query. Optimization problem has following structure:

1. Objective / Goal: To minimize the average call time between callers and agents
2. Decision Variables: Are the connection between callers and agents.



- There is a total of 16 possibilities of connections (i.e.  $4 \times 4$ ) hence there are 16 decision variables.
- These variables can only take the value of either 0 or 1

3. Constraints

- Each caller can be connected to just one agent
- Similarly, each agent needs to be connected to just one caller

A simple problem with three agents and three callers, one from each segment, is considered for illustration. A caller can be attended by one of three agents; hence 9 combinations are possible. The table contains the time taken in seconds by agents in resolving queries of the callers. Linear programming is used to optimize these connections and minimize average call time between caller and agent

	Caller Profile 1	Caller Profile 2	Caller Profile 3
Agent 1	110	296	340
Agent 2	156	405	500
Agent 3	310	488	603

Linear optimization solver prescribes to following connection to minimize average call time

	Caller Profile 1	Caller Profile 2	Caller Profile 3
Agent 1	110	296	340
Agent 2	156	405	500
Agent 3	310	488	603

- The solution minimize total time is  $340+156+488 = 984$  Seconds
- New average call time equal to 5 minutes and 28 seconds

The above optimization is relatively simple, as the number of callers is equal to number of agents. Hence, one to one mapping based on the call time is possible. But in real world operations, the number of callers at a time will be always more than the number of agents available to attend the callers. To solve this type of agent allocation problem, queues can be deployed. Based on the number of agent segments, three queues can be deployed where callers

will be pushed in to recommended queue by the optimization process. If any queue becomes empty, agents from those queues can pull the callers from other busy queues.

#### 4.5 Designing Queue Simulation to support Resource Planning

In section 4.4, application of linear programming was discussed for agent assignment to callers. The problem was deterministic, as number of agents and numbers callers were known. But, there is great amount of uncertainty in the real world. The number of customers making call to customer care center varies thru the day. Numbers of callers may be high in the afternoon or low in the mornings. There is also variation in the amount of time taken by agents to handle customers calls. All these factors lead to complexity in resource planning for call handling.

Optimization models are used for problems that are deterministic in nature, while simulation models are used for problems that are probabilistic in nature. Simulation model can be designed to construct real-world scenario using empirical distribution. Historical data can be used to develop the probability distribution function for uncertainties in arrival time, type of customer, type of call, time of the day etc. Also, Queuing theory principles can be used if all data is not available.

A simulation can be developed to model the EasyCredit call center operations using queuing theory principles and historical data. Simulation model of EasyCredit call center can be developed in iterations, by understanding the complexity of the call center operations. The advantage of simulation, is that it allows the decision maker to analyze the impact of any changes in the call center operations like increase/decrease of agents, low/high call volumes etc. Decision maker can then simulate the system with different policy decisions and understand the overall impact of the new policies.

#### 4.6 Technology Stack

CoreCompete has strong experience in building and deploying custom analytical solution using latest cutting-edge technologies in the world. CoreCompete has deployed multiple analytical solution using data mining techniques on proprietary platform (ex. SAS) as well as in open source platforms (Python, Pyspark). CoreCompete with strong base in domain, analytics, technology and agile services has delivered custom products to solve the business problem in short periods. CoreCompete, married SAS Visual Analytics with Apache Hadoop data processing at Travis Perkin, to develop supply chain intelligence solution on Amazon Web Service (AWS) cloud platform. A hybrid solution, inventory optimization system was developed for leading Auto Parts retailer in US using SAS Inventory Optimization, SAS Visual Analytics and Microsoft Excel by integrating data processing in Apache Hadoop and optimization on AWS batch services.

Voice of Customer solution was developed for Lenovo by marrying SAS sentiment analysis with Python object-oriented programming on AWS cloud platform.

Hybrid approach for analytical solution has been cost effective and efficient. Using open source technology such Spark on Hadoop, for processing data of the size in the order of terabytes is efficient and less time consuming. Proprietary software products such as SAS Enterprise Miner or Optimization are more business-friendly tool and easy in implementation.

## 5 CONCLUSION

---

Use of analytics techniques and machine learning algorithm can help in reducing operating cost by 5% to 7% and increase in first call resolution rate by 10% to 15% in a larger set ups. Based on business problem, right analytical approach should be identified from hundreds of analytical techniques available. For EasyCredit call center problem unsupervised machine learning model; K-means clustering was used for caller and agent profiling, while optimization for allocating agent to a caller. Analytics can be leveraged for other problems such as call volume forecasting, drop in first call resolution rate etc. And to measure the customer satisfaction. For call volume forecasting, time series models or regression models can be used. Businesses should be agile in adopting latest cutting edge open source technology and proprietary software such SAS for solving the business problem. There are several open source platforms available to process unstructured and structured efficiently. Based on the size and nature of the business problem hybrid technology platform can be deployed.

This paper presents elevated level conceptual design for optimizing call handle time using the principle of skilled based allocation. Machine learning algorithms can be used efficiently to profile callers providing opportunity to better manage their complaints/queries.

## 6 ABOUT CORE COMPETE

---

CoreCompete brings a unique blend of Domain, Analytics, and Technology expertise combined with Agile methodologies to deliver rapid value to our clients. Our experience with enterprise customers and strong partnerships with major technology vendors allows us to bring expertise and access critical resources that will accelerate our client's success. We work with our clients to:

- Identify opportunities that have a measurable business impact
- Translate opportunities into analytics-based solutions
- Operationalize them using technology for sustainable benefits

## 7 REFERENCES

---

1. Transforming Range Management to Reduce Inventory and Maximize Availability ( <https://corecompete.com/case-studies/digital-transformation-transforming-range-management-to-reduce-inventory-and-maximize-availability/> )
2. Responding Faster to the Voice of the Customer in Product Planning ( <https://corecompete.com/case-studies/digital-transformation-responding-faster-to-the-voice-of-the-customer-in-product-planning/> )
3. Creating an Endless Aisle of Parts with Fast Availability ( <https://corecompete.com/case-studies/digital-transformation-responding-faster-to-the-voice-of-the-customer-in-product-planning/> )