

# The use of Decision Tree in Breast Cancer-Related Research: a Scoping Analysis Based on Scopus-Indexed Articles

Iffah Syafiqah Meor Badi'auzzaman, Moey Soo Foon, Mohd. Zulfaezal Che Azemin, Mohd. Izzuddin Mohd. Tamrin

**Abstract:** Breast cancer is the leading cancer that occurs in women globally. The use of machine learning has been introduced to supplement the work in breast cancer studies. There are undisputed pieces of evidence of the existence of publications pertaining to the use of decision tree in breast cancer-related research. However, little is known regarding the types and frequencies of the searched articles. The main objective of this paper is to unearth the broad variety of articles related to breast cancer research that utilized decision trees. The Scopus database was chosen to examine the trend, frequencies and themes of the related publications from the year 2013 until 2018. The study was also intended to disclose the categories of articles based on the areas of breast cancer that have employed the decision trees method. A total of 259 articles from Scopus database were found to meet the inclusion criteria. The analysis of the frequency of published articles generally shows an upward trend. The majority of articles targeted diagnosis of breast cancer (37.8%) in comparisons with other categories. Even though the number of articles found is adequate, several categories of breast cancer are lacking in publications specifically the survivability, incidence, and recurrence of breast cancer among patients. There is a need to redirect the focus of breast cancer research on these categories for future efforts.

**Index Terms:** Breast Cancer, Mammography, Machine Learning, Decision Tree

## I. INTRODUCTION

Breast cancer is the leading cancer that occurs in women globally. The disease takes place when the breast cells undergo an abnormal growth that eventually leads to the state of malignancy. Breast cancer constitutes 11.6% of the total global burden of cancer incidence [1]. The American Institute for Cancer Research revealed that in 2018 alone, over 2

### Revised Manuscript Received on July 12, 2019.

**Iffah Syafiqah Meor Badi'auzzaman**, Kulliyah of Allied Health Sciences, International Islamic University, Kuantan, Malaysia.

**Moey Soo Foon**, Kulliyah of Allied Health Sciences, International Islamic University, Kuantan, Malaysia.

**Mohd. Zulfaezal Bin Che Azemin**, Kulliyah of Allied Health Sciences, International Islamic University Kuantan, Malaysia.

**Mohd. Izzuddin Bin Mohd. Tamrin**, Kulliyah of Information and Communication Technology, International Islamic University, Gombak, Malaysia.

million new cases of breast cancer have been discovered. The data from GLOBOCAN 2018 affirmed that nearly one in 4 of new cases of cancer detected in women across the world meet the criteria of breast cancer [2]. The disease is also listed in the top five of the largest contributing factor of deaths due to cancer. The financial expenditures for cancer care are significantly high to the patients and their family, as well as the nation as a whole. The National Cancer Institute reported that the economic costs accounted for the care of cancer survivors in the United States are 19.3 billion in 2017 and the amount is predicted to gradually increase over the years [3].

To find the remedies towards these issues, a greater effort in conducting breast cancer research will be a golden rule to understand, prevent, treat and in due course eradicate the disease. Breast cancer researchers are striving to not only finding a cure for breast cancer. Instead, attention has also been given towards early detection as well as predicting the risk of women getting the disease. With respect to this matter, the National Breast Cancer Foundation has played its role by funding a number of research concerning breast cancer [4]. The research involves the study on the effect of lifestyle on having the risk of breast cancer, the development of more effective treatment in regards with the breast cancer risk, in addition, to seek for better knowledge of the genetic and non-genetic factors that contribute to breast cancer.

The imminence of technology has caused an impact in revolutionizing the healthcare domain, and indubitably the areas of breast cancer. The present author, acknowledges the sufficient use of machine learning to supplement the work in breast cancer studies. Machine learning is a data analysis method which entails the application of artificial intelligence (AI) to provide the systems the ability to “learn from data, identify patterns and make decisions with minimal human intervention” [5]. The use of machine learning that utilizes data mining techniques alongside with algorithms is also for the purpose of building analytical model from the dataset in order to predict future outcomes [6]. Decision tree is one of the methods in machine learning where the likely outcomes of decisions or choices from the predictions are portrayed in the form of tree representation. In breast cancer-related research, decision tree is surging as a widely held machine learning tool to predict the outcome of the disease from the patterns and relationships in a dataset such as in diagnosis [7] and prognosis [8] of breast cancer.

There are undisputed shreds of evidence of the existence of publications pertaining to the use of decision tree in breast cancer-related research. A scoping review on related literature that would be the highlight in the study is conducted on a scientific database named Scopus since it is the largest abstract and citation database to seek for quality-assessed peer-reviewed academic publications [9]. Scopus is also known to index multidisciplinary scientific literature which cover the fields of science, technology, medicine, social sciences, and arts and humanities. This study presents the status of the decision tree algorithm presented in journal articles indexed in Scopus as an indicator of the extent of its usage in the research involving breast cancer.

**II. RESEARCH OBJECTIVES**

The main objective of this paper is to unearth the availability of numerous articles related to breast cancer research that utilized decision trees. This is parallel with the objective of a scoping review which is “to map the body of literature on a topic area” [10]. The study examined the trend, frequencies and themes of the articles pertaining to the subject matter in the Scopus database over 5-years period. A scoping review was also conducted in this study since it “can be of particular use when the topic has not yet been extensively reviewed or is of a complex or heterogeneous nature” [11]. Essentially, the study was intended to categorize the articles based on the areas of breast cancer that have employed the decision trees method. The categorization of breast cancer-related areas as mentioned was made on articles that are available in the Scopus indexed journal.

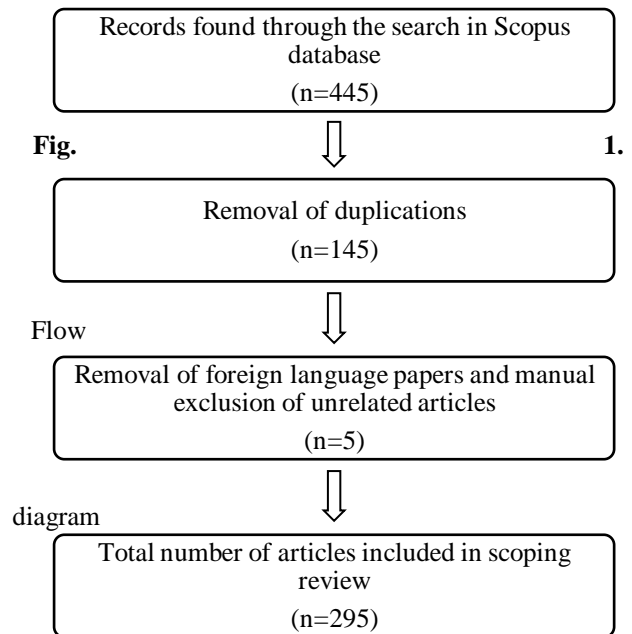
**III. METHODOLOGY**

This study examines research related to the use of decision tree in the area of breast cancer conducted from the year 2013 until 2018. The cited works were considered to include articles that are less than five years old to be on a par with the recent knowledge in science and technology. The inclusion criteria applied to the literature search were the availability of the articles that involve the use of decision tree in the aspect of breast cancer research. Articles were selected should the full text or at least their abstracts were available in English. The inclusion criteria were not highly restrictive for the aim of emphasizing the numbers and variation of topics of the searched articles concerning the interest of this study. The Scopus database <http://www.scopus.com/> was chosen to search for relevant research papers. Scopus allows the users to do basic and advanced search using the fill-in and drop-down boxes with additional filters to search for various fields by using keyword. In this work, the Boolean operator “AND” was used and the terms ‘Decision Tree’ and ‘Breast Cancer’ were individually entered into the database to conduct search in the title, abstract and keywords of the articles. The filtering function is applied to select only journal articles which thereby exclude the books and conference proceedings. The end results of the search in Scopus was exported into CSV (Comma Separated Value) format which therefore allows the data to be further analyzed in form of Excel document. The resultant data consists of information pertaining to the searched publications namely the authors,

titles of articles, year of publication, journal title, abstract and keywords. The data was then filtered for:

1. The number of articles present in each year
2. The number of publication sites of the articles
3. The number of articles related to each category.

The process of literature search is summarized in the flowchart (Fig. 1).



summarizes the selection of related articles.

**IV. RESULTS**

The search using the selected terms using the Scopus search engine had identified 445 articles that comprised the terms ‘Decision Tree’ and ‘Breast Cancer’ either in the title, abstract or keywords.

**A. Subject Area**

Each article are catalogued to specific subject areas. Table 1 presents the articles distribution per subject areas. The highest articles that were recorded are from the field of Medicine (137, 30.8%). Other subject areas that followed are Biochemistry, Genetics and Molecular Biology (86, 19.3%), Computer Science (86, 19.3%) and Engineering (48, 10.8%).

**Table 1**

The number of publications that were identified in Scopus per subject areas.

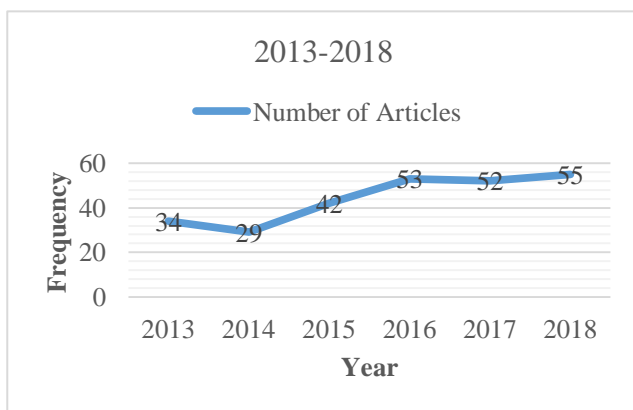
Subject Areas	Number Of Papers
Medicine	137
Biochemistry, Genetics and Molecular Biology	86
Computer Science	86
Engineering	48
Mathematics	24
Health profession	17
Materials Science	13



Chemical Engineering	12
Agricultural and Biological Sciences	11
Pharmacology, Toxicology and Pharmaceutics	11
<b>Total</b>	<b>445</b>

## B. Frequency

The above results were filtered to remove duplication in publications. The list of articles was further screened one by one to ensure the finalized data only entailed the articles pertaining to the use of decision tree in the breast cancer-related areas. This step was taken to exclude the articles in which their title or abstract were not about breast cancer such as breastfeeding, Diameter-at-breast heights of tree, etc.). Out of the 445 searched articles, 295 were the articles that consist of the subject matters specifically.



**Fig. 2.** The trend in journal articles in Scopus regarding the use of decision tree in the breast cancer-related research in 2013 till 2018.

The analysis of the frequency of published articles generally shows an upward trend between 2014 and 2018 as depicted in Fig. 2. In 2014, there is a considerable drop in the number of publications pertaining to the use of decision tree in the breast cancer-related areas. With regards to this matter, the researcher could not offer any deduction to enlighten the scenario at this moment. A significant increase in this work had occurred in the year 2015 and 2016. The highest number of articles was published during 2018 and this rising trend is expected to continue with more articles are anticipated to be published in the upcoming years.

## C. Publication Site

The studies, although being dispersed across 182 journals, had appeared more frequently in journals such as Plastic and Reconstructive Surgery (8, 4.40%), PLoS ONE (7, 3.85%), Computers in Biology and Medicine (7, 3.85%), Journal of Biomedical Informatics (6, 3.30%), Computer Methods and Programs in Biomedicine (5, 2.75%) and Neural Computing and Applications (5, 2.75%). This indicates that the research utilizing decision tree is wide-ranging in terms of the field of journals contributing to the articles related to breast cancer research.

## D. Categories

The results of the literature search are categorized by taking into account the main scope of breast cancer research that have been focused in the articles. Despite much of the included literature encompasses a number of themes, the best efforts have been made to categorize them according to the subject matter focused on most comprehensively. The analysis of the articles' title and abstract to determine the purpose of decision tree usage in breast cancer-related areas has brought towards several categories of articles. These results are outlined in Table 2.

### Diagnosis of Breast Cancer

There is a presence of almost a hundred articles (98/259) related to the use of decision trees in the diagnosis of breast cancer. The amount of related articles found in this category is nearly 40 percent of the total searched articles in Scopus. This denotes the breast cancer diagnosis as the most predominant area that uses decision trees in comparisons with other categories. Based on the author's observation, the use of decision trees for breast cancer diagnosis was acknowledged primarily in the field of medical imaging and biomedicine or bioinformatics. Therefore, the search is limited towards several input terms related to medical imaging field such as "Mammography", "Mammogram", "Ultrasound", "Ultrasonography", "MRI" and "CT".

The prominent use of decision tree was to assist the detection of masses in mammography images as presented in the articles by Sajan and Germanus [7] and Gupta, Banerjee and Rubin [12]. The use of decision tree in Magnetic Resonance Imaging (MRI) in breast imaging can be seen in a publication by Sun et al. [13]. The use of machine learning in breast ultrasound to detect breast cancer based on BI-RADS Features is also evident in several papers by Rodríguez-Cristerna, Gómez-Flores and de Albuquerque Pereira [14] and Shan, Alam, Garra, Zhang and Ahmed [15]. An intriguing finding is the existence of research that utilized decision tree in the classification of thermographic images to indicate breast cancer patients [16], [17]. In comparison with the traditional way of detecting breast mass visually, the decision tree has been utilized specifically in mammography, MRI, ultrasound and thermography to detect abnormalities in the medical images.

The use of decision tree for breast cancer detection is also prevalent in molecular diagnostic tests which is related to the field of biomedical or bioinformatics. The usage of DNA alongside with decision tree is recognized in several articles that document the use of DNA sequences [18] and plasma DNA [19] for cancer detection. The profiling of gene expression using decision tree to search the gene signature of breast cancer cell is depicted in a number of studies where microRNA [20] and RNA-Seq datasets [21], [22] are used in cancer classifiers. The potential of using protein amino acid features for classification of breast cancer in cancer detection is presented in publications by Ali and Majid [23] and Ali, Majid, Javed and Sattar [24].

### Classification of Breast Lesions

The categorization of articles had yielded 48 articles that are relevant to the use of decision tree to classify breast lesion into benign or malignant. Again, limiting the search towards several input term related to medical imaging field such as “Mammography”, “Mammogram”, “Ultrasound”, “Ultrasonography”, “MRI” and “CT” results in 20 articles related to the use of decision tree in the field. The findings signify that almost half of the articles in this category (20/48) are also predominantly related to medical imaging as portrayed in the first category mentioned before.

An interesting finding is that several techniques were utilized to extract features from the medical images for the classification input in the decision tree. The use of decision trees was largely seen in mammography as identified in 12 articles. Several publications have used textural-based feature from mammography images as the input in decision tree classifier to identify and characterize the abnormalities in breast tissue [25], [26]. Vadivel and Surendiran [27] proposed a new method using geometrical shape and margin features to distinguish breast mass lesion based on the regularity or irregularity of its shape in mammogram. For ultrasound-based classification, Uniyal et al. [28] had performed the analysis on ultrasound radio frequency (RF) time series on breast lesions to estimate the likelihood of malignancy. Another study on ultrasound depicted the use of various quantitative parameters for classification purposes in decision tree to properly discriminate benign and malignant breast lesions [29]. In Computed Tomography (CT) area, Hammer and Mortani Barbosa [30] had presented an evidence-based decision algorithm for management of patients with breast cancer and incidental pulmonary nodules. The algorithm permits several factors such as number, size and shape of nodules to be taken into account to predict the malignancy of the nodules which are commonly visualized on CT among breast cancer patients.

Besides medical imaging, the assessment of the type of lesion in breast tissue can also be done in a histopathological manner by examining the tissue under the microscope. This method appears in several publications with the main purpose to determine the characterization of lesion into benign or carcinoma as well as to identify their subtypes by using machine learning [31], [32]. A number of studies related to histopathological grading of breast cancer which represents the grade of tumor based on their morphological assessment were presented by Dalton [33], Chang, Kuo, Wu, Huang and Chen [34] and Paul and Mukherjee [35]. These findings signify the role of microscopic examination in classifying breast lesion by using decision tree beside the sole utilization by standard diagnostic imaging procedure.

### *Treatment of Breast Cancer*

A considerable number of journal articles had captured the role of decision tree in the treatment decision for breast cancer (36/259). Vos et al. [36] in their paper had shown the use of decision tree in guiding breast cancer patients to opt between two different surgical procedures as the treatment options. Catanuto et al. [37] in his review had discussed the role of decision trees in assisting the decisional process for primary breast surgery. Zolfagharnasab et al. [38] had

developed a regression model using decision tree to predict the deformation of breast shape after surgery. The usage of decision tree is evident in an article by Hanna et al. [39] to evaluate the benefit of radiotherapy. The development of radiotherapy demand models by using a mathematical modelling software that uses decision tree called the Malthus programme had been depicted in several studies as well [40], [41], [42]. In drug discovery, a few of studies had evaluated the best model ensemble by using decision trees during the breast cancer drug design stage in order to overcome the drug-drug interactions issues [43], [44]. Decision trees have been utilized in the prediction of aromatase inhibitors (AIs) which contributes for effective treatment of breast cancer as portrayed in several different studies [45-47].

### *Cost-effectiveness*

The articles on decision tree were also affiliated with the cost-effectiveness analysis to investigate the effect of health care actions based on their relative costs and health outcomes (37/259). In Thailand, a cost-utility study was done by using decision tree to evaluate the sentinel lymph node biopsy (SNB) as the alternative procedure to the axillary lymph node dissection (ALND) for the treatment of breast cancer [48]. The Society of Surgical Oncology undertook a cost-effectiveness analysis on the different type of mastectomy procedures to compare their expenses in the long run [49]. The Journal of Plastic and Reconstructive Surgery has presented an article related to cost-effectiveness evaluation in five standardized methods for breast reconstruction using decision tree [50]. In Colombia, Perea and Rosselli [51] have presented a comparative study in comparing between immediate or delayed breast reconstructions using decision tree model to verify the best breast cancer treatment strategy in term of cost-effectiveness. For breast cancer medication, a cost-utility analysis using decision tree was performed to allow the comparison between two chemotherapy drugs [52]. In breast cancer screening procedure, a paper by Sato et al. [53] had considered the most cost-effective method for breast cancer detection either by performing double reading or the standard double reading by two radiologists. This was followed by another publication that focuses on breast cancer screening where the cost-effectiveness of tomosynthesis as the additional test in annual screening routine beside mammography was examined using decision tree analysis [54]. The results suggest that there is a substantial role of decision tree for cost-effectiveness analysis in the aspect of breast screening and treatment.

### *Survivability of Breast Cancer Patients*

The analysis of the use of decision tree in terms of breast cancer-related publications contributed to 15 out of 259 articles. A review by Kate and Nadig [55] had unearthed that the survival rate of breast cancer patient was different across each stage of breast cancer. In the journal of BMC Bioinformatics, an article had depicted the use of decision tree algorithm in the prediction of cancer survival outcome by using data from gene expression [56]. In more specific content of the paper, the researchers had shown that decision trees can support a platform called PCM-SABRE

(Precision Cancer Medicine - Survival Analysis  
Benchmarking, Reporting and Evaluation) to help in  
producing predictive models of breast cancer recurrence.  
Different data mining

**Table 2**

Summary on the thematic analysis of the usage of decision tree in breast cancer research.

Categories	Scope	Subcategories	Number of Papers
Diagnosis of Breast Cancer	Identify the presence of breast cancer in patients.	Mammography [7],[12] MRI [13] Ultrasound [14],[15] Thermography [16], [17] DNA sequences [18] Plasma DNA [19] MicroRNA [20] RNA-Seq [21], [22] Amino acid features [23], [24]	98
Classification of Breast Lesions	Characterize the lesion into benign or malignant.	Mammography [25], [26], [27] Ultrasound [28], [29] CT [30] Histopathology [31], [32] Carcinoma subtypes [31], [32] Histopathological grading [33],[34],[35]	49
Treatment of Breast Cancer	Involve the treatment related to breast cancer such as radiation therapy, surgery, chemotherapy drug study etc. as well as postsurgical management such as breast reconstruction.	Surgical procedures [36], [37], [38] Radiotherapy [39], [40], [41], [42] Drug-drug interactions [43], [44] Aromatase inhibitors [45], [46], [47]	36
Cost-Effectiveness	Cost-utility analysis such as in screening, treatment, breast reconstruction surgery etc.	Surgical procedures [48], [49] Breast reconstruction [50], [51] Chemotherapy drugs [52] Breast cancer screening [53], [54]	37
Survivability of Breast Cancer Patients	Predict the life expectancy of breast cancer patients.	Stage of breast cancer [55] Gene expression [56]	15
Incidence of Breast Cancer	Predict the possibility of the occurrence of the disease	Hereditary factors [63], [66], [67] Genomic data [64] Non-genetic factors [65], [66], [67]	11
Recurrence of Breast Cancer	Predict the likelihood of relapse of the disease.	Feature selection [59], [60] Regulatory mutations [61] Microarray databases [62]	8
Others			5
<b>TOTAL</b>			<b>259</b>

strategies were utilized in a document to construct a classification model for breast cancer survival patterns [57]. They tested three methods to establish the model namely Support Vector Machine, Logistic Regression and Decision Tree to find the best approach that will contribute as a reference in decision-making frame. In 2018, a group of different authors carried out the same method by comparing decision tree and logistic regression for survival prediction of breast cancer patients [58].

#### *Recurrence of Breast Cancer*

An observation to be underlined is the presence of a minority number of articles (8/259) regarding the usage of decision tree for prediction of breast cancer recurrence. Early recurrence prediction is significant for breast cancer patients in allowing them to receive treatment earlier. The accessibility of the vast amount of data has allowed different feature selection technique to be used in decision tree to predict breast cancer recurrence. A group of researchers has

utilized Particle Swarm Optimization (PSO) to assist with feature selection in predicting 5-year breast cancer recurrence [59]. The same feature selection method was also used by different authors for the inputs in decision tree [60]. In the BMC Bioinformatics journal, a publication had evaluated different induction strategies for decision trees called the random forest on a breast cancer model to predict the probability of regulatory mutations to recur [61]. In the same journal, another paper by Chou et al. [62] had shown the profiling of gene expression to predict five-year recurrence of breast cancer by using microarray databases. The study had tested three classical approaches in data mining namely artificial neural networks (ANN), decision trees (DT) and logistic regression (LR) as well as two hybrid models (DT-ANN and DT-LR) to compare their performance in making the best prediction.

#### *Incidence of Breast Cancer*

The use of decision tree

was further acknowledged in breast cancer areas where 11 articles had utilized the algorithm in predicting the incidence of breast cancer among women in a population. The decision tree model produced acts as a baseline in risk assessment of breast cancer by considering several factors that could affect the persons to get the disease either by genetic or sporadic cause. This is evidence in a number of articles by several groups of authors that analyze possible interactions of hereditary factors such as family history [63] as well as genomic data [64] with the risk of having breast cancer. The non-genetic factors for example age, the age of first delivery and number of live births were also involved in risk factor analysis that assembled by decision trees to predict the occurrence of breast cancer [65]. The combination of both hereditary and sporadic factors in assessing the breast cancer incidence appears in the publications by Garrisi [66] and Kitsantas and Wu [67] to examine whether these factors independently affect breast cancer incidence. These findings suggest that decision tree is applicable in predicting breast cancer that might occur genetically or non-genetically.

## V. DISCUSSION

### *General findings of the study*

The highest amount of papers constituted by the field of medicine is tantamount to approximately 30 percent of the total number of related articles indexed in Scopus. It can be postulated that a possible factor could be due to the specific nature of the field that involves countless decision-making process. Physicians use decision support systems to help in the decisional process particularly in situations in which decision must be made efficiently and reliably [68]. According to Kaplan and Frosch [69], there are two categories of decisional process in healthcare where the decision is made by healthcare providers on patients' behalf and shared decisions between patients and healthcare providers. Despite some evidence portrays that patients prefer to have their own authority in decisional making, yet other evidence claims that some patient would rather play a passive role. This is presented in a study by Lamanna and Byrne [70] where machine learning is utilized on geriatric and psychiatric patients to estimate the reliability of prediction that the patient will agree with the prescribed treatment.

Another possible factor of the usage of decision tree in medicine is the availability of the large medical databases that store patient information. It is common for patients to register themselves before getting medical consultation which permits the hospital database to store basic information such as their name, age, gender, marital status, blood type, etc. There is also a larger database in the healthcare setting called Hospital Information System (HIS) which is linked to more advanced databases that stores more detailed and classified information of patients such as pathology lab result, medical images and patient tracking system. This stored information in these databases may comprise useful patterns, trends or rules that seemingly hidden and unconnected in the data. This is the point where machine learning algorithms could be harnessed to extract the potential information of the data to be applied in making

medical decisions such as in the process of giving diagnosis or treatment.

Pertaining to the overall number of articles that depict the use of decision tree in research related to breast cancer, the findings are found to be satisfactory in comparison with the total number of journal articles available in Scopus database. A positive finding is that the general observation on the frequency of publications had shown a steady increase throughout the years. This finding is in line with a study done by Kolachalama [71] who reported the growing usage of machine learning based on the trend of publications in PubMed database. The finding in the said study revealed that the number of searched papers related to machine learning, when used as the medical subject heading (MeSH) term, has found to be largely escalating in the journal database since 2010. In the upcoming years, the popularity of machine learning is expected to rise in scientific writings. It is interesting to ponder diverse fields of study that could implement machine learning in the future. This move will increase the appreciation of researchers towards the use of machine learning that later will affect the trend of the publications.

Based on the publication sites identified, a number of journals related to the use of decision tree in breast cancer research remain diversified across several fields namely medicine, sciences, bioinformatics, computer science, etc. The decision tree which is a part of machine learning is originated from the field of computer science. Due to its versatility, it is significant to denote that machine learning has infiltrated various science fields where the data and systems require more than simple statistical analysis [72]. In relation to the fact, the integration of machine learning element can be found in the top journals of science and engineering principally Science, Nature, Cell and PNAS (Proceedings of the National Academy of Sciences of the United States of America).

### *Categories of breast cancer-related research*

The majority of articles related to breast cancer research depict the role of decision tree in diagnosing breast cancer in patients as well as in classifying the lesion into normal or malignant. The field of medical imaging which is a branch in medicine has portrayed the major use of decision tree in detecting breast cancer. Medical imaging is a process of producing medical images of human anatomy to visualize the internal structure for the aim of diagnosis and treatment of patient. In the medical imaging field, several imaging modalities have utilized the decision tree algorithm in depicting abnormalities in the breast tissues. Mammography is the gold standard for breast cancer screening which involves essentially a capture of x-ray images of the breasts. Besides mammography, MRI, ultrasound and thermography are among the complementary imaging techniques to detect lesions in breast. By using specific principles of these technologies, the modality will produce medical images that would be analyzed by radiologists to determine the existence of the disease. Reaching a precise and correct diagnosis is key for the management of patient. Nevertheless, having the analysis of medical images by

merely visual assessment might induce flaws in term of “variability, subjectivity and irreproducibility” [73]. To overcome these shortcomings, Computer Aided Diagnosis (CAD) system is employed alongside with imaging technologies to improve diagnostic accuracy. The CAD system acts as a second set of eyes to check the presence of abnormal areas of mass, density or calcification in medical images by using computer software. Machine learning plays a significant role in the application of CAD system. The use of sophisticated machine learning algorithms such as decision tree would assist in modeling of data based on the cases of breast cancer that would later be used to identify breast lesion as well as distinguish the findings into normal or cancerous lesion.

Besides imaging procedure, biopsy is an extended procedure to analyze the abnormalities in breast cancer tissues under microscopic evaluation. The specimen of breast tissue observed in stained tissue sections will be monitored for cancer parameters that originate from histopathology. The sample may further need other tests namely molecular, protein or genetic test to improve diagnostic capabilities. The potential use of machine learning is highlighted where the decision tree algorithm is used as a classifier for the informative features obtained from computational tools such as feature selection and extraction methods when conducting the test. This approach is of the aim to improve the breast cancer prediction through microscopy procedure.

A remarkable number of breast cancer research has taken advantage of the predictive performance in decision tree to determine the best pathway of treatment for patients. Generally, patients with early-stage breast cancer require two different types of treatment which are local treatment (direct treatment to the affected breast) and systemic treatment (whole-body treatment). Local treatment includes radiotherapy and surgery. For surgical option, the patients might opt for either mastectomy (removal of the whole breast) or lumpectomy (removal of the cancer area along with the margin of normal tissue around it). After the surgical procedure, systemic adjuvant therapy will be given which involve either endocrine therapy, chemotherapy, targeted therapy or radiation therapy [74]. These treatment options are to eliminate the remaining tumor cells that might be responsible for breast cancer recurrence. Despite having the same aim, the treatment options must be considered thoroughly to tailor the best treatment that fits the patients’ needs. This is in light of the fact that each treatment pathways has a different mechanism of action, risk and side effects to each breast cancer patients. This is when decision tree can be implemented in assisting physician to select for the best treatment plan. To guide the choice, decision tree can be employed by using the data based on previous study and literature to build decision models that compare the considered treatment plan. Hence, the best treatment option could be indicated from the model that provide the best treatment outcomes based on patients’ need. Similarly, a document portrayed a systematic review of the use of decision tree in treatment guidelines for radiotherapy in lung cancer [75].

Considering the use of decision tree in deciding the best

treatment plan, the extensive use of the algorithm can be observed in cost-effectiveness analysis of breast cancer treatment. This is in the point of fact that decision tree is a method that assists in providing good options, ultimately for decisions that consist of high risks and costs [76]. Cost-effectiveness incorporates not only the analysis of the expenditures on drugs or treatment but also reflects the cost of opting alternative decision and how it influences the quality of life [77]. With regards to this matter, decision tree is used to guide in indicating best options that consider the cost of treatment in addition to the treatment outcomes. It uses graphical approach to make comparisons of “competing alternatives and assign values to those alternatives by combining uncertainties, costs, and payoffs into specific numerical values” [76]. Similarly, a study presented the cost-effectiveness assessment for the treatment of ulcerative colitis by using decision tree in a clinical trial [78].

The utility of decision tree was evident in journal articles that described breast cancer prognosis. Prognosis of breast cancer is the prediction of the disease’s outcome which includes the life expectancy (survivability) of breast cancer patient and possibilities of recurrence of breast cancer [79]. The prognostic factors of breast cancer are either chronological or biological [80]. The first category is based on the period of the cancer existence (the size of tumor) and the second category is the “indicators of the metastatic potential behavior of a tumor (tumor grade or axillary node status)” (p.137).

There is a quantity of evidence of the researchers’ involvement in studies that emphasize the use of decision tree in predicting the survivability of breast cancer patients. The traditional way in ascertaining the prognostic factors related with the survival time of breast cancer patients is by doing statistical methods. The statistical analysis assists in determining the important predictors with regards to the outcome of patients’ survivability or recurrence of time [58]. However, with the machine learning method, more accurate prognostic factors of breast cancer can be identified for more effective management of cancer. In a study by Momenyan et al. [58], it is found that the use of decision tree is more accurate in predicting the probability of death due to breast cancer. Likewise, the adoption of decision tree in the cases of non-metastatic invasive breast cancer is found to be 70% accurate to predict the early relapse of the disease [81].

On the contrary, the results in this study depicted that the use of decision tree in predicting breast cancer recurrence is not satisfactory. The inadequate amount of research related to breast cancer recurrence, in general, is supported by research by Song et al. [82] which claimed that less attention is given in establishing factors related to the timing of breast cancer recurrence. The study did not have a definite justification concerning this matter. However, it can be deduced from a meta-analysis study which stated that the radiotherapy-induced benefits have significant contribution towards the reduced breast cancer recurrence [83]. Further, the researcher cited that the local recurrence of breast cancer is low in patients that underwent mastectomy with breast reconstruction.





Nevertheless, concerning breast cancer prognosis, the author feels that this area should be gaining more attention since these predictions can aid in designing appropriate treatment on breast cancer patients that are in line with the projected outcomes.

It can be observed that decision tree has been used to investigate the relationship of breast cancer incidence with genetic factors. This is due to the fact that inherited gene mutations contribute to high risk of breast cancer due to family history. It is of the absolute fact that the implication of the family history factor in developing breast cancer upsurges with having a first-degree female relative precisely the mother, daughter or sister that is identified with the disease. The risk is even greater when women inherit a faulty gene specifically the BRCA1 or BRCA2 gene which is linked to the high risk of having breast cancer [84]. Conversely, there is also evidence of the application of decision tree in estimating the risk of having breast cancer due to non-genetic factors. The data from U.S Breast Cancer Statistics highlighted that almost 85% of breast cancer cases affecting women with no prior family history of the disease [85]. This is corresponding to the fact that the several lifestyle habits such as lack of exercise, smoking and the intake of alcohol could contribute to genetic mutations which increase the chance of having cancer. Despite both factors have potential

**VII. CONCLUSION**

This paper presents the state-of-the-art of the utilization of decision tree in breast cancer-related research. The results had unveiled publications from the Scopus-indexed

in causing breast cancer in women, the use of decision tree could help in predicting the possibility of breast cancer occurrence before it happens. This is crucial to provide knowledge to women of their risk and subsequently encourage them in prevention and early detection of breast cancer. In addition, the findings have brought attention to investigate whether there is a relationship between occupational factors and the risk of developing breast cancer. As such, it is interesting that future study should incorporate decision tree in assessing the breast cancer incidence on people that recurrently exposed to radiation namely the radiographers, radiologists, pilots, flight attendants, etc.

**VI. RECOMMENDATIONS**

There is a need to fully take advantage of the decision tree to be applied in breast cancer research. For future research that incorporate the idea, it is proposed that the use of decision tree can be applied into several domains of breast cancer-related research. Based on the latest research in breast cancer, the present author suggested that a number of studies can be conducted by utilizing decision tree as shown in Table 3.

breast cancer areas. The number of articles related to the implementation of decision tree in research is found to be sufficient, yet, attention should be given towards several categories of breast cancer that have inadequacy in publications. It is highly suggestive for future study to

**Table 3**

The possible applications of decision tree in future breast cancer-related research.

Categories	Research related to Breast Cancer	Suggested Application Of Decision Tree In Breast Cancer-Related Research
Diagnosis of Breast Cancer	The use of scintimammography (molecular breast imaging) for finding breast cancer [86]	Applying decision tree method to improve breast cancer diagnosis in molecular breast imaging.
Classification of Breast Lesions	Tests for circulating tumor cells (CTCs) for detecting metastatic breast cancer using sensitive lab tests [86]	Classification of circulating tumor cells (CTCs) by using decision tree to indicate metastatic breast cancer.
Treatment of Breast Cancer	The use of hormone therapy drugs or non-hormonal drugs in lowering breast cancer risk [86]	A comparison of the effectiveness of hormone therapy drugs or non-hormonal drugs in breast cancer reduction using decision tree analysis.
Cost-Effectiveness of Breast Cancer Management	Oncoplastic surgery (combination of techniques for cancer surgery and plastic surgery) [86]	The cost-effectiveness analysis of oncoplastic surgery to standard breast-conserving surgery utilizing decision tree.
Survivability of Breast Cancer Patients	Whole- and Partial-Breast Radiation for prevention the returning of breast cancer after surgery [87]	Utilizing decision tree models for predicting the survivability of breast cancer patients receiving the whole- and partial-breast radiation.
Incidence of Breast Cancer	The relationship between environmental factors and the risk of developing breast cancer [88]	Applying decision trees in determining breast cancer incidence based on environmental factors.
Recurrence of Breast Cancer	Exploring gene variations that contribute to breast cancer recurrence [86]	Predicting breast cancer recurrence based on gene variations data using decision tree.
Prevention of Breast Cancer	The use of topical drugs to help in preventing breast cancer [87]	A decision tree analysis on the patient adherence to topical drugs for breast cancer prevention.
Early Detection of Breast Cancer	The relationship between stage of adopting breast screening with concern of women in developing breast cancer [89]	A decision tree analysis on the relationship of stage of behavioral adoption towards breast screening with early detection of breast cancer.

journal that were classified in seven categories related to

incorporate decision tree in



more areas of breast cancer such as in prevention and early detection of breast cancer.

## ACKNOWLEDGMENT

This research was supported by the Research Management Centre, International Islamic University Malaysia, under Research Initiatives Grant Scheme (RIGS)-Publication (P-RIGS 18-035-0035). The author is grateful to Assoc. Prof. Dr. Ahmad Aidil Arafat Dzulkarnain for the review and valuable suggestions that improved the paper.

## REFERENCES

1. International Agency for Research on Cancer, Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018, 2018.
2. American Institute for Cancer Research, New Global Cancer Data: GLOBOCAN 2018 | UICC, Uicc.Org. (2018).
3. National Cancer Institute, Financial Burden of Cancer Care Cancer Trends Progress Report, Progressreport.Cancer.Gov. (2018).
4. National Breast Cancer Foundation, How research is making a difference to breast cancer risk, (n.d.).
5. Machine Learning: What it is and why it matters, Sas.Com. (2019).
6. Hackartist, What is the difference between data mining, statistics, machine learning and AI? [Msg 2]. Message posted at <https://stats.stackexchange.com/questions/5026/what-is-the-difference-between-data-mining-statistics-machine-learning-and-ai>. (2017).
7. S. K. Sajan, M. Germanus Alex, A Comparative Study to Evaluate the Performance of Classification Algorithms in Mammogram Analysis, International Journal Of Engineering & Technology. 7 (2018) 154.
8. H. Shin, Y. Nam, A coupling approach of a predictor and a descriptor for breast cancer prognosis, BMC Medical Genomics. 7 (2014) S4.
9. D. Rew, An Introduction to the Scopus Content Selection and Advisory Board. (2015).
10. H. Arksey, L. O'Malley, Scoping studies: towards a methodological framework, International Journal of Social Research Methodology. 8 (2005) 19-32.
11. N. Mays, E. Roberts, J. Popay, Synthesizing research evidence. In: Fulop N, Allen P, Clarke A, Black N, editors. Studying the Organisation and Delivery of Health Services: Research methods. London: Routledge; pp. 188-219. (2001).
12. A. Gupta, I. Banerjee, D. Rubin, Automatic information extraction from unstructured mammography reports using distributed semantics, Journal of Biomedical Informatics. 78 (2018) 78-86.
13. L. Sun, J. He, X. Yin, Y. Zhang, J. Chen, T. Kron et al., An image segmentation framework for extracting tumors from breast magnetic resonance images, Journal of Innovative Optical Health Sciences. 11 (2018) 1850014.
14. A. Rodríguez-Cristerna, W. Gómez-Flores, W. de Albuquerque Pereira, A computer-aided diagnosis system for breast ultrasound based on weighted BI-RADS classes, Computer Methods and Programs in Biomedicine. 153 (2018) 33-40.
15. J. Shan, S. Alam, B. Garra, Y. Zhang, T. Ahmed, Computer-Aided Diagnosis for Breast Ultrasound Using Computerized BI-RADS Features and Machine Learning Methods, Ultrasound In Medicine & Biology. 42 (2016) 980-988.
16. L. Silva, A. Santos, R. Bravo, A. Silva, D. Muchaluat-Saade, A. Conci, Hybrid analysis for indicating patients with breast cancer using temperature time series. Computer Methods and Programs in Biomedicine, 130, (2016).142-153.
17. M. Santana, J. Pereira, F. Silva, N. Lima, F. Sousa, G. Arruda et al., Breast cancer diagnosis based on mammary thermography and extreme learning machines, Research On Biomedical Engineering. 34 (2018) 45-53.
18. Q. Liu, M. Gan, R. Jiang, A sequence-based method to predict the impact of regulatory variants using random forest, BMC Systems Biology. 11 (2017).
19. S. Wang, Y. Cai, Identification of the functional alteration signatures across different cancer types with support vector machine and feature analysis, Biochimica Et Biophysica Acta (BBA) - Molecular Basis Of Disease. 1864 (2018) 2218-2227.
20. A.G.Floares, G.A.Calin, F.B. Manolache, Bigger Data Is Better for Molecular Diagnosis Tests Based on Decision Trees. In: Tan Y., Shi Y. (eds) Data Mining and Big Data. DMBD 2016. Lecture Notes in Computer Science, vol 9714. Springer, Cham. (2016).
21. D. Castillo, J. Gálvez, L. Herrera, B. Román, F. Rojas, I. Rojas, Integration of RNA-Seq data with heterogeneous microarray data for breast cancer profiling, BMC Bioinformatics. 18 (2017).
22. S. Kim, H. Lee, M. Kon, Comparisons of cancer classifiers based on RNA\_seq and miRNA\_seq, International Journal of Data Mining and Bioinformatics. 17 (2017) 359.
23. S. Ali, A. Majid, Can-Evo-Ens: Classifier stacking based evolutionary ensemble system for prediction of human breast cancer using amino acid sequences, Journal Of Biomedical Informatics. 54 (2015) 256-269.
24. S. Ali, A. Majid, S. Javed, M. Sattar, Can-CSC-GBE: Developing Cost-sensitive Classifier with Gentleboost Ensemble for breast cancer classification using protein amino acids and imbalanced data, Computers in Biology and Medicine. 73 (2016) 38-46.
25. A. Mohanty, M. Senapati, S. Beberta, S. Lenka, Texture-based features for classification of mammograms using decision tree, Neural Computing And Applications. 23 (2012) 1011-1017.
26. A. Midya, R. Rabidas, A. Sadhu, J. Chakraborty, Edge Weighted Local Texture Features for the Categorization of Mammographic Masses, Journal Of Medical And Biological Engineering. 38 (2017) 457-468.
27. A. Vadivel, B. Surendiran, A fuzzy rule-based approach for characterization of mammogram masses into BI-RADS shape categories, Computers in Biology and Medicine. 43 (2013) 259-267.
28. N. Uniyal, H. Eskandari, P. Abolmaesumi, S. Sojoudi, P. Gordon, L. Warren et al., Ultrasound RF Time Series for Classification of Breast Lesions, IEEE Transactions On Medical Imaging. 34 (2015) 652-661.
29. P. Kapetas, R. Woitek, P. Clauser, M. Bernathova, K. Pinker, T. Helbich et al., A Simple Ultrasound Based Classification Algorithm Allows Differentiation of Benign from Malignant Breast Lesions by Using Only Quantitative Parameters, Molecular Imaging And Biology. 20 (2019) 1053-1060.
30. M. Hammer, E. Mortani Barbosa, Predictive factors for malignancy in incidental pulmonary nodules detected in breast cancer patients at baseline CT, European Radiology. 27 (2016) 2802-2809.
31. Y. Li, X. Tang, Z. Bai, X. Dai, Exploring the intrinsic differences among breast tumor subtypes defined using immunohistochemistry markers based on the decision tree, Scientific Reports. 6 (2016).
32. Z. Gandomkar, P. Brennan, C. Mello-Thoms, MuDeRN: Multi-category classification of breast histopathological image using deep residual networks, Artificial Intelligence in Medicine. 88 (2018) 14-24.
33. L. Dalton, Invasive breast cancer: stratification of histological grade by gene-based assays: a still relevant example from an older data set, Histopathology. 65 (2014) 429-433.
34. C. Chang, S. Kuo, H. Wu, Y. Huang, D. Chen, Stellate Masses and Histologic Grades in Breast Cancer, Ultrasound In Medicine & Biology. 40 (2014) 904-916.
35. A. Paul, D. Mukherjee, Mitosis Detection for Invasive Breast Cancer Grading in Histopathological Images, IEEE Transactions on Image Processing. 24 (2015) 4041-4054.
36. E. Vos, L. Koppert, W. van Lankeren, C. Verhoef, B. Koerkamp, M. Hunink, A preliminary prediction model for potentially guiding patient choices between breast conserving surgery and mastectomy in early breast cancer patients; a Dutch experience, Quality Of Life Research. 27 (2017) 545-553.
37. G. Catanuto, F. Pappalardo, N. Rocco, M. Leotta, V. Ursino, P. Chiodini et al., Formal analysis of the surgical pathway and development of a new software tool to assist surgeons in the decision making in primary breast surgery, The Breast. 29 (2016) 74-81.
38. H. Zolfagharnasab, S. Bessa, S. Oliveira, P. Faria, J. Teixeira, J. Cardoso et al., A Regression Model for Predicting Shape Deformation after Breast Conserving Surgery, Sensors. 18 (2018) 167.
39. T. Hanna, J. Shafiq, G. Delaney, S. Vinod, S. Thompson, M. Barton, The

- population benefit of evidence-based radiotherapy: 5-Year local control and overall survival benefits, *Radiotherapy And Oncology*. 126 (2018) 191-197.
40. C. Round, T. Mee, N. Kirkby, T. Cooper, M. Williams, R. Jena, The Malthus Programme: Developing Radiotherapy Demand Models for Breast and Prostate Cancer at the Local, Regional and National Level, *Clinical Oncology*. 25 (2013) 538-545.
  41. J. Borrás, Y. Lievens, P. Dunscombe, M. Coffey, J. Malicki, J. Corral et al., The optimal utilization proportion of external beam radiotherapy in European countries: An ESTRO-HERO analysis, *Radiotherapy And Oncology*. 116 (2015) 38-44.
  42. R. Jena, T. Mee, N. Kirkby, M. Williams, Quantifying Uncertainty in Radiotherapy Demand at the Local and National Level using the Malthus Model, *Clinical Oncology*. 27 (2015) 92-98.
  43. M. Gantner, R. Peroni, J. Morales, M. Villalba, M. Ruiz, A. Talevi, Development and Validation of a Computational Model Ensemble for the Early Detection of BCRP/ABCG2 Substrates during the Drug Design Stage, *Journal Of Chemical Information And Modeling*. 57 (2017) 1868-1880.
  44. N. Aniceto, A. Freitas, A. Bender, T. Ghafourian, Simultaneous Prediction of four ATP-binding Cassette Transporters' Substrates Using Multi-label QSAR, *Molecular Informatics*. 35 (2016) 514-528.
  45. C. Nantasenamat, H. Li, P. Mandi, A. Worachartcheewan, T. Monnor, C. Isarankura-Na-Ayudhya et al., Exploring the chemical space of aromatase inhibitors, *Molecular Diversity*. 17 (2013) 661-677.
  46. C. Nantasenamat, A. Worachartcheewan, P. Mandi, T. Monnor, C. Isarankura-Na-Ayudhya, V. Prachayasittikul, QSAR modeling of aromatase inhibition by flavonoids using machine learning approaches, *Chemical Papers*. 68 (2014).
  47. W. Shoombuatong, V. Prachayasittikul, V. Prachayasittikul, C. Nantasenamat, Prediction of aromatase inhibitory activity using the efficient linear method (ELM), *EXCLI journal*, 14 (2015) 452-64.
  48. D. Songtish, N. Praditsithikorn, Y. Teerawattananon, A Cost-Utility Analysis Comparing Standard Axillary Lymph Node Dissection with Sentinel Lymph Node Biopsy in Patients with Early Stage Breast Cancer in Thailand, *Value In Health Regional Issues*. 3 (2014) 59-66.
  49. R. Keskey, A. LaJoie, B. Sutton, I. Kim, W. Cheadle, K. McMasters et al., Cost-effectiveness Analysis of Contralateral Prophylactic Mastectomy Compared to Unilateral Mastectomy with Routine Surveillance for Unilateral, Sporadic Breast Cancer, *Annals Of Surgical Oncology*. 24 (2017) 3903-3910.
  50. W. Padula, R. Grover, M. Van Vliet, E. Ridgway, Comparing Five Alternative Methods of Breast Reconstruction Surgery: A Cost-Effectiveness Analysis, *Value in Health*. 17 (2014) A88.
  51. A. Perea, D. Rosselli, Análisis de costo-utilidad en Colombia del tratamiento integral del carcinoma de seno con reconstrucción inmediata frente a la reconstrucción diferida, *Biomédica*. 38 (2018) 363-378.
  52. N. Hatam, M. Askarian, J. Javan-Noghabi, N. Ahmadloo, M. Mohammadianpanah, Cost-Utility of "Doxorubicin and Cyclophosphamide" versus "Gemcitabine and Paclitaxel" for Treatment of Patients with Breast Cancer in Iran, *Asian Pacific Journal Of Cancer Prevention*. 16 (2016) 8265-8270.
  53. M. Sato, M. Kawai, Y. Nishino, D. Shibuya, N. Ohuchi, T. Ishibashi, Cost-effectiveness analysis for breast cancer screening: double reading versus single + CAD reading, *Breast Cancer*. 21 (2012) 532-541.
  54. V. Kalra, X. Wu, B. Haas, H. Forman, L. Philpotts, Cost-Effectiveness of Tomosynthesis in Annual Screening Mammography, *American Journal Of Roentgenology*. 207 (2016) 1152-1155.
  55. R. Kate, R. Nadig, Stage-specific predictive models for breast cancer survivability, *International Journal of Medical Informatics*. 97 (2017) 304-311.
  56. N. Eyal-Altman, M. Last, E. Rubin, PCM-SABRE: a platform for benchmarking and comparing outcome prediction methods in precision cancer medicine, *BMC Bioinformatics*. 18 (2017).
  57. C. Chao, Y. Yu, B. Cheng, Y. Kuo, Construction the Model on the Breast Cancer Survival Analysis Use Support Vector Machine, Logistic Regression and Decision Tree, *Journal Of Medical Systems*. 38 (2014).
  58. S. Momenyan, A. Baghestani, N. Momenyan, P. Naseri, M. Akbari, Survival Prediction of Patients with Breast Cancer: Comparisons of Decision Tree and Logistic Regression Analysis, *International Journal Of Cancer Management*. 11 (2018).
  59. M. Mohebian, H. Marateb, M. Mansourian, M. Mañanas, F. Mokarian, A Hybrid Computer-aided-diagnosis System for Prediction of Breast Cancer Recurrence (HPBCR) Using Optimized Ensemble Learning, *Computational And Structural Biotechnology Journal*. 15 (2017) 75-85.
  60. S. Sakri, N. Abdul Rashid, Z. Muhammad Zain, Particle Swarm Optimization Feature Selection for Breast Cancer Recurrence Prediction, *IEEE Access*. 6 (2018) 29637-29647.
  61. W. Yang, H. Bang, K. Jang, M. Sung, J. Choi, Predicting the recurrence of noncoding regulatory mutations in cancer, *BMC Bioinformatics*. 17 (2016).
  62. H. Chou, C. Yao, S. Su, C. Lee, K. Hu, H. Terng et al., Gene expression profiling of breast cancer survivability by pooled cDNA microarray analysis using logistic regression, artificial neural networks and decision trees, *BMC Bioinformatics*. 14 (2013) 100.
  63. A. Malehi, Diagnostic Classification Scheme in Iranian Breast Cancer Patients using a Decision Tree, *Asian Pacific Journal Of Cancer Prevention*. 15 (2014) 5593-5596.
  64. A. Bochare, A. Gangopadhyay, Y. Yesha, A. Joshi, Y. Yesha, M. Brady et al., Integrating domain knowledge in supervised machine learning to assess the risk of breast cancer, *International Journal Of Medical Engineering And Informatics*. 6 (2014) 87.
  65. Q. Li, S. Diao, H. Li, H. He, J.Y. Li, Applying decision trees to establish risk rating model of breast cancer incidence based on non-genetic factors among Southwest China females. *Zhonghua Zhong Liu Za Zhi*, 40(11) (2018) 872-877.
  66. V. Garrisi, S. Tommasi, A. Facchiano, I. Bongarzone, M. De Bortoli, M. Cremona et al., Proteomic profile in familial breast cancer patients, *Clinical Biochemistry*. 46 (2013) 259-265.
  67. P. Kitsantas, H. Wu, Body mass index, smoking, age and cancer mortality among women: A classification tree analysis, *Journal of Obstetrics and Gynaecology Research*. 39 (2013) 1330-1338.
  68. V. Podgorelec, P. Kokol, B. Stiglic, I. Rozman, *Decision Trees: An Overview and Their Use in Medicine* (2002).
  69. R. Kaplan, D. Frosch, *Decision Making in Medicine and Health Care*, *Annual Review of Clinical Psychology*. 1 (2005) 525-556.
  70. C. Lamanna, L. Byrne, Should Artificial Intelligence Augment Medical Decision Making? The Case for an Autonomy Algorithm, *AMA Journal of Ethics*. 20 (2018) E902-910.
  71. V.B. Kolachalama, P. Garg, Machine learning and medical education, *Npj Digital Medicine*. 1 (2018).
  72. C. Bingham, Which are the top scientific journals to publish applied machine learning research? [Msg 1]. Message posted at <https://www.quora.com/Which-are-the-top-scientific-journals-to-publish-applied-machine-learning-research> (2017).
  73. H. Ji, Assessing the assessments: Radiologists vs. automated breast density software, *Health Imaging*. (2016).
  74. Breast Cancer - Types of Treatment, *Cancer.Net*. (2018).
  75. G. Delaney, M. Barton, S. Jacob, B. Jalaludin, A model for decision making for the use of radiotherapy in lung cancer, *The Lancet Oncology*. 4 (2003) 120-128.
  76. R. Olivas, *Decision Trees A Primer for Decision-making Professionals*. (2007)
  77. E. F. Werner, S. Wheeler, I. Burd, Creating Decision Trees to Assess Cost-Effectiveness in Clinical Research, *Journal of Biometrics & Biostatistics*. 01 (2012).
  78. A. Nishikawa, L. Paladini, R. Delfini, P. Kotze, O. Clark, Decision Tree Construction And Cost-Effectiveness Analysis Of Treatment Of Ulcerative Colitis With Pentasa®/Mesalazine 2 G Sachet, (2013).
  79. Breast Cancer Care, Prognosis, (2016).
  80. N. Bundred, Prognostic and predictive factors in breast cancer, *Cancer Treatment Reviews*. 27 (2001) 137-142.
  81. J. Guo, B. Fung, F. Iqbal, P. Kuppen, R. Tollenaar, W. Mesker et al., Revealing determinant factors for early breast cancer recurrence by decision tree, *Information Systems Frontiers*. 19 (2017) 1233-1241.
  82. W. Song, K. Kim, S. Park, M. Kwon, T. Lee, H. Park, J. An, The Risk Factors Influencing between the Early and Late Recurrence in Systemic Recurrent Breast Cancer. *Journal of Breast Cancer*, 15(2), (2012) 218.

83. A. Ahmad, Pathways to Breast Cancer Recurrence. *ISRN Oncology*, (2013) 1-16.
84. Cancer Australia, (2017).
85. U.S. Breast Cancer Statistics | Breastcancer.org, Breastcancer.Org. (2019).
86. American Cancer Society, What's New in Breast Cancer Research?, Cancer.Org. (2018).
87. National Cancer Institute, Breast Cancer Research, (2018).
88. J. Cauchi, L. Camilleri, C. Scerri, Environmental and lifestyle risk factors of breast cancer in Malta—a retrospective case-control study, *EPMA Journal*. 7 (2016).
89. E. Choi, Y. Lee, H. Yoon, S. Lee, M. Suh, B. Park et al., Relationship between Cancer Worry and Stages of Adoption for Breast Cancer Screening among Korean Women, *PLOS ONE*. 10 (2015) e0132351.

## AUTHORS PROFILE



**Iffah Syafiqah Meor Badi'auzzaman**, Master of Health Sciences, Kulliyah of Allied Health Sciences, International Islamic University Malaysia.



**Moey Soo Foon**, Doctor of Philosophy, Assistant Professor, Kulliyah of Allied Health Sciences, International Islamic University Malaysia.



**Mohd. Zulfaezal Bin Che Azemin**, Doctor of Philosophy, Associate Professor, Kulliyah of Allied Health Sciences, International Islamic University Malaysia.



**Mohd. Izzuddin Bin Mohd. Tamrin**, Doctor of Philosophy, Assistant Professor, Kulliyah of Information and Communication Technology, International Islamic University Malaysia.