CS 4202 - FINALISED RESEARCH PROPOSAL

ANANTA - SMARTER INSIGHTS FOR CUSTOMER DATA ANALYSIS

By

Project Group - ANANTA

Madusanka W. A. T (100324X)

Madusanka W. A. T (100324X)

Madusanka W. A. T (100324X)

Madusanka W. A. T (100324X)

Project Supervisors

Dr. Sumith Matharage

Dr. Shehan Perera

Contents

Introduction

1.1 General Introduction

Customers are a vital part in any organization or business. With the advent of modern communication technologies, customers are more empowered, exposed to and endless variety of markets and are hence more dynamic in behavior than ever before. Although companies have been able to capture enormous amounts of data owing to the modern storage and communication capabilities, that alone would not suffice to conduct a proper analysis of the customer behavior. To analyze the customer buying habits and gain an understanding of their customers, companies need a more sophisticated technique. This has led to a new era of customer analytics or customer intelligence which has become the backbone of most marketing strategies of companies.

At the core of customer analytics lie data-mining techniques. In detail it can be said that predictive analysis, data segmentation and visualizations form the foundation of customer analytics techniques. Despite the many data-mining and visualization techniques, the techniques allocated for customer data analytics are relatively new the full potential of new data-mining techniques have not been utilized in the said field.

As the first objective of the project, a survey around the existing ensemble methods in customer data analysis as well as the other domains are carried out. The feasibility of implementing existing other domain specific ensemble methods in the field of customer data analysis is tested. It is targeted to initiate a framework to develop ensemble data mining methods using the existing non-ensemble methods and ensemble methods as well.

The project is targeted to visualize all the results obtained in the data mining techniques with more graphically advanced techniques to improve the user experience. The project is targeted to perform a thorough comparison between the performance of the existing methods and the ensemble methods which are going to be developed.

1.2 Motivation

Successful Marketing is a key to a successful organization. Early days marketing was mainly based on qualitative market research. However in this world of big data, traditional marketing alone is not good enough to remain competitive. This has lead to an era of data driven marketing where organizations use advanced data mining techniques to generate insights that helps organizations to attain, retain and grow the customer base. Therefore, customer analytics has attracted the attention of most of the businesses and is becoming a key function within organizations. This very current, highly useful and practical nature of customer analytics is the main motivation to work in this domain.

As customer analytics is an emerging topic, from a computer science perspective there is a lot of potential to explore new directions in customer analytics. Data mining research has been there for few decades with the data scientists, however the applicability of these techniques in customer analytics space is very limited. This motivated us to carry out research on investigating the applicability of advanced data mining techniques in customer analytics applications. Also, ensemble methods are becoming more promising in many different data mining tasks. Therefore we are very much motivated to try out ensemble methods in the field of customer analytics.

Moreover, according to our knowledge availability of customer analytics tools is very limited. Most of the organisations use different data mining tools to accomplish different customer analytics tasks. This needs highly technical people with specialized knowledge in advanced data mining techniques to use customer analytics functions. However as marketing is not a very technical domain, having a tool that makes marketers to easily use customer analytics functions without having a technical barrier is essential in todays world. This motivated us to design and development of a customer analytics tool with great visualization to bridge the technical gap between marketing and data mining.

1.3 Objectives

The main objectives of the project are;

1. Understand the customer analytics functionalities within an organization and, design and develop a customer analytics tool that provides different customer analytics functionalities.

2. Investigate the use of existing data mining techniques in the field of customer analytics and integrate them into the tool.

3. Research on ensemble methods in data mining and look for possible enhancements to improve the accuracy of the customer analytics techniques.

4. Develop visualizations to better demonstrate the customer analytics tool to improve the usability of the tool among the non-technical users.

Proposed work and contributions

2.1 Our approach / Proposed Work

Our approach on achieving above objectives is illustrated in Figure 1.1.

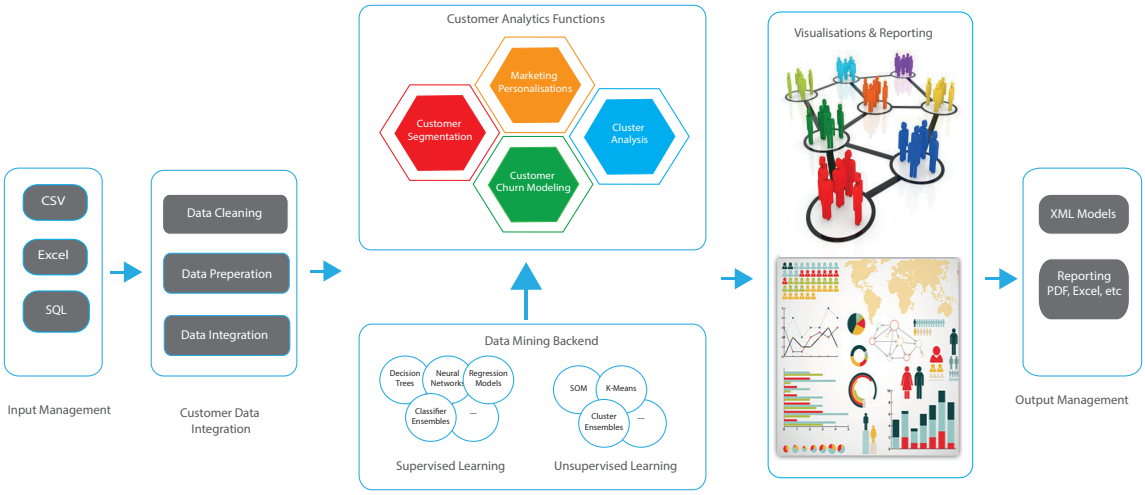

Figure 2.1: GSOM model for text analysis with proposed enhancements

As shown in the figure we will develop a complete customer analytics tool, which enables customers to perform various customer analytics tasks. The tool will enable the following marketing functions.

- Customer Segmentation Group customers based on the customer characteristics. This will be mainly based on customer behavior data to identify the different customer behaviors. Also this will enable build customer profiles incorporating other aspects such as customer demographics, financial patterns and any other combinations of customer data.

- Cluster Analysis This will enable analyzing customer behaviors across time. Different customer behavior movements can be tracked and will be prompted for different marketing actions as required.

- Customer Churn Models Tool can be used to identify customers with possibility of churning in advance. This will enables to take preventive actions by the marketing division.

- Marketing personalization Tool will enable to identify the potential customers who would positively respond for different types of offers.

Output of all above functions is supported with enhanced visualisations.

Data-mining framework is the backbone of all the above customer analytics functionalities. A comprehensive research on existing supervised and unsupervised techniques will be carried out to decide the suitability of existing data mining techniques in different customer analytics functionalities. The best-identified algorithms will be incorporated to the tool at the initial stage of the customer analytics tool. Furthermore research will be carried out to develop new data mining algorithms specially in the areas of classifier ensembles and cluster ensembles and those will be integrated to further enrich the customer analytics capabilities of the proposed tool.

2.2 Contributions

The main contributions of the project would benefit both theoretical and practical domains. The proposed customer analytics tool would contribute to the practical aspects where organizations can use the tool to support their data driven marketing functions without the need of data mining experts.

Research carry out on customer analytics space would also contribute to the data mining and customer analytics research. This would align the existing data mining techniques with the customer analytics functions and also possible contributions to the data mining research with new ensemble data mining methods.

Background

Project Plan

4.1 Project Scope

4.2 Assumptions and Limitations

4.3 Deliverables

4.4 Project Milestones

4.5 Risks and Challenges

4.6 Timeline

References