



ABSTRACT

This paper explores the implementation of a newsletter recommendation system for Leimberg Services Inc. (LISI) to enhance customer satisfaction. The research leverages machine learning algorithms and data analysis to segment subscribers and predict their newsletter interests. By collecting demographic information and consumption records, along with newsletter content, the study identifies clusters of subscribers and matches them with relevant newsletters based on keyword similarity. The findings provide valuable insights for LISI to personalize email marketing campaigns, improve customer segmentation, and conduct A/B testing for optimization. The proposed strategies encompass recommendation systems, customer segmentation, and A/B testing, which can contribute to increased subscriber engagement, retention, and revenue. Additionally, suggestions are given for future improvements, including adjusting the registration system, providing clearer categorization options, and establishing a comprehensive database with interactive data collection. These recommendations aim to enhance personalization, improve user experience, and enable data-driven decision-making to optimize LISI's service and content over time.

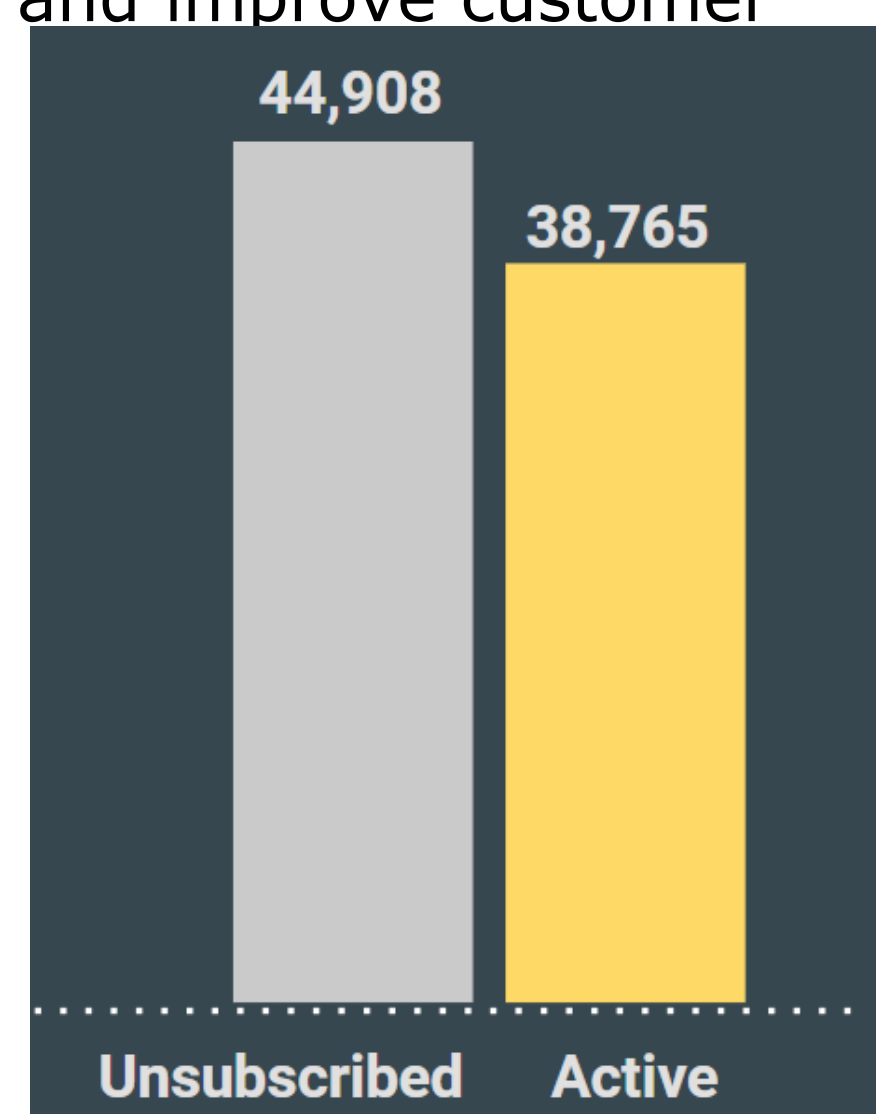
INTRODUCTION

Leimberg Information Services Inc. (LISI) is a subscription-based service that offers newsletters and podcasts to customers in various sectors, including attorneys, real estate, and financial planning. Although the company has been undergoing changes to implement a more focused business model, they currently lack a marketing strategy to engage existing customers and attract new subscribers.

The main challenge lies in how subscriber information is collected during the sign-up process, where customers are not required to specify their industry interests, leading to a lack of personalized email communication. This results in a high volume of generic emails, causing customer disengagement and unsubscribing. The objective of this project was to analyze customer data using machine learning algorithms and develop a marketing strategy based on customer segmentation and predicted newsletter interests. By understanding customer needs and preferences, the study aims to provide recommendations to Leimberg Services on how to implement personalized email campaigns and improve customer satisfaction.

In 2022, a total of 12,325 subscribers chose to unsubscribe due to receiving untargeted newsletter emails.

By successfully reducing the yearly churn by half, resulting in a decrease of 6,163 subscribers, the company can generate an additional monthly revenue of \$215,705 or an annual revenue of \$2,588,460.



METHODS AND MATERIALS

Data collection involved obtaining consumption records, and newsletter content from our sponsor. However, we faced challenges during data processing due to insufficient user data for segmentation. To overcome this, we removed irrelevant features, selected subscribers with comprehensive data, and manually categorized records based on company names. The project employed four models: a web crawler for extracting executive summaries, a text rank model for keyword extraction, a clustering model, and a similarity calculation model. This workflow allowed us to extract keywords, cluster subscribers, and match them with suitable newsletters.

The TextRank algorithm played a crucial role, employing an iterative process to assign scores to nodes based on their connections. Keywords and keyphrases were extracted from nodes with the highest scores, making TextRank effective for longer texts. Additionally, as the data consisted of text, word embedding techniques were employed to convert strings into numerical features for clustering. Word embeddings represent words as numerical vectors in a high-dimensional space, capturing their semantic meaning and facilitating clustering.

To match subscribers and newsletters, the Jaccard similarity index was used to calculate the similarity between their respective keywords. This index measures the similarity and diversity of two sets by comparing the common elements to the total number of elements. The resulting Jaccard similarity index ranges from 0 to 1, indicating the degree of similarity between sets.

RESULTS

Throughout the process, we selected 500 subscribers from the original data and ensured that they had valid inputs for name, email, address, zip code, and company affiliation, and that they met the criteria specified by the algorithms we discussed above. In the matching process, each subscriber was assigned two keywords that describe the industry they belong to, as shown in Table 1. Next, these subscribers were separated into five clusters: Investment Management, Legal Services, Financial Services, Accounting and Tax Services, Consulting Services and Others. Based on the clusters, Leimberg was able to match up subscribers with different newsletters.

Next, we further categorized the subscribers based on the key words associated with their respective industries, as presented in Table 2. These items were the top 200 words associated with the following industries: real estate, tax and accounting, legal services, investment and wealth management and consulting. Our focus has been on the sectors that were identified in Table 1. However, we recognize the need for further refinement of the key words associated with different industries in future editions. After determining the common key words between the subscribers and the newsletters, we have matched each subscriber with the relevant newsletters as shown in Table 3 and provided personalized recommendations accordingly.

Unnamed: 0	id	Name_x	Status	city	state	zip	company	First Industry	Second Industry
1	21.0	cus_NEKqzDf3wRw	Joshua Miller	active	Boston	Massachusetts	2108	Acadia Management Company, Inc.	Investment Management
2	166.0	cus_MjYy1K9PqEIT	James Provenza PC	active	Glenview	IL	60025	James C Provenza, PC	Legal Services Attorneys
3	185.0	cus_MdmpN4Zc12N4	Kathy Skinner	active	San Rafael	CA	94901	White & Case LLP	Legal Services Attorneys
4	213.0	cus_MY06aLc74BdP	Justin Gilbert	active	Franklin	Tennessee	37067	Music City Estate Law	Legal Services Attorneys
7	275.0	cus_MicAluLzQ8jyC5	Tim Meinhart	active	Chicago	IL	60631	Willamette Management Associates	Finance Consulting Services

Table 1. Example of subset subscribers who are assigned with industry interests

company	First Industry	Second Industry	combined_text	cluster_label	Subscribers_keywords
Acadia Management Company, Inc.	Investment	Investment Management	Boston Massachusetts Acadia Management Company...	0	[Accredited investor, Active management, Alpha...
James C Provenza, PC	Legal Services	Attorneys	Glenview IL James C Provenza, PC Legal Service...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...
White & Case LLP	Legal Services	Attorneys	San Rafael CA White & Case LLP Legal Services ...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...

Table 2. Example data of the first three subscribers' associated key words.

Name_x	Status	city	state	zip	company	First Industry	Second Industry	combined_text	cluster_label	Subscribers_keywords	news_id
Joshua Miller	active	Boston	Massachusetts	2108	Acadia Management Company, Inc.	Investment	Investment Management	Boston Massachusetts Acadia Management Company...	0	[Accredited investor, Active management, Alpha...	2
James Provenza PC	active	Glenview	IL	60025	James C Provenza, PC	Legal Services	Attorneys	Glenview IL James C Provenza, PC Legal Service...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...	1
Kathy Skinner	active	San Rafael	CA	94901	White & Case LLP	Legal Services	Attorneys	San Rafael CA White & Case LLP Legal Services ...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...	1
Justin Gilbert	active	Franklin	Tennessee	37067	Music City Estate Law	Legal Services	Attorneys	Franklin Tennessee Music City Estate Law Legal...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...	1
Tim Meinhart	active	Chicago	IL	60631	Willamette Management Associates	Finance	Consulting Services	Chicago IL Willamette Management Associates FL...	4	[Accountant, Accounting, Accounting cycle, Acc...	1
Evan G. Mathew	active	Davis	CA	95616	BCJ Financial Group	Investment	Investment Management	Davis CA BCJ Financial Group Investment Invest...	0	[Accredited investor, Active management, Alpha...	2
Steven M. Wyatt	active	Atlanta	GA	30080	Chamberlain Hedlicka	Legal Services	Consulting Services	Atlanta GA Chamberlain Hedlicka Legal Services...	1	[Abetment, Abrogate, Acquittal, Adjudicate, Ad...	1

Table 3. Final result of the matching and clustering. Each subscriber is assigned with a cluster (cluster_label) and an optimal recommendation of newsletter (news_id).

DISCUSSION

Based on our analysis, three recommended strategies for Leimberg include utilizing recommendation systems, implementing customer segmentation, and conducting A/B testing. Recommendation systems powered by AI can enable personalized email marketing campaigns, increasing subscriber engagement and reducing churn. Customer segmentation, achieved through clustering based on demographic and consumption data, allows tailored content and communications for improved engagement and retention.

A/B testing helps optimize email campaigns by testing variables like subject lines, call-to-actions, and send times. Despite requiring additional resources, the benefits of personalized content, targeted messaging, and optimized campaigns can drive long-term growth and customer loyalty. Continuous analysis of subscriber data and experimentation with different strategies allow Leimberg to build stronger relationships with subscribers, ultimately boosting revenue.

CONCLUSIONS

To enhance personalization, we recommend that Leimberg adjusts its registration system by offering clearer categorization of newsletter topics and allowing subscribers to select only the topics of interest. This would reduce the bulk of untargeted emails and provide more personalized content. Additionally, providing options for subscribers to customize the frequency and timing of newsletter emails would improve the user experience and prevent overwhelming amounts of communication.

We also suggest that Leimberg establishes a steady database by regularly sending surveys to gather customer feedback on satisfaction, preferences, and desired content. Tracking metrics like click rates and reading time in both the website and emails can provide valuable insights for improving engagement. By analyzing subscriber data using machine learning and natural language processing techniques, Leimberg can personalize newsletter recommendations and tailor content to individual interests. These adjustments and data-driven strategies will enable Leimberg to better understand its target audience and continuously optimize their service and user experience.