The Use of Machine Learning in Digital Marketing

[1] Pathan Noumankhan Sayeedkhan, [2] Dr. Sharvari C. Tamane
[1] Research Scholar, Dr.Babasaheb Ambedkar Marathwada University, Aurangabad
[2] Professor, HOD IT Department, MGM's Jawaharlal Nehru Engineering College, Aurangabad.
[1] noumanps@gmail.com, [2] hodit@jnec.ac.in

Abstract: The constant use of the internet by adults of this era is increasing, and it will increase in near future. Since majority of potential customers are online, using at least one social media channel. The data generated through social networks can provide significant insights into today's customers. The data can be analyzed to identify needs and requirements of the customers as well as to create personalized marketing content. By means of Persona creation many marketers identify their potential customers and prepare personalized content to target them. In this paper we have done a literature survey to identify potential use of machine learning technology to create marketing personas. The advantage of using machine learning is that it reduces the need for supervision by a human being if any change in processing is required. Hence it can be used to generate automated marketing persona, further this persona can also help in creating personalized marketing content which can be again automated with machine learning.

Keywords— Buyer Persona, Digital marketing, Machine learning, Social media

INTRODUCTION

The internet has become a basic need for the people of the current era, the technology is readily available, and the number of people who are online on a daily basis is increasing regularly. According to a survey conducted by "Pew Research Center"[1], the constant internet usage among adults has been increased by 5% in the last three years. The increase in the use of internet technology has also shaped the way people buy and shop for things in today's market.

"It's important to be where your audience of potential customers is today, and where they might be tomorrow." — Andrew Delaney, Senior Manager of Social Media, HubSpot[2]. Since the majority of people are active online, marketers for a band or service are advertising more and more using online channels. These channels include but are not limited to social media, search engines, e-mails, websites, web applications, and mobile applications. These efforts that use an electronic device or the internet, to connect with current and prospective customers [3] are called *Digital Marketing*.

When digital marketing is used as a brand or product marketing tool the result of a marketing campaign is measured in terms of "Return on Investment" in short ROI. The ROI is maximized when there is a tone of research done about the customer, where they (the customers) most spend time online and what triggers them to buy a product of particular brand or subscribe a particular service. This research needs to be done before launching a marketing campaign.

For research purposes, the data can be obtained from the internet as most of the people are using social media and search engines to find the products according to their needs. The data obtained from the internet which consists of a user's search patterns and decisions is called clickstream data [4].

This clickstream data has wide variety, it is generating continuously over time and therefore it is very big in volume. Due to these properties, clickstream data can be classified as Big Data. To process this big data and gather insight from it, methods such as Machine Learning can be used. The advantage of using machine learning is that it reduces the need for supervision by a human being if any change in processing is required. This is possible since machine learning algorithms learn and adapt the changes based on data provided to them [5]. So the use of machine learning technology can be done for researching the customer's data before launching a digital marketing campaign.

DIGITAL MARKETING

The terminology "*Digital Marketing*" has evolved over the course of time, that describe the marketing of products, brands, and services, using digital channels to an umbrella of terminology that encompasses the processes, such as using digital technologies to acquire customers, and build customer preferences, promote brands, retain customers and increase sales[6].

Digital marketing is all about personalizing the marketing content and marketing efforts for potential customers. To prepare for personalized content and marketing content a marketer needs to identify the buyer's journey [7] stage. A buyer's journey or consumer journey

[8] is a five-stage process from which a potential customer goes through to become aware of available products according to his need, compare and review products of different brands then the customer finally decides to buy a product or service and lastly based on the customer's evaluation of the product or service the brand or service provider engage with the customer to retain him or to invoke repurchase of different product or services..

Based on the buyer's journey stage [9] the Buyer Persona is created which is again mapped on to the customer journey life cycle.

BUYER PERSONA

A persona is generated to tell a story of the product, brand or services [10, 11]. Alan Cooper in his book, "The Inmates Are Running the Asylum" [12] has presented the idea for creating user personas for the software industry. He was the first to create a persona for the software industry. Later in his book he presented the idea of buyer persona creation for sales and marketing.

The marketing personas are created such that it should be familiar and easily recognizable, that connects to the emotional bond and personality of a customer group for which a persona is being created [10]. In digital marketing the buyer personas play a very vital role. Based on a buyer persona the customers are segmented into different groups and mapped onto the customer journey phase.

We want to know whether a buyer's persona can be created with the help of data present online. Since most people are constantly active online on social media channels which seem to be a prominent place to start, to better understand the people.

MACHINE LEARNING

In this section we look for different methods that can be used to create a buyer persona with the help of machine learning.

To create a buyer persona we want to know the customer's needs and what influences our customers to make a decision of buying a product or service. To understand the customers present online, we first need to understand the personalities of our online customers, and then identify their personality traits and behaviour in an online environment.

Based on the literature survey done by us we found that machine learning is being used exhaustively for the analysis of social media data. To better understand the different aspects of analysis of social media data for which machine learning has been employed we will

categories the machine learning approaches as given below.

i.   Analysing social media data using ML
ii.   Personality traits recognition using ML
iii.   Identifying the behaviour of user on Social networks using ML.
iv.   Persona creation through ML

The remaining part of this paper will discuss the research work done in the above categories from analysing social media data.

A. Analysing social media data using ML:

During the literature survey, we found that machine learning can be used on social media data to classify, predict, learn and forecast certain aspects of the customers. Machine learning can help analysing different aspects of social media data for instance sentiment analysis can be done on social media data that suggest the general mood of the public active in online circles. Such sentiment analysis can be done for a brand or service. F. M. Takbir Hossain [13] has classified the text reviews either positive or negative for restaurant. Such reviews can be helpful in the post-purchase phase of customer journey to retain an already existing customer by solving his issues and thereby attracting new customers. Likewise the use of supervised and semi-supervised learning is employed by B. Billal et al[14], to classify multi-label data of social media. J Hu et al[15] in his paper presents a popularity prediction model of images on social media, which can be a useful tool to predict product popularity in the near future. S. H. Ali[16] in his paper introduces a user profile learning model to provide accurate recommendations, that make targeting the customer much easier. H. I Alsaadi et al [17] in his paper proposed a forecasting system for hash tag time series, which is again very useful in determining the brand or service's popularity for future aspects.

B. Personality traits recognition using ML

For better development of buyer persona, we need to understand the nature of customers present on social networks. Understanding what type of personality users are interested in which type of product, brand or services will help us understand, what attributes contribute towards making a purchase decision by the customer. This information can be used to determine customer's journey phase the user is currently in [18]. This understanding will lead to persona creation which is practical and can be used to develop targeted content for the targeted audience. B. Y. Pratama et al[19] in his paper proposed the use of text classification to predict

personality of twitter user based on text tweets of users. V. Varshney et al[20] in his paper explores the techniques of various machine learning algorithms in order to deduce the personality of a user from his activities on social media. D. Xue et al[21] proposed a method of big five personality recognition from micro-blog in Chinese language environment with machine learning paradigm. D. Sewwandi et al[22] proposed a system that has focus on detecting the personality of person in three dimensions, namely ontology-based, linguistic analysis, questionnaire-based. T. Yo et al[23] discussed inference of personal attributes from tweets using machine learning. L. Asadzadeh et al[24], proposed a model of automatic personality recognition on Facebook.

C. Identifying the behaviour of user on Social networks using ML

We are gathering our source data from social media networks, and for understanding what drives a need for a product or service for the user before he/she actually takes an interest in that product or service, we need to first understand people behave on social media.

Y. Pan et al [25], have found that people on social media have two types of behaviours that are inseparably intertwined. The two types mentioned are functional behaviours and social behaviours. Zou1 et al [26], in his paper proposed a model of mining user behaviour patterns based on location. Similarly C. Tsai et al [27], discuss the user's personal preferences behaviour on social media. The obtained patterns can be used in recommendation systems and personalization services for the users of same area having similar behaviour.

D. Persona creation through ML

Lastly, we wanted to find research related to persona development which has made use of machine learning. We found the following research work that can be used as the starting grounds for automated persona creation from social media data. Although all the papers do not make use of machine learning technology and social media data, they are worth mentioning here.

A. Shiga et al [28], proposed a support system for making personas using Bayesian network analysis. It uses Bayesian network analysis (Bayesian-net) to organize and analyses the data. J. similarly An et al [29], has presented an approach for creating personas based on automated analysis of actual social media data, integrating data from Facebook, Twitter, and YouTube channels for a large commercial organization. Y. Watanabe et al [30], proposed Iterative Data-Driven Development of Personas

(ID3P) for practical applications of data-driven persona development on real services.

EXPERIMENTAL ANALYSIS

For the purpose of exploring the data present online and study the results this data yields under different Machine Learning algorithms, we have prepared three data sets from different sources.

The first dataset consists of reviews posted on Amazon for different products under category baby products. We have selected only one product category for sake understanding how machine learning models are applied to the real world data and what type of prediction can we do using the prepared model.

For Amazon dataset we have done sentiment analysis on the product reviews. We have chosen sentimental analysis due to the fact that similar personality people are most likely to choose the product in case of positive sentiment. This should narrow the personalities which are suitable for product marketing in our product persona.

The second dataset is obtained from twitter and used to study product popularity. The data is gathered by using hashtag and a linear regression model (ordinary least square) is used to predict the popularity of Sprite in the product category soft drinks.

The third dataset is obtained from YouTube, and it is also used to study the product popularity of Sprite using the same regression algorithm. Additionally we have also gathered data of product 7up which is same flavoured product of different brand and compared the predicted popularity.

Lastly the same products with different online platform are chosen to compare the product popularity on text based social media and video based social media.

EXPERIMENT RESULTS

For sentiment analysis we have used LogisticClassifier from turicreate library. We have taken very simple approach to decide positive and negative sentiment. We have taken all the reviews with 4 or higher rating as positive sentiment reviews and 3 or less rating as negative sentiment. We have than trained our model by supplying word count of positive and negative sentiment as features. The data is divided as 80 percent training data and 20 percent test data, the seed value is kept 0.

Our model got 92 percent accuracy on test data. The results suggest that sentiment analysis will be very useful for deciding the customer personality for the product persona.
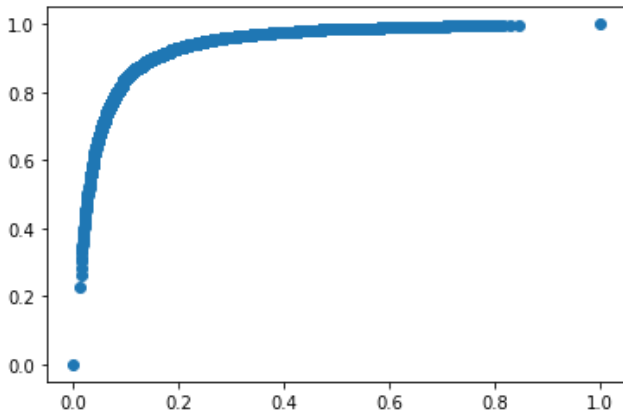
ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

Figure 1: Sentiment analysis ROC curve fpr vs tpr

We have used ordinary least square algorithm for preparing the linear regression model for twitter and YouTube data.

For the twitter data we have chosen retweets as a parameter to measure popularity. We found that retweets are directly affected by the number of followers a user has on twitter. But our regression model has predicted constant number of retweets irrespective of number of followers.
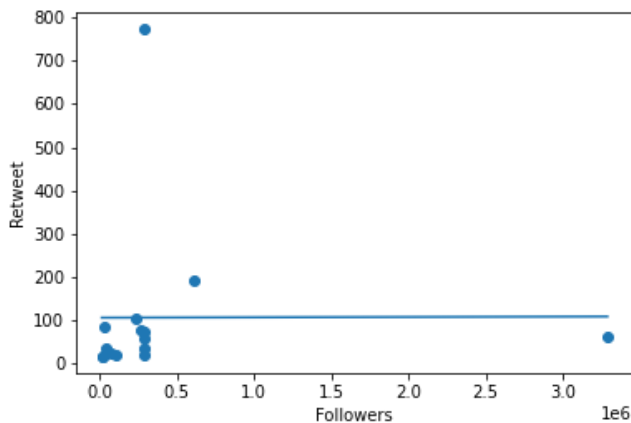

Figure 2: Prediction model for twitter results of Sprite

For the YouTube data we have chosen like count of the video as a parameter to measure popularity. We found that like counts are directly affected by the number of views by the users for a video. Our regression model has also predicted the rise in number of likes for the video as number of views increases. This observation was true for both sprite as well as 7up.

Looking at the figure 3 and 4 it seems our model has predicted slightly more popularity for 7up as compared to sprite. This is due to the fact that one video has great

much number of views resulting many like. If we consider this point on the graph as outlier then we can say that sprite is more popular among YouTube users then 7up.
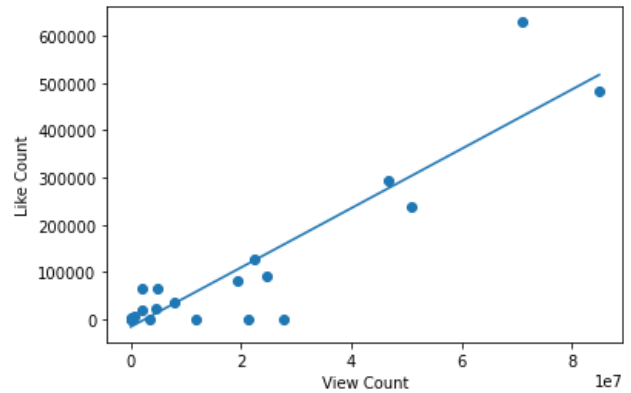

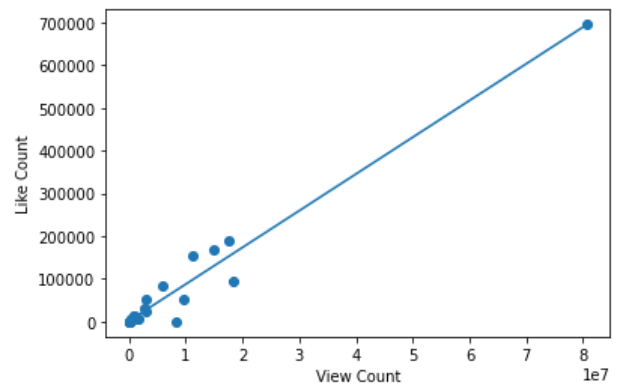**Figure 3: Prediction model for YouTube results of Sprite**

Figure 4: Prediction model for YouTube results of 7up

CONCLUSION

In this paper, we have gone through digital marketing and the potential use of machine learning technology to create digital marketing persona. We have carried out simple experiments to gain the insights about the social media data and find out whether this data can be used to build user personas. We have found that machine learning can be applied for analysis, classification, and prediction on social media data. Our results shows that this data indeed can be used for the purpose of gaining insights about the customers that will help in carrying out effective digital marketing campaign.

REFERENCES

[1] A. Perrin, M. Kumar, "About three in ten US adults say they are almost constantly online", www.pewresearch.org

[2] K. Baker, "The ultimate list of marketing quotes for digital inspiration, get inspired and motivated with the help of the best marketing quotes, blog.hubspot.com.

[3] L. Alexander, "What Is Digital Marketing?" https://blog.hubspot.com/marketing/what-is-digital-marketing

[4] [4]. R. E. Bucklina, C. Sismeiro, "Click Here for Internet Insight Advances in Clickstream Data Analysis in Marketing", Journal of Interactive Marketing 23 (2009) 35–48

[5] N. Singh, D. P. Singh, B. Pant, "A comprehensive study of Big Data Machine Learning Approaches and Challenges", International Conference on Next Generation Computing and Information Systems (ICNGCIS) 2017

[6] P.K. Kannan, H. Li, "Digital marketing: A framework, review and research agenda", International Journal of Research in Marketing, 34(1), pp. 22-45, 2017.

[7] L. Hintz, "What Is the Buyer's Journey?" https://blog.hubspot.com/sales/what-is-the-buyers-journey.

[8] J. Kietzmann, J. Paschen, E. Treen, "Artificial Intelligence in Advertising How Marketers can Leverage Artificial Intelligence Along the Consumer Journey", Journal of Advertising Research, September 2018.

[9] A. Micheaux, B. Bosio, "Customer Journey mapping as a New Way to Teach Data-Driven Marketing as a Service", Journal of Marketing Education, Volume: 41 issue: 2, page(s): 127-140, November 25, 2018.

[10] S. Herskovitz, M. Crystal, "The essentials brand persona storytelling and branding", Volume. 31, Issue. 3 of Journal of Business Strategy, pp. 21-28, 2010.

[11] A. Revella, "Buyer Personas Manifesto", 2011

[12] A. Cooper, P. Saffo, "The Inmates Are Running the Asylum", Macmillan Publishing Co., Inc., 1999.

[13] F.M.T. Hossain, Md. I. Hossain, S. Nawshin, "Machine Learning Based Class Level Prediction of Restaurant Reviews", IEEE Region 10 Humanitarian Technology Conference Dec 2017.

[14] B. Billal, A. Fonseca, F. Sadat, and H. Lounis, "Semi-supervised learning and Social Media Text Analysis towards multi-labeling categorization", IEEE International Conference on Big Data, 2017.

[15] J. Hu, T. Yamasaki, K. Aizawa, "Multimodal Learning For Image Popularity Prediction On Social Media", International Conference On Consumer Electronics, Taiwan-2016.

[16] S. H Ali, Ali I El Desouky, A. Saleh, "A new profile learning model for recommendation system based on machine learning technique" , JITSE.

[17] H. I. Alsaadi, L. K. Almajmaie, W. A. Mahmood, "Forecasting of Twitter Hashtahg Temporal Dynamics Using Locally Weighted Projection Regression" ICET-2017.

[18] Z. Yuzdepski, "Needs-based selling: Following five phases of the modern customer journey", https://www.vendasta.com

[19] B. Y. Pratama, R. Sarno, "Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM", International Conference on Data and Software Engineering, 2015.

[20] V. Varshney, A. Varshney, T. Ahmad, A. M. Khan, "recognizing personality traits using social media", Ieee international conference on power, control, Signals and Instrumentation Engineering ICPCSI-2017.

[21] D. Xue, Z. Hong, S. Guo, L. Gao, L. Wu, Jinghua, Zheng, N. Zhao, "Personality Recognition on Social Media With Label Distribution Learning", Digital Object Identifier.

[22] D. Sewwandi, K. Perera, S. Sandaruwan, O. Lakchani, A. Nugaliyadde, S. Thelijjagoda, "Linguistic Features Based Personality Recognition Using Social Media Data.", National Conference on Technology and Management, 2017.

[23] T. Yo, K. Sasahara. "Inference of personal attributes from tweets using machine learning", ieee international conference on big data, 2017.

[24] L. Asadzadeh, S. Rahimi, "Analyzing Facebook Activities for Personality Recognition", IEEE International Conference on Machine Learning and Applications, 2017.

[25] Y. Pan, Z. Wang, "Data Collection and Modelling of Personal Social Behaviors", International Conference on Services Science. 2015.

[26] Z. Zou1, X. Xie1, C. Sha, "Mining User Behaviour and Similarity in Location-based Social Networks", Seventh International Symposium on

Parallel Architectures, Algorithms and Programming, 2015.

[27] C. H. Tsai, H. W. Liu, T. Ku, W.F. Chien, "Personal Preferences Analysis of User Interaction based on Social Networks", International Conference on Computing, Communication and Security, 2015.

[28] A. Shiga, N. Nishiuchi, "A support system for making persona using bayesian network analysis", international conference on biometrics and kansei engineering, 2013

[29] J. An, H. Cho, H. Kwak, B. J. Jansen, M. Z. Hassen "towards automatic persona generation using social media", international conference on future internet of things and cloud workshops, 2016.

[30] Y. Watanabe, H. Washizaki, K. Honda, Y. Noyori, Y. Fukazawa, A. Morizuki, H. Shibata, K. Ogawa, M. Ishigaki, S. Shiizaki, T. Yamaguchi, T. Yagi, "ID3P: Iterative Data-Driven Development of Persona-based on Quantitative Evaluation and Revision", International Workshop on Cooperative and Human Aspects of Software Engineering, 2017.