|  |
| --- |
|  |
| MAFS6010U ARTIFICIAL INTELLIGENCE IN FINANCEFinal Assignment |
|  |

**Taekyun Ku**

Student ID: 20477924

HKUST

*tku@connect.ust.hk*

|  |
| --- |
|  |

Abstract

This report suggests the decent way to enhance the telemarketing operation in insurance business via decision tree, one of Artificial Intelligence and Machine Learning Algorithm.

**1. Progress and Learning from Group Project**

## Introduction (Industry Overview)

## Modern life insurance was introduced in 17th century England to seafaring traders who wished a measure of financial protection for their families should an accident befall them. The sale of life insurance in the United States began in the late 1760s. The Life Insurance Industry is mature, and growth opportunities are essentially limited to increasing the geographic footprint and developing new offerings. The industry is comprised of large multinational corporations, as well as smaller regional and niche carriers.[[1]](#footnote-1)

## Barriers to enter this industry are quite high because of the strict reserve requirements for life insurance and related financial products. Even so, there is ample competition among those within the sector. The term life segment is the most competitive. Such policies are essentially generic, with price and convenience of purchase being the main differentiators. Still, customers might be inclined to pay a modest premium for a simple term policy if they have already done business with a particular company. The breadth of product offerings and degree of innovation can determine the market share and degree of success of an insurer.[[2]](#footnote-2)

## Current Business Model

# Since the life insurance industry is B2C-oriented, the most important part is marketing and sales to attract new customers and those to maintain existing contracts and cross-selling, which is usually through telemarketing and agents. My company also focuses on these two areas, each of which has conflicting characteristics, which helps the company to protect from concentration risk. For instance, the number of customers whom agents meet within a period is small, but the rate of contract success is high. In contrast, the number of customers whom telemarketers contact through the call is significant, but the rate of contract success is quite low. This is because customers prefer contracts through trustworthy agents that they can receive counselling in person since terms and conditions of insurance products are tricky and the length of the contract are longer, so customers are more careful when purchasing insurance products than when purchasing other products in general. However, telemarketing cannot be taken lightly in terms of this sales performance as telemarketing operations has been proving to achieve its target although the portion of telemarketing operations is far low compared to that of agent operations. Furthermore, the group of new customers induced by agent operations will be utilized as potential customers to cross-selling marketing for telemarketing operations in the future. On top of that, the agent sales have a serious downside that faces always a risk of outstanding agents poached by competitors despite high compensation from the company. Therefore, it is more desirable to better support telemarketing operations in a more efficient manner as Telemarketing provides stable performance.

## Problem Statement

In this business, sales type is mainly categorized in original contract with potential customers and upselling/cross-selling contract with both potential and existing customers. The key to encourage consumers to purchase our products is how quickly, correctly we can recommend an appropriate product for individual customers according to their requirements or introduce attractive products which they even did not recognize it was necessary but seems quite fit to their situation. For that, the current telemarketing team is approaching to existing customers under the predefined selection criteria or randomly with its stored data about customers by call. For example, telemarketers make a call to existing customers for cross-selling simply based on the product types that the customers already have. Before telemarketers turn the dial, they usually collect all information about the customers to introduce other products besides their incumbent contracts. However, what telemarketers have in advance is only unclassified raw data such as demographics, income, date of incumbent contract, the rate of contract success and so on, which they need to analyze to find appropriate products for recommendation to customers. Hence, their rate of contract success is low since it is not sure whether the recommended products are fit to the customers or not and customers usually do not wait until they recommend more than two products.

## Proposed Solution & Expected Benefit

To enhance telemarketer’s rate of contract success per call, I propose a system which recommends the appropriate products to the individual customers with their existing data base by one of classification Machine Leaning Algorithm, decision tree. The structure of decision tree is referred to Figure 1.



Figure 1: Decision Tree Algorithm

Decision trees are flowchart graphs or diagrams that help explore all of the decision alternatives and their possible outcomes. Each "branch" of the tree represents one of the possible options that are available when making a decision. The branches can be extended when one alternative outcome leads to another decision that must be made. Added into each branch are the costs associated with each choice and the probability each is likely to occur. With these numbers, leaders can calculate the value of each set of branches to determine the best choice.[[3]](#footnote-3)

Decision trees based on inductive reasoning are one of the most commonly used machine learning models in practice and is a map learning model. This is a very robust model that primarily deals with discontinuous data and does not stop or display misleading results in the event of noise. The basic concept is “divide and conquer” algorithm which split the answer of each question into two repeatedly. Therefore, it is an effective strategy to proceed first with attributes that have a great effect on repeated dividing. Also, optimal structure of decision tree is forged when the number of decision attribute is minimized so that it encourages to make a decision quickly by the shortest path from the root node to the terminal node. To make it happen, let me introduce the most standard decision tree model, Iterative Dishotomiser 3(ID3). In this model, we need to understand two important concepts, information gain and entropy first. Information function shows how valuable each information is to me by a probability, which means the higher the probability of an event happening is, the smaller the value of information function is and vice versa. That is because high probability implies the information is nothing special since the event of the information commonly happens. The relationship between the value of information and the probability that it will happen is represented by as follows:

$$l\left(x\right)≔log\_{2}\frac{1}{P(x)}$$

That is, if probability that x will happen, p(x), is close to 1, the value of the information, l(x), converges with zero and if the probability of x is close to zero, l(x) increases to infinity. This figure can be utilized to explain entropy, which is generally expressed as follows:

$$E\left(S\right)≔\sum\_{i=1}^{c}Pil(Xi)= \sum\_{i=1}^{c}Pilog\_{2}\frac{1}{P\left(Xi\right)}=\sum\_{i=1}^{c}Pilog\_{2}\frac{1}{Pi}$$

In the equation, S is the collection of all events already happened and it is also the one of all data in Machine Learning. Now, I explain how to interpret the value of information with this entropy equation. First, the acquisition of information is expressed in formulas as follows:

$$G\left(S,A\right)≔E\left(S\right) - \sum\_{a\in value(A)}^{}\frac{\left|Sa\right|}{\left|S\right|}E(Sa)$$

Here, S is the collection of all events, E(S) is the entropy of all events. A is attribute and a is the value of attribute. |S| and |Sa| means the probability of all events and the probability of event that has attribute a, respectively. E(Sa) is the entropy of the event that has attribute a. Through this formula, we can find that when G is big, we need to put the attribute a in the high level of decision tree node since it implies that the attribute a has a number of events to be classified.

What I am looking forward to is a customer group with the high contract success rate, through a decision tree, for each product analysed based on the relevance, needs, interest, etc. of each customer and all the products in the company. Thus, telemarketers will be able to increase the rate of contract success more efficiently and effectively by calling only to the customers who are included in the customer group with the high contract success rate in their list rather than calling randomly everyone in their list. Also, in customer’s perspectives, it would be more helpful to both two different groups who are quite interest in our products and are not interest in them, respectively, as the latter group won’t be bothered from the redundant calls.

To be more specific, let me assume that I have a list of ten existing customers who already signed the contract a couple of years ago and still maintain their contracts with my company as well as new customers who recently purchased our products. Now I, a telemarketer, would like to do cross-selling to the customers with five relevant products for them. In prior to make a call to them, I want to put a priority to call rather than make a call randomly in order to work efficiently. In this regard, the rate of contract success by call and appropriate recommending products based on their individual characteristic data base should be prepared in my hands in advance. Then, I access the system to get the report of each individual’s name with rank of recommended products as well as call success rate. Here are the example of decision tree process and the report from the system I expect.

No

Yes

Yes

No

No

Yes

## Figure 1: Example of Decision Tree

## Formulation of Database

Typically, Telemarking department maintained the good set of database because their sales activities are being recorded and monitored in order to check the sales quality as well as performance. To get the better analysis, it is very important to secure the large set of data. Regarding the customer data set, it is maintained by Operations department (especially Policy Administration department). However, one of the opportunity is we may have other data set in other functions which may not be understood as the data that we can leverage. Each function may keep their data separately, it would be very helpful if we can conduct some analysis to find out useful data across the company.

# Reflection on An AI Article and A Finance One

**Article 1**

Harvard Business Review “The Legal and Ethical Implications of Using AI in Hiring”[[4]](#footnote-4)

To sum up the article, many of modern technologies such as Big Data, Machine Learning and AI now help organizations improve their ability to find the right person for the right job and screen out the wrong people for the wrong jobs, faster and cheaper than ever before. While these tools are disrupting the recruitment and assessment space, they leave many yet-unanswered questions about their accuracy, and the ethical, legal, and privacy implications that they introduce. Hence, the author focuses on the potential repercussions of new technologies on the privacy of job candidates, as well as the implications for candidates’ protections under the Americans with Disabilities Act and other federal and state employment laws.

First, the writer criticizes the companies’ access to the candidate’s private information without their proper (or even) consent in the pre-employment stage. I also totally agree with his criticism since it is apparent infringement of privacy. No matter it is critical or minor, the private information should be collected under the candidate’s consent as only the candidate has a right to allow to provide the information. Some people may argue that the private information posted in public place such as the open updates of individual’s private post via social media would not require the consent of the information owner as the owner of the post is likely to open it to the public. I also agree with the argument in general situation. But my point is that it should not be used in employment process without the candidate’s consent as most candidates are not aware their personal information will be used in such process because most people commonly think their private information is relevant to their job hunting. If they are told from the company in advance that their personal activities via social media is to be utilized in the pre-employment stage, they would have prepared their all personal posts to look good as much as they prepare their resumes. This is because the collection of private information of candidates should be executed under candidates’ agreement in advance.

On top of that, I am seriously concerned about the writer’s statement that unlike previous legal sanctions and limited use to prevent potential discrimination even though psychometrical test, which was widely used for candidate evaluation in the recruitment process in the past and actually helpful in the personnel recruitment process, current new technologies used in recruiting human resources are not even scientifically driven, nor are they legally regulated.

Nonetheless, like the author said, it is undeniable that new technologies can already cross the lines between public and private attributes. I also believe that it is now AI’s era and AI will be spread everywhere as time goes by because of its decent capability in various fields. The ethical and legal issues will be more serious and controversial. However, I believe it is a kind of common downside of anything new when it is introduced in human society. Many infringements of privacy these days will be resolved by new regulation and legal sanction sooner and later since most people are already aware of its stern necessity.

In conclusion, if it is inevitable that new technologies are broadly applied in job hiring system, it would be better idea to accept the current situation and try to find preventive measures such as facilitating the relevant regulation and finding out how to use it reversely to take an advantage rather than only siting and criticizing since we cannot stop a running cow.

**Article 2**

Harvard Business Review “Insurance Companies’ Untapped Digital Opportunity”[[5]](#footnote-5)

The article argues that many insurers, struggling with legacy technology and outmoded organizational structures, are playing catch up whereas retailers and cable companies moved their businesses online years ago. It also introduces the improvement of insurance field by utilizing digital gadget in their business strategy.

Most current developed digital applications and their advantages in insurance business that the writer depicted are agreeable in my business experience. Actually, data mining to find out customer’s feature for marketing and parsing big data has been started decades ago. Also, the digitized claims are now very common in insurance business since mobile app is necessary in this smartphone-oriented world.

However, I firmly disagree with the statement that the digital carrier takes benefits of Progress app. According to the author’s description, the app boosts customers to buy insurance simple by generating quotes for auto insurance from just a photo of the applicant’s license since Progressive has been very focused on the initial consideration and moment-of-purchase battlegrounds. This can be true for only certain insurance products with standardized terms and conditions and very simple product structure. For other insurance products, many people are less likely to be attracted by purchasing insurance products via online platform without counselling in person since they have tendency of risk aversion when they buy a long-term payment product without enough explanation from anyone in person. Moreover, if the online purchasing insurance product looks attractive, it implies the product has little profit margin to the insure, which is not quite attractive to the insurers. At last, the most serious reason is that the online product can be a target from fraudulent or defective customers as it is vulnerable to filter them. Namely, the product can be anti-selection which customers select insurance products unlike most common insurance products select good customers in the underwriting process. Therefore, I, unfortunately, articulate the implementation of buying insurance simply like other commercial products explicitly need more time until more advanced technology is developed.

In brief, like the author said, it is time to adapt the evolution of industries as Darwin found that the survival of the fittest was their adaptability rather than a function of the strength or intelligence of species. Although the insurance business has started to change their legacy, it is still behind the curve compared to what other industries has done. Therefore, insurers should figure out to go beyond their safe zone to keep face with customers who is more likely to evolve as much as technology does.

# Synthesis and Suggestions for Further Study

## Model Challenges

Although decision tree should be optimized by minimum number of nodes to encourage the speed of decision process and to drive the result more accurate, it does not work well as we expect when we put all the success cases with the relevant attributes of each customer since the attributes of personnel is significantly different and various for all success customers. Hence, the model could face up to the challenges of how to make the tree more efficient.

Moreover, the feature of decision tree is yes-or-no binary decision, which can lead the loss of opportunity to access to the potential customers who has 99% of contract success rate but the result shows ‘No’. In other words, since decision tree only consider the customers as contractor or non-contractor, no one easily know the probability of contract success rate.

## Suggestion to Tackle the Model Challenges

I suggest building multiple trees from multiple trials of sampling who are randomly collected from the whole papulation. For instance, I can make more than a thousand of sample groups which consist of a thousand or more of people randomly collected from, let’s say, million customers in our data base. Then, I put each sample group to each decision tree. Once new customer purchased our product and when we’d like to figure out how likely this customer would purchase additional products, then using the multiple decision tress can provide us the probability instead of simple ‘Yes’ or ‘No’ answer.

## Suggestion for Difference Usage

The model that I suggested is mainly to generate the business. But in the business context, to keep the existing customers is also as important as to have new customers on board. With the exact same approaches, we can also build the decision tree to figure out what kinds of customers would likely to lapse or surrender their policies. If we have such predicted information, then the company can make efforts to retain such customers, for example, by calling them to ask their needs or difficulties or by asking agents to talk to the customer if they have strong relationship.

# Note Individual Contribution

100%: Taekyun Ku (Student ID: 20477924)

1. http://www.valueline.com/Stocks/Industries/Industry\_Overview\_\_Life\_Insurance.aspx#.XOUB-e97khc [↑](#footnote-ref-1)
2. http://www.valueline.com/Stocks/Industries/Industry\_Overview\_\_Life\_Insurance.aspx#.XOUB-e97khc [↑](#footnote-ref-2)
3. https://www.businessnewsdaily.com/6147-decision-tree.html [↑](#footnote-ref-3)
4. https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in-hiring [↑](#footnote-ref-4)
5. https://hbr.org/2014/03/insurance-companies-untapped-digital-opportunity?autocomplete

=true [↑](#footnote-ref-5)