

Possibility of Preventing Cancer using Big Data Analytics and Data Mining Technology

Prof. Dr. Satyajit Chakrabarti,
University of Engineering & Management, Kolkata, India

Abstract: This paper presents a case study of how Big Data analysis & Data Mining technology can provide an indicative solution to humanity for early detection and prevention of cancer.

Keywords: *Prevention of Cancer, Big Data Analytics, Data Mining*

I. INTRODUCTION

This paper discusses the use of Big Data Analytics and Data Mining technology for determining personal habits that may lead to the development of cancer tissues in humans. The University of Engineering & Management (UEM), Kolkata, India, got access to the medical history of around 1,50,000 patients, preserved over a period of 25 years at Dr. Sudarsan Chakrabarti Memorial Centre, Ultadanga, Kolkata, India.

The concept of this close interdisciplinary co-operation between UEM and Dr. Sudarsan Chakrabarti Memorial Centre provides a methodology for clinical diagnosis based on the patients' data which has been maintained and categorized as follows:

- (a) Life-style
- (b) Habits
- (c) Promptness of accepting adequate treatment
- (d) Family history of disease
- (e) Profession/livelihood pattern

This data may help in determining the possibility of developing cancer in patients.

The above-mentioned sections - (a), (b), (c), (d) and (e) - of the medical data of patients can be treated as non-confidential. These fields can be designed to help the software perform desired data mining.

II. PROBLEM FACED IN DATA MINING

(A) A patient's medical history is generally confidential. In order to maintain this confidentially, the data must be maintained in a separate field so that required data can be mined without disturbing the confidential medical history of the patient.

(B) The period for which the study was done comprises of 25 years. The period should ideally be for 50 years or more.

(C) Data card was not uniformly maintained. On many occasions, the physicians did not bother to maintain the medical history as required for the present study. And data mining required long hours and laborious effort.

(D) Out of the huge volume of patient historical data only a few data points could be used as many of the data available were having inadequate information about patients' identity and/or contact address.

(E) Researcher had to use telecallers to get permission to use the available data without disclosing the confidential medical history.

(F) Detail study could only be conducted for the patients who suffered from the dreaded cancer disease and reported the same to the medical centre.

III. THROAT CANCER & PROSTRATE CANCER - A CASE STUDY

By conducting Big Data analysis and using Data Mining techniques we noted interesting facts:

(i) A person who remained careless about throat problems did not take immediate remedial measures and allowed that problem to persist on as a minor and insignificant irritation has higher probability of turning regular minor irritants into major disease of cancer.

In the patient history sheet, the clinical habits like the following were noted:

- (i) Does the patient suffer from frequent cold & cough?
- (ii) Does the patient practice regular saline water gargle?
- (iii) Does the patient take anti-allergic drug immediately?
- (iv) Does the patient take anti-biotic drug immediately or wait for symptoms of the problem to increase?

The history of the patients' data sheet showed that patients

(A) who suffered from chronic bronchial problem and/or regular allergic cold and cough syndrome or have the habit of taking too hot or too cold meals and/or smoking etc. for a long period of time faces the risk of converting the healthy throat tissues into cancerous tissues.

(B) who have the habit of drinking less water or frequently suffer from “urinary tract infection” or hold urine for a long time due to bad habit, and/or professional demands (e.g. security guard cannot leave the post etc.) faces higher threat of prostatic cancer.

The historical data analysis showed that if a person takes immediate action to mitigate the problem with adequate measures, the problem of turning a chronic disease into cancer is less.

This was demonstrated while we did data mining and Big Data analysis. Total data available was the medical history of 1,50,000 patients. Out of that around 350 patients medical history could be studied and patients who developed cancer, was found to

(i) be medically negligent,

(ii) have not continued with the anti-allergic drug/antibiotic drug promptly and

(iii) have avoided medicine administration as a good practice to avoid “side effects” of strong medicine.

This finding gives a strong indication that cancer is an irreversible rebellion of healthy tissues of the body. These tissues, time & again, suffer from external ill treatment like smoking, taking extra hot or extra cold meal, inadequate attention to the suffering of the tissues due to the untreated cold and coughing, unfriendly environment, inadequate protection against allergy attack,

etc. These rebellious tissues subsequently start killing red blood cells (RBC).

CONCLUSION OF THE STUDY

The maintenance of the patients’ data for a longer time of say 50 years, and Big data analysis & Data Mining technology will help firmly determine the truth in finding of this article that cancer can be prevented by taking some prompt and simple medical attentions throughout life.

Cancer is an irreversible rebellion of the healthy tissues resulting from prolonged negligence of the natural function of the tissues. The healthy tissues seem to dramatically turn into irreversible and mutinous as a result of continuous torture and negligence.

Some types of cancer, as per our study, shows similar trait i.e. the good healthy tissues become irreversibly cancerous tissues. The old saying “A stitch in time, saves the nine” is valid in medical field also. Our study showed that immediate medical action taken throughout life can protect the body from developing cancerous cell. Result is probabilistic and not deterministic.

FURTHER SCOPE OF STUDY

Our study was limited to the local area of Kolkata, India. The study needs to be done on various ethnic groups and in many countries with different food habits, life-style and social behavior. These can give definite proof of the conclusion made in this article.

References

This work is a unique experiment. We did not have any reference of any earlier work on this line of thought.