

Necessity of Machine Learning and Data Visualization Principles in Marketing Investment Management

Akshay Sreekant, Senthilnathan P, Gopichand G, Manoov Rajapandy, Naresh Kannan

Abstract— Since these sudden bursts of Technological advancements have taken over and are making forays into the corporate worlds research and development, it is important for the marketing investment sectors to embrace and adapt with the changes, and thus to move forward, need to implement concepts such as machine learning to keep up with the pace. Thus in this paper, we will discuss why it is necessary to do so, and what the desired output has been on some sample applications of machine learning.

Keywords—machine learning, forecast, return on investment, spends, medium, marketing, dashboard, word cloud.

I. INTRODUCTION

Marketing forms an integral part of any product based company's strategy to boost their sales and get the word out to the consumers about their new products and brands. Now until a few years ago, the analysis of which sections of the public are showing an increased demand for a particular product and thus more spending is needed, has been largely instinct based and inaccurate. Leading to an integral part of the money being spent needlessly and budgeted incorrectly. Now in the last few years, there has been a considerable improvement with the arrival and implementation of Data visualization. Even so, the marketing industry has not used it to its optimal usage.

Thus with these new approaches in question it is important to discuss what separates one from other, and what are the individual repercussions of each of them and ultimately what is the USP of each concept.

Data Visualization helps simplifying the current scenario that the products in a company is in, compare with the fiscal results of the past, and not only on how its performing sales wise, but also on which fronts is money getting spent. Secondly Data Visualization, is in principle, what the fancy new Machine Learning people call as Exploratory Data Analysis, thus Data Visualization actually forms a very integral part of Machine Learning as well. Secondly, Data Visualization is essential for users to provide a simpler report of the findings, instead of having to explain it in long paragraphs.

Machine learning and Artificial Intelligence go hand in hand. Concepts like Natural Language Processing, Feature Engineering, etc, assist the core principles of Machine Learning and help carry the driving technology forward into the modern world.

Using Machine Learning, one can literally teach a

computer, which has a far higher processing and computational power compared to humans, the ability to learn about various data in the past and then apply this logic to help the humans in a far more accurate predictive estimates.

Also, using the unsupervised techniques, unlabelled features can be utilized in activities such as clustering and sorting and filtering of data can be done more efficiently.

Also, integrating these sections of algorithms with an existing web application code or a mobile application code, can be done seamlessly and thus can assist with a user friendly, GUI based system, which on clicks of buttons provides ground breaking insights about the products.

Also, a predictive or a forecasting application is not the end. Machine Learning and its concepts can also be used to make the work of the users easier and more convenient. Creating a smart module, that uses machine learning to understand user behaviour and then implements suggestive measures to help the user finish their job quickly is just another feather in the cap.

Finally, the whole corporate industry is moving towards an automated side. Using Natural Language Processing, chatbots and smartreply systems go a lot into increasing the quality of customer service, and largely reducing the response time. Moreover, NLP can be used in tandem with Machine Learning to devastating effect. Customer reviews can never be analysed any better, and the general sentiment of the customer can be easily comprehended using sentiment analysis technique.

Thus we can see how important it is for Marketing Investment Management companies to incorporate Machine Learning and similar progressive tactics in their system to assist in better performances, ease of use, and quicker response time.

II. RELATED WORKS

[1] Discusses the role that AI and ML is going to play in Sales and Market. It tells us how the robots will replace the sales people and how using AI and other technological advancements, there will be self updating websites and automated transactions on based on user behaviour.

[2] This paper discusses how machine learning can be used to make user custom marketing strategy, and how a real time data recording

allows the developers to manipulate the strategies instantly and not have a lag period. Secondly it touches how the data available, lowkey transcends to Big Dat, and how the modern day paradigm has a convenient mechanism to handle it.

[4] Delving specifically into one medium, as these days social media marketing is on the rise, it is imperative that the money spent on marketing via social media is done to the maximum level of optimisation. Thus, it is important, as this paper discusses the methods to undertake sentiment analysis and understand how the consumers are reacting on the different products. Thus it emphasizes on the fact how using techniques like NLP and ML, the estimate to gauge the public reaction is much more reliable and accurate.

[5] This paper talks about how the advertisements can be analysed and response be judged, as adverts on TV are an integral part of marketing our products. Thus helps in identifying appropriate target groups and time of response to be at an optimal level for efficient marketing.

III. WORKING METHODOLOGY AND RESULTS

We decided that the best way to enunciate the necessity and how data visualisation and machine learning has transformed the way marketing investments can be managed, is to actually implement these techniques over a sample data set and project the results.

Thus, we took dummy data of a company that was the spends data, which had the in depth records of how the money being spent in marketing different brands of the product in different sectors at different periods over different mediums

Thus this data could report the following values:-

- 1. Date
- 2. Brand
- 3. Medium
- 4. Location

This data also had other irrelevant features recorded like Remarks, and Activities etc, which were deemed useless for the current activity.

Similarly a Secondary Sales Data was arranged for which reported factors that could tell you, which brand sold how much in what month.

Thus with all this raw data available for the past 2 years, we carried out Cleaning and Pre processing, that gave a structured data format.

Once this cleaning is done, we are left with the refined dump of data, that was not used at all until that point for any analysis. Which means that future marketing is done without any foresight of reference from the past data.

Our first activity, was a simple Data Visualization that charted up plots that showed the numbers for how the statewise spends and sales were, and what were the mix of media for the spends. Thus we can already see how Data Visualization is proving to be a game changer in deciding the budget for your spends and also the distribution

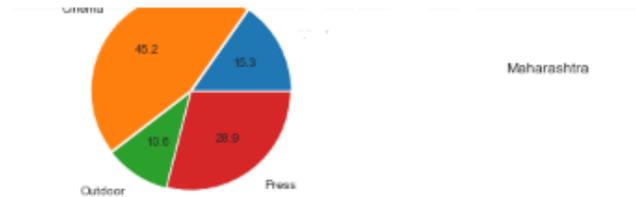


Fig 1 Shows the medium wise spends distribution for Maharashtra

Thus, interesting useful insights can be drawn from Data Analysis like this, as it gives a picturesque idea about how I am spending my money across the state, and say if the sales aren't that good, how should the money be rechanneled to meet the demands.

Also the advantage of such visualization reports is that they can be used as a briefing report that one hands to the higherups. So that they don't have to read all the numbers and stuff, and instead look at bar graphs and pie plots to get the gist.

Moreover, using Time Series forecasting methods, a timeline can be plotted for which an estimate spends or estimate sales is being recorded, and then using Holts Winter a forecast can be made which is correlated with the spends, to create trimmer, more efficient plans.

Thus, a time series forecast, helps in mapping out the timeline of how the product has performed, and will continue to perform. is plotted.

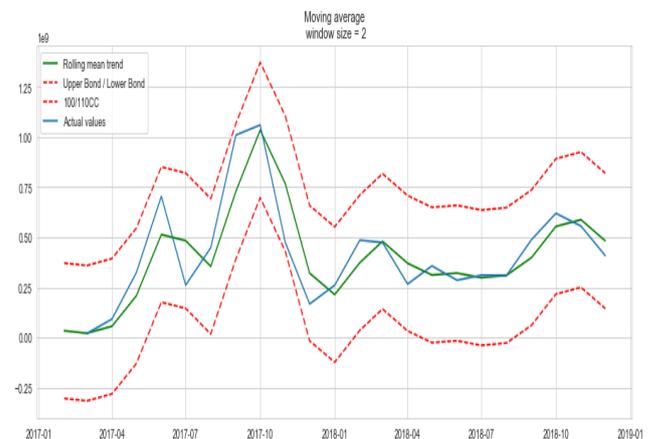


Fig 2. Actual and Predicted forecast of the spends across the time

Also anomaly detection that clearly indicated where the spends go out of a stipulated window, etc can be done with time series forecast, and these statistics are a huge indication of what was not being done.



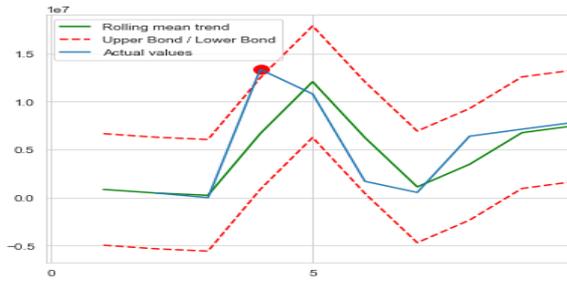


Fig 3. Anomaly detection for when the spends has exceeded the given window, or may exceed

Thus with all the visualization that is going on, it is evident that these techniques are proving a huge help in allowing the users to see and observe stuff that was not being done earlier, and so far Machine Learning has not even entered the scenario.

Another game changer is Natural Language Processing, that is an essential strategy to use when we need to get inference out of texts. It is common knowledge that when you are involved in marketing, involuntarily a lot of captions, remarks, reviews, comments etc come into play. This means, an abundance and a gold mine of data in which Natural Language Processing can play a pivotal role with.

Using functions like difflib and sequence matchers scripts that can match texts with similar spellings can be used for correlating the remarks with a specific proper noun or location they are talking about.

Thus using word clouds, customer feedback and employee feedback can be analysed in a user friendly way, and the customer service department, need not have to sift through pages and pages of reviews and forms thereby saving time, and saving the environment as well. The bigger the word is, more frequently it has been used in the reviews.



Fig 4. A wordcloud of the words used frequently in amazon alexa review

Thus, these visualizations are important in the build up to making algorithms and applications that assist in sentiment analysis.

Here we can see that when a customer is happy with the product, they use adjectives as a means of appreciations. like love, great, good, etc.

But, on the other hand, when the negative reviews are analysed, and the word cloud is created which is as follows

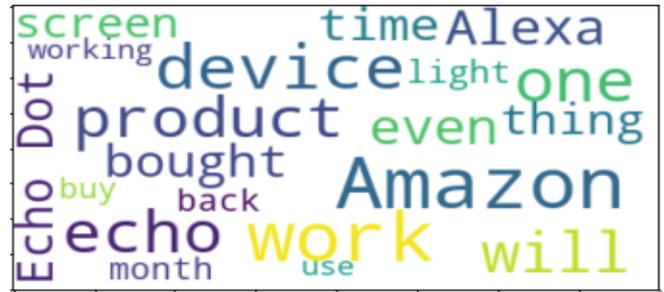


Fig 5. A word cloud of the words frequently used in the negative reviews for Alexa.

The reviews are more about the different features and aspects that went wrong like device, product, screen, light etc,

Thus using such techniques, this information can be surmised and presented to our higher ups and clients, which is obviously not possible at this level of efficiency otherwise.

Data Visualization, word clouds, graphs and plots, are all very fancy and useful for simpler reports, briefs and summaries.

Predictive Machine Learning Models, assist in creating better planned budgets, where the algorithm, learns from the way the sales and spends have been related in the past few years, and then predicts what the sales could be if the given expenditure is so and so.

Thus we considered a dummy data, where each feature is a medium of marketing, i.e. TV, Radio, Newspaper, and the predictor is the Sales Variable.

Using Linear Regression, the equation, for how each feature affects the final sales can be modelled after training and testing.

Thus in tests we got this equation for the dataset,

$$y = 2.88 + 0.0466 \times TV + 0.179 \times Radio + 0.00345 \times Newspaper$$

Also, we plotted linear regression plots that show how each feature correlates to the predictor.

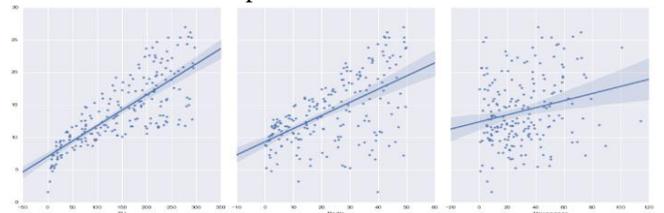


Fig 6: Linreg plot showing each medium against sales.

Individually analysing the TV Medium spends alone, using the data that is retrieved from BARC, information about the individual spots at which the ads are broadcasted is made available. Now each spot has a unique id, which can be linked back to every Purchase Order made by the company wanting to market its product at that ad spot. Thus for each spot in the add, the TRP is calculated and recorded by BARC, this value gives us an indication as to how my money spent on advertising my products are being watched across different sections and whether my money is being spent wisely.

This application is only possible with proper data cleaning, structuring, linking and merging across



multiple platforms .

Secondly, this simple piece of information from BARC, about how my ads are being impressed upon, is in reality a gold mine of analysis to unearth

The tip of the iceberg is in itself such a big revelation that wouldn't have been possible without the machine learning and data visualization techniques.

The tip of the iceberg, is that using these algorithms and techniques, coupled with the immense potential of the BARC Data. investment management companies can get a record, of which ads that I have spent money on, have unfortunately resulted in ZERO impressions across any state, over any target audience, and also which channel is the culprit for this.

impressions- A figure that shows how many people have seen that advertisement.

Thus, I can receive, what portion of my total money invested on marketing over TV, has resulted into no impact at all and thus can be deemed as money wasted. This data, can then be used into predictive modelling, to see if the powerful algorithms can unearth any pattern, or a particular link point, leading to a more cautious and probably wiser spending, the next time.

Thus, BARC Data, for an unnamed product was brought in by us, and after applying all the necessary algorithms,

this zero impressions report was generated for the 2nd quarter of 2018

APRIL			MAY			JUNE		
Channel	Spends	%	Channel	Spends	%	Channel	Spends	%
Disney International HD	306720	100	Manoranjan TV	122820.8	100	Bangla Time	38492.63	100
Housefull Movies	27115	100	News Time Bangla	21823.75	100	Bflix Movies	8032.5	100
Tollywood TV	15963	100	RT Movies	382.5	100	Housefull Action	6778.75	100
Music Fatafati	9735	52.33	Disney International HD	232880	99.49	Khushboo Bangla	12852	100
Sangeet Bhojuri	51616.25	51.64	Raj Music Kannada	50363.25	45.49	RT Movies	573.75	100
Raj Music Kannada	37156.13	47.24	Epic	22525	42.91	RUPOSHI BANGLA	22984	100
Epic	9562.5	43.27	Music Fatafati	10725	38.12	Studio One+	48195	100
PTC Punjabi	12750	41.67	BHOJPURI DHAMAKA DISHUM	36337.5	36.31	WOW CINEMA	86411	74.16
Kasthuri TV	92225	37.58	STAR Gold HD	354375	35.35	Raj Music Kannada	22589.18	49.41
STAR Gold HD	249750	35.24	Zee Cinema HD	254000	32.44	STAR Gold HD	222075	44.16
Kolkata TV	302923.5	32.23	Maiboli	40111.5	29.59	Epic	25245	43.2

Fig 7. Shows a month wise report of which channels the adverts have had no impressions at all anywhere in India

IV. CONCLUSION

Thus every procedure, conceptual topic and algorithm that have been touched upon earlier in the paper, give a clear indication as to what Machine Learning and Data Visualization is capable of . Now the discussion needs to focus on what is the need of deviating from the old methods.

What Machine Learning brings to the table is a higher quality of pre emption,where the people making the big decisions can be ahead of the bandwagon, and make informed and more efficient decisions on how to spend the marketing money wisely.Also, machine learning means the algorithm if carefully written and smartly written can make the computer into recognizing trends, patterns and seasonality in the data, that even the most esteemed of human eye cannot.

Also, Machine Learning is necessary to keep up with the rapid pace at which the industry and the other sectors are moving, and in the end, everyone's in the industry to sell something, and to get the word out, you have to market.Thus, Marketing Investment management companies would do

well to heed on the advice that machine learning and its associate principles need to be incorporated as soon as possible.

V. FUTURE SCOPE & RESULTS

What has been discussed in the paper is just a drop in the vast ocean of potential applications that this topic has. Not being constricted by the very 'RnD' nature , one can easily use such visualization concepts in tandem with practical, web applications as well.

We Have developed a web framework in flask that acts like a dashboard, where people can input their set of filters and select what kind of data visualization they want to generate, and the app will do it and present it to them.This app can be deployed for all those people on a public server, and also gives the option of a live table.

The dashboard controls snapshot includes

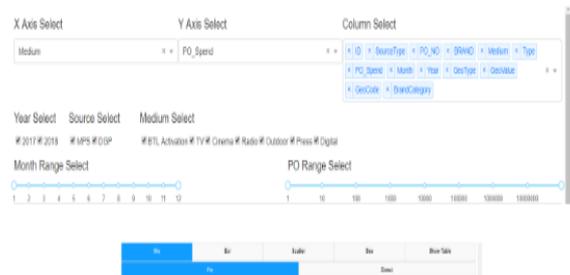


Fig 7. The basic dashboard that allows user to apply various filters on their reports and visualize it in graphs.

This is a very useful, low load visualization feature that the investment recording companies can use in their systems. It is called as a point to click software , and since all the data is real time, with retrieval from a real time database, the data reports are always performed on the latest data available.

This is something , that can thus be thought of as a future scope.

REFERENCES

1. Siau, K., and Yin Yang. "Impact of artificial intelligence, robotics, and machine learning on sales and marketing." Twelve Annual Midwest Association for Information Systems Conference (MWAIS 2017). 2017.
2. Chintagunta, Pradeep, Dominique M. Hanssens, and John R. Hauser. "Marketing Science and Big Data." (2016): 341-342.
3. Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.
4. Nguyen, Thien Hai, Kiyooki Shirai, and Julien Velcin. "Sentiment analysis on social media for stock movement prediction." Expert Systems with Applications 42.24 (2015): 9603-9611.
5. Freeman, Tom, and Mike Kennewick. "System and method for delivering targeted advertisements and/or providing natural language processing based on advertisements." U.S. Patent No. 9,406,078. 2 Aug. 2016.
6. Chatterjee, Samprit, and Ali S. Hadi. Regression analysis by example. John Wiley & Sons, 2015.
7. <https://www.ritchieng.com/machine-learning-linear-regression/>

