University of Technology Sydney

FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY

DECISION MANAGEMENT FOR NEXT BEST ACTION MARKETING

HOW TO BRING TOGETHER BUSINESS PROCESSES, BUSINESS RULES AND ANALYTICS TO DELIGHT YOUR CUSTOMERS

Constantin R Codis

a thesis submitted for the Degree of Masters of Analytics at the University of Technology Sydney, September 2019

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I Constantin R Codis declare that this thesis, is submitted in fulfilment of the requirements for the award of Master of Analytics, in the Faculty of Engineering and Information Technology at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature: Signature removed prior to publication.

Date: 03.05.2020

ACKNOWLEDGMENTS

First of all, I would like to acknowledge the very patient and efficient guidance of Professor Guandong Xu who as my principal supervisor has always responded to all my queries regardless of the time of the day or day of the week. Indeed, Dr Guandong Xu has helped me shaping out my initial research proposal into a proper thesis. I also would like to extend my warm thanks to my co-supervisor Dr Shaowu Liu, who answered all my emails promptly and always made sure I understood all the requirements correctly. It would not be fair not to acknowledge the state-of-the-art facilities at UTS and particularly the library and the kind and helpful staff members. The staff members of the Faculty of Engineering and Information Technology who work extremely hard in the background for everything to run smoothly.

ABSTRACT

This thesis examines the concept of Next Best Action (NBA) Marketing and its uses within the greater context of Customer Relationship Management (CRM). The presented complete NBA Framework consisting of 1. Architecture, 2. Analytics and 3. Project Development and Implementation.

The methodology adopted in this thesis is a combination of both theory and praxis. From an architecture perspective the point is made that 21st century enterprises need to group their resources around their customers. NBA by definition is an orchestration of technology, process and people to deliver a differentiated customer proposition. Therefore, business architecture has to lead the way. To this end, the personalisation of propositions requires a structure or hierarchy. The business issues and groups provide such a structure on which the decision logic can be build. On the analytics side the thesis explores the use of adaptive models powered by Naïve Bayes which provide real-time models in the field of marketing. In this context, posterior probabilities are explored with different types of data grouping techniques such as Platt scaling, Binning and Pool Adjacent Violators (PAV). A model to calculate the final customer score for an online proposition is presented. Unlike conventional marketing, NBA marketing is putting the customer at the centre. NBA is using "propositions," i.e. a proposed course of action, to understand what type of problems the customer is trying to solve. Customer Lifetime Value (CLV) is a crucial baseline metric. The thesis is also exploring various NBA delivery methodologies and applications to other areas of the business such as risk. The last part of the thesis focuses on the implementation of NBA Marketing in the form of project management.

NBA marketing enables the organisation to establish a line of sight between its high-level strategic priorities and its more tactical marketing activities. The decisioning hub is the place where all the decision data and other inputs inform the decision logic. Extending CLV increased the enterprise profitability. Also by using a more extensive model to calculate the customer score for a proposition the more relevant the proposition is and the customer derives a unique value out of interacting with the organisation. This paper shows that Naïve Bayes (NB) is a very appropriate model to work out the customer affinity for a particular marketing outcome. The binning method used in the data analysis is also relevant and appropriate in building the NB classifier. All four experiments performed on a banking dataset show to Binning and PAV are robust methods of grouping the data for analysis. Binning is a robust way to deal with both numeric and non-numeric data. The bins create an ordinal

view of the data. Once the data variables are binned Coefficient of Concordance (CoC) is used to measure how good the model discriminates between positive and negative cases. CoC is a convenient way of measuring the power of discrimination of the model because its high granularity and because is cut-off independent. From a project implementation perspective, it is shown that NBA implementations allow the business the change the way they do business rather than an IT driven re-implementation of how the business is currently done. A centralised decisioning system can provide a balanced NBA in real-time across multiple channels will ensure consistency between customer needs and business objectives. Real-time adaptive models should support the decision system by scoring customers affinity for a particular outcome.

Table of Contents

TABLE OF FIGURES

1 INTRODUCTION

Next Best Action (NBA) marketing is the process of making sense of, anticipating and fulfilling the needs and wants of individuals and organisations. Traditional marketing relies on the "four Ps," i.e. product, price, promotion and place to entice individuals and organisations into profitable transactions. Unlike conventional marketing, NBA marketing is putting the customer at the centre. NBA is using "propositions," i.e. a proposed course of action, to understand what type of problems the customer is trying to solve. Customer Lifetime Value (CLV) is a crucial baseline metric. Marketing efforts measure how much customer lifetime value is extended as the result of engaging the customer at the moment of their choice. NBA is, therefore, a very efficient way of aligning organisational strategies with marketing initiatives. NBA marketing enables the organisation to establish a line of sight between its high-level strategic priorities and its more tactical marketing activities. To achieve this balance, the NBA should combine good process with the right decisions. With NBA marketing, each customer interaction becomes a mini business case in which all the inputs are considered to decide the best course of action. For this to occur, the organisation requires a system to make decisions. The decisioning hub is such a place where all the decision data and other inputs inform the decision logic. The most critical factor in the decision system for NBA marketing is to provide information in a way that allows self-organisation.

Customer Relationship Management (CRM) is both a methodology and a technology, businesses are using to interact intelligently with their customers. The concept of CRM is based on the customer-centric idea that different strategies are required to interact with customers during their lifecycle. A typical customer lifecycle comprises four distinct phases; first phase is focused on finding and identifying potential customers or prospects for the business, the second phase, acquisition is the phase in which the business is focused on attracting the right customer, during the third phase the business concentrates its efforts on developing a relationship with the customer by providing the right product or service, whilst the fourth phase is concerned with retaining the right customers. In describing the customer lifecycle notice the use of word “right” to qualify the customer for the product or service offered. This is a very important distinction in doing marketing in 21st century. The use of Artificial Intelligence (AI) and Machine Learning (ML) techniques has come a long way in the past five years to facilitate the exploration of massive data sets to help the businesses understand how to attract, sell and retain the customers likely to stick around for a very long time. This thesis is exploring what Next Best Action (NBA) is in the context of marketing. Essentially, the concept that there is a Next Best Action for everything customer interaction had grown out of the discipline of direct marketing which later became known as database marketing[1]. The principle of direct marketing or direct response marketing are well covered by notables in the field such as David Ogilvy[2] of Ogilvy & Mather, Bob Stone[3] in his Methods, and Claude C Hopkins[4] the father of Scientific Advertising. However, the NBA for marketing is different and this thesis is going to present what are some of the critical considerations organisations need to be aware of when adopting the NBA for marketing. This paper will look at the reasons why NBA marketing is a paradigm shift in customer relationship management (CRM). The information presented here represents the author's research and experience in implementing Next Best Action marketing in banking, telecommunications and superannuation industry. This thesis presents original materials in the area of decision making in Customer Relationship Management (CRM) and proposes a new model in CRM decisioning for customer-centric organisations, such as decision framework and final score calculation for customer interaction in the online channel in the context of Next Best Action marketing. Furthermore, this paper is going to provide a fresh perspective on Naïve Bayes and its usage in adaptive models. Although automated decision making has been around for quite some time for example airlines have been using them in yield management of their ticketing system, it is in the last four years that automated decision has taken hold in the CRM space with the introduction of frameworks like Next Best Action Marketing from companies like Pegasystems[5] ('Next-best-action marketing: a customer-centric approach' 2012). These systems are the perfect combination business process, business rules and predictive analytics. Bob Stone, one of the fathers of direct marketing in his classic Successful Direct Marketing Methods (Stone & Jacobs 2008), a book that has been in print since 1975, talks about building a customer centre where *"All activities are coordinated and flow through the Customer Center"*. The Customer Information Center is built around a database combining information from

[1] https://en.wikipedia.org/wiki/Database_marketing, accessed 8 February 2020.
[2] https://en.wikipedia.org/wiki/David_Ogilvy_(businessman), accessed 8 February 2020.
[3] https://www.goodreads.com/book/show/867521.Successful_Direct_Marketing_Methods , accessed 8 February 2020.
[4] https://en.wikipedia.org/wiki/Claude_C._Hopkins, accessed 8 February 2020.
[5] https://en.wikipedia.org/wiki/Pegasystems ,accessed 19 September 2019

the various sources – lead management, sales activities, customer service, and product. This is the central nervous system of the decisioning hub where all the decisions originate, just like in the human brain.

1.1 Background, Research Questions and Motivation

1.1.1 The Problem

Doing business in the 21st century is fraught with challenges both without and within the organisation. Markets are global. Economic conditions intertwine into the tight grip of market forces.

The conventional way of incremental cost-cutting and revenue growth are fading into irrelevance. Re-grouping organisational resources around customer experience to increase customer loyalty and reduce customer defection has become critical (*Adaptive BPM | Pega.* 2019). The issue is that organisations need to stay "agile" to thrive in an environment where the only constant is change. Therefore, focusing on customer experience is more important than providing a product or a service. The challenge is to continually adapt, leveraging customer transactional data, changes in the customer as well as market behaviour, and to create innovation opportunities with iterative and measurable improvements. With this end in mind organisations need to embrace a more holistic approach to adaptability (*Adaptive BPM | Pega.* 2019).

This research is answering the following questions:

1. What are the critical capabilities enterprises need to focus on to stay relevant and gain competitive advantage in 21st century global economy?
2. Why Next Best Action (NBA) Marketing is the "new" marketing where technology, processes and people are orchestrated to deliver differentiated customer propositions through multiple channels?
3. How should organisations approach, plan, implement and ultimately deliver martech[6] programs across the enterprise on-time and within budget?

[6] https://martechtoday.com/library/what-is-martech, accessed on 8 February 2020.

1.1.2 Objectives

The objective of this research is to develop a methodology for optimising customer value. Next Best Action Marketing provides an enterprise decision system to guide and respond to each customer interaction across every channel and line of business (*Leveraging Predictive Analytics to engage your customer* 2014).

With Next Best Action marketing, all types of transactional businesses such as carriers, agents and brokers can exploit and extract deep levels of understanding about each customer. Therefore, be more to engaging and relevant to the customer. With NBA, the business executes on customer behaviour and manage the relationship in a way that optimises the value of the business to the customer, and the customer to the business (*Leveraging Predictive Analytics to engage your customer* 2014).

There are two critical aspects of this research:

1. Explore and understand what type of analytics are required to execute NBA marketing. Specifically look at how adaptive analytics can learn from real-time customer interactions.
2. Provide a detailed overview of the NBA Marketing Delivery Model and propose a process to implement "decisioning" projects such as NBA Marketing.

1.2 Significance

Significance 1: Selling is one of the essential parts of a business. Sometimes sales teams can get a bad rap because they focus on selling the product instead of caring for the customer experience. A lot of the business models in the 21st century revolve around customer experience, for example, Spotify and Netflix are more focused on providing an experience and fulfil a need instead of selling a product (Tzuo 2018). Traditional marketing relies on the four Ps, which stands for the product, price, promotion and place. Marketing is the science of understanding what people want — then making the product or service in a way that is commercially viable. Then *"advertising them intelligently and selling them in the right places"* (Tzuo 2018). So, promotion is the push-and-pull factors within marketing. The question is, how to personalise an offer? Also, how to know it is the right offer for that customer to make at the right time in the right channel? Next Best Action marketing is designed to answer these questions. Nowadays crunching thousands of records in a split second is the new normal. So how to use this immense computing power to recognise and satisfy customers better? This research is putting forward a solution.

Significance 2: Another key point here is that although organisations face a situation where they have lots of data and access to lots of algorithms packaged via open-source software such as R[7] and Python[8], there is a real gap on how to go about operationalising "the smarts". This thesis provides valuable insights and answers in this somewhat “controversial” area. Furthermore, the thesis is showing how binning Naïve Bayes classifiers[9] is an effective way to adapt and learn in real-time. The distinct advantage is that even in situations where historical data is not available, adaptive models (Hu, Bodyanskiy & Tyshchenko 2019) are an excellent way to provide the probability of an event based on real-time data. The research is proposing and equidistance binning method to transform model score to posterior probabilities. Furthermore, this research is proposing a process based on various inputs to optimise what advertising message (propositions) online.

Significance 3: This research documents the empirical observations of the author. Also, the author has had direct experience with different project management methodologies and delivery models. This research is valuable, as is showing the way an organisation could go about implementing Next Best Action for marketing. What are some of the critical factors to consider in the initial stages? Why are these factors important, and what are the benefits of approaching a project in a specific way? This research answers a very pressing question as to how should an organisation go about implementing the Next Best Action for marketing.

The area of commercial applications, such as marketing and CRM, has recently experienced a total convergence of skills and abilities. McCormack (McCormack 2007) was remarking back in 2007 that information technology has moved on with more business oriented and increasingly sophisticated tools.

[7] https://en.wikipedia.org/wiki/R_(programming_language) ,accessed on 7 July 2019
[8] https://en.wikipedia.org/wiki/Python_(programming_language) ,accessed on 7 July 2019
[9] https://en.wikipedia.org/wiki/Naive_Bayes_classifier, accessed on 3 July 2019.

1.3 CONTRIBUTION TO KNOWLEDGE

The contribution to knowledge this thesis makes can be categorised into theoretical contribution and practical contribution.

The following points address the theoretical contribution:

1. Clarification of what Next Best Marketing is from business, Technology and Analytics perspectives. There are no publications that explain Next Best Action Marketing holistically.
2. There are no publications explaining a very key ingredient into structuring decisioning for Next Best Action. The Business Issue and Group hierarchy. This hierarchy holds in place and provide an audit of the decisions made in the system.
3. Providing a brief literature review of papers dealing with Next Best Action Marketing. This can be a good starting point for anyone wanting to study the subject from a different angle like a more theoretical one.

From a practical standpoint the thesis makes the following contribution:

1. The final score calculation is a singular formula for optimisation Next Best Action score.
2. In the Analytics this paper proposes binning of the predictor in the data analysis. This turns out to be a very efficient way of dealing with real-world data.
3. Explanation on how an Adaptive Algorithm can be build and applied to Next Best Action Marketing. While there are thousands of algorithms out there, this particular one is very well suited to deal with marketing data.

The key limitation to this approach is that a broader perspective is proposed than it is normally the case with this type of research. Namely any one of the above points could be zoomed into both theoretically and practically. However, it is the broader perspective this thesis is addressing as the discipline of database marketing is being reshaped into the digital economy without a clear framework and strategic motivation.

1.4 Thesis Structure

The thesis building blocks are the framework which is the overall structure on which the capability rests. With the foundation in place, the research is going to look in turn at each of the building blocks starting with architecture followed by analytics and finishing up with the implementation of Next Best Action Marketing for agile enterprises.

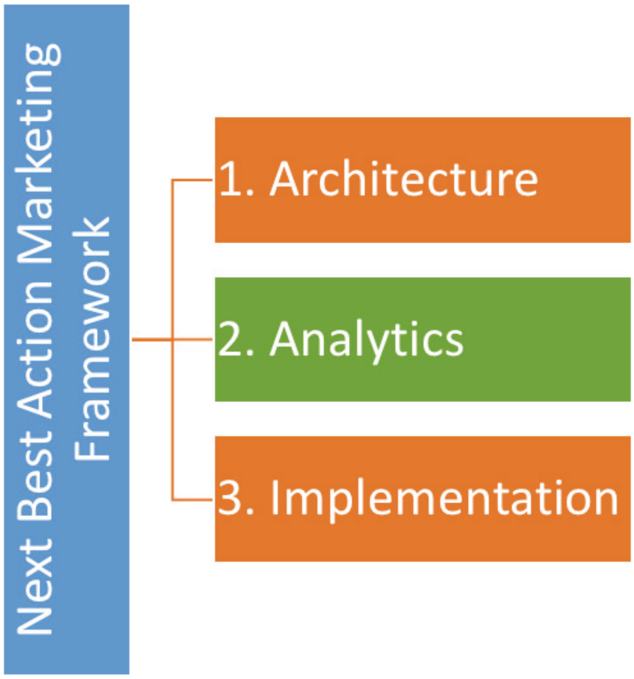

FIGURE 1 THE HIGH-LEVEL STRUCTURE OF THE THESIS

The Chapters on Architecture and Implementation explain the technology development and deal with the scope and business requirements for Next Best Action Marketing. As the objective of the thesis is to provide knowledge across all the elements of Next Best Action Marketing. As indicated in the Figure 1 the orange colour indicated elements that are typically part of the IT while the green colour indicated the analytics side of Next Best Action Marketing.

The analytics part of the thesis is the largest part of the work and provides new knowledge in the area of real-time adaptive algorithms with a very unique treatment to the data analysis.

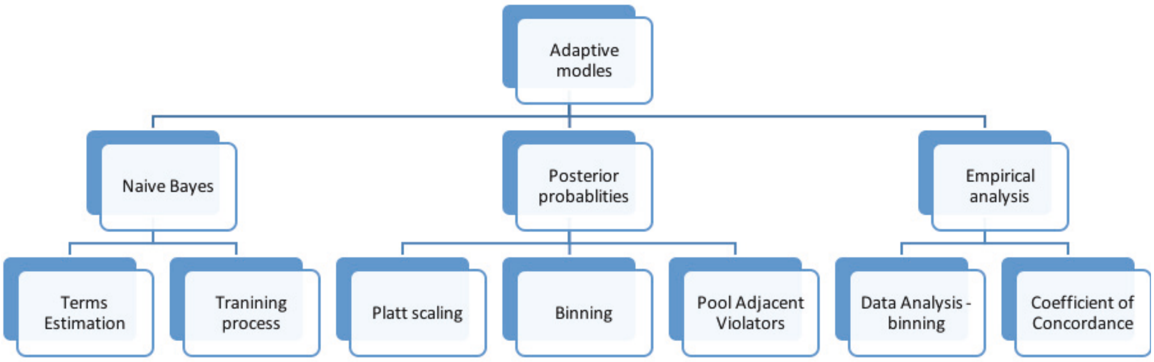

FIGURE 2 THE STRUCTURE OF THE TOPICS IN ANALYTICS

2 LITERATURE REVIEW

Next Best Action marketing is a methodology within customer relationship management (CRM). NBA paradigm is possible due to three crucial IT breakthroughs:

1. *The adoption of enterprise-wide application servers like IBM's WebSphere™* [10] *and BEA's WebLogic™* [11]*, and the Service-Oriented Architecture (SOA)* [12] (Chordiant NextBestAction Marketing White Paper 2016)
2. Unprecedented breakthroughs into hardware development. Even a consumer laptop can meet the needs for real-time AI and decisioning.
3. The wide-spread availability of open-source AI capabilities in R, Python coupled with massive gains in data management such as NoSQL[13] and Spark™ [14]

The papers reviewed in this section show that there is a need for more research in the area of operationalising decision making for CRM. Also, this literature review provides a somewhat historical reference of what has been published in the past 20 years on the subject of decisioing for Next Best Action marketing.

2.1 Next Best Action Papers

Two key papers are explaining the mechanism of Next Best Action, one of which is a white paper by K.R. Sanjiv of Wipro (Sanjiv 2012). In his paper, the author is making the point that the NBA is about targeting a segment of one and also presents a right technology agonistic overview of what type of technologies are enablers of Next Best Action. Sanjiv illustrates very well that the *"decision management hub and the data streams into the Next Best Action is 800-pound gorilla"(Sanjiv 2012)*.

[10] https://en.wikipedia.org/wiki/IBM_WebSphere ,accessed on 7 July 2019

[11] https://en.wikipedia.org/wiki/BEA_Systems ,accessed on 7 July 2019

[12] https://en.wikipedia.org/wiki/Service-oriented_architecture , accessed on 7 July 2019

[13] https://en.wikipedia.org/wiki/NoSQL ,accessed on 7 July 2019

[14] https://spark.apache.org/ ,accessed on 7 July 2019

The second paper is a press release from Pegasystems talking about Next-best-action marketing: a customer-centric approach ('Next-best-action marketing: a customer-centric approach' 2012). Without getting into too much detail, the paper outlines the business benefits of Next Best Action Marketing.

They are as follows:

- **Channel Consistency**. As the decision system coordinates the best decision for the customer considering all the information to avoid channel inconsistencies.
- **Cost Alignment**. NBA marketing is using predictive models to determine the most significant impact on the organisation.
- **Optimisation of Interaction Time**. Considering all the past interaction and current context NBA values customer time by engaging only with the most relevant information. NBA is a value creation interaction.
- **Natural Conversations**. From a marketing perspective, natural conversations are perceived by the customer as creating more value which generates greater satisfaction.
- **Effective Monitoring**. The decision system or central hub captures the outcome and logic of every decision; therefore, the basis and the result of that decision can be measured and accounted rigorously.

2.2 Segmentation Of Next Best Action Marketing

In his paper, "*What happened to strategic segmentation?"* (Jenkinson 2009) Angus Jenkinson puts the case forward for a reduced segmentation as opposed to full-scale strategic segmentation backed up by research. In the section titled Predictive Analytics, Next Best Action (NBA) describes a form of affinity modelling. Using customer historical information, an organisation can present the customer with the best offer. Jenkinson argues that this technique works even better in the context of a segmentation framework. However, based on the model presented in this paper, Next Best Action marketing is a unique offer presented to a segment of one. Besides, segmentation is the ultimate bottom-up product centric view of the customer, therefore not very useful when doing marketing in the 21st century.

To some extent *"Next product to offer for bank marketers"* (Kin-nam et al. 2003)., is yet another example where the authors argue that Standard Chartered Bank in Hong Kong could cater better to the needs and priorities of each customer by presenting each customer with a product as determined by a model. This paper presents a customer-centric approach in contrast to more traditional marketing campaign management. The model presented still relies on segmentation to make single product offers to a "list" of customers. The authors recognise that understanding customers should form the basis for marketing offers. To this end, they glean information such as life stage, occupation, wealth, income class, social class, external products holding, loyalty prices sensitively, gender, behaviours variables such as gambling and travelling. The next product to offer is using several deterministic inputs to the form “if this, then that” (IFTTT) statements. Also, this is still a current model for some large organisations in Australia.

2.3 Data Management Of Next Best Action Marketing

Closer to mark Bill Franks in his paper *“Analytics on Web Data: The original Big Data”* (Davenport 2013) presents the case for understanding customer intent. To capture customer intent organisations need to integrate detailed, customer-level behavioural data from their websites to their enterprise analytics data sources. Particularly actions that customer takes during their interactions with the organisation such as:

- Purchases
- Product reviews
- Shopping basket additions
- Watching a video
- Accessing a download
- Reading/writing a review
- Requesting help
- Forwarding a link
- Posting a comment
- Registering for a webinar
- Doing a search

For decisions to be effective, they need to be automated. As pointed out in the Automated Decisions (Davenport 2013), Davenport shows that optimal decision making occurs when business rules and analytics are embedded into business processes. This way, a clear line of sight can be achieved between how decisions are made and their outcomes.

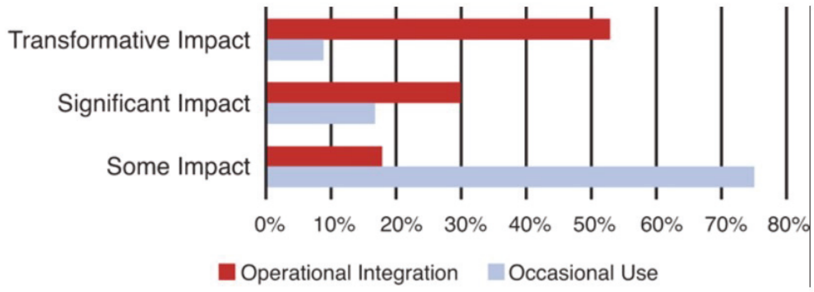

FIGURE 3 - THE IMPACT OF PREDICTIVE ANALYTICS ADAPTED FROMM (DAVENPORT 2013)

It has been shown that decision management is the way organisations can leverage on and capitalise their customer base with analytics. The chart in Figure 3 is based on a survey where respondents have reported on the level of transformational impact from integrating predictive analytics into their operations. It shows that more impact is experienced in operational integration than in other areas of the business.

In terms of Next Best Offer web, data provides a unique perspective on customer needs that otherwise would have been unknown (Jacopo Tagliabue et al. 2019). Considering the profile of a banking customer who has several accounts, and a high balance unless there is a clear indicator that the customer is shopping for a home loan a bank would not usually present the customer with such an offer. However, upon collecting real-time data from the web searching behaviour, such an offer becomes the Next Best Offer considering that the customer looked at various home loan options online.

2.4 Implementation Of Next Best Action Marketing Challenges

Valos J.M. provided a more interesting example in his paper (Valos 2010), which talks about the challenges of implementing a multichannel marketing strategy. By definition, a multichannel implies a product or channel-centric organisation. As explained in her paper, Lin Pophal (Pophal 2015) shows that organisations with multichannel strategies have siloed structures where each channel has its reporting structure and revenue goals. Siloed structures are by far the biggest hurdle in implementing an omnichannel marketing strategy as supported by the NBA model. This paper explains in the section titled Next Best Action conceptual vs. system based process that having a robust business architecture is a prerequisite for a customer-centric business architecture supported by NBA decisioning.

2.5 Possible Reason For Paucity of Literature In The Area Of NBA Marketing

Next Best Action Marketing can be viewed as the next logical iteration in the evolution of database marketing. Whilst this is great news for both marketers and technologists alike the discipline of marketing has moved more into technology in the past ten years or so. In the past marketers used to be very good copy writers and great communicators and story tellers. These days the new generation of story tellers paint pictures with numbers in the form of graphs and visual maps of customer journeys based on large data sets. On this basis NBA Marketing as a subset or technique of database marketing can be easily overlooked by both groups due to its technology heavy side and emphasis on command and control of customer interactions with data and facts. Over the past few years the author has noticed marketers embracing a more fact data-driven approach to marketing however this has delivered mixed results as marketers are ill-equipped with the necessary knowledge to make sense of the data collected. This brings into sharp focus the third point, NBA Marketing is reliant on good data for the models to work properly. Despite all the advances in the data space getting “good data” is still a challenge for most organisations. This is due to the product-centric siloed approach enterprises have been working with for many years. However, this is about to change.

3 ARCHITECTURE FOR NEXT BEST ACTION MARKETING

3.1 A Conceptual Framework For Next Best Action Marketing

Next Best Action Marketing is a paradigm shift;

A glance at a textbook definition of marketing (Kotler 1998) shows that the discipline started from attending to human needs, however, over time it has grown into a more extensive process of exchange in which individuals and organisations fulfil their needs and wants through value-creating transactions. In an ideal world, these transactions are moving back and forth with minimum waste (Naumenko 2015).

A definition of Next Best Action Marketing can be enunciated as the process of understanding, predicting and fulfilling the needs and wants of individuals and organisations using historical data and real-time contextual interactions. In other words offering the right product or service, to the right customer, at the right time in the right channel.

Therefore, the processing of information based on business rules and predictive analytics fulfilling the needs and wants of individuals and organisations are the key ingredients to the CRM method of Next Best Action Marketing (Sanjiv 2012).

The technological improvements over the past 10 to 15 years, namely computer power in the form of cloud computing as well as the proliferation of new technological patterns, has ushered in more effective ways in which the value-creating exchanges in the form of transactions ('Next-best action marketing: a customer-centric approach' 2012).

Next Best Action marketing falls into the broad spectrum of the methodologies for Customer Relationship Management (CRM). CRM technical applications involve functional solutions to deal with customer needs across the full customer lifecycle (Pāvels 2017). CRM includes Direct Marketing, defined as a measurable customer engagement process to generate a positive return on marketing investment (Stone & Jacobs 2008).

In his book, Bob Stone explains that *"to generate a customer response, direct marketers use four key variables called the Elements of Promotion"*(Stone & Jacobs 2008). According to direct marketing experts, customer responses comprise the following:

- selection of target audience determines 40 per cent of the response
- the offer, a proposed course of action accounts for 20 per cent
- while the language and communication (copy) accounts for 15 per cent
- creative layout 15 per cent and
- the timing of the offer represents the remaining 10 per cent (Stone & Jacobs 2008).

Next Best Action marketing is making use of very granular data to deliver the most relevant proposition perfectly timed at the right place or channel to the right customer. In effect creating a paradigm shift in how direct marketing is used, instead of using marketing as a demand generation tool, the “four Ps” and “push and pull” to target customers (Tzuo 2018). Next Best Action marketing focus is on understanding what type of problems the customers are trying to solve, then providing them with the right solution at the right time in the right channel (Gončarovs & Grabis 2017).

Next Best Action marketing is not a single product-centric transaction. Instead is a continuous customer-centric mini-case using predictive analytics and business rules to balance off customer needs to business priorities. Therefore when the interaction occurs between the organisation and customer, a unique case runs based on the parameters of that interaction which can consistently generate for that customer the correct outcomes (*Pega Marketing Product Overview 7.31* 2018).

For example, when a telecommunication customer applies for broadband services, the organisation can predict from the actual transactions of customers with similar features whether likely delays or issues are leading to customer dissatisfaction and low Net Promoter Score (NPS). Next Best Action mitigates this with proactive communication such as outbound Next Best Action, e.g. SMS, to inform the customer of possible delays or technical issues. The monitoring of these interactions happens in real-time. Therefore, the system takes corrective action, and the customers are experiencing better service. The business value behind NBA is an increase in Customer Lifetime Value (CLV) of the customer, which represent the net profit an organisation derives from the entire relationship with the customer (Stone & Jacobs 2008). Using CLV to determine Next

Best Action can provide long term benefit to the organisation ('Next-best-action marketing: a customer-centric approach' 2012). Also, satisfied customers are more likely to stay longer with the organisation. So CLV becomes a key measure in determining the success of the Next Best Action, but it is not limited to it. Likelihood to recommend is an essential input in NPS, so the NBA can be used to increase customer advocacy (Reichheld 2006).

Looking at another example in the casino industry, Next Best Action for a gaming customer can be the recommendation for a new game. Also, the casino gives credits for trial incentives on new games (Davenport 2013).

In banking and finance, leveraging real-time information is key to providing a differentiated highly tailored offer (Todd 2008). A predictive model for a customer with a low savings account balance would predict the customer is not likely to be interested in an investment product. However, customers with certificates of deposit (CDs) are not likely to have a high balance in their savings account. The customer may not be aware of investment opportunities with the bank, however, if the customer is calling the bank, a customer service representative (CSR) can present the customer with investment opportunities based on Next Best Action capability, using business rules and predictive models. So by building a model based on CSR engagement, the model can be very predictive using context to drive the interaction. Next Best Action marketing is not just about predetermined products and services but also about the ability of the system to consume real-time data. This way the CSR can select from a range of optimised offers, such as the top three offers cross-selling or upselling, to make the right offer at the right time in the right channel to the right customer.

What does Decisioning Framework do?

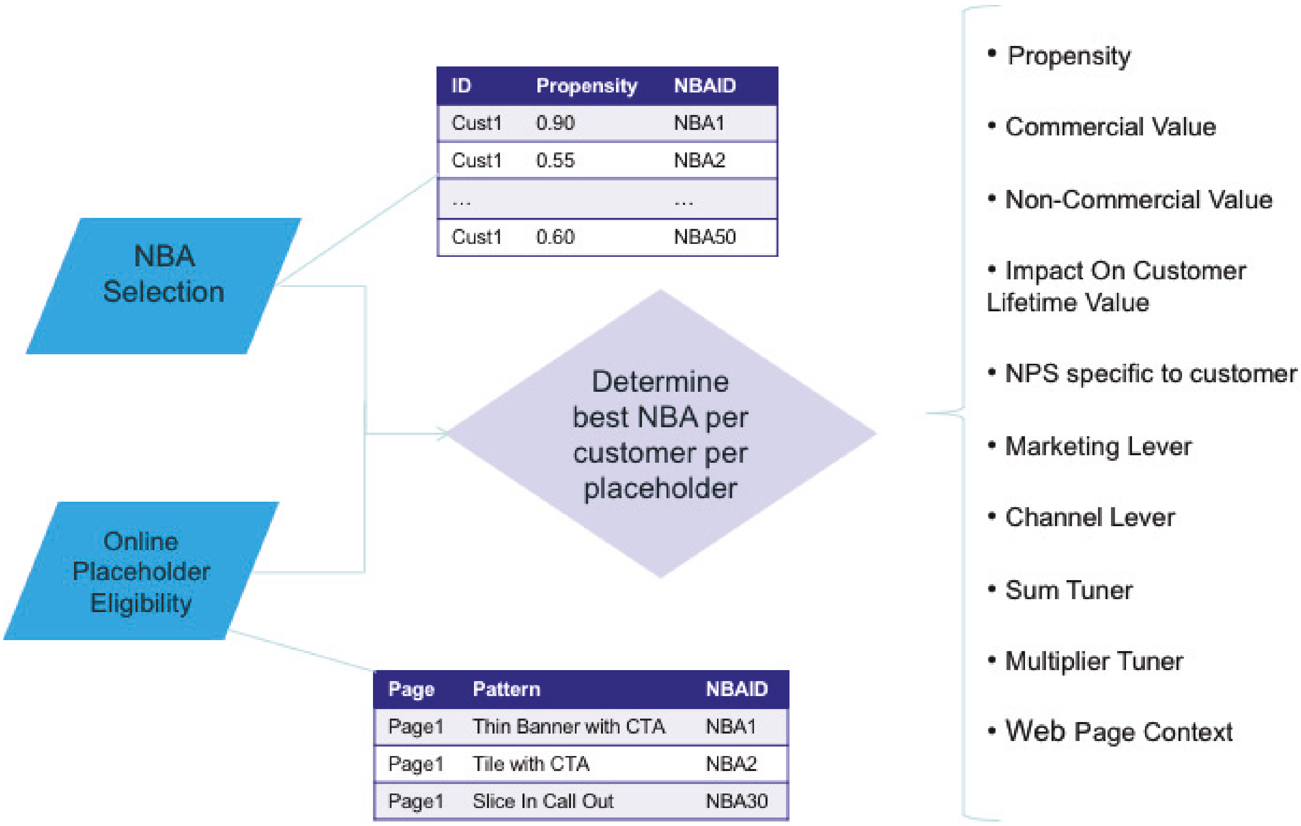

FIGURE 4 NEXT BEST ACTION FRAMEWORK

A Next Best Action for marketing decisioing framework is an optimisation process (Copulsky, Richardson & Simone 2017). The framework is combining several inputs to present the customer with the right message at the right time in the right channel. So how to know that is the right message at the right time in the right channel for the right customer? This work will delve deeper into the details covering the above figure into a dedicated chapter.

3.2 Next Best Action Marketing Conceptual Versus System-Based Process

The concept of Next Best Action marketing needs to be adopted by the entire organisation. A customer-centric organisation takes a holistic approach by grouping its resources around customers(Vaid 2017). Organisations with robust business architectures can capitalise efficiently on their data assets to optimise decision making (Whelan & Meaden 2012) for Next Best Action decisioning.

In Australia, the Commonwealth Bank has implemented and optimised its customer engagement model where all customer interactions across all its products and services are used as inputs to drive the next best conversation (Crozier 2018). This Next Best Action model searches horizontally looking for “the best” value creation interaction for the customer, thus bypassing the idiosyncrasies associated with a siloed organisation. The organisation is running one continuous marketing campaign for every single customer where customer responses are inputs into the next best action proposition.

Next Best Action marketing is more than just technology. Whelan and Meaden have explained that business architecture brings the "big picture" thinking into play and provides “the 20,000-feet view, that can bring those things into sharp focus that cannot be observed in daily business transactions of the silos” (Whelan & Meaden 2012).

Next Best Action marketing is a way of aligning organisational strategies with marketing activities. Therefore, the organisation is establishing a line of sight between its high-level strategic priorities to the marketing activities which represent the various tactics for achieving the strategic intent. Under these circumstances, monitoring of actions becomes a crucial tool for readjusting the customer conversations to become more relevant.

4 RESEARCH METHODOLOGY

This thesis is the product of both the theory and practice of the author's experience. The author has delivered several Next Best Action Marketing projects successfully in Telecommunication and Banking Industry and more recently in the Automotive Industry. The framework proposed here is technology agnostic; therefore, it provides a high degree of flexibility in terms of its implementation. In the case of the Next Best Action algorithm, this has been tested using a limited number of customer records. However, customer lifetime value is the ultimate measure of how impactful Next Best Action Marketing is. Next Best Action marketing is subject to further research over a more extended period, possibly 3 to 4 years.

The methodology used in section 4.4.2 illustrate the analysis of algorithms for transforming the model scores to posterior probabilities is as follows:

A real customer banking data set with 11,301 is used create a real-time Adaptive Model using Naïve Bayes. The following process is used to generate the results in section 4.4.3 Empirical Evaluation

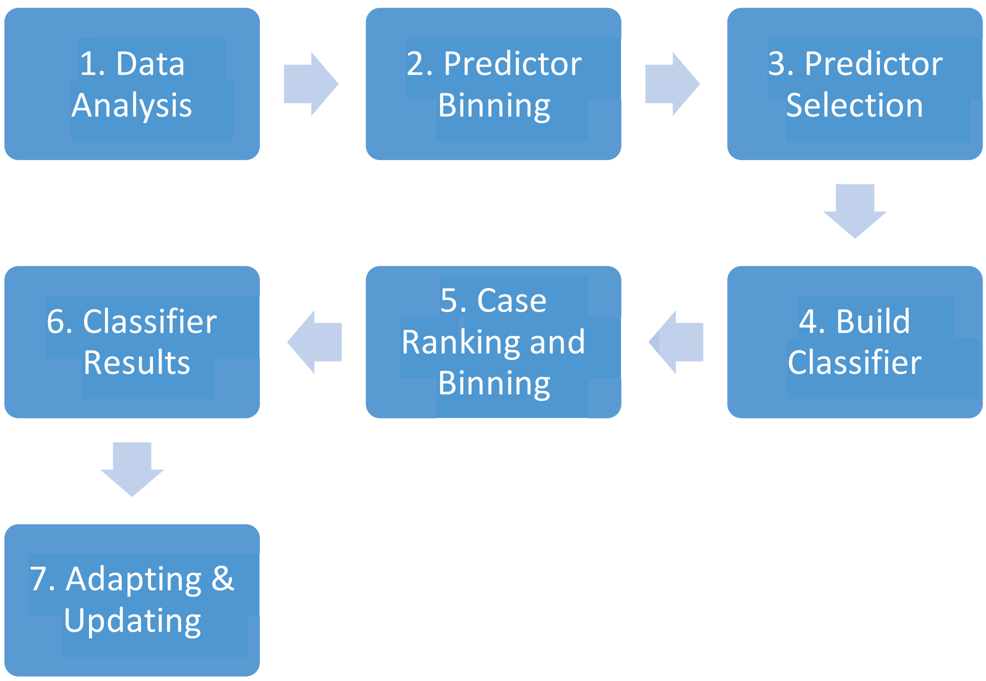

FIGURE 5 REAL-TIME ADAPTIVE MODEL PROCESS

Predictor binning is an important treatment of predictions in the Adaptive Model presented here. Therefore, two method of grouping predictors are compared and contrasted with how the Pool of Adjacent Violator (PAV) is grouping data.

The treatment of predictors proposed in this paper are the powerhouse of the Adaptive Algorithm. There are thousands of algorithms to create models there are only a few algorithms that deal with real world data and prepare it for models.

4.1 Business Architecture For NBA Marketing

NBA should combine good process with the right decisions. This relationship is illustrated with a matrix below exploring the relationship between process and decisions (Mauboussin 2006).

FIGURE 6 RELATIONSHIP BETWEEN PROCESS AND DECISIONS

Being both effective and efficient is dependent on the ability of the organisation to understand what proper processes are and making decisions by using inputs from all parts of the organisation. For example, finance can provide the payment history and credit score, and sales data can be used to leverage purchase history while marketing data can provide the history of the responses to different offers (Pāvels 2017). Besides historical data, the input streams can be websites, mobile, credit card, loyalty card, CRM, business process management, ERP, transactional data, social data, unstructured data, e.g. call centre text data, location data and syndicated data. All this data can be leveraged to calculate a one-to-one customer business case to generate “the best” proposition or action.

The place where the decisions are made based on a multitude of inputs acts like a crucible for fusing a variety of data inputs, on the one hand, generating NBA (propositions) on the other running the decision logic for operationalising them. Decisioning hub is the decision system for NBA marketing (Sanjiv 2012). Architecturally, this is where the decision framework is configured.

4.2 Data Used In Decisioning System

From a CRM perspective, it needs to store case history and conversations in the form of textual data for text analytics.

Engaging customers in the right place is vital. Therefore, decision hub can leverage channel information both outbound and inbound, e.g. email, direct call, point to URL, point to POS, direct mail, mobile message (SMS/MMS) to work out the best channel for each customer. By consuming real-time context from questions answered by the customer during the interaction, the NBA decision framework can be refined to become more relevant to the customer needs and wants at the point of interaction (Dietrich 2014).

Architecture for NBA decision requires the ability of the system to forecast and simulate trends, run predictive models, leverage big data technologies such as Kafka[15] Real-Time data, use adaptive learning, apply business rules and use real-time event processing as well as behaviour models, sentiment analysis and apply organisational knowledge to refine the proposition continuously.

The information required for the NBA decision system needs to be ready for consumption in the form of business rules or denormalised inputs into predictive models. For this purpose, input data needs to be denormalised to be decision-ready. Predictive models consume denormalised data as inputs into decision trees (Bertone et al. 2001), so that a score is generated on the fly.

[15] https://en.wikipedia.org/wiki/Apache_Kafka ,accessed on 10 August 2019

The decisioning table is designed to have all the features ready at hand to be consumed by business rules and predictive models at runtime. In marketing, it is critical to consider:

- recency, this is a business consideration as to how many times an offer is shown to a customer before the message wears out and it becomes ineffective
- offer eligibility takes into account product considerations like age and region, e.g. some products are no available for customers under 18 years of age
- the offer can be refined based on a specific margin which means only products above a specified margin will be offered customer propensity to accept a particular product out of several products and
- Lastly, calculate a priority score which on a fundamental level should be the product of margin times propensity to accept.

This way, the organisation captures in one compact score the business side of the equation, i.e. product margin and customer side, i.e. propensity to accept.

Score = product margin X propensity to accept

An NBA proposition is “best” because it takes into consideration the customer lifecycle along with a set of competing priorities when selecting the best proposition for a customer but also is sensitive to the business priorities in the process.

The customer lifecycle can be defined in four critical stages as follows: 1) customer identification, 2) customer attraction, 3) customer retention and 4) customer development (Pāvels 2017). These stages as it will be seen will become the scaffolding on which the architecture for NBA marketing is resting.

4.3 Structure For The Next Best Action Marketing System

A well-functioning system should be resilient, self-organised and hierarchical (Wright & Meadows 2012). Building a robust Next Best Action system should comply with three fundamental principles.

1. Next Best Action decision system is like a child who needs to learn what is right and what is wrong. For example, how many of the decisions are based on historical information as opposed to real-time data. Should the system evolve with every real time interaction and therefore learn and adapt to these, or be overwritten with deterministic business rules? The answer to these questions mostly depends on the overarching organisational objectives and vision (Whelan & Meaden 2012). A well- designed system should have the ability to self-organise (Wright & Meadows, 2012). As the system matures, it can learn and enlarge the range of the issues it can tackle. For example, it can build its offers based on specific inputs. For instance, for a customer who is experiencing mobile phone drop-outs, an NBA should be an offer to compensate for the inconvenience, something like free calls between specific hours. The critical factor is for the system to have enough information and the ability to self-organise. Also, allowing it to select from a ready-made library of modular propositions to make optimal decisions when required.
2. Both resilience and self-organisation are primarily dependent on a hierarchy. As the decisioning system becomes more sophisticated, it will need to rely on a hierarchy to store information. Hierarchies are good at reducing the level of information but preserving the core information.
3. Typically, Next Best Action system should have at least two levels of hierarchy ('Next-best action marketing: a customer-centric approach' 2012). The first level is built around what business issues need to be solved or dealt with at the point of customer interaction.

For customer-centric organisations, the business issues represent the customer lifecycles, such as identifying who the customer is, customer acquisition, customer growth and customer retention. Depending on which stage the customer is with the organisation, the appropriate decision logic mediates the transaction with the customer.

Each business issue should have a logical way to group information. As illustrated in Figure 7, under each business issue like a decision tree, there are branches which represent how the business wants to tackle a specific issue. The individual propositions are presented appropriately to the customer at the point of interaction.

Thus, by employing at least a three-level classification, such as business issues, strategic groups and propositions, the system is flexible enough to deliver the most relevant proposition, i.e. the most compatible with customer needs and wants and organisational objectives. Therefore, the elements that allow for the decision logic to flow in a Next-Best-Action system are:

- Business Issues and Proposition Group hierarchy. Represent the minimum structure on which the decisions rest.
- A proposition is a proposed course of action for a customer.
- A strategy represents the decision logic i.e. the combination of business rules and predictive models that forms the pattern for serving propositions to customers.
- Prioritisation is the way propositions are ranked within the decision logic.
- Arbitration is a sophisticated form of proposition prioritisation for a customer across several business issues and groups.

Therefore, a customer-centric model requires a good understanding of the customer lifecycle. The customer lifecycle is the process of understanding at which point the prospects are likely to become customers, then how to attract these prospects to become customers, once the customers are on-boarded understand how to retain them then and finally develop a deeper relationship with the customer (Pāvels 2017).

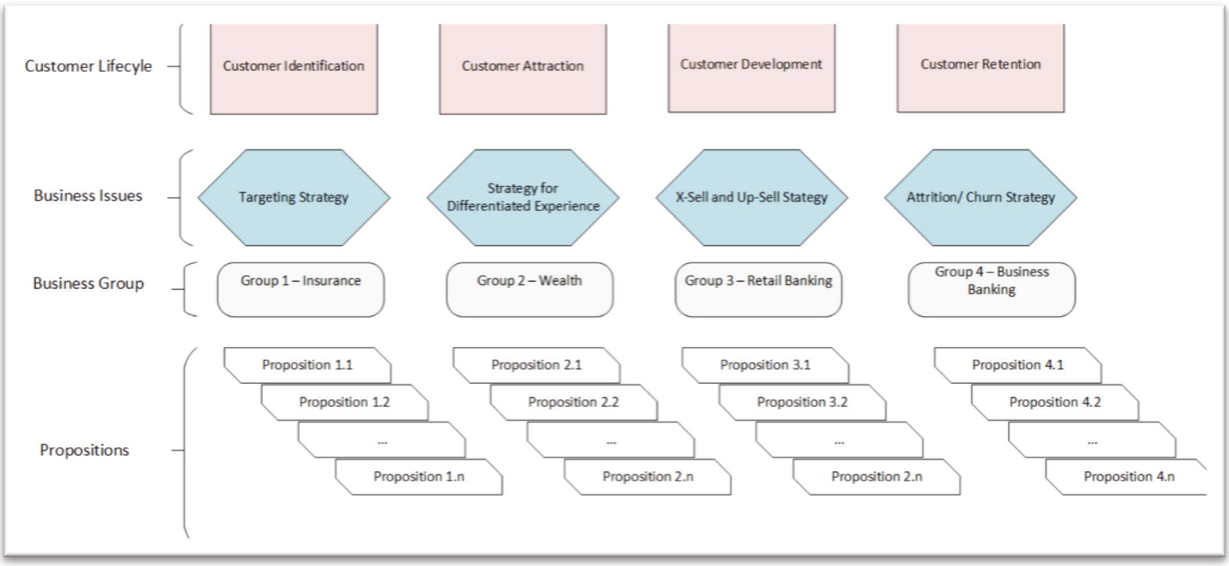

FIGURE 7 BUSINESS ISSUES AND GROUPS HIERARCHY

4.4 Analytics Of NBA Marketing – Adaptive Models

Introduction

This paper is intended to provide optimisation ideas and knowledge about the implementation of the Next Best Action model.

What is the point of an adaptive model?

Adaptive models have been around for quite a while and in fact, are a part of our daily lives since they can provide a reliable automated decisioning capability. Without the utility of this automated decisioning we would typically be confronted with a lot more tedious decision making and the inconvenience/time waste/inconsistency of having to manage things ourselves, examples of adaptive model use are:

- email spam filters
- document classification
- sentiment analysis
- proposition acceptance

THE CUSTOMER INTERACTION SCENARIO – USE CASE

In the most general/relevant use-case is working out whether a particular customer is most likely to either accept (A) or reject (R) a proposition offered to them via an arbitrary marketing channel. Which is greater: the probability of accepting or the probability of rejecting the proposition?

We have available to us a data set (features) relating to the customer, which doubtless cover their demographics, current status, previous behaviour and likely a whole raft of other features 'describing' the customer to us. The sub-set features that we know or believe have some 'power' to discriminate between acceptance and rejection, become our predictors.

So, with problem statement well defined and useful insights (Zhu & Xu 2016). How to combine these inputs into a solution that adapts to changes, requires little maintenance and is reliable in selecting the appropriate proposition for the customer?

4.4.1 THE NAÏVE BAYES MODEL

The family of naïve Bayes model is ideally suited to this situation. This intrinsic capability enables the simple and effective deployment of reliable decisioning models in many everyday marketing situations (Muler & Guido 2016).

Strengths

- Computationally efficient and effortless setup
- Tune performance by adjusting binning of predictors and suppression of ineffective predictors
- Adapts to change in behaviour
- No ongoing up-keep
- Empirical evidence shows strong performance
- Integrated performance monitoring/reporting

Weaknesses

- Makes strong independence assumptions
- Maybe outperformed by more 'technical' methods

THE UNDERLYING IDEA AND RATIONALE – BAYES THEOREM

It is rare to find an approach so well suited to its intended purpose and that it is both simple and intuitively appealing.

The following explanation is going to introduce the underlying ideas in sufficient detail for the concepts to be understood.

Consider that there are two events relating to the weather, we can express an idea; 'what is the probability that it is cold AND raining' in a shorthand form as; p(A and B) where A is the event that it is cold and B is the event that it is raining

Extending this a little, it is not hard to accept, applying the commutative property of probabilities that:

$p(A \text{ and } B) = p\ (B \text{ and } A)$ **(a)**
It Is also true that in general;
$p(A \text{ and } B) = p(A) * p(B|A)$ **(b)**

In words; 'the probability of it being cold AND it is also raining is equal to the product of the probability that it is cold, and the probability of it is raining GIVEN that it is cold'. We have specified that it is cold, so we can be more specific about it raining since we are only interested in the chance of rain when it is cold.

This equation is conditional probability – the probability of one event being conditional on another event.

Reorganising the equations (a) and (b) into:

$$p(A|B) = p(A) * p(B|A)$$

$$p(A|B) = \frac{p(A) * p(B|A)}{p(B)}$$ **(c)** Voila! Bayes Theorem

The equation is an essential idea because it enables a 'connection' between observable behaviour – the stuff on the right and an outcome – the left side.

If the event of the outcome is represented as O and the observed customer features as D we can re-write Bayes Theorem as;

$$p(O|D) = \frac{p(O) * p(D|O)}{p(D)}$$ **(d)**

In words this says;

$p(O|D)$ The probability of a customer accepting (or rejecting) a proposition (event O) GIVEN their feature set (event D) – this is called the POSTERIOR probability since it is the 'output' is equal to the product of:

$p(O)$ The 'global' probability of customers accepting (or rejecting) the proposition – this is called PRIOR probability, since it is known in advance, from training or business understanding

$p(D|O)$ The probability of the customer feature set for customers who accept (or reject) – this is called the LIKELIHOOD since it is the likelihood of a customer accepting/rejecting the proposition. Normalised by;

$p(D)$ The probability of all customers having the data values they have. The number is the NORMALISING factor since it scales (normalises) probabilities to fall between 0 and 1 i.e. values $0 < p < 1$

The equation is an essential idea because it enables a "connection" between observable behaviour (e.g. the features of a customer – the stuff on the right) and an outcome (e.g. likelihood of accepting/ reject – the left side) with the connection being made using data previously collected (the "training").

Making connections between behaviour and outcome has real value in the marketing domain, especially when the process can adapt and "learn" from observed behaviour – almost too good to be true!

APPLICATION OF BAYES THEOREM – THE ADAPTIVE MODEL

Via a brief discussion, the idea of conditional probability and Bayes Theorem have been introduced. So, the notional connection between predicting an outcome from observed behaviour was made in the last section; the question remains: how is it applied?

In practice, we are only interested in calculating the top line (numerator) of equation D, since the bottom line (denominator) does not depend on the outcome (O) and is constant, it depends only on the attribute values for each case (customer). The next section has a quotation from Wikipedia substituting O for C (the outcome) and d for x (the data attributes). The numerator is equivalent to the joint probability model (*Naive Bayes classifier* 2019):

$p(C_k, x_1, \ldots, x_n)$ where C_1 is the accept outcome, C_2 is the reject outcome and $x_1,\ldots,x_n$ are the data values for each feature $F_1,\ldots,F_n$ which can be rewritten as follows, using the chain rule for repeated applications of the definition of conditional probability:

$$
\begin{aligned}
p(C_k, x_1, \dots, x_n) &= p(C_k)\ p(x_1, \dots, x_n|C_k) \\
&= p(C_k)\ p(x_1|C_k)\ p(x_2, \dots, x_n|C_k, x_1) \\
&= p(C_k)\ p(x_1|C_k)\ p(x_2|C_k, x_1)\ p(x_3, \dots, x_n|C_k, x_1, x_2) \\
&= p(C_k)\ p(x_1|C_k)\ p(x_2|C_k, x_1)\ \dots p(x_n|C_k, x_1, x_2, x_3, \dots, x_{n-1})
\end{aligned}
$$

Now the "naive" conditional independence assumptions come into play (*Naive Bayes classifier* 2019): assume that each feature F is conditionally independent of every other feature $F\ for\ j\ \neq i$, given the outcome C This means that;

$$
\begin{aligned}
&p(x_i|C_k, x_j) = p(x_i|C_k) \\
&p(x_i|C_k, x_j, x_q) = p(x_i|C_k) \\
&p(x_i|C_k, x_j, x_q, x_l) = p(x_i|C_k)
\end{aligned}
$$

And so on, for $i\ \neq j, q, l$. Thus, the joint model can be expressed as

over the outcome variable C is;

$$p(C_k|x_1, \dots, x_n) = \frac{1}{Z} p(C_k) \prod_{i=1}^{n} p(x_i|C_k)$$

This means that under the above independence assumptions, the conditional distribution over the outcome variable C is ;

$$
\begin{aligned}
p(C_k|x_1, \dots, x_n) &\propto p(C_k, x_1, \dots, x_n) \\
&\propto p(C_k)\ p(x_1|C_k)\ p(x_2|C_k)\ p(x_3|C_k)\ \cdots \\
&\propto p(C_k) \prod_{i=1}^{n} p(x_i|C_k)\,.
\end{aligned} \quad \textbf{(e)}
$$

where; $C_1$is the accept outcome, C_2 is the reject outcome, $x_1, \dots, x_n$are the data attributes and $Z = p\ (x)$ is the NORMALISING factor, a constant for a known set of data attributes.

In general, Naïve Bayes is used for identifying the most likely outcome; this is achieved by evaluating (e) for all outcomes and assigning a classification to the outcome with the highest value. Our model varies from this procedure by calculating the log odds, and this is possible since we are only working with binary outcomes – accept or reject the proposition.

Log odds is a sensible choice since we are dealing with the product of values between 0 and 1, this product will very quickly become pretty small and in a practical case, ultimately lead to an arithmetic overflow; taking logs enables working with sums which will not vanish. The odds ratio (the ratio of p(accept) / p(reject)) is a right choice because it completely removes the dependence on the NORMALISING constant, but most importantly gives a single 'score' which increases with the likelihood of acceptance.

$$LogOdds(O) = \log\left[\frac{\frac{1}{Z}p(O_1)\prod_{j=1}^{n}p(d_j|O_1)}{\frac{1}{Z}p(O_2)\prod_{j=1}^{n}p(d_j|O_2)}\right]$$

$$= \log[p(O_1)] + \log\left[p\left(d_j|O_1\right)\right] + \cdots + \log[p(d_n|O_1)] \qquad \textbf{(f)}$$

$$-\log[p(O_2)] - \log\left[p\left(d_j|O_2\right)\right] - \cdots - \log[p(d_n|O_2)]$$

As long as we observe the convention $C_1$is the accept outcome, C_2 is the reject outcome, **(f)** provides a scoring measure for the acceptance of the proposition, since log odds is a monotonic transformation, the higher the score, the more likely acceptance of the proposition.

4.4.1.1 Estimating The Terms

To calculate the score, we must estimate each of the terms in F; estimating the terms is pretty straight forward and achieved by counting the number of cases in the bins for each feature.

The setup of the bins and determination of counts is completed by the data analysis process using the learning data to establish the initial estimates.

Estimating $p(O)$

Evaluating (O) is straight forward, the probability of customers accepting (or rejecting) the proposition. It can be estimated from historical data (the training set) or known from previous experience, to figure it out count the number propositions accepted and reject by customers to give the following table:

	Accept (O_1)	Reject (O_2)	Total
Count	n	m	n + m

The estimates for accepting/rejecting the proposition are straightforward to evaluate giving:

$$p(O_1) = \frac{n}{n+m} \quad \text{and} \quad p(O_2) = \frac{m}{n+m}$$

The upkeep of these estimates is pretty crucial since they exert leverage over the outcome and can 'drift' due to changes in marketing and economic conditions.

Our model allows the frequency, and 'memory' used to 'learn' these estimates to be adjusted and adapted to the circumstances.

Estimating $p(\frac{d_j}{O_j})$

The probability of the customer feature set for a customer who accepts (or rejects) the proposition, consider more carefully what this means: what is the probability of the customer having the observed value feature (d_j) GIVEN they accept/reject (O_1 or O_2) the proposition?

The frequency table for the feature (e.g. income level or recency of the last transaction) will look like the table below; there will be ‘k’ bins, and for each bin, the number of cases that have accepted or rejected the proposition will be identified.

Bin	Accept (O_1)	Reject(O_2)
Bin 1	q_1	r_1
.	.	.
Bin k	q_k	r_k
Total	q	r

The estimated probability of a feature having a value in bin ‘k’ when the proposition is accepting or rejected can be easily calculated as:

$$p(Feature\ having\ a\ vaule\ in\ bin\ \frac{k}{O_1} = \frac{q_k}{q}\left(Fature\ having\ a\ value\ in\ bin\ \frac{k}{O_2}\right) = \frac{r_k}{r}$$

The process is repeated for all features (i.e. all the *d’s*) and the expression, given in (f), is calculated for each case, giving each case a score for the likelihood of acceptance of the proposition. Cases can be ordered on the likelihood of acceptance.

4.4.1.2 ADAPTIVE MODEL TRAINING PROCESS

The adaptive modelling procedure is an implementation of the ideas described above – a Naïve Bayes classifier.

An outline of the steps followed by the NBA Adaptive Model to build a classifier from training data is detailed below. The treatment is not extensively detailed but explains the steps and outcomes at each point in the procedure.

DATA ANALYSIS

Each attribute (predictor) specified for inclusion in the classifier is subject to a preliminary analysis.

❶	Data Analysis	Each attribute (predictor) specified for inclusion in the classifier is subject to a preliminary analysis.
❷	Predictor Binning	A critical step in the procedure is the 'binning' (grouping) of each predictor, and this is the process of reducing continuous and categorical variables to an optimised set of categories.
❸	Predictor Selection	The Coefficient of Concordance (CoC) (Legendre 2005) for each predictor is calculated, predictors below the specified threshold are excluded. Pairwise correlations are made, and the lowest-performing predictor from any pair with a correlation over a given threshold value is excluded. This is an essential step since Naïve Bayesian classifiers can be derailed by multi-collinearity and this is a pretty good precaution.
❹	Classifier Build	The classifier is built from the frequency tables (based on the binned categories) for each selected predictor. The calculations and methods are described in the section above. The classifier is applied to each case (customer) in the learning/training dataset, the log-odds measure described above (f) is calculated for each case
❺	Case Ranking and Binning	The cases are ranked on the log-odds score from the classifier; this enables the classifier to be assessed from its ability to discriminate acceptance/rejection of cases via the coefficient of concordance Secondly, the cases are grouped using the log-odds score into a given number of bins. The observed propensity for acceptance is then calculated for the learning dataset. This is the 'baseline' classifier which can be used to reject/accept (classify) new cases based on the values of the predictors.
❻	Classifier Results	When the classifier is completed and assessed, summary details of overall classifier performance (CoC), and the performance of each predictor independently. The results for each predictor details the binning for each predictor (all predictors whether numeric or categorical are binned – see step 2), showing the frequency count of cases in the bin and a statistical test (Z-score) for the significance of the bin split.
❼	Adapting and Updating	The 'baseline' classifier can be applied to new cases within the context of the implementation, but the whole point of having an adaptive classifier is that it can adapt. Adapting in this situation means allowing the classifier to be re-calculated using new cases

4.4.2 ANALYSIS OF ALGORITHMS FOR TRANSFORMING MODEL SCORES TO POSTERIOR PROBABILITIES

INTRODUCTION

Most supervised learning methods produce classifiers that output scores s(x), which can be used to rank cases based on their probability to a certain class c. That is, for case x and y, if s(x) < s(y) then we have P(c|x) < P(c|y). For example, Naive Bayes (NB) classifier tends to produce scores that rank cases well. Under the hood, Adaptive Algorithms uses a Naive Bayes-like model for scoring (the difference is that we compute log-odds instead of loglikelihood).

Typically, un-calibrated scores produced by the classifier cannot be used directly as probabilities. For example, the "Naive" assumption in an NB classifier is that all attributes (or, predictors in our terminology) are conditionally independent given the class of cases. Because attributes tend to be correlated in real-life data, the scores produced by Naive Bayes are typically too extreme: say either 0s or 1s in a zero-to-one scale. Therefore, we need to come up with ways to transform ranking scores into accurate probability estimates.

There are practical reasons that we want to compute well-calibrated probabilities as well. In many applications, the ranking of cases is not enough. Probability estimates are needed for decision making when outputs are not used in isolation but are combined with other sources of information. Probabilities are more meaningful and more accessible to interpret than un-calibrated scores.

In the following sections, we look at three algorithms for score-to-probability mapping.

4.4.2.1 PLATT SCALING:

Platt scaling is a parametric approach in finding parameters for a sigmoid function that fit SVM scores. This method is motivated by the fact that the relationship between scores produced by an SVM model and the empirical probabilities appears to be sigmoidal for many datasets (Wang 2013). We do not investigate this approach further as it is shown in the literature that the sigmoidal assumption is broken for models such as Naive Bayes (Shalev-Shwartz & Ben-David 2014). Also, for such problems, we generally prefer a fast and non-parametric approach without any assumption on the distribution of the dataset.

4.4.2.2 BINNING:

Binning is a non-parametric approach that works for any shape of a mapping function (Racine 2019). An equidistance (aka Histogram) binning method is proposed. Firstly, the scores are sorted, and the score band (B = Smax - Smin) is divided into N bins so that each bin has a width of B/N. The probability is then calculated as the fraction of cases belongs to class c (c=1 in a binary problem) in that bin.

In Adaptive Algorithm modelling for Next Best Action Marketing, we employ a more advanced form of binning, namely, equi-weight binning. We have a max of 800 histogram bins of counts, and then we use an equi-weight binning algorithm to reduce them to 20 bins further.

Probabilities, in turn, are calculated from the counts in the bins.

The disadvantage of binning is that we have to pre-define the number of bins. Such a constant is not likely to be optimal for a wide range of datasets. Also, the size of histogram bins is fixed, and if the boundaries are chosen in such a way that we average the set of cases that should have different probability estimates, the estimations will not be accurate. This is one of the reasons we see that probability estimates are not monotonically increasing along with the scores, in the experimental results shown in the next section.

4.4.2.3 PAV (Pool Adjacent Violators):

"The PAV algorithm is conceptually straightforward. Given a set of training cases ordered by the scores assigned by the classifier, it first assigns a probability of one to each positive instance and a probability of zero to each negative instance and puts each instance in its group. It then looks, at each iteration, for adjacent violators: adjacent groups whose probabilities locally decrease rather than increase. When it finds such groups, it pools them and replaces their probability estimates with the average of the group's values. It continues this process of averaging and replacement until the entire sequence is monotonically increasing. The result is a sequence of instances, each of which has a score and an associated probability estimate" (Fawcett & Niculescu-Mizil 2007).

Algorithm 1 Basic PAV method for generating probability estimates

Input: Scored training set (f_i, y_i), where f_i is the score assigned by the classifier and y_i is the correct class.

Output: Stepwise constant function generated by m

begin
Sort training set instances increasing by f_i
Put each training instance in its own group, $G_{i,i}$ and predict $m_{i,i} = y_i$ for it
while $\exists G_{k,i-1}$ and $G_{i,l}$ ST $m_{k,i-1} \geq m_{i,l}$ **do**
    Pool the instances in $G_{k,i-1}$ and $G_{i,l}$ into one group, $G_{k,l}$
    $m_{k,l} = (\sum_{i=k}^{l} y_i)/(l - k + 1)$
    Predict $m_{k,l}$ for all instances in $G_{k,l}$
end while
Output the stepwise constant function generated by m
end

OUTPUTS OF PAV ALGORITHM:

#	Score	Probabilities						
		Initial	**a1**	**a2**	**b**	**c1**	**c2**	**d**
0	.9	1	1	1	1	1	1	1
1	.8	1	1	1	1	1	1	1
2	.7	0	0	0	0	0	0	3/4
3	.6	1	1	1	1	1	1	3/4
4	.55	1	1	1	1	1	1	3/4
5	.5	1	1	1	1	1	1	3/4
6	.45	0	0	0	0	1/2	2/3	2/3
7	.4	1	1	1	1	1/2	2/3	2/3
8	.35	1	1	1	1	1	2/3	2/3
9	.3	0	0	0	1/2	1/2	1/2	1/2
10	.27	1	1	1	1/2	1/2	1/2	1/2
11	.2	0	0	1/3	1/3	1/3	1/3	1/3
12	.18	0	1/2	1/3	1/3	1/3	1/3	1/3
13	.1	1	1/2	1/3	1/3	1/3	1/3	1/3
14	.02	0	0	0	0	0	0	0

It has been shown in the table above that for binary pattern classifiers, the non-parametric optimisation of calibration, subject to a monotonicity constraint, can be solved by PAV and that this solution is optimal for all regular binary proper scoring rules.

4.4.3 Empirical Evaluation

<u>Introduction</u>

The question is whether the binning method proposed for data analysis in the case of adaptive NB model is robust enough when compared to the PAV algorithm. In this section, we present the experimental results based on real customer dataset (banking industry).

The dataset has 11,000 records with the following features:

TABLE 1 - DATA SET USED IN EXPERIMENTS

Input data variable	Description
CustomerID	Unique customer identifier
Age	Age of customer
App	HasRegForApp
MaritalStatus	Married or not
IsBusinessCustomer	Yes/No
hasBadQualityCreditAccount	Yes/No
HomeLoanAccountBalance	Yes/No
HLApptLastReqTimeWeb	Home loan application online
HLApptLastReqStatusApp	HLApptLastReqStatusApp
Job Title	Classification of the job
PLNotClosedInLast3Months	Personal loan not closed in the last three months
PLApplicationChannel	On which channel has the personal loan been applied on
PLCount	Duration in days of personal loan
IsLargeDepositInLast7Days	Show the number of deposits
SumMonthlyAccountKeepingFees	SumMonthlyAccountKeepingFees
TxAccountSumDishonourFees	Transaction account dishonour fees
Accept	Acceptance of the offer

The experiments below are comparing the binning classification versus PAV classification. As explained responses are assigned to a bin, and the bins are combined to derive significant grouping of the response data.

Experiments

EXPERIMENT 1

In experiment 1 compares the binning method versus the PAV algorithm for the duration of a personal loan (PLCount) as can be seen below the probability of assigned a response to a specific range is very similar.

TABLE 2 - FREQUENCY BINS AND CLASSIFIER FOR DURATION OF PERSOANL LOAN

Frequency Bins	Responses (in %)	Positives count	Positives percentage	Negative count	Negative percentage	Propensity (in %)	Z-Ratio	Lift
[2, 98.24>	23.9	3	0.55	2,698	25.07	0.11	-46.63	2
[98.24, 180.16>	27.28	0	0	3,083	28.65	0	-65.74	0
[180.16, 364.48>	28.87	9	1.66	3,254	30.24	0.28	-40.48	6
[364.48, 477.12>	7.25	42	7.76	777	7.22	5.13	0.46	107
[477.12, 640.96>	5.34	106	19.59	497	4.62	17.58	8.71	367
[640.96, 3881]	7.36	381	70.43	451	4.19	45.79	33.59	957
Total Responses	100.00%	541	100.00%	10,760	100.00%	4.79%		

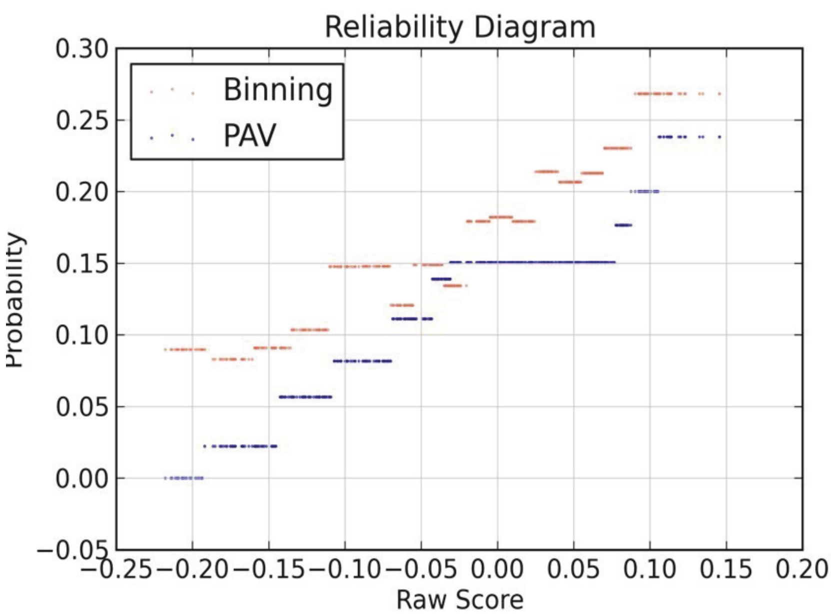

FIGURE 8 BINNING VS PAV (PL COUNT)

EXPERIMENT 2

Experiment 2 is looking at the classification of home loan account balances. This experiment shows that the PAV algorithm is slightly more reliable; however, the results are very similar.

TABLE 3 - FREQUENCY BINS AND CLASSIFIER FOR HOME LOAN BALANCES

Frequency Bins	Responses (in %)	Positives count	Positives percentage	Negative count	Negative percentage	Propensity (in %)	Z-Ratio	Lift
[-8019, -52.92>	9.31	33	6.1	1,019	9.47	3.14	-3.16	66
[-52.92, 274.76>	36.55	181	33.46	3,949	36.7	4.38	-1.56	92
[274.76, 66721]	54.15	327	60.44	5,792	53.83	5.34	3.07	112
Total Responses	100.00%	541	100.00%	10,760	100.00%	4.79%		

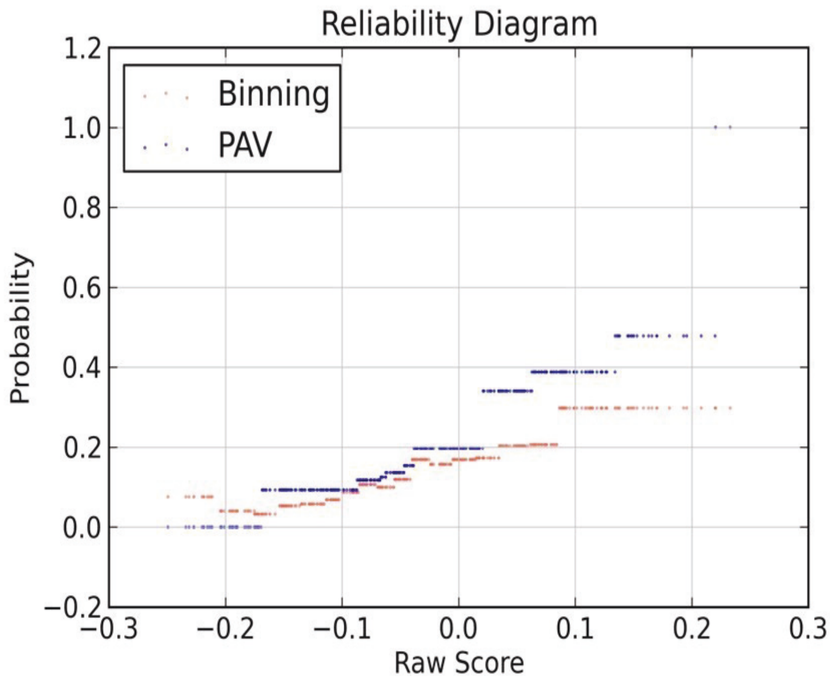

FIGURE 9 - BINNING VS PAV (HOME LOAN BALANCE)

EXPERIMENT 3

Similar results are gleaned from experiment 3 where the binning of the variable IsLargeDepositLast7days

TABLE 4 - FREQUENCY BINS AND CLASSIFIER FOR LARGE DEPOSIT IN THE LAST 7 DAYS

Frequency Bins	Responses (in %)	Positives count	Positives percentage	Negative count	Negative percentage	Propensity (in %)	Z-Ratio	Lift
1	31.05	151	27.91	3,358	31.21	4.3	-1.67	90
3, 2	43.12	239	44.18	4,634	43.07	4.9	0.51	102
4	9.34	72	13.31	984	9.14	6.82	2.8	142
19, 55, 35, 17, 36, 18, 33, 15, 34, 16, 13, 14, 11, 38, 12, 21, 20, 43, 41, 46, 22, 23, 24, 25, 26, 27, 28, 29, 10, 30, 7, 6, 32, 5, 31, 51, 9, 8, 50	16.49	79	14.6	1,784	16.58	4.24	-1.27	89
Total Responses	100.00%	541	100.00%	10,760	100.00%	4.79%		

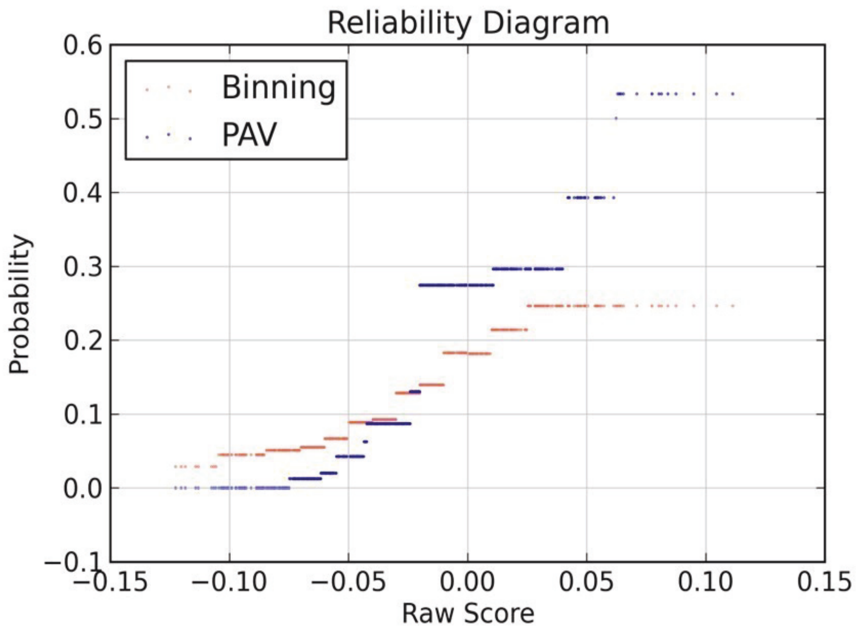

FIGURE 10 - BINNING VS PAV (LARGE DEPOSIT IN THE LAST 7 DAYS)

EXPERIMENT 4

Lastly looking at similarly looking that binning for the acceptance rate for we find that binning is robust enough to be used in classification for the adaptive algorithm.

TABLE 5 - FREQUENCY BINS AND CLASSIFIER FOR JOB TITLE

Frequency Bins	Responses (in %)	Positives count	Positives percentage	Negative count	Negative percentage	Propensity (in %)	Z-Ratio	Lift
admin.	10.83%	52	9.61%	1,172	10.89%	4.25%	-0.98	89
blue-collar	24.98%	125	23.11%	2,698	25.07%	4.43%	-1.06	92
technician	17.50%	95	17.56%	1,883	17.50%	4.80%	0.04	100
services, management	29.18%	176	32.53%	3,122	29.01%	5.34%	1.71	111
unemployed, housemaid, retired, unknown, self-employed, entrepreneur, student	17.50%	93	17.19%	1,885	17.52%	4.70%	-0.20	98
Total Responses	100.00%	541	100.00%	10,760	100.00%	4.79%		

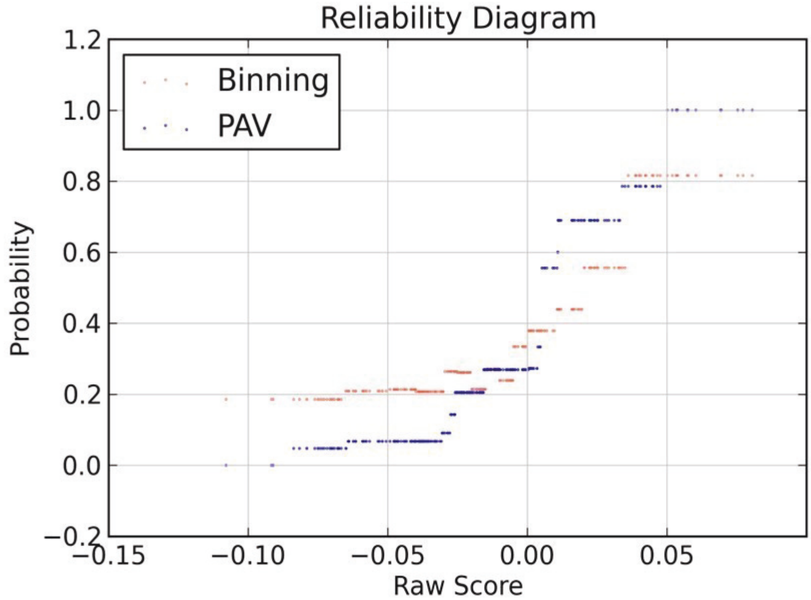

FIGURE 11 - BINNING VS PAV (JOB TITLE)

The above experiments are conducted using Python script the logic for the PAV algorithm is illustrated in figure 3, while the logic for assigned responses to bins is covered in the Binning section. In conclusion, it has been shown that both PAV and Binning are effective ways of determining the posterior probabilities for a Naïve Bays classifier. The following section will cover in greater detail the proposed process for data anlaysis.

4.4.3.1 DATA ANALYSIS BINNING

The data analysis involves the creation of a binned, ordinal view of both character and numeric predictors(Larose & Larose 2015). This is an ordinal problem. The reason is an ordinal problem is because it discriminates between positive and negative cases.

Preliminary binning Numerical Values

- Missing values (NaN) treated separately (in a separate bin)
- Special values (for instance -999 indicates missing) treated separately Preliminary binning: use 100 bins to:
    1. Equi-weight: make sure each bin contains an equal number of weighted cases
    2. Equi-distance: split boundaries evenly over the range (histogram)

Equi-weight is preferred. Take for instance "wealth", make a histogram of 100 bins and include someone who owns a large company in the sample. For example the wealth is split into 100 bins and Jeff Bezos who is worth

111.8 billion USD is part of the sample then the first bin will be between 0 and 55 million that's everybody in the sample so it would not work. With equi-weight Jeff Bezos, Bill Gates, Elon Musk will go into the final bin because there are not many of them.

Preliminary Binning (Non-Numeric) Character

This involves the creating of a list of the most 100 most frequent characters. For example if email address is considered a predictor then there are quite a few possible predictors. Therefore the 100 most frequent makes sense. The rest are assigned to "remaining character" bin

The missing values are treated separately

Then compute the propensity for each of the bins. Now there are numeric (ordinal) fields because the character is replaced with propensities

Grouping of predictors into bins

For numeric and character we start with maximum of 100 bins with:

1. An ordinal value (either the numeric value or probability)
2. The weighted number of positive cases falling in that bin
3. The weighted number of negative cased falling in that bine

The case count (1 per actual record) represents the sample, and (2+3) represents the population

The process of grouping will take the 100 bins of data and try to create a statistically robust grouping of, say, 3 to 10 bins of data. It does this trough merging (instead of partitioning)

	P(weighted)	*N(counted)*
admin.	10.83%	1224
blue-collar	24.98%	2823
technician	18%	1978
management	22%	2186
services	11%	1112
entrepreneur	4%	362
self-employed	4%	383
retired	4%	447
housemaid	3%	402
unemployed	3%	236
student	0.00	62
unknown	0.00	86

For example, if we take two similar bin we can calculate the Z-Ratio for merging bins (Jekel, Elmore & Katz 1996)

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2}}}$$

(Null) Hypothesis: Assume that Management and services have the same behaviour estimated at 16.5%) associated with them. How likely is that if we look at the samples of 2186 and 1112 respectively, that we observe management at 22% and services at 11%? **The Bootstrap Method** One possible way of doing this

1. Create a big pile of 3298 cases, of which 544 are positive 2. Repeat 5000 times:
    a. Draw 2186 at random (with replacement) and call this stack "management"
    b. Draw 1112 case at random (with replacement) and call this stack "services"
    c. Measure the difference in behaviour between stack "management" and stack "services", and attached to the result array
2. Sort the result array, and check what the rank of the "observed" difference of 1% is 4. Divide the rank by 5000, and here is the probability

Algorithm 1 Proposed method for generating numeric bins

Input: Two parameters: minimum % of cases + probability threshold

Output: Statistically significant binned predictors

1: **begin**

2: find the bin with the minimum number of sample cases

3: if the number of cases < threshold, merge with the neighbour (lowest (absolute) Z- ratio)

4: if the number of cases > = threshold, do Z-ratio test, then merge if the probability is larger than the threshold

5: then we merge if we don't have any reason to assume that's anything else but chance behind the difference in intervals

6: continue until no new bins can be found to merge

7: **end**

The minimum percentage of cases represents **the threshold**. For example, if minimum percentage of cases is 10% and the number of cases found in one bin is less then 10% then merge the bin. The minimum percentage of cases in one bin is found by counting the number of cases in each bin and diving by the total number of cases e.g. if we have a total of 100 cases out of which 10 cases are found in bin A then we have 10% of cases in bin A. **The probability threshold** is set by the Z -ratio.

The idea is to start with 100 bins and then merge it down to 5 to 10 bins

For example, after numeric grouping

Interval	P
Missing	0.2
-999	0.3
<0.1	0.0
<0.5	0.6
<10	0.3
Higher	0.4

Missing Values

Special Values

Lowest bin + underflow

Highest bin + overflow

CHARACTER GROUPING

Algorithm 2 Proposed method for generating non-numeric (character) bins

---

Input: Two parameters: minimum % of cases + probability threshold

Output: Statistically significant binned predictors

1: **begin**

2: group all intervals below minimum % of cases into a residuals bin

3: find the bin with the minimum number of sample cases.

4: find the bin with minimal (absolute) Z-ratio against the bin in (3), test, merge if probability is larger then the threshold

5: continue until no new bins can be found to merge

6: the bin with the missing value is not merged

7: **end**

The minimum percentage of cased can be set to 5% so everything below 5% can go into the residuals bin.

After character grouping:

Interval	P
Missing	0.2
Wide of Scheme	0.4
Remaining characters	0.3
Residual	0.0
Axy, SCB	0.6
Ee,..	0.4

Missing values

Unknown (never encountered before character get expected propensity (i.e. average propensity)

Cut out because of the parameter

4.4.3.2 How to measure the quality of the models:

It turns out that the Coefficient of Concordance (CoC) offers a good measures “discrimination”

$$\frac{\sum_i^N \left[p_i \sum_j^{i-1} n_j + \frac{n_j p_j}{2}\right]}{\sum_i n_i \sum_i p_i}$$

Suppose you are scoring two cases, one is 'positive', one is 'negative'. So, 'CoC' is the probability that you give the 'good' case a higher score than the 'bad' case.

N	**N**	**P**	**N**	**N**	**P**	**N**	**N**	**P**	**P**
0.1	0.2	0.3	0.32	0.4	0.43	0.55	0.60	0.8	0.9

We have several cases on which we compute the scores, now we want to assess how good that model is in discriminating between positive and negative. What is the probability that the model will give a higher score to a positive case rather than a negative case.

Possible pairs = 6 negative x 4 positive = 24 – this means that we have 24 possible discrimination points in the data.

Hits = 4*2 + 3*2 + 2*2 = 18/24 = 75%

In the table we can see that all 4 positive have a higher score than the first two negative.

N	**N**	**P**	**P**	**P**	**P**
0.1	0.2	0.3	0.43	0.8	0.9

Hits = $\boxed{4*2}$ + 3*2 + 2*2 = 18/24 = 75%

Next we can see that we have three positives with higher scores than the next two negatives. This means 6 extra pairs for which we are correct.

N	**N**	**P**	**P**	**P**
0.32	0.4	0.43	0.8	0.9

Hits = 4*2 ++ 2*2 = $\boxed{3*2}$ 18/24 = 75%

Lastly, the two positives and two negatives

N	N	P	P
0.55	0.60	0.8	0.9

Hits = 4*2 + 3*2 += 18/24 = 75% 2*2

So out of all 24 possible ways we have 18 correct which means we have a coefficient of concordance of 75%. This means that both positives and negatives are in proper score range.

Back in the 90s a lot of academic literature used accuracy as a measure of discrimination this means that you are trying to find a model that is accurate as possible(Diebold & Mariano 2002). The reason accuracy many not be as robust as CoC is that accuracy assumes some kind of segmentation. With accuracy you decide where the cut-off is then call what is bad and what is good what is the percentage of times that I am correct. With CoC you postpone the segmentation question until you have a good scoring model.

With ranked scores, and n the sum of negative for that ranked score, and p the sum of positive (Yitzhaki & Schechtman 2013):

$$CoC = \frac{\sum_i^N \left[p_i \sum_j^{i-1} n_j + \frac{n_j p_j}{2}\right]}{\sum_i n_i \sum_i p_i}$$

So, we have an array of positives and an array of negatives they are the same lengths and that's the sum of the positives at a particular score times or the negatives that have a higher score. What does it mean if you have a model and you plot some positive cases and some negative case you want the higher score for more and more positive cases you want to have 100% positive cases before you start selecting negative cases? That's impossible to achieve so in reality you select a couple of positive an a couple of negative cases so the CoC measure the area above the curve in the image below. So the coefficient of concordance can be expressed as CoC = 2 * (Gini Coefficient – 1).

CoC is also known as:

1. Gini coefficient (different normalisation) (Warrens 2018)

2. Receiver Operator Character Curve (ROC-Curve) (Krzanowski & Hand 2009)

3. Somers Dxy Rank Correlation (Newson 2006)

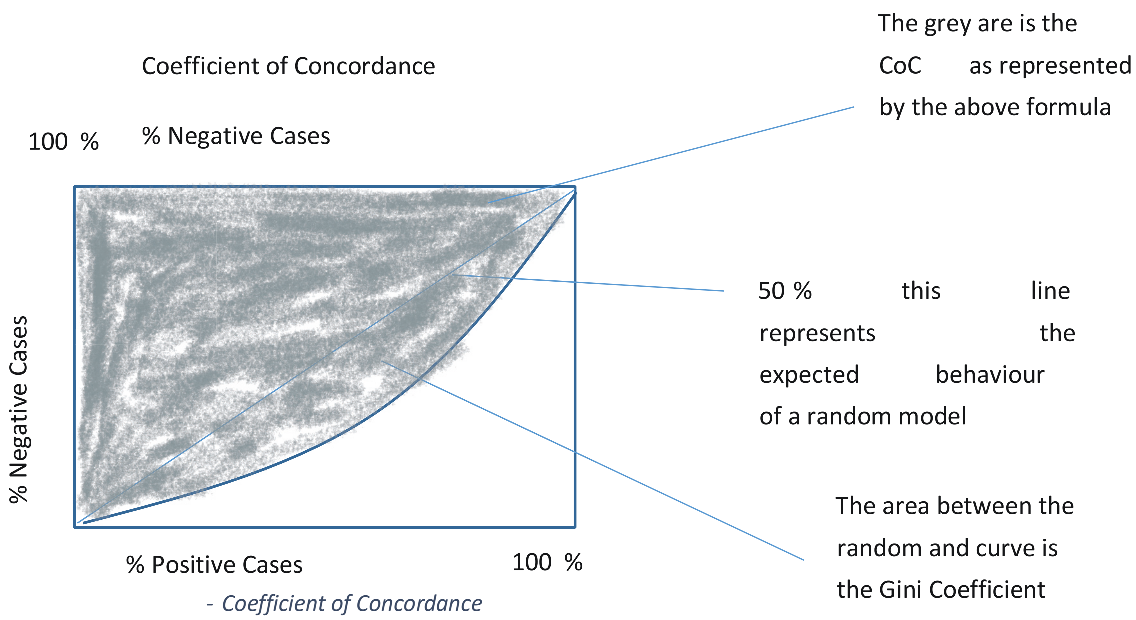

FIGURE 12 - COEFFICIENT OF CONCORDANCE

CoC is a convenient because of high granularity, and because it is ‘cutoff independent’

So in conclusion the coefficient of concordance measures how likely it is that we give a positive case a higher score than a negative case and this makes it a very robust measure. With this information in mind now we can turn to grouping of predictors into bins.

4.5 A Model To Calculate The Final Score For A Next Best Action Proposition

Introduction

Based on the extensive practical experience of the author, the following model is proposed to calculate the score of one Next Best Action per customer per online advertising space. $Next\ Best\ Action = f(x)\ where\ x = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10})$

As discussed earlier in the paper, Next Best Action marketing is a unique proposition that a customer is presented with at the most appropriate time and place during their interaction with the enterprise. Moreover, a Next Best Action should be the result of an arbitration process across several business issues and groups. To this end, an NBA is a function of ten distinct inputs.

The model illustrated below shows the processes, calculations and decisions that the ten inputs are used to calculate at the final score for a customer per advertising space.

4.5.1 Inputs into the final score calculation

Propensity (X_1) is the proxy for customer preference for one action over the other. Both supervised and unsupervised models calculate the propensity. For example, this could be using a Multi-Aspect Recommender System for Point-of-Interest as explained in their paper (Li et al. 2015) MARS model can be used to harness multi-aspect user preferences. This model captures both sides the business side, e.g. for a merchant and a user. This way, user preferences can be matched to merchant's quality providing increase customer utility and NPS.

Commercial Value (X_2) is the value to the business (merchant) of a customer taking up an offer. It is used to calculate the overall value of the NBA, which is then used to rank customers on value (subsequently used to multiply with P (propensity) to obtain a final ranking).

Non-Commercial Value (X_3) is the value to the business of a customer responding to a recommendation. These values can be estimated, e.g. a value between \$1 to \$10 per click.

Impact on Customer Lifetime Value (X_4) is the dollar impact on customer lifetime value when responding to a proposition. In effect, this is a crucial input into the final score as a measure of how effective the decision is. Right decisions result in propositions that increase customer lifetime value (CLV) bad decisions have no impact or even lead to churn, i.e. customer leaving the business for better deals elsewhere.

NPS Specific to Customer (X_5) is the dollar impact on customer net promoter score (NPS) when responding to a proposition. This can be used to determine a way to calculate NPS for propositions at the NBA level. The expectation is that some propositions will drive more customer advocacy over the others.

Marketing Lever (X_6) controls how propositions can be boosted or down-weighted based on marketing priorities. This can be used as an exception and should be based on planned marketing calendar, for example, a bank in September may want to boost Financial Advice as part of integrated Investment campaign.

Channel Lever (X_7) provides how propositions can be up-weighted or down-weighted based on channel priorities. Channel Lever should be used as an exception, and it applies to an organisation that has not implemented an omnichannel strategy. Therefore, it should be based on a planned channel strategy. For example, up-weight credit cards this month as the online channel is not meeting credit card sales targets.

Sum Tuner (X_8) provides a lever to deal with ad-hoc requests. Sum Tuner can be used to increase or decrease the first score values based on ad-hoc requirements to meet business emergencies. For example, to push credit card campaigns by 10% or move home loans down by 10%. At all times, the value in this field should be zero unless there is a good reason not to be. The sum tuner works by adding or subtracting percentages directly to the score.

Multiplier Tuner (X_9) provides a lever to deal with ad-hoc requests. Just like the Sum Tuner, this can be used to increase or decrease the first score values according to situational requirements like pushing credit card campaigns by 10% or move home loans down by 10%. At all times, this field should be one unless there is a good reason not to be. An analyst should use the multiplier tuner only in response to an urgent business request. The analyst should determine whether sum tuner or multiplier tuner is most appropriate based on the nature of the request.

Web Page Context (X_{10}) In the digital space, some pages are more appropriate for showing certain NBAs than others. This way, the optimiser should up-weight or down weight NBAs based on the web page. The web page context values are made up of values from -5 to 5. For example, a value of 5 will boost the score for appropriate NBAs by a factor of 0.6. This way an optimised proposition will take precedence over other propositions for which the customer may be eligible but are not context-sensitive.

The process flow below describes in detail the calculations carried out to arrive at the final score, which will determine which offer the customer will be presented with. The starting point of the process flow is a Master Customer Information File (MCIF) for which the customer propensities for each proposition have already been determined using supervised and unsupervised predictive models.

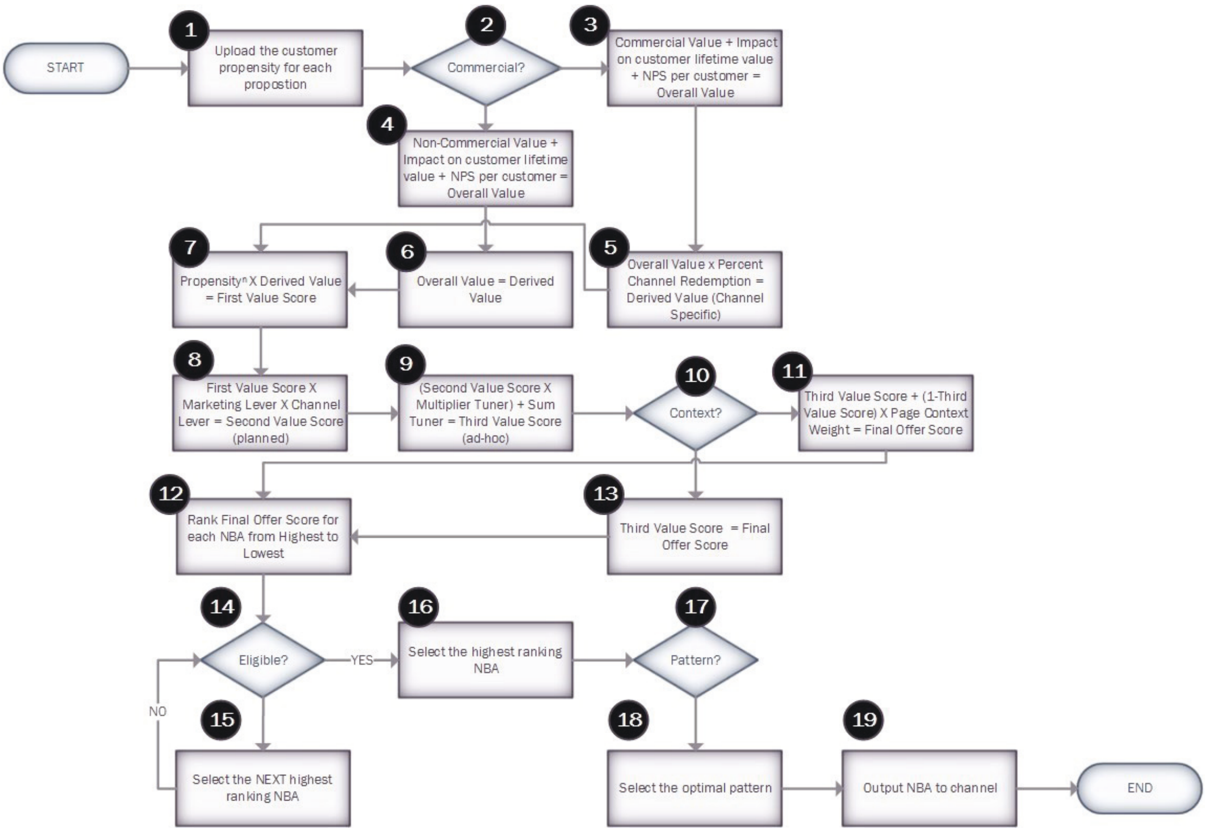

FIGURE 13 - NEXT BEST ACTION DETERMINATION PATTERN

4.5.2 NBA Process Description

TABLE 6 - STEPS FOR DETERMINATION OF NEXT BEST ACTION FOR A DIGITAL CUSTOMER

Step	Type	Description of the calculation
1	Process	Upload a list of propositions propensities for each customer o the file.
2	Decision	If the NBA is commercial then a sale would be made send it down the path where commercial value is used in the calculation Else use a non-commercial value in the calculation for NBAs not resulting in a sale
3	Process	Overall value to be calculated as follows: Commercial value of the product or service plus impact on customer lifetime value expressed in dollars plus normalised NPS value specific to a customer.
4	Process	Overall value to be calculated as follows: the Non-commercial value of the proposition plus impact on customer lifetime value expressed in dollars plus normalised NPS value specific to a customer.
5	Process	For non-commercial propositions, the Overall value equals Derived Value
6	Process	For commercial propositions, the Derived Value is a channel specific function. Is using the Overall Value calculation in step 4 times the percentage in redemption in the channel. This is to distinguish between online propositions with a redemption rate of 10% versus a call centre with a redemption rate of 50%.
7	Process	The first value score is calculated as the product between propensity score at the power of n and Derived value. The power of n used in the propensity score is optional, and it is used to control the distribution of the propensity for a particular proposition. This provides a form of control over how propositions may be biased towards sales-type propositions.

8	Process	Second Value Score is calculated for planned interventions, i.e. Marketing and Channel. It is the product of First Value Score and Marketing Lever and Channel Lever.
9	Process	Second Value Score is calculated for ad-hoc emergency interventions. This is the product of Second Value Score and Multiplier Tuner plus Sum Tuner.
10	Decision	If the Page Context is different from zero, then calculate the Final Offer Score otherwise use the Third Value Score
11	Process	Use the following formula to calculate the Final Offer Score. Third Value Score plus one minus Third Value Score times Page Context Weight
12	Process	Make Third Value Score equal to Final Offer Score
13	Process	Rank Final Offer score in descending order
14	Decision	If the highest-ranking proposition is eligible to go on page N the select otherwise go back to the next highest proposition
15	Process	Select the highest-ranking proposition
16	Process	Select the next highest-ranking proposition
17	Decision	Based on external analytics determine the optimal pattern, i.e. advertising space to place the proposition
18	Process	Select the optimal pattern
19	Process	Output selection for the Master Customer Information File

Repeated simulations of the above algorithm have shown that context, as a rule, is the main booster of the final score as it should be because it is based on real-time feedback.

ANALYSIS AND JUSTIFICATION OF THE MODEL

The above model has grown out of the praxis of implementing NBA scoring algorithms. As outlined earlier in the paper a straight forward score should take into consideration the balancing of the business objectives and customer needs. And this can be addressed with the simple formula - *Score = product margin X propensity to accept.* However, enterprises are complex ecosystems where various interests interact and one key requirement is to capture and measure the interaction between different teams and their priorities. For example, how to account for a channel priority? This is the situation where for example the call centre has prioritised a specific type of call like home loans enquiries, this is going to skew the score in favour of home loan enquiries through the call centre channel. While this may bias the equation in the favour of the business it is also important to track what decision was made in a specific context. While the model may come across as having “too many moving parts” this is a design feature. The levers in the model can help the business understand the impact it has on the customer interaction when the customer needs are overwhelmed by business priorities. Over many such interactions with the help of a control group some very completing insights can be garnered.

4.6 Next Best Action Marketing Delivery Methodology

INTRODUCTION

This document provides a detailed overview of a Next Best Action Marketing Delivery Model and how it should be used for decisioning projects. It is based on empirical studies for project delivery and provides a framework for Next Best Action Marketing projects and associated materials such as deliverable templates and tools. Together in partnership between business, system integrators and technology, this model creates a foundation for success for Next Best Action Marketing projects.

4.6.1 Managing Customer Experience Through Next Best Action

Businesses around the world understand that the customer-centric model is the way business is done in the 21st century. Therefore, they need to ensure that their interactions with their customers are always relevant, timely, cost-efficient and compliant. Big enterprises, with millions of customers, across multiple brands and channels, and complex products, need to ensure that every interaction with their customer is the best it can be. With increasing regulation, intense competition and a need to market well while continuously balancing risk, the quality of the millions of decisions businesses make when interacting with their customers will determine success or failure (Taylor 2012).

Next Best Action Marketing decisioning systems provide organisations with the ability to control the experience of customers by centralising, managing, monitoring and optimising the decisions made at the time of interaction. Creating a mini-business case for each customer to run in real-time, the decisioning system can ensure that the customer is always offered the best or most relevant offer. Whether that action is a sales offer, a retention strategy or a risk mitigation action, businesses using the Next Best Action Marketing architecture can ensure that their actions during interaction time with customers is always the best thing to do.

4.6.2 Centralised Next Best Action Decisioning For Marketing

In order to provide consistent Next Best Actions to customers, company objectives must be balanced across multiple, often conflicting, demands (Medina, Altschuler & Kosoglow 2019). The customer will have their own needs when interacting. These stated, and unstated needs must be weighed against the priorities of the company around risk, retention strategies, sales needs and service. Making decisions in isolation for each of these areas creates clashes that increase the risk to the company and poor customer experience. A marketing initiative to sell more loans may not be appropriate for a customer who is struggling financially. A centralised decisioning system that can provide balanced Next Best Action decisions in real-time and across multiple channels (Henry 2017) would ensure consistency between customer and business objectives. To provide this consistency we need to consider the three pillars of proper enterprise decision management as follows:

4.6.2.1 Decision Quality

All dimensions of a decision need to be taken with the most up-to-date data and taking into account all influencing factors. A marketing decision to cross-sell needs a favourable view of the risk posed by the sale. Waiting for batch runs to refresh customer profiles may result in lost sales opportunities or increase business risk. A customer's dialogue with an agent during an interview may significantly change the best strategy to adopt with a customer. Decision quality requires decisions to be made in real-time, with all decision dimensions taken into account. The logic driving decisions may include business rules, policies and strategies to express business priority and robust and accurate models to predict customer behaviour, e.g. propensity to accept, churn, default (Taylor 2012).

4.6.2.2 Decision Quantity

Customers are used to interacting across multiple channels, often for the same transaction. Decisions need to be consistent and available wherever the customer chooses to be. As the customer interacts, new decision evaluations need to reflect their new status. The requirement can result in significant decision volumes that are consistent and ubiquitous (Tzuo 2018).

4.6.3 Command And Control – Business Agility

In competitive markets with continually evolving regulation and compliance needs, command and control over decisions are imperative. Businesses need to be able to see what decisions are being made, predict how strategy changes will affect their business and to be able to deploy new decision logic rapidly and without disruption. Changes in the parameters being used in decisions and the strategies employed cannot wait for the periods demanded by IT control. The business requires agility and flexibility concerning strategy change and in control of its own decisions (Mundra et al. 2019).

4.6.4 What Is Driving Next Best Action Decision Management?

Next Best Action Marketing decision system consists of logic that simultaneously improve business performance and customer satisfaction by enabling enterprises to consider and act upon both customer and business needs during every customer interaction. This logic can be monitored in real-time, and the way the decisions are made can be audited if required.

With the Next Best Action Marketing decision system, business users can plan, design, deploy and manage and control the Next Best Action logic. The decision logic consists of rules and policies and predictive models assembled into business managed process (King 2015).

The resulting decision logic can be centrally deployed in the run-time environment for batch decisioning, real-time decisioning or both. The decision logic can then be tested, deployed, and monitored, across all channels. Once the decision logic is deployed, it also combines with adaptive models that learn from real-time customer behaviour and modify the underlying process flow on-the-fly. Customer responses to the next best action recommendations and answers to dynamic questions are captured, immediately fed back into the next decision. Decisions and their outputs are centrally stored for measurement, reporting and what-if simulation (Thomopoulos 2013).

With this powerful centrally deployed decision hub for enterprise-wide decisioning, companies can automatically ensure that every customer interaction is unique, appropriate, and consistent at all times and that Next Best Action decisions can occur dynamically. Furthermore, the dashboards tracking customer actions are "live," i.e. the business can see what the customer is doing in real-time.

The solution allows business executives in marketing, sales, risk and customer contact channels to take direct control of interaction strategies, managing their success in real-time and measuring that success in terms of impact on the bottom line.

Leveraging all of these capabilities, the Next Best Action Marketing decision system allows companies to optimise customer relationships by implementing a personalised strategy for each customer that treats them as an individual while furthering corporate objectives. Leveraging technologies like Hadoop[16] Next Best Action Marketing is highly scalable as it enables this functionality across thousands of users and hundreds of thousands of simultaneous interactions (*Apache Hadoop* 2019).

The benefits of centralised Next Best Action Marketing include:

- All (customer) decisions are delivered through a single decision system across channels
- Per definition, consistent across channels
- Optimised interaction time
- 'Reprogramming' the decision system reprograms the company's customer experience (with regards to branding, service, sales, risk)
- Calculate a 'mini-business case' per interaction and recommend the action that will maximise the bottom line
- Comprehensive real-time monitoring
- Total control by the business

4.6.5 Examples Marketing Applications

One of the best uses for Next Best Action Marketing is in inbound marketing. Traditional outbound marketing suffers from abysmal response rates. For every satisfied and fulfilled customer, a hundred may exist who found the marketing communication untimely, inappropriate and irrelevant to their needs.

By initiating sales discussions during inbound contact from customers, Next Best Action Marketing decision system allows only relevant conversations to be had and at a time that is convenient for the customer. These conversations can be re-enforced across channels with appropriate online material backing up call centre discussion and, depending on responses, and resultant outbound communication can be made knowing that

[16] https://en.wikipedia.org/wiki/Apache_Hadoop accessed on 10 August 2019

it fits with customer's needs. Marketing driven in this manner can result in a significantly higher response rate and customer satisfaction levels.

Making a sale to a customer may not always be the most appropriate action. Customers may show a propensity to leave, and appropriate retention strategies must be employed. Decision system for NBA marketing allows the needs for retention strategies to be arbitrated against sales strategies and employed when most required. Decisioning can evaluate retention budgets based on future customer profitability and construct bundles of offers to create a package of incentives for the customer to remain. Agents can negotiate with these bundles knowing that they are relevant to the customer and appropriate for the business

4.6.6 Examples Of Risk Applications

Next Best Action Marketing decision system can uniquely combine marketing decisions with real-time risk decisions across all areas of the business. This could include risks such as fraud, operational risk, claims risk & leakage, but a crucial particular risk area is credit risk (Taylor 2012).

Scores from multi-outcome risk decision logic are used that go beyond the simplistic accept/decline type of decisions and lead to risk-based pricing, negotiation tactics, decline reason resolution and down selling. Risk decisions can absorb data from multiple sources, both batch and real-time and take into account both predictive models, adaptive models and scorecards. Multiple bureau feeds can be used to create comprehensive risk views of customers as input for decisions. Specific applicant scoring predictive analytics tools are used for reject inference (outcome inferencing) which is a common problem in credit risk – inferring behaviour for customers that were rejected or declined themselves as the use of random controls for application scoring is prohibitively costly and typically not compliant with internal and external regulations (Bücker, van Kampen & Krämer 2013).

For existing customers, portfolio risk management is an important application area. Risk decision logic based on behavioural scoring models and business rules is used to calculate measures for internal and external reporting, such as Expected Loss (IRB Basel II) and Economic Capital (Engelmann & Rauhmeier 2011). Besides, risk is an essential dimension in needs analysis and financial planning.

Finally, the Next Best Action Marketing decision system can also be used in the collection stage, such as collections segmentation. Referring a customer for collections is part of the customer lifecycle.

5 IMPLEMENTATION OF NEXT BEST ACTION MARKETING DECISION SYSTEMS

The proposed implementation model represents the combination of the author's "knowhow" together with refined best practice and years of field implementation experience. The approach not only reduces risk through assessment, tracking and mitigation planning but also by providing a comprehensive set of prerequisites, processes, tasks, deliverables are assured.

5.1 Project Strategy For A Decisioning Project

Whereas there are many commonalities with other projects, centralised decisioning for Next Best Action Marketing projects can be defined by several typical characteristics. For instance:

5.1.1 Business Transformation.

The purpose of a decisioning project to give business control of customer interactions by designing, deploying and controlling the customer experience, marketing and risk strategies ultimately across all channels and customer processes. So it allows the business to change the way they do business, rather than an IT reimplementation of how business is done today (even if this is the focus of a first phase). This means that the program should be approached as a business transformation project. Business sponsorship is not sufficient, the program is typically business-led, and it is not uncommon to balance the amount of IT activity with at least a similar amount of activity in the business stream of the project.

5.1.2 Enterprise next best action marketing program

A Next Best Action Marketing program of work should be scoped for and implemented within a phased approach as it can provide benefits within a relatively short timescale. The typical dimensions for scoping are the issues to be addressed (sales versus retention versus risk), the channels to be connected (inbound and outbound, batch and real-time, contact centre, web, internet etc) and the major customer segments to be served (different strategies for small business versus private versus consumer banking; pay monthly versus prepaid Telco customers). Roll out and program planning is driven by business opportunities and technical and operational constraints, which can change frequently, so short phases with clear benefits are critical.

5.1.3 Focus on Strategic Next Best Action Architecture.

Beyond the project or a project phase, Next Best Action Marketing will primarily be driven by the business as usual (BAU) business configuration, within the constraints of that phase (channels customer segments, issues). This means that from an IT perspective, the focus and scope need to be on creating or extending a generic strategic decisioning architecture. For a specific phrase 'just enough' business requirements need to be gathered (and not more) to drive the IT decisions on what data items, data sources and channels need to be hooked up to decisioning for this particular phase. No further assumptions should be made on what customer strategies are going to be built by the business as these will be ever-changing as part of business as usual. For example, a first phase could be cross-selling propositions in the contact centre and web or loan applications through brokers and branches. The IT part of the project should not be concerned with the actual requirements for the decision logic (what products, rules, models, strategies) as the technical environment should be robust to change – it should focus on making internal and external data sources available and integrating with contact centre, web, brokers and branches. The business stream will likely create a first version of the logic with enough scope (but rigorous logic testing) for go live and then change it on an ongoing basis, without impact on the technical data architecture. IT can then focus on extending the architecture for the next phase. Future-proofing, the design for multichannel centralised decisioning, is essential as such, which could be guaranteed by a central program team.

5.1.4 How is a Next Best Action Marketing Project Similar to Other Projects?

A Next Best Action Marketing project is similar to other IT projects in that it uses the development approach of Scope, Design, Build and Test, and Deployment. It also requires the project infrastructure, management and control that would be required for any project.

5.1.4.1 A Typical Decisioning Project

A typical Next Best Action Marketing project will have the following phases:

1. Inception Phase
2. Scoping
3. Design
4. Build and Test
5. Deployment

The business stream of the project will have a more iterative, continuous nature than the IT stream starting from the scoping phase and will carry on beyond project launch into BAU.

5.1.4.2 Inception Phase

The Inception phase, sometimes labelled a scoping study, is primarily focused on the investigation and understanding of the domain of interest to a sufficient level such that well educated and defensible decisions can be made around business priorities, estimation, scope, increment definition and planning. A typical Inception phase usually takes between 4-6 weeks elapsed duration dependent upon the scope of the domain of interest for the client, and the accuracy level of estimates required. The resultant deliverables of the Inception phase, in addition to providing input for planning and estimating, also form the basis for ongoing development activities in the initiation phase and beyond. It is therefore crucial that deliverables are clear, complete and fit-for-purpose as they form the foundation for ongoing development activities.

5.1.4.3 Scoping

Defining Business Requirements for Next Best Action Marketing is less critical than for a typical technology implementation project. The decision system solution allows business users to change the way they do business and therefore change the business requirements. To determine the scope of the project (channels, business issues, groups and propositions) and business priorities the business needs to indicate how many channels are going to be connected to the decision system, how many business issues, how many propositions.

5.1.4.4 Design

The Real-Time Decision Services can be set-up as a general decision service for all channels, processes and systems making decisions about customers (Brezillon, Burstein & Zaslavsky 2011). It is therefore essential but not a huge task to determine the overall architecture. Solution Architecture show is focused on showing how decisioning will be integrated into the corporate architecture.

5.1.4.5 BUILD AND TEST

Points that need to be determined for testing are:

- Confirm what is the typical duration of a user acceptance test at?
- Confirm a specific test approach?
- What phases are expected

5.2 DELIVERY WORK PACKAGES (HIGH-LEVEL PLAN)

Below is a high-level plan showing the work packages and the phases. The section below goes on to describe each work package.

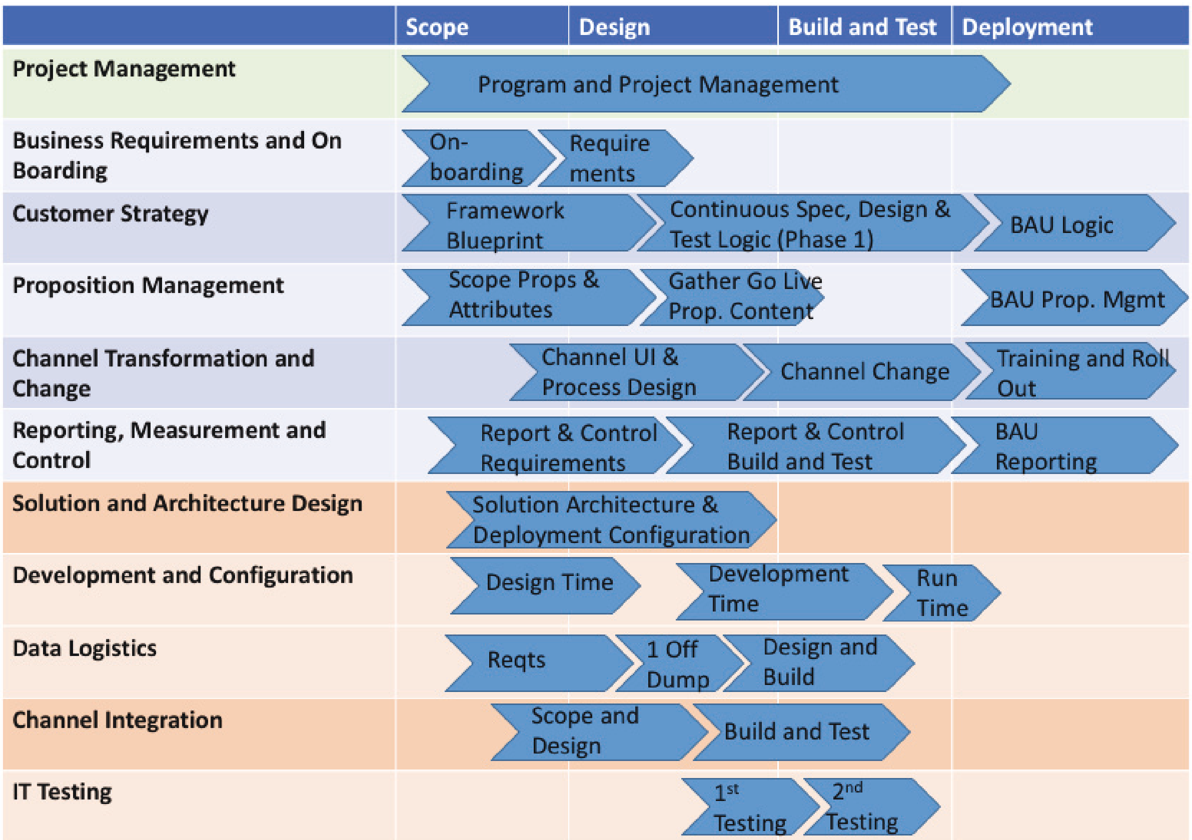

FIGURE 14 - NBA PROJECT DELIVERY ROADMAP

5.2.1 Business Work Packages

5.2.1.1 Business Requirements and on-boarding

These tasks define the high-level business requirements for the generic capability of the system. They include high-level goals, scope and functional architecture; they are not detailed requirements as the logic development will further develop this. The delivery phases will be defined at a high level, and more specific requirements for the first phase will be documented.

Example: Phase 1 could be Credit Application Scoring; Phase 2 could be Marketing Web prompts.

On-boarding ensures that all the areas within the organisations that are affected by Next Best Action Marketing projects understand the projects and the impact and business transformation processes.

5.2.2 Customer Strategy

This work package defines the Strategy, Blueprint framework, Logic development, Propositions management and Processes for the project. Next Best Action Marketing projects are interactive in the sense that does not require a full set of very detailed requirements from the start as the logic build process is an iterative process that provides more functionality per iteration.

Strategy - The business team defines the requirements and provides specifications in a template format.

Blueprint Framework – Define a product blueprint and define how it works. Typically, a blueprint template for some of the products would be defined and prototyped until the template was finalised.

Logic Development – Builds the logic to support the blueprint and the project. Again, this is an iterative process with iterative simulation and testing.

Proposition Management – Defining and collecting the content associated with propositions (scripts, visuals) including legal (and other areas) approval if required.

Process Design – Definition of processes within the organisation for managing requirements, developing new strategies and driving real-time decisioning. It also defines how the end to end processes and areas are impacted (Taylor 2012).

Predictive Analytics - Decision logic should include predictive models to predict probability to take up a particular offer, probability of default or expected loss, probability of churn. The amount of time spent on building models will differ based on the type of model required.

5.2.3 Channel And Process Transformation And Change

Channel Transformation and Change – This work package is concerned with the adoption of the Next Best Action Marketing solution in the channels in scope for a particular phase – for instance, sales through service for contact centre and retail. Especially in case of agent operated channels as opposed to customer operated channels, special attention needs to be paid to engaging the channels and driving adoption and cultural change.

This includes involving the channels in developing the solution, creating a to-be process where decisioning fits in, incentive management, internal communication, agent or broker training, roll-out planning, functional user interface (UI) requirements and any change management tasks relevant for the specific channels or applications that are in scope.

5.2.4 Reporting, Measurement And Control

Decisioning should not take place blindfolded, so transparent reporting is vital for business control and compliance. Typically, capability is added to steer the strategy through control parameters (i.e. controls outside the decision logic) and simulate changes.

Reporting Prototyping and Development. Proposition Monitoring: Reporting requirements are defined, and closed-loop reports will be implemented. This can include batch reports as well as real-time dashboards by using an off the shelf package such as Tableau[17].

Control and Simulation. The customer strategy/decision logic can contain control parameters to control the relative importance of issues (sales, retention, risk), propositions, contextual factors (contact centre queues). Business must be made aware of what controls exist and with what purpose. Control parameters can be set up in the logic and the values controlled through the model, such as the one presented earlier in how to calculate the Next Best Action Score. Results can be visualised in Tableau.

5.3 Technical Work Packages

5.3.1 Architecture And Solution Design

Specifies the Solution Architecture Design and provides an impact analysis of the solution. This package also plans for the deployment configurations, e.g. Development, Production, Test, Training, Performance Test, User Acceptance Testing (UAT), Staging, and Pilot.

5.3.2 Development and Configuration

Depending on what application is chosen for the Marketing Automation Platform, for example, Marketo or Salesforce, there will be some development and configuration required to meet the business requirements.

5.3.3 Data Logistics

Next Best Action Marketing System is data-driven, and therefore the data logistics are critical. Data can be provided through a data warehouse environment, and in some cases, it is recommended to extend the existing data warehouse environment. Data can also be provided by the calling/hosting application and or real-time data adapters can be created to obtain data from operational systems. Typically a decisioning project will involve the combination of multiple data sources, internal or external. The data used by Decisioning is preferably made available in a strongly de-normalised format.

[17] https://en.wikipedia.org/wiki/Tableau_Software ,accessed on 20 July 2019

Therefore, the recommendation is to create a data mart with so-called spine tables (Ponniah 2001). A spine table holds one row per customer with more or less static information. Spine tables typically have between 250 and 800 columns of information for each customer, and these spine tables contain in most cases between 80 and 90% of the necessary information and sometimes even more. The spine table with the most recent information needs to be made available to Decision Management environment to execute Decision Logic in real-time or batch. For predictive modelling, it is also recommended having historical spine tables.

5.3.4 Channel Integration

For each channel to be integrated, the channel has to be designed, built with integration and tested.

5.3.5 Testing

IT testing is constrained to the technical components in the Next Best Action Marketing Architecture, i.e. the data sources, installed software, channel integration. Business is in control of and owns the decision logic, so logic testing is a business responsibility within the customer strategy work package(Hambling & van Goathem 2013).

The required phases of testing are similar to any project:

Unit Testing (performed under Customer Strategy – Logic Build) Technical components, data, application programming interface (APIs), UI.

System Integration testing – to prove that the system works from end to end. The "work" flows from one end to the other.

UAT Functional Testing – to ensure that all the business functions work with the data

Performance testing – to ensure that the system performs to stated requirements, e.g. the ability to make one decision per second

6 CONCLUSION AND FURTHER RESEARCH

6.1 CONCLUSIONS

A conceptual framework for NBA marketing provided the reasons why NBA marketing is a paradigm shift which requires organisations to reorganise their resources around their customers. As it has been shown every enterprise has a unique source of competitive advantage and that is the intelligence around their customers. As opposed to products or services which can be easily reversed engineered the unique intelligence about one's customers is not replicable. On this basis, NBA Marketing is defined as the process of understanding, predicting and fulfilling the needs and wants of individuals and organisation using historical data and real-time contextual transaction data. Therefore, ensuring that the customer is offered the right value proposition at the right time in the right channel. An NBA framework is an optimisation process whereby various inputs such as propensity, CLV, NPS and context are used to present the customer with the right proposition in the right channel at the right time. For this to occur organisations need robust business architectures that can capitalise efficiently on their data assets to optimise decision making. In fact, NBA is very effective way to align organisational strategies with more tactical marketing activities. So, NBA is not only a mechanism for decision making but also a way to continuously monitor and measure the outcomes from different customer interactions. To this end, NBA should combine good processes with the right decisions. For example, finance can provide the payment history and credit score and sales data can be used to leverage purchase history while marketing data can provide the historical responses to different offers. Engaging customers in the right place at the right time is vital. Therefore, a decision hub can leverage channel information both outbound and inbound such as email, direct call, point to URL, POS, direct mail and mobile messages to work out the best channel for each customer. Also, the technical architecture of NBA decision hub requires the ability of the system to forecast and simulate trends, run predictive models, leverage big data technologies such as Kafka to extract real-time data and use adaptive learning models. A decision-ready data inputs are needed to feed the decision logic in real-time. The decision logic consists of both deterministic business rules as well as predictive models therefore, the decision-ready data need to be denormalised in the form of one row per customer. A well-functioning system should be resilient and have the ability to self-organise. To this end it was found that business issues and groups provide the ideal hierarchy from which to execute the NBA propositions.

One of the key requirements for NBA marketing is understand customer intent and executing on it in an intelligent manner. To this end the analytics of NBA focused on Naïve Bayes classifier as the key model, this is despite its shortcomings NB provides a key advantage namely ability to classify real-time data as per NBA use case. Specifically, this paper focused on the data analysis i.e. the inputs into the NB model. Grouping of data was proposed as robust way to deal with variety of data. To this end after we compared and contrasted three data grouping methods: Platt scaling, Binning and PAV. It was shown in the experimental evaluations that Binning is a robust technique in dealing with both numeric and non-numeric data. The creating of the binned view is to help the adaptive model based on NB classifier to discriminate between positive and negative cases. It was shown that splitting the bins using equi-weight ensuring an equal number of weighted cases per bin is preferable in marketing where a predictor like "wealth" can skew the results. Also, it has been shown that grouping predictors into bins can be done using the bootstrap method which provides an intuitive result to the Z-score used in deciding which bins to merge. It has also been shown that to measure the quality of the models Coefficient of Concordance (CoC) was the preferred method to measure discrimination of positive cases from negative cases. CoC is convenient because of high granularity and because it is cut-off independent there is no minimum or maximum number of cases required. A unique model to calculate the final score was presented.

The model is using ten distinct inputs:

- Propensity (X_1),
- Commercial Value (X_2)
- Non-Commercial Value (X_3)
- Impact on Customer Lifetime Value (X_4)
- NPS Specific to Customer (X_5)
- Marketing Lever (X_6)
- Channel Lever (X_7)
- Sum Tuner (X_8)
- Multiplier Tuner (X_9)

- Web Page Context (X_{10})

These inputs are modelled into a process for calculating the final score per customer per online advertising space. The model is comprehensive and has a number of practical inputs such as Marketing and Channel lever for cases where there are two different teams dealing with the channel and marketing communication. As it has been shown earlier in the paper that this is not an ideal situation however the model offers flexibility in this area.

Businesses around the world understand that the customer-centric model is the way business is conducted in 21st century. Therefore, they need to ensure that their interactions with their customers are always relevant, timely, cost-efficient and compliant. Enterprises with millions of customers, across multiple brands and channels, and complex products, need to ensure that every interaction with their customer is the best it can be. With increasing regulation, intense competition and a need to market well while continuously balancing risk, the quality of the millions of decisions businesses make when interacting with their customers will determine the success or failure.

The benefits of centralised Next Best Action Marketing are:

- All customer decisions are delivered through a single decision system across all channels
- Consistent across all channels
- Optimised interaction time
- Reconfiguring the decision system, reconfigures the company's customer experience with regards to branding service, sales, and risk
- Calculate a "mini-business case" per interaction and recommend the action that will maximise the bottom line
- Comprehensive real-time monitoring
- Total control by the business

A typical Next Best Action project comprises the following stages:

- Inception Phase is focused on the investigation and understanding of the domain of interest to a sufficient level such that well-educated and defensible decisions can be made around business priorities, estimation, scope, increment definition and planning.
- Scoping determines which channels, how many business issues and propositions will be connected to the decision system.
- Design is the task to determine the overall solution architecture which is showing how the decision system will be integrated into the corporate architecture.
- Build and test which looks specifically at acceptance criteria, system integration testing, user acceptance testing, business verification testing and warranty periods.

From a delivery perspective an NBA project delivers:

- Business work packages comprising business requirements, customer strategy, channel and process transformation and change management, also reporting measurement and control.
- Technical work packages comprising architecture and solution design, development and configuration, data logistics, channel integration and testing.

6.2 FURTHER RESEARCH:

There are some critical elements to Next Best Action Marketing.

- Such as ensuring that the organisational architecture aligns to business objectives. Further research should be done in the area of business architecture. For example, many organisations confuse business architecture with business requirements. Business architecture is concerned with the strategic direction of the business, while business requirements are used to gather requirements to build a specific application. Enterprises in the 21st century require forward-looking attitude there is a need to continuously work on the next iteration of the product or service and more research is required into how predictive analytics can help achieve that. It is easy for a business to ask for an algorithm to predict the stock market or replicate google or Nextflix, but the challenge is how to re-invent the business using predictive analytics? What areas of the business are currently opaque and could use more transparency to dispel some of the "witch-hunting" going on?
- The scope of this paper is to show how decision management works in marketing in the form of Next Best Action; however, future research can look into the centralised decision system in the area of operations. What type of predictive models are used in the area of operations? Is there more demand for real-time, less precise models or more in-depth, particular predictive model to solve one business problem.
- Another area of automated decision making is credit risk and insurance premiums; these are sources of competitive advantage; they have a direct impact on the bottom line. The most common practice is to test two processes side by side with, e.g. 90 per cent of the time run the established process while 10 per cent of the time run the new test process.

The model presented used to calculate the final score for customer interaction is a starting model.

- What other optimisation models are there and what is the basis for calculating a Next Best Action?

- Research can focus on long term study where a control group should be kept aside Gauge improvements into CLV, and NPS or other measures of customer satisfaction. Also, calculating the score differently can be compared and contrasted with the one proposed in this paper.
- On the analytical side, more research can be done on the binning algorithms? What are the benefits of binning in data analysis? What are other models possible and even better than the ones presented in this paper?
- Also, how to build and deploy adaptive algorithm into processes more research is required into process flows leveraging probabilistic decisions as opposed to the deterministic. There is still much doubt in the business community about probabilistic decision models in the process flow.
- On the data side, this research has shown that denormalisation of data is a crucial requirement for a real-time decision. What other structures are possible, and what is the best way to structure data for decision systems?
- Finally form the project management discipline more research needs to be done in the motivation and management of multiskilled teams like solution architects, integration architects, business analysts and data scientists and how to keep the projects sponsors up to date with sufficient level of detail. There seems to two distinct views of the world the IT with very definite true/false outcomes and the business and data science world with less well-defined outcomes but through successive iterations honing in on a workable solution.

As a parting note, it is worth suggesting the path of self-experimentation which nowadays is so easy to implement. Consider your person there lots of wearable devices that can track anything sleep, activity, e.g. Oura ring provides rich data on sleep and activity this can be easily downloaded, and predictive models can be built like sleep as a function of several inputs; food, water, coffee and other inputs. This can be combined with a bicycle computer like Wahoo to track fitness as a function of inputs. There are a whole plethora of experiments awaiting the curious and inquisitive.

REFERENCE

[1] *Adaptive BPM | Pega.* 2019, viewed 12 May 2019, <https://www.pega.com/insights/resources/adaptive-bpm>.

[2] *Apache Hadoop* 2019, wikipedia, viewed 30 June 2019, <https://en.wikipedia.org/wiki/Apache_Hadoop>.

[3] Academy of Marketing Science. Annual Conference (1991: Fort Lauderdale Florida), King, R.L. & Ohio Library and Information Network. 2015, *Proceedings of the 1991 Academy of Marketing Science (AMS) Annual Conference*.

[4] Bertone, P., Kluger, Y., Lan, N., Zheng, D., Christendat, D., Yee, A., Edwards, A.M., Arrowsmith, C.H., Montelione, G.T. & Gerstein, M. 2001, 'SPINE: an integrated tracking database and data mining approach for identifying feasible targets in highthroughput structural proteomics', *Nucleic Acids Res*, vol. 29, no. 13, pp. 2884-98.

[5] Bücker, M., van Kampen, M. & Krämer, W. 2013, 'Reject inference in consumer credit scoring with nonignorable missing data', *Journal of Banking & Finance*, vol. 37, no. 3, pp. 1040-5.

[6] Brezillon, P., Burstein, F. & Zaslavsky, A. 2011, *Supporting real time decision-making : the role of context in decision support on the move*.

[7] Crozier, R. 2018, 'CBA system suggests 20m customer conversations a day', *itnews*, 2018/06/15.

[8] *Chordiant Next-Best-Action Marketing White Paper* 2016, docshare.tips, viewed 5 December 2018,<http://docshare.tips/chordiant-next-best-action-marketingwhitepaper_586133b5b6d87fcd1d8b71f7.html>.

[9] Copulsky, J., Richardson, S. & Simone, M. 2017, 'Marketing technologies, customer data and analytics: Enabling responsive customer journeys and scalable marketing processes', *Applied Marketing Analytics*, vol. 3, no. 2.

[10] Diebold, F.X. & Mariano, R.S. 2002, 'Comparing predictive accuracy', *Journal of Business & Economic Statistics*, vol. 20, no. 1, pp. 134-44.

[11] Dietrich, B. 2014, *Analytics across the enterprise how IBM realizes business value from big data and analytics*, IBM Press,, <https://virtual.anu.edu.au/login/?url=https://www.safaribooksonline.com/library/view/title/9780133835922/?ar>.

[12] Davenport, T.H. 2013, Enterprise analytics: optimize performance, process, and decisions through big data, FT Press, Upper Saddle River, N.J.

[13] Darmont & S. Rizzi, Springer International Publishing, Cham, pp. 313-23.

[14] Engelmann, B., Rauhmeier, R. & SpringerLink 2011, *The Basel II Risk Parameters: Estimation,*Validation, Stress Testing - with Applications to Loan Risk Management.

[15] Fawcett, T. & Niculescu-Mizil, A. 2007, 'PAV and the ROC convex hull', *Machine Learning*, vol. 68, no. 1, pp. 97-106.

[16]487Gončarovs, P. & Grabis, J. 2017, 'Using Data Analytics for Continuous Improvement of CRM Processes: Case of Financial Institution', *New Trends in Databases and Information Systems*, eds M. Kirikova, K. Nørvåg, G.A. Papadopoulos, J. Gamper, R. Wrembel, J.

[17] Henry, P. 2017, 'Success in OMNICHANNEL Marketing', *Printing Impressions*, vol. 60, no. 5, pp. 20-4.

[18] Hu, Z., Bodyanskiy, Y.V. & Tyshchenko, O.K. 2019, Self-Learning and Adaptive Algorithms for Business Applications: A Guide to Adaptive Neuro-Fuzzy Systems for Fuzzy Clustering under Uncertainty Conditions, Emerald Publishing Limited, Bingley, UNITED KINGDOM.

[19] Hambling, B. & van Goathem, P. 2013, *User Acceptance Testing A Step-By-step Guide*, BCS, The Chartered Institute for IT Imprint BCS Learning & Development LimitedTurpin

Distribution Services Limited Distributor,

<https://virtual.anu.edu.au/login/?url=https://www.safaribooksonline.com/library/view/title/9781780171678/?ar>.

[20] Jenkinson, A. 2009, 'What happened to strategic segmentation?', *Journal of Direct, Data and Digital Marketing Practice*, vol. 11, no. 2, pp. 124-39.

[21] Jacopo Tagliabue, Lucas Lacasa, Ciro Greco, Mattia Pavoni & Polonioli, A. 2019, 'Predicting e-commerce customer conversion from minimal temporal patterns on symbolized clickstream trajectories', *ArXiv*, vol. abs/1907.02797.

[22] Jekel, J.F., Elmore, J.G. & Katz, D.L. 1996, *Epidemiology, biostatistics, and preventive medicine*, W.B. Saunders, Philadelphia.

[23] Kin-nam, L., Wong, S., Ma, M. & Liu, C. 2003, ''Next product to offer' for bank marketers', *Journal of Database Marketing*, vol. 10, no. 4, p. 353.

[24] Kotler, P. 1998, *Marketing*, 4th edn, Prentice Hall, Sydney.

[25] King, R.L. 2015, *Proceedings of the 1991 Academy of Marketing Science (AMS) Annual Conference*.

[26] Krzanowski, W.J. & Hand, D.J. 2009, *ROC curves for continuous data*, CRC Press, Boca Raton.

[27] Larose, D.T. & Larose, C.D. 2015, *Data mining and predictive analytics*, Second edition.

edn, John Wiley & Sons Inc.,,

<https://virtual.anu.edu.au/login/?url=https://www.safaribooksonline.com/library/view/title/9781118868706/?ar>.

Li, X., Xu, G., Chen, E. & Li, L. 2015, 'MARS: A multi-aspect Recommender system for PointofInterest', *2015 IEEE 31st International Conference on Data Engineering*, pp. 14369.

[28] Legendre, P. 2005, 'Species associations: the Kendall coefficient of concordance revisited', *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 10, no. 2, p. 226.

[29] *Leveraging Predictive Analytics to engage your customer* 2014, The Digital Insurer, viewed 25 March 2019, <https://www.the-digital-insurer.com/wp- content/uploads/2014/10/362-2014-5-Steps-to-Optimizing-Customer-Value-inInsurance-July14.pdf>.

[30] Mauboussin, M.J. 2006, More than you know : finding financial wisdom in unconventional places, Columbia University Press, New York.

[31] McCormack, A. 2007, *The IT value stack : a boardroom guide to IT leadership*, John Wiley, Chichester, England ; Hoboken, NJ.

[32] Medina, M., Altschuler, M. & Kosoglow, M. 2019, Sales engagement : how the world's fastest growing companies are modernizing sales through humanization at scale.

[33] Muler, A.C., Guido, S. & Ohio Library and Information Network 2016, Introduction to machine learning with Python : a guide for data scientists.

[34] Mundra, S., Rising, L., Guo, X. & Ohio Library and Information Network 2019, *Enterprise agility : being agile in a changing world*.

[35] Naumenko, L. 2015, 'The subordinate concept ‘marketing’ in the contemporary business English discourse', *ScienceRise; Том 5, № 1(10) (2015)*.

[37 Newson, R. 2006, 'Confidence intervals for rank statistics: Somers' D and extensions', vol. 06, pp. 309-34, viewed 2006,

<http://ageconsearch.umn.edu/record/117583/files/sjart_snp15_6.pdf>.

[38] *Naive Bayes classifier* 2019, Wikipedia, viewed 30 May 2019, <https://en.wikipedia.org/wiki/Naive_Bayes_classifier>.

[39] 'Next-best-action marketing: a customer-centric approach' 2012, *CRM Magazine*, vol. 16, no. 5, 2012/05/, p. 42.

[40] *Pega Marketing Product Overview 7.31* 2018, Pegasystems, viewed 20 October 2018, <https://www1.pega.com/system/files/resources/201806/Pega_Marketing_Product__Overview_V7_31.pdf>.

[41] Pāvels, G. 2017, 'Data Analytics in CRM Processes: A Literature Review', *Information Technology and Management Science.*, vol. 20, no. 1, pp. 103-8

[42] Pophal, L.I.N. 2015, 'Multichannel vs. Omnichannel Marketing: Is There a Difference, and What Does It Mean to You?', *EContent*, vol. 38, no. 2, pp. 15-20.

[43]Ponniah, P. 2001, *Data warehousing fundamentals a comprehensive guide for IT professionals*, Wiley,,

https://virtual.anu.edu.au/login/?url=https://www.safaribooksonline.com/library/view/title/9780471412540/?ar>.

[44] Racine, J.S. 2019, An Introduction to the Advanced Theory and Practice of Nonparametric Econometrics: A Replicable Approach Using R, Cambridge University Press, Cambridge.

[45] Reichheld, F. 2006, 'The Microeconomics of Customer Relationships', *MIT Sloan Management Review*, vol. 47, no. 2, p. 73.

[46] Sanjiv, K.R. 2012, *Next Best Action: The One-To-One Future*, viewed 25/09/2018 2018, <www.wipro.com>.

[47] Shalev-Shwartz, S. & Ben-David, S. 2014, Understanding machine learning : from theory to algorithms.

[48] Stone, B. & Jacobs, R. 2008, Successful direct marketing methods : interactive, database, and customer-based marketing for digital age, 8th edn, McGraw-Hill, New York.

[49] Taylor, J. 2012, Decision management systems : a practical guide to using business rules and predictive analytics.

[50] Thomopoulos, N.T. & SpringerLink 2013, Essentials of Monte Carlo Simulation : Statistical Methods for Building Simulation Models.

[51] Todd, G. 2008, 'Using Predictive Analytics to Determine the Best Next Offer', *DM Review*, vol. 18, no. 10, pp. 35-.

[52] Tzuo, T. 2018, Subscribed : why the subscription model will be your company's future - and what to do about it, Portfolio/Penguin, New York, New York.

[53] Valos, M.J. 2010, 'Structure, people and process challenges of multichannel marketing: insights from marketers', *Strategic Direction*, vol. 26, no. 3.

[54] Vaid, S. 2017, 'The Four Pillars of Customer Centric Marketing', *Business World*, 2017 Aug 08.

[55] Wang, M. & Ohio Library and Information Network., Knowledge science, engineering and management : 6th international conference, KSEM 2013, Dalian, China, August 1012, 2013 : proceedings.

[56] Wright, D. & Meadows, D.H. 2012, *Thinking in Systems A Primer*, Taylor and Francis, <http://UWA.eblib.com.au/patron/FullRecord.aspx?p=430143>.

[57] Whelan, J. & Meaden, G. 2012, *Business architecture a practical guide*, Gower, <https://ebookcentral.proquest.com/lib/unimelb/detail.action?docID=944776

[58] Warrens, M.J. 2018, 'On the Negative Bias of the Gini Coefficient due to Grouping', *Journal of Classification*, vol. 35, no. 3, pp. 580-6.

[59] Yitzhaki, S. & Schechtman, E. 2013, *The Gini methodology : a primer on a statistical methodology*, Springer, New York ; Heidelberg.

[60] Zhu, X. & Xu, G. 2016, 'Applying Visual Analytics on Traditional Data Mining Process: Quick Prototype, Simple Expertise Transformation, and Better Interpretation', *2016 4th International Conference on Enterprise Systems (ES)*, pp. 208-13.