

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/258789725>

# Early Prevention and Detection of Skin Cancer Risk using Data Mining

Article in *International Journal of Computer Applications* · January 2013

DOI: 10.5120/10065-4662

CITATIONS

21

READS

170

3 authors, including:



**Kawsar Ahmed**

Mawlana Bhashani Science and Technology ...

90 PUBLICATIONS 405 CITATIONS

[SEE PROFILE](#)



**Tasnuba Jesmin**

9 PUBLICATIONS 98 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Bio-informatics [View project](#)



Surface Plasmon Resonance (SPR) Based Biosensor Design [View project](#)

# Early Prevention and Detection of Skin Cancer Risk using Data Mining

Kawsar Ahmed

Department of Information and Communication Technology,  
Mawlana Bhashani Science and Technology University,  
Santosh, Tangail-1902

Tasnuba Jesmin

Department of Information and Communication Technology,  
Mawlana Bhashani Science and Technology University,  
Santosh, Tangail-1902

Md. Zamilur Rahman

Department of Information and Communication Technology,  
Mawlana Bhashani Science and Technology University,  
Santosh, Tangail-1902

## ABSTRACT

Till now Cancer is a big question for scientific community cause of no existing treatments could solve the problems related to this dreadful disease. Research is in well progress since half century but it failed to give an accurate solution to fight against it. The development of technology in science day night tries to develop new methods of treatment. One such mile stone treatment for cancer that is giving good hope to the people is cancer treatment based on genome sequencing. With respect to Bangladesh, Skin Cancer is a fatal, deadly, disabling and costly disease whose risk is increasing at alarming rate because of unconsciousness. Like other cancer Skin Cancer also depends on some factors that are known risk factors of skin cancer. So the detection of Skin Cancer from some important risk factors is a multi-layered problem. Initially according to those risk factors 200 people's data is obtained from different diagnostic centre which contains both cancer and non-cancer patients' information and collected data is pre-processed for duplicate and missing information. After pre-processing data is clustered using K-means clustering algorithm for separating relevant and non-relevant data to Skin Cancer. Next significant frequent patterns are discovered using MAFA algorithm shown in Table 1. Finally implement a system using Lotus Notes to predict Skin Cancer risk level with suggestions which is easier, cost reducible and time saveable.

## General Terms

Computer Science, Data Mining, Bioinformatics.

## Keywords

Skin Cancer, Data Pre-processing, Disease Diagnosis, Classification, MAXimal Frequent Itemset Algorithm (MAFA) algorithm, K-means clustering and significant frequent pattern.

## 1. INTRODUCTION

Skin is that, which helps to protects the body from infection, heat, injury, and any type of damage which caused by ultraviolet (UV) radiation (such as from the sun or sunlamps), germs, chemical. Skin stores water and fat. It helps body to control heat and water. Also, helps skin to make vitamin D. So protection of skin from diseases is the significant and intricate job in medicine.

Sometime skin is affected by different kinds of diseases where skin cancer is most one. Skin cancers are named for the type of

cells that become malignant (cancer). The three most common types are [1]:

- **Melanoma:** Melanoma begins in melanocytes (pigment cells). Most melanocytes are in the skin. Melanoma can occur on any skin surface. In men, it's often found on the skin on the head, on the neck, or between the shoulders and the hips. In women, it's often found on the skin on the lower legs or between the shoulders and the hips. Melanoma is rare in people with dark skin. When it does develop in people with dark skin, it's usually found under the fingernails, under the toenails, on the palms of the hands, or on the soles of the feet.

- **Basal cell skin cancer:** Basal cell skin cancer begins in the basal cell layer of the skin. It usually occurs in places that have been in the sun. For example, the face is the most common place to find basal cell skin cancer. In people with fair skin, basal cell skin cancer is the most common type of skin cancer.

- **Squamous cell skin cancer:** In people with dark skin, squamous cell skin cancer is the most common type of skin cancer, and it's usually found in places that are not in the sun, such as the legs or feet. However, in people with fair skin, squamous cell skin cancer usually occurs on parts of the skin that have been in the sun, such as the head, face, ears, and neck.

Unlike moles, skin cancer can invade the normal tissue nearby. Also, skin cancer can spread throughout the body. Melanoma is more likely than other skin cancers to spread to other parts of the body. Squamous cell skin cancer sometimes spreads to other parts of the body, but basal cell skin cancer rarely does. When skin cancer cells do spread, they break away from the original growth and enter blood vessels or lymph vessels. The cancer cells may be found in nearby lymph nodes. The cancer cells can also spread to other tissues and attach there to form new tumours that may damage those tissues. The spread of cancer is called metastasis.

When skin cancer spreads from its original place to another part of the body, the new tumor has the same kind of abnormal cells and the same name as the primary (original) tumor. For example, if skin cancer spreads to the lung, the cancer cells in the lung are actually skin cancer cells. The disease is metastatic skin cancer, not lung cancer. For that reason, it's treated as skin cancer, not as lung cancer.

Treatment for skin cancer depends on the type and stage of the disease, the size and place of the tumor, and your general health and medical history. In most cases, the goal of treatment is to remove or destroy the cancer completely. Most skin cancers can be cured if found and treated early.

Skin cancer treatment may damage healthy cells and tissues, unwanted side effects sometimes occur. Side effects depend mainly on the type and extent of the treatment. Side effects may not be the same for each person. Before treatment starts, your health care team will tell you about possible side effects and suggest ways to help you manage them. Many skin cancers can be removed quickly and easily. But some people may need supportive care to control pain and other symptoms, to relieve the side effects of treatment, and to help them cope with the feelings that a diagnosis of cancer can bring.

## 2. BACKGROUND

When one got cancer, it's natural to wonder what may have caused the disease that means what is the risk factor. Risk factor means those reasons which increase the chance of getting a disease. The main risk factor for skin cancer is exposure to sunlight (UV radiation). Approximately 65% to 90% Skin cancer is caused by UV radiation [2]. There are also another some risk factors of skin cancer. The diagnosis of Skin Cancer is a significant and tedious task. The detection of Skin Cancer from some important risk factors is a multi-layered problem.

Children have more opportunities and time than adults to be exposed to sunlight and thus more opportunities to increase their risk of developing skin cancer. At least 25% of a person's lifetime UV exposure occurs during childhood [2]. Most parents remembered hearing about the importance of protecting their children from the sun, yet children are still playing in the sun without sunscreen or protective clothing. Many people these days are not taking skin cancer seriously after knowing its significance.

Skin cancer is more common where the sun is strong. Sunlight can be reflected by sand, water, snow, ice, and pavement. The sun's rays can get through clouds, windshields, windows, and light clothing. So we encourage people to limit their exposure to sunlight.

Another risk factor of Skin cancer as well as other diseases is Obesity. A person who has more weight than height that means more BMI is responsible for occurring different kinds of disease [2].

A person who was affected by any kind of cancer has an increased risk of developing another skin cancer of any type. A person who has two or more close relatives (mother, father, sister, brother, or child) who are responsible for developing Skin cancer has a risk factor of developing skin cancer for his own. Rarely, members of a family will have an inherited disorder that makes the skin more sensitive to the sun and increases the risk of skin cancer. Having fair (pale) skin that burns in the sun easily, blue or gray eyes, red or blond hair, or many freckles increases the risk of skin cancer.

Skin cancer assessment in Bangladeshi population, which includes- age, gender, hereditary, outdoor activities, working period in any industries, colour of body/skin, serious experienced in childhood, previous health examination, use of anti-hypersensitive drugs, smoking, food habit, obesity, genetic risk, environment, excessive alcohol, radiation therapy and use chemical for body without thinking its quality. Those Skin cancer risk factors we get with the help of some people and study.

A widely recognized formal definition of data mining can be defined as "Data mining is the non-trivial extraction of implicit previously unknown and potentially useful information about data" [3]. Data mining has some fields to analysis of data such as classification, clustering, correlations, association rule etc. Now-a-days data mining has been used intensively and extensively by many organizations. In-healthcare, data mining

is becoming increasingly popular [4]. Data mining provides the methodology and technology to analysis the useful information of data for decision making.

Data pre-processing is an important task of data mining. It mainly used for making analysis appropriate and also making data appropriate for clustering by avoiding duplicate records and adding missing data according to past recorded data. The main benefits of data pre-processing reduces memory.

Clustering is a process of separating dataset into subgroups according to the unique feature. Clustering separated the dataset into relevant and non-relevant dataset to Lung Cancer. AprioriTid [5] and Decision Tree algorithm [6] are mainly used to find out frequent patterns of dataset. But we use MAFIA algorithm [7] that is very easy and effective to find out frequent patterns. Frequent patterns, the sets of data are frequently occurred into data warehouse. Significant frequent pattern, the set of data are mostly responsible to Skin Cancer. Using this significant pattern implemented a prediction system for Skin Cancer that is capable to provide Skin cancer risk level and some suggestions to reduce Skin cancer risk.

Day by day the number of Skin cancer person is increasing rapidly because of unconsciousness about Skin cancer and risk factors of Skin cancer. The main goal of our research is to develop a system that can be used by a person for testing his/her Skin Cancer risk level and provides suggestions according to his/her skin cancer risk factor.

## 3. METHODOLOGY

The skin is the largest organ in your body. It does many different things [8]:

- Covers the internal organs and helps protect them from injury
- Serves as a barrier to germs such as bacteria
- Prevents the loss of too much water and other fluids
- Helps control body temperature
- Protects the rest of the body from ultraviolet (UV) rays
- Helps the body make vitamin D.

So protect our body from Skin cancer is must to lead a healthy life. Here we discuss about the whole process of discovering Skin cancer risk level.

### 3.1 Data Collection

200 persons' data is collected from different diagnostic centre where both male and female are present of different age. Obtained data also contains both patient and non patient information. From the previous studies 18 risk factors were considered for Skin cancer assessment in Bangladeshi population, which includes- age, gender, hereditary, outdoor activities, working period in any industries, colour of body/skin, serious experienced in childhood, previous health examination, use of anti-hypersensitive drugs, smoking, food habit, obesity, genetic risk, environment, excessive alcohol, radiation therapy and use chemical for body without thinking its quality.

### 3.2 Data Pre-processing

Sometime obtained data contains duplicate information and missing any values of information. So data pre-processing is an essential which is the tedious task of data mining. Making an appropriate analysis and suitable for clustering of collected data is the main goal of data pre-processing. Data pre-processing deletes the duplicates data and supplies the missing values according to the past recorded data. It also reduces the memory and normalizes the values used to represent information in database.

### 3.3 Clustering of Collected Data

The process of partitioning and category of collected data into different subgroups where each groups have a unique feature is called clustering [9]. Clustering is another tedious term of data mining. The clustering problem has been addressed in numerous contents besides being proven beneficial in many applications [10]. The goal of clustering is to classify objects or data into a number of categories or classes where each class contains identical feature. The main benefits of clustering are that the data object is assigned to an unknown class that have unique feature and reduces the memory.

The K-means clustering [11] is a widely recognized clustering tool that is used for robotics, diseases and artificial intelligence application purposes [10]. Here  $k$  is a positive integer representing the number of clusters. The pre-processed data is clustered using the K-means clustering algorithm with the value of  $k$  equal to 2. This represents there is two clusters where one cluster contains relevant data to Lung Cancer and another contains remaining data that means non relevant data.

### 3.4 Discover Frequent Pattern

Discover frequent pattern is the most useful, significant and vital topics of data mining. It is known as the principle data mining problem that intends to find out the frequent items or patterns from the data warehouse [12]. There are different kinds of algorithms, used to mine interesting frequent patterns from databases like association rules, clusters, classifications and correlations etc such as Apriori, AprioriTid, Decision Tree, and FP-Tree.

After clustering MAXimal Frequent Itemset Algorithm (MAFIA) is used to mine the frequent patterns. The MAFIA algorithm is efficient algorithm than AprioriTid and Decision Tree algorithms of extracting the frequent patterns from clustered dataset.

Now one may question? Why MAFIA algorithm is selected to find out frequent patterns, but not others? Then answer is that MAFIA algorithm is chosen for its efficiency and simplicity than others for large data set. And MAFIA algorithm is treated as better algorithm than others because it combines diverse old and new algorithmic ideas to form a practical algorithm [7] [13].

**Pseudo code for MAFIA [14] [7]:**

```
MAFIA(C, MFI, Boolean IsHUT) {
  name HUT = C.head C.tail;
  if HUT is in MFI
    stop generation of children and return
  Count all children, use PEP to trim the tail, and recorder by
  increasing support,
  For each item i in C, trimmed_tail {
    IsHUT = whether i is the first item in the tail
    newNode = C I
    MAFIA (newNode, MFI, IsHUT)}
  if (IsHUT and all extensions are frequent)
    Stop search and go back up subtree
  If (C is a leaf and C.head is not in MFI)
    Add C.head to MFI
}
```

The cluster that holds most relevant data to Skin cancer is fed as input to MAFIA algorithm to mine the frequent patterns present in it. Then the significance weightage of each pattern is calculated using the following approach that described in the below subsection.

### 3.5 Significant Pattern Find-Out

After discovering the frequent patterns using MAFIA algorithm, the weightage significant patterns are mined by using the Equation (1) [10]

$$Sw(i) = \sum (W_i * F_i) \quad (1)$$

Where  $W_i$  is the weightage of each attribute and  $F_i$  denotes the number of frequency for each rule.

And significant Frequent Pattern is selected by using the following Equation (2)

$$SFP = Sw(n) \geq \phi \text{ for all values of } n \quad (2)$$

Where SFP shows significant frequent pattern and  $\phi$  represents significant weightage.

## 4. RESULTS

Cancer can endanger life and is the leading cause of death in many countries. Skin cancer is the most prevalent of all cancers, and it's increasingly on the rise in Bangladesh. Skin cancer has been a growing problem in Bangladesh and millions of people have suffered from it every year. The experimental results are separated into two sections. One is significant frequent patterns discover and another is represents prediction tools to Skin Cancer. Finally compare it to others existing techniques.

### 4.1 Result for significant frequent pattern

Using data from data warehouse, the significant patterns are extracted for Skin cancer prediction. The collected data are pre-processed by avoiding duplicate records and adding missing values. Then pre-processed data is clustered using K-means cluster algorithm with  $k$  equal to 2. And finally significant frequent patterns are mined using MAXimal Frequent Itemset Algorithm (MAFIA) shown in Table 1.

**Table 1. Significant Pattern and their corresponding Weightage value using MAFIA Algorithm**

Significant Patterns	Weight age
Age -Sex-Outdoor activities –Affected family members - Childhood serious experienced - Working period in industries – Affected any cancer before	185.50
Age- Use unhealthy chemical -Sex- Affected family members - Childhood serious experienced - Working period in industries - Skin colour	170.00
Age- Green foods – Skin colour –Sex- Affected family members – Childhood serious experienced – Working period in industries	155.50
Outdoor activities – Sex- Obesity – Affected family members –Green foods– Use unhealthy chemical- Affected any cancer before	150.00
Outdoor activities –Obesity- Childhood serious experienced – Green foods-Affected any cancer before	135.50

### 4.2 Result for Prediction to Skin cancer

Finally using the significant patterns that are taken by using MAFIA Algorithm the prediction tool to Skin Cancer is

implemented. The frequent pattern parameters and their corresponding score are shown at Table 2 and Figure-1 represents the risk level of Skin Cancer which is implemented using Table 2.

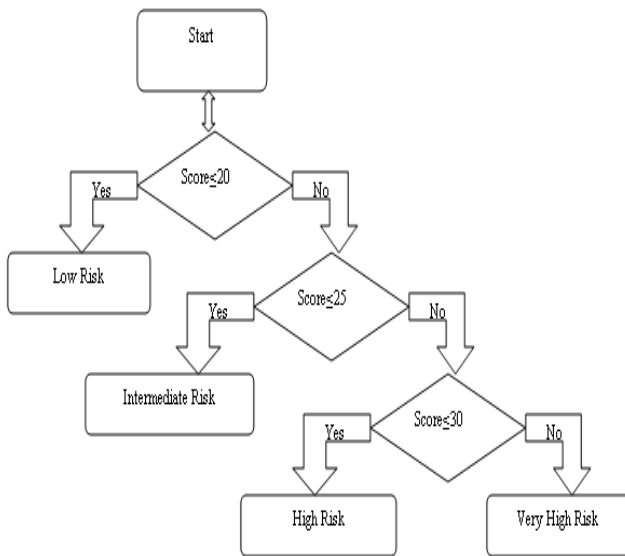


Fig 1: Flow diagram of decision tree algorithm

Table 2. Significant Pattern and their corresponding Weight age and Score

WelCome To EveryBody

Name	<input type="text" value="Kona Akter Mou"/>	Affected Family Members?	<input type="text" value="None"/>
Age	<input type="text" value="45 Years"/>	Working Period in Industries?	<input type="text" value="None"/>
Gender	<input checked="" type="radio"/> Female <input type="radio"/> Others <input type="radio"/> Male	Color of Body/Skin?	<input type="text" value="Swart"/>
Out door Activities?	<input type="text" value="Medium"/>	Childhood Series Experienced?	<input checked="" type="radio"/> No <input type="radio"/> Yes
Eat Green foods/vegetables?	<input checked="" type="radio"/> No <input type="radio"/> Yes	Use chemical for body with low quality or without check?	<input type="text" value="Frequently"/>
Take Radiation Treatment Before?	<input checked="" type="radio"/> No <input type="radio"/> Yes	Affected By any Cancer before?	<input checked="" type="radio"/> No <input type="radio"/> Yes
Weight	<input type="text" value="50 Kg"/>	Result/Risk Level of Skin Cancer	<input type="text" value="Low Risk"/>
Height	<input type="text" value="1.54 Meter"/>	Press Submit Button for Result	<input type="button" value="Submit"/>
Suggestions	<p>Hi Kona Akter Mou. 1. Plz take green foods. 2. Protect Skin from Sunlight.</p> <p>3. Please take care of Weight. Your BMI is Ok. 4. Don't use chemical without doctor advice.</p> <p>4. Ok. Keep free from Unhealthy Industries.</p>		

Fig 2: Our implemented software using Lotus Notes

WelCome To EveryBody

Name	<input type="text" value="Kona Akter Mou"/>	Affected Family Members?	<input type="text" value="None"/>
Age	<input type="text" value="45 Years"/>	Working Period in Industries?	<input type="text" value="None"/>
Gender	<input checked="" type="radio"/> Female <input type="radio"/> Others <input type="radio"/> Male	Color of Body/Skin?	<input type="text" value="Swart"/>
Out door Activities?	<input type="text" value="Medium"/>	Childhood Series Experienced?	<input checked="" type="radio"/> No <input type="radio"/> Yes
Eat Green foods/vegetables?	<input checked="" type="radio"/> No <input type="radio"/> Yes	Use chemical for body with low quality or without check?	<input type="text" value="Frequently"/>
Take Radiation Treatment Before?	<input checked="" type="radio"/> No <input type="radio"/> Yes	Affected By any Cancer before?	<input checked="" type="radio"/> No <input type="radio"/> Yes
Weight	<input type="text" value="50 Kg"/>	Result/Risk Level of Skin Cancer	<input type="text" value="Low Risk"/>
Height	<input type="text" value="1.54 Meter"/>	Press Submit Button for Result	<input type="button" value="Submit"/>
Suggestions	<p>Hi Kona Akter Mou. 1. Plz take green foods. 2. Protect Skin from Sunlight.</p> <p>3. Please take care of Weight. Your BMI is Ok. 4. Don't use chemical without doctor advice.</p> <p>4. Ok. Keep free from Unhealthy Industries.</p>		

Fig 3: Skin cancer prediction with Low risk level

Parameters	Weight age	Score
Age	Age ≤ 13	4
	13 < Age ≤ 30	1
	30 < Age ≤ 45	3
	Age > 45	4
Sex	Male	3
	Female	1
Out door activities?	Frequent	3
	Medium	2
	No	1
Affected Family members?	No	1
	Less than 2	3
	More than 2	4
Working period in industries?	No	1
	Less than 5 hours	2
	More than 5 hours	4
	More than 7 hours	5
Colour of Skin?	Black	1
	Swart	2
	White	3
Eat green foods/vegetables?	No	3
	Yes	1
BMI	BMI ≤ 24	1
	24 < BMI ≤ 28	2
	BMI > 28	3
Childhood serious experienced?	No	1
	Yes	3
Use unhealthy chemical?	Frequent	4
	Medium	2
	No	1
Affected any cancer before?	Yes	3
	No	1



WelCome To EveryBody

Name	Ali Heider	Affected Family Members?	None
Age	30 Years	Working Period in Industries?	Less than 7 hour
Gender	Female Male	Color of Body/Skin?	White
Out door Activities?	Medium	Childhood Series Experienced?	No Yes
Eat Green foods/vegetables?	No Yes	Use chemecal for body with low quality or without check?	Medium
Take Radiation Treatment Before?	No Yes	Affected By any Cancer before?	No Yes
Weight	55 Kg.	Result/Risk Level of Skin Cancer	Intermediate Risk
Height	1.57 Meter	Press Submit Button for Result	Submit
Suggestions	<p>Hi Ali Heider. 1. Plz take green foods. 2. Protect Skin from Sunlight</p> <p>3. Please take care of Weight Your BMI is Ok. 4. Don't use chemecal without doctor advice.</p> <p>4. Keep free from Unhealthy Industries and if possible decrease your Working Time</p>		

Fig 4: Skin cancer prediction with Intermediate risk level.

WelCome To EveryBody

Name	Afroza Akter	Affected Family Members?	More than to 2
Age	13 Years	Working Period in Industries?	None
Gender	Female Male	Color of Body/Skin?	White
Out door Activities?	None	Childhood Series Experienced?	No Yes
Eat Green foods/vegetables?	No Yes	Use chemecal for body with low quality or without check?	Frequently
Take Radiation Treatment Before?	No Yes	Affected By any Cancer before?	No Yes
Weight	24 Kg.	Result/Risk Level of Skin Cancer	High Risk
Height	1.01 Meter	Press Submit Button for Result	Submit
Suggestions	<p>Hi Afroza Akter. 1. Plz take green foods. 2. Protect Skin from Sunlight</p> <p>3. Please take care of Weight Your BMI is Ok. 4. Don't use chemecal without doctor advice.</p> <p>4. Ok. Keep free from Unhealthy Industries</p>		

Fig 5: Skin cancer prediction with High risk level

WelCome To EveryBody

Name	Shahinur Islam Sagor	Affected Family Members?	Less than to 2
Age	50 Years	Working Period in Industries?	More than 7 hour
Gender	Female Male	Color of Body/Skin?	Black
Out door Activities?	Frequently	Childhood Series Experienced?	No Yes
Eat Green foods/vegetables?	No Yes	Use chemecal for body with low quality or without check?	None
Take Radiation Treatment Before?	No Yes	Affected By any Cancer before?	No Yes
Weight	72 Kg.	Result/Risk Level of Skin Cancer	Very High Risk
Height	1.68 Meter	Press Submit Button for Result	Submit
Suggestions	<p>Hi Shahinur Islam Sagor. 1. Plz take green foods. 2. Protect Skin from Sunlight</p> <p>3. Please take care of Weight Your BMI is High. 4. Don't use chemecal without doctor advice.</p> <p>4. Keep free from Unhealthy Industries and Immediately decrease your Working Time &amp; meet with doctor</p>		

Fig 6: Skin cancer prediction with Very High risk level

### 4.3 Comparison with Existing System

The existing paper published in 2000 from American Family Physician [15] use 13 risk factors but this paper used only 11 risk factors shown at Table 2 through those easily detect skin cancer risk level. Here MAFIA algorithm is used to pick up significant pattern shown at Table 1 that reduces time and memory compare to others. And this implemented software shown at Figure. 2 is very time saveable as well as cost saveable compare to others. Poor person can easily check his/her risk level. This paper mainly follows the diabetic prediction system [16] [17] [18] and provide Skin Cancer prediction system. This implemented software is very suitable for Bangladeshi people skin cancer risk find out as well as others people.

## 5. CONCLUSION

Large numbers of people in the Bangladesh and the world have Skin cancer because of his/her unconsciousness about Skin cancer as well as risk factors of Skin cancer. Many of them are uneducated and poor. They can't go doctor and do must out door activities lack of money. Most of them do not even know they have Skin cancer. So the ability to predict Skin cancer with minimum cost plays an important role in the diagnosis process. In this paper is proposed an effective Skin cancer prediction system based on data mining and provided an efficient approach for the extraction of significant pattern from data warehouse for efficient prediction of Skin cancer. The proposed method is implemented using Lotus Notes. The proposed method can efficiently and successfully predict the Skin cancer. And implemented software will be provided through online so that any person can easily check their Skin cancer risk level.

## 6. REFERENCES

- [1] U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, National Institutes of Health, NIH Publication No.10-7625, 2010, pp.1-55.
- [2] National Library of Australia Cataloguing-in-Publication data: Lifestyle and cancer: knowledge, attitudes and behavior in NSW 2009 SHPN (CI) 120203 ISBN 978-1-74187-760-1, Published by the Cancer Institute NSW, 2012, pp.1-29.
- [3] Frawley and Piatetsky-Shapiro, Knowledge Discovery in Databases: An Overview. The AAAI/MIT Press, Menlo Park, C.A, 1996.
- [4] Hian Chye Koh and Gerald Tan, "Data Mining Applications in Healthcare", Journal of Healthcare Information Management, Vol. 19, No. 2, pp. 64-72.
- [5] Dr. Ilias Petrolis and Quan Nguyen, "Association rule tool an implementation of AprioriTID Algorithm", ID 2429851.
- [6] Jiang Su and Harry Zhang, "A Fast Decision Tree Learning Algorithm", American Association for Artificial Intelligence, 2006
- [7] Douglas Burdick, Manuel Calimlim and Johannes Gehrke, "MAFIA: A Maximal Frequent Itemset Algorithm for Transactional Databases", Proceedings of the 17th International Conference on Data Engineering, pp.443-452, April 02-06, 2001.
- [8] "Skin Cancer Prevention and Early Detection", Published by American cancer society 1-800-ACS-2345, 2012, pp.1-16.
- [9] Zakaria Nour, Berna Sayrac, Benoît Fourestié, Walid Tabbara, and Françoise Brouaye, "Generalization Capabilities Enhancement of a Learning System by Fuzzy Space Clustering," Journal of Communications, Vol. 2, No. 6, pp. 30-37, November 2007.
- [10] Shantakumar B.Patil and Y.S.Kumaraswamy, "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network", European Journal of Scientific Research ISSN 1450-216X, Vol.31, No.4 , pp.642-656, Inc. 2009.
- [11] C. Ordonez, "Programming the K-Means Clustering Algorithm in SQL," Proc. ACM Int'l Conf. Knowledge Discovery and Data Mining, pp. 823-828, 2004.
- [12] Eric Li and Li Liu, Optimization of Frequent Itemset Mining on Multiple-Core Processor, pp. 1275-1285, 2007.
- [13] Doug Burdick, Manuel Calimlim, Jason Flannick, Johannes Gehrke and Tomi Yiu, MAFIA: A Performance Study of Mining Maximal Frequent Itemsets, Proceedings of the 17<sup>th</sup> International Conference on Data Engineering, pp.443-452, April 02-06, 2001.
- [14] Shantakumar B.Patil and Y.S.Kumaraswamy, "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network", European Journal of Scientific Research ISSN 1450-216X, Vol.31, No.4 , pp.642-656, Inc. 2009.
- [15] Anthony F. Jetant, Jennifer T. Johnson, Catherine Demastes Sheridan and Timothy J. Caffrey, "Early Detection and Treatment of Skin Cancer", American Family Physician, Vol. 62, No. 2, pp. 357-368, July 15, 2000.
- [16] Muhammad Akmal Sapon , Khadijah Ismail and Suehazlyn Zainudin , Prediction of Diabetes by using Artificial Neural Network, 2011 International Conference on Circuits, System and Simulation IPCSIT, Vol.7, pp. 299-303, Singapore, 2011.
- [17] Huy Nguyen Anh Pham and Evangelos Triantaphyllou, Prediction of Diabetes by Employing a New Data Mining Approach Which Balances Fitting and Generalization , R. Lee and H.-K. Kim (Eds.): Computer and Information Science, SCI 131, pp. 11–26, 2008.
- [18] Manaswini Pradhan and Dr. Ranjit Kumar Sahu, "Predict the onset of diabetes disease using Artificial Neural Network (ANN)", International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004), pp.303-311 Volume 2, Issue 2, April 2011.