Prospects on Predictive Analysis & their Applications

Sinu Mathew[1] Manisha Giri[2] Neha Singh[3] Neha Ghaisas[4]
[1,2,3,4]Assistant Professor
[1,2,3,4]Atharva College of Engineering and Technology, Mumbai

Abstract— Predictive Analytics is a cross disciplinary domain that roots across cognitive computing, cognitive Informatics, data science, knowledge science ad Artificial Intelligence. The growth and significance of this field is clearer in this century, especially in the light of social networks, location tracking and tailored services. The paper aims to preset a comprehensive study on the prospects on the prediction technologies, cover various applications ranging from industrial, medical to personal behavior. The implications that the analytics can have are also analyzed in detail. Few points that need to be addressed by organizations and companies while employing prediction techniques are also presented.

Key words: Prospects on Predictive Analysis**,** Predictive Analysis

I. INTRODUCTION

Data is considered to be the new oil. The interest-ed in data has increase so much in the recent years despite the fact that the word in itself sounds mundane and familiar we can't be fooled by the truth that it is an in valuable collection of experiences from which we need to learn. Any kind of negative or positive results like a fraudulent act, medical diagnosis tweets, and posts in face book-everything is encoded as data and they are ware-housed. The growth of this vast amount of data is estimated to grow by 2.5quintillion bytes per day. If the computers are efficiently they can easily process this ocean as if it is a drop of water.

Predictive technology is very common and it effects the experiences of every one every day. In terms of business, prediction can empower an organisation in such a way that it can come up with a completely new competitive edge. Businesses can secure an incomparable competitive strong hold by foreseeing and predicting the future and the value of assets.

Computerised prediction can bring fourth valuable achievements and it is made possible by the people behind it and the science that fuels it. But the toughest challenge is to make such prediction which depend on many factors though the data required about a person, an organisation is known, it is quite a difficult task to put all these pieces together to make powerful prediction.

Though it might seem simple it is not so easy. It is by systematic and scientific means that we can face the challenge of developing and improving the prediction. Machine learning is the answer .computers need to develop new knowledge, capabilities by continuously feeding on the unnatural resource called Data.

Some of the examples of what can be predicted:
- Barack Obama was re-elected in 2012 with the help of predicting the voters. The campaign "Obama for America" predicted who would be persuaded in a positive way and who would be discernibly influenced to vote adversely by con-tact. The prediction method proved to be successful enough to convince a large numbers of voters to choose Obama. This would have been otherwise less successful if the campaign targeting in thru traditional way.
- The American public university system predicted the students that may drop out and employed these prediction results to intervene in a successful manner.
- $1 million was awarded to a group of scientists by Netflix who enhanced the ability of their recommendation system to predict the movie that people will like.
- Most of the leading social networks that are focused on career like LinkedIn predicts the jobs skills of the people.

In short, Predictive Analysis (PA)is a technology that learns from experience, which is nothing but data, to predict the future behaviour of people, which in turn can help make good decisions. Predictive analysis differs very much from fore-casting the field spans across domains like analytics, data science, big data, artificial intelligent, cognitive computing and business intelligent.PA is not synonymous with data mining but the latter is commonly used in broader aspects.

II. LITERATURE SURVEY

In "Predictive Analytics, The power to predict who will click, buy, lie or die", Eric Siegel, the founder of Predictive analytics presents an extensive study on the field ad their ever growing significance spanning every field. Various case studies of organisations that already employed PA tools to increase their efficiency, turnover and prevent adverse effects are made.

The analytics system that are for social media are Whisper [1], which aims on information propagated in the social platform Twitter, SensePlace2 [2], that analyses the weightage of tweets geo-graphically and TweetXplorer [3] which unifies the visualisation of tweets geographically as well as their social networks. Some other applications rely on the idea of using such analytic tools in case of emergency response. Various data sources of various organisations are integrated for disaster management through a analytic tool developed by Thom et al. [4]

A variety of applications employs the predictive analytics tools utilized by both novice and expert analysts. (e.g., R [5], SAS [6], Weka [7], JMP[8], Excel). These technologies provide a variety of machine learning procedures that are employed for analytics tasks[9][10]. The following section presents a basic observation on predictive Analytics.

III. OBSERVATION

Exponentially increasing loads of data, the welcoming approach of organisations to embrace predictive technology and well advanced software solutions to integrate PA to the organisation are opening ways for the expansion of this field. The launching of predictive analytics means to act on the prediction and apply what has been learned, and what has been discovered with in data .the sheer value the results from the prediction effects straightforward. For example while targeting direct marketing can be predicted who are worth contacting and who should be left alone. The dealing, for example, can be stated cutting costs by three quarters in exchange for losing only quarter of sales. This is not just some notion of abstraction rather it mean business

A predictive model is mechanism that predicts and individuals' behaviour. The characteristic of individual as taken as input and predictive score provided as the output. The higher the score, the more likely by the individual to do the predicted behaviour .Machine learning built the predictive model .predictive model generate the whole model right from the scratch. The rules of the model are created by the computer automatically. The PA develop this predictive power through this (pg no.26) automation.

The ever increasing need and on utilising data to make decisions can influence in several ways; at times it is all about expecting the results that can generate more positive outcomes, manipulating risks that can prevent adverse outcomes. Author Ian Kerr explains the different needs of using predictive analytics and the technologies. Preferential predictions, a kind of prediction explained, are the attempt to study and learn the preferences of the individual and their respective behaviour in order to tailor the services provided to them and under-stand their needs in an efficient way. Preemptive predictions are another kind that prevents some kind of actions which were most likely to lead to some social complications.

Preferential and preemptive predictive analysis forms the foundation of basic knowledge of predictive analytics. Both of them generate some predicted outcomes which are sought by the vary-ing needs of the predictive analytics. In the next section, we will try to cover some applications where predictive analytics are and will be widely used in order to enhance their performance, be it public or private sectors.

IV. APPLICATIONS

A. Law Enforcement and Intelligence

Data mining and profiling technologies has for long been employed and incorporated by the law enforcement agencies in the western countries especially, in order to anticipate, predict and identify social risks and potential criminal activities that can occur. The contemporary world which is ever increasing crime rates, the significance of such technologies are highly considered. PA is becoming more prominent among the law enforcement agencies and in many Western countries, like U.S where they predict potential threats based on previous crimes, historical records and patterns of criminal behaviour.

Some of the crime detecting technologies claims to be well advanced enough to predict when and what crimes will be detected before anything could ever happen. Rather than knowing it after the crime is done, the new way of law enforcement will be to predict and then act upon it. Behavioural patterns of criminals and terrorists are analysed in order to predict actions and the public safety and the national security may be more into predicting which person is more likely to commit crimes.

Depending on the level of success rate of predictions, technologies can be developed enough to come up with more accurate predictions to enforce law and order in an effective manner.

B. Tracking the Location

The explosive rate of growth of the mobile technology has enabled them to grow leaps and bounds concerning the personalisation and customisation of the way they reach out to the users. Location tracking has become the backbone to reach out to the users with tailored services. The services available online and the apps in the mobile are popularising this trend.

The rate of accuracy of predicting a future journey or travel by a person based on their location tracking at current time has proved to be more precise and accurate compared to traditional methods of assumptions. In one of the recent studies, an algorithm developed was able to predict the GPS co-ordinates of where a person might travel in the future with an accuracy of within 1000 square meters. When an additional information of asking friend is done, the accuracy rate become more i.e., within 20 meters. Even without any of these in-formation if the mobile tower locations are tracked the future GPS coordinates can be predicted, according to the study. This is definitely an interesting knowledge for companies or markets to tailor their advertisement according to the pre-dictions made. Law enforcement agencies can employ this technology so that they can predict the location of criminals.

C. Social Media

In case of social media, PA is a well renowned domain since long, to collect the data that can tell about the behaviour of the human beings, and the patterns of their communication, analyse their sentiments and their level of social influence. Tweets in twitter, posts in Facebook, Job searches in LinkedIn, Google searches, enables to predict the future behaviour and sentiments.

With respect to collecting such a voluminous data and studying on them, especially mining the data, Facebook is considered to be the pioneer. It has a vast set of data which has ever been assembled in order to study the behaviour of human and their scientists are working on finding innovative ways to mine these data to get information on the communication of

human beings. As the social plat-form is collecting data from the users in real time, the scientists are in a position wherein they have to analyse the behavioural patterns of the human beings, their interactions and their influence in real time collects data from users interacting in real time.

D. Fraud Prevention

Fraud detection is another area where the power of Predictive Analytics power can be highly capitalised. In order to prevent fraud, government and private sectors could utilise PA. Fraudulent claims of insurance, benefits of the government can analyse the patter and trends to detect and prevent such claims. Multiple databases are analysed and identify the applicants who have been overpaid. The files which are flagged for review are further investigated. Instead of using analytical tools to detect frauds ad manage risks, this will provide a shift towards automation of these processes

The applications said above can further be narrowed down. There are many others apart from these where PA can play a vital role in improving their efficiency. Some of them are covered in Ta-ble.1.

V. IMPLICATIONS OF PREDICTIVE ANALYTICS

Given so many advantages across a large number of fields, the implications it can have on the privacy will differ depending on the context in which it is being used, how it is used ad how it is implemented. Though the field can trigger advanced research, innovative techniques and enable us to look at the world in a new way, societal implications o privacy and the perspective of the individuals that can arise when the PA is implemented is contemplated in the following section.

A. Privacy Issues

Though at many times the predictions may go wrong and can be inaccurate, if the predictions prove right, especially if the prediction is about the behaviour, beliefs and inferences of an individual, it is widely considered as invasion of core privacy. The collection of data is not harmful or violating, but piecing it together to acquire some important information and using it to monitor an individual is a violation of privacy. The predictions that are made about an individual based on their social activities and interactions are not considered alarming. It is those presumptions that are obtained by inspecting our activities in a road ad deep manner and using technical assistance for the same can be termed a complete violation of individual privacy.

B. Vague Terms and Results

Many companies acknowledge the use of the employee's personal information for the sake of analytics but the process is quite opaque. Most of the individuals do not understand the complex legal terms in which they are specified. Many of the private policies in internet are quite vague. Though some of them read the policy well, the impact of the consequences in future are not properly weighed. The more information an individual discloses, the more the risks can be.

Though the users give information seemingly well aware of the control they have, they do not really know how much information is provided, how they are going to be used or who is able to access those. They might have some control at the front end but there is a lack of comprehensive under-standing about the manipulation of their information behind the scenes. The organizations and companies are in the process of acquiring every intricate detail of individuals, whereas individuals know very much less about the former. In case of PA techniques which are employed by government, the process and the results are hidden from the public for public safety and for reasons of national security.

C. Prejudice and Reputation Concerns

While employing PA techniques to infer the future actions and behavior of a person, there are risks that the person can be discriminated and be prejudiced based on the predictions. The problem is not about collecting the information, but what is deduced from the information. If the information obtained from predictions is inaccurate, the person upon whom the wrong reputation is sealed, it will be an issue of unfairness. Based on the misinformation, it will be an injustice to exclude the individual from the society and deprives him of his benefits

Predictions	Organizations that use PA
Target marketing	Mutual fund investment management firm: Clients who were more likely to make additional investments where identified. PREMIER Bankcard: Mailing cost was reduced by $12 million.
Increase the sales rate	Hewlett-Packard: Sales staff is alerted with business opportunities with outcomes for 91% of sales efforts and the accuracy is 95%.
Choice of prodts	Amazon.com: Through product recommendation, they acquire 35% of sales
Mouse clicks	Google: The search feature of Google is improved through predictions about which pages will satisfy the users standards if it is shown as search results
Friendship	Facebook: Improves the precision of suggested people you might know and would like to link to.
Death	Life insurance companies: The age of death of a policy holder is predicted to make decision on application approval and pricing
HIV Progression	Researchers: The accuracy of predicting the disease improved from 70 to 78%
fraudulent activities	Aviva Insurance (UK): Detection of fraudulent claims amounted to savings of half million pounds in a month
Defective Items	Washing machine manufacturer : 87% increase in the performance of Fault detection

Dropped calls	Nokia Siemens: To improve the availability of services and service continuity by 70%accuracy on a 4Gwireless network.
Voter persuasion	US 2012 campaign "Obama for America": Prediction on which voter would be positively persuaded by campaign contact and who would actually be negative influence to vote adversely by contact.
Driver inattentiveness	Ford motor company: When a driver is not alert due to fatigue or if he is distracted, if there is an alert and non-alert driver, the latter can be identified with an accuracy of 86%.

Table 1: Predictions extensively being used in Organizations

Key areas to be addressed by Organizations in the utilization of PA

- Consent – Each individual should be imparted a basic understanding of how their information will be used.
- Transparency – Sharing information handling practices with individuals in such a way that it is neither too complex for them to understand or too much summarized in such a way that important details are lost.
- Accountability & Ethics – Companies should respect the cultural and social standards of acceptable behavior and make decisions accordingly.

VI. CONCLUSION

The field of predictive analytics in the light of social networks, tracking location and many more advanced technologies prompts a extensive re-search and study explicitly on its own. The power of prediction can take the growth of organisations to a different level, enable better healthcare services ad especially aid better law enforcement. The various issues of such prediction technologies, especially their effect on privacy is also studied.

REFERENCES

[1] Millar, J. (2009). "Core Privacy: A Problem for Predictive Data Mining", in Lessons from the Identity Trail: Anonymity, Privacy and Identity in a Networked Society, Ian Kerr, Val Steeves, Carole Lucock (Eds.), Oxford University Press. 103 – 119.

[2] Gutwirth, S. and Hildebrandt, M. 2010. 'Some Caveats on Profiling'. In Data Protection in a Profiled World. S. Gutwirth et al. (eds.), Springer Science & Business Media B.V. 2010.

[3] CNN Tech. "Should we fear mind-reading fu-ture tech?" By Andrew Keen, posted June 19, 2012.

[4] Schwartz, Paul.M (2010) Data Protection Law and the Ethical Use of Analytics. The Centre for Information Policy Leadership Hunton & Williams LLP

[5] The Vancouver Sun. The erosion of anonymity: Today's digital world forces us to share more of our personal information. By Misty Harris, Published April 2, 2012.

[6] Terrence Craig and Mary E. Ludloff. 2011. Privacy and Big Data. O'Reilly Media Inc. p.5.

[7] Pridmore, Jason and Detlev Zwick. 2011. Edi-torial: Marketing and the Rise of Commercial Consumer Surveillance . Surveillance & Socie-ty 8(3): 269-277.

[8] The New York Times. How Companies Learn your Secrets by Charles Duhigg, Published: Feb-ruary 2012.

[9] Duhigg, C. 2012. The Power of Habit: Why we do what we do in life and business. Doubleday Canada, Random House Canada Ltd.

[10] Dellarocas C, Awad NF, and Zhang X. 2007. Exploring the Value of Online Product Ratings in Revenue Forecasting: The Case of Motion Pic-tures. Journal of Interactive Marketing, 21(4):23-45