Artificial Neural Networks: Utilising Machine Learning for Equitable Breast Cancer Diagnosis

Dr. Roy Jafari

Scientia

ARTIFICIAL NEURAL NETWORKS: UTILISING MACHINE LEARNING FOR EQUITABLE BREAST CANCER DIAGNOSIS

Breast cancer is the most commonly occurring cancer in women, and it is pivotal that it is diagnosed correctly and promptly. Artificial Intelligence (AI) is now widely used in diagnosis to produce more accurate results. **Dr. Roy Jafari** from the University of Redlands is training a type of AI called an Artificial Neural Network to make more equitable diagnosis decisions for patients. He hopes that focussing on decision-making will reduce stress on patients and create a better care experience.

Identifying and Treating Breast Cancer

Breast cancer is the second most common cancer in the world and the most commonly occurring cancer in women. It can present as several types and in different parts of the breast, and a patient is usually diagnosed with either the non-invasive or invasive form. Non-invasive breast cancer is characterized by the tumor being confined to the ducts, without spread to the breast tissue. More commonly, invasive breast cancer is diagnosed, where cancer cells have spread to the surrounding tissue from the ducts. Other, more unusual forms are invasive lobular breast cancer, inflammatory breast cancer, and Paget's disease of the breast.

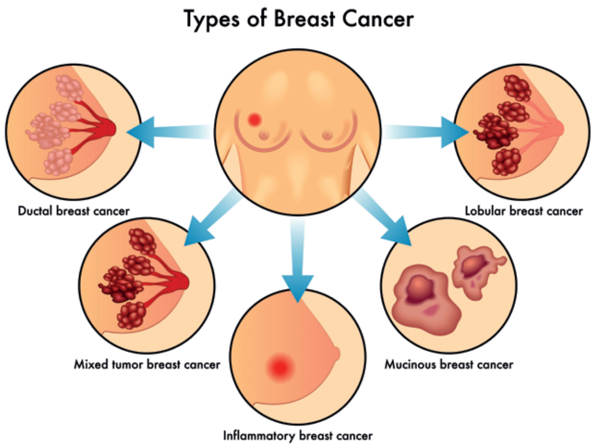
Symptoms of the disease may include a lump, tissue thickening, a rash, and other changes to the appearance of the breast. Finding any of these symptoms can prompt a person to visit their doctor for tests, which leads to a diagnosis. However, many countries including the USA, also provide screening services to women over the age of 50 or those perceived to be at high risk. This service helps to find people who are unaware that they either have or may be at risk of developing breast cancer because they are showing no symptoms.

The first test carried out is usually a mammogram, which involves taking x-rays to identify and quantify abnormalities within the tissue. If any are found, the mammogram may be followed by an ultrasound and a biopsy, which is a sample of cells taken from the suspicious area. The cells are then sent to a laboratory to be examined by a histopathologist, who can establish whether or not cancer is present in the tissue.

Following a positive diagnosis, one or a combination of interventions including surgery, radiotherapy, chemotherapy, targeted therapy, and hormone therapy are used to combat cancer. Thanks to decades of dedicated research, the survival rate of breast cancer after 10 years is now 76%.

The Role of Technology in Breast Cancer Diagnosis

One interesting tool that can be used to diagnose breast cancer is a machine-learning subset of Artificial Intelligence (AI), known as an Artificial Neural Network (ANN). These fascinating computer systems are inspired by human brains, which contain millions of interconnected neurons that fire to allow function and learning. An ANN is made up of layers of units (nodes) called artificial neurons. Each artificial neuron is connected and sends signals to all the neurons in the next layer, which are ordered starting with the input layer, then the hidden layers, and finally the output layer.



The networks can be trained, and they learn to recognize and correct errors using supervised and unsupervised learning. Supervised learning involves inputting labeled data into the first layer and providing the 'answer' to the data so that when similar data is next put in, the algorithm will produce accurate results by itself. It has learned what you taught it. Unsupervised learning is the process by which the algorithm extracts its own patterns from unlabelled data.

Diagnosis of breast cancer can be achieved through inputting data of a person's test results and life circumstances into an ANN. The accuracy of this method has proven to be very high but there remains a margin of error. A false negative is the term for a diagnosis of no cancer when, in fact, cancer is present. Conversely, a false positive occurs when someone is diagnosed with cancer when they are actually cancer-free.

Most ANN protocols run on the assumption that receiving a false negative result has the potential to be significantly more damaging to a person than a false positive. The logic here is that that patient would rather take further measures to investigate their suspected cancer and eventually rule it out than cancer go undetected and worsen over time. Dr. Roy Jafari from the University of Redlands wants to challenge this expectation and investigate how the assumptions programmed into an ANN can improve its decision making.

Should We Prioritise Life Span or Life Quality?

This is an interesting question – are the consequences of a patient missing out on treatment due to a false negative worse than the psychological consequences of a false positive. The answer may seem obvious at first; surely it is better to be tested as thoroughly as possible and retested as necessary to ensure that a patient receives treatment if they need it. However, this premise, which is also coded into the ANN, is not necessarily correct. Some studies have shown that the fear and anxiety that wrong or premature cancer diagnosis can create, may have real psychological effects, not to mention the physical pain that unnecessary mammograms cause.

Dr. Jafari emphasizes that the consequences of a false positive will differ from person to person. For instance, a 45-year-old woman, with immediate family care and high emotional resilience would be encouraged to undergo further tests if her results showed she had a 10% risk of cancer. This woman would have a good support system and would be able to bounce back quicker from any emotional stress. On the other

hand, a 75-year-old woman with no immediate family care and low emotional resilience would perhaps not benefit from further testing if her cancer risk was only 10%. She may experience serious stress that could impact her mental and physical health, in addition to not having access to adequate support.

Of course, if either risk was 95%, both individuals would be tested for breast cancer further because the threat to life is significantly higher. Dr. Jafari argues that all of these factors and more (such as quality of health insurance) should be taken into account when deciding if continued tests would actually dimmish a patient's quality of life when their risk is very low anyway. Currently, most ANN approaches use the assumption that all patients would want further testing, even if their risk of breast cancer is as low as 10% – but this is not necessarily the case.

Following on from this, Dr. Jafari believes that improving ANN functioning should focus more on decision-making, rather than accuracy so that patients receive the correct care, personal to them. As he explains, 'In my research, I investigate and develop algorithmic approaches to decision-making that are both data-driven and equitable.'

Self-Organizing Error-Driven Learning

To study how altering an ANN's priorities towards better decision-making could influence outcomes, Dr. Jafari and his team developed a new method of teaching for the networks called Life-Sensitive Self-Organizing Error-Driven (LS-SOED). This method takes a patient's uniqueness into account by recognizing that if the data analytic results are inconclusive, a false positive or a false negative result will have varying consequences depending on their circumstances.

An ANN learns from its errors, so if the calculations to rectify these are driven by accuracy, the networks will focus on improving accuracy, not improving decision making. The LS-SOED method aims to put the ANN's power into making better decisions for patients, using both supervised and unsupervised learning. Data from patients is organized into a base, where their personal similarities and differences are mapped and recognized by the system (this is the self-organizing part of the title). This base allows LS-SOED to use decision error-driven learning along with clearer decision goals so that the ANN decision making improves.

A combination of underlying methods assists the new programming. Self-Organizing Map recognizes patterns and maps them, Multi-Layered Perceptron uses this map to predict patient types and Fluid Genetic Algorithm deals with inconclusive data.

Findings from the New Method

The new LS-SOED method applied to an ANN provided some interesting results. In line with their original thinking, Dr. Jafari and his team showed that a highly accurate ANN does not always give the best diagnosis decisions. They compared two sets of data analyzed using LS-SOED with three other advanced data analytic techniques. The first group of 253 patients diagnosed using decisions made by the LS-SOED method were predicted to save 30 years of human life between them. The second group of 57 were predicted to save 8 years, meaning collectively 38 years would be saved between 340 people.

Dr. Jafari notes that implementing these decision-making improvements comes with a great computational cost and is very complex. But this is only the beginning, the complexity of these systems will pave the way for more research into the promising possibilities for AI and machine learning.

Meet the researcher

Dr. Roy Jafari
School of Business
University of Redlands
Redlands, CA
USA

Dr. Roy Jafari completed his Bachelor of Science in Industrial Engineering at Tafresh University in Iran and then went on to obtain a Master of Science in Industrial Engineering from the University of Tehran. After moving to the USA to continue his studies, he graduated with a Ph.D. from Mississippi State University in 2018. He served as Assistant Professor of Industrial and Manufacturing at California Polytechnic State University from September 2018 to July 2020, and now he holds the role of Assistant Professor of Business Analytics at the University of Redlands. In his academic positions, he researches diagnosis decisions and data science. His current work investigates how machines can use data analytics to make smarter and more equitable decisions for real-world applications such as breast cancer diagnosis.

CONTACT

E: roy_jafari@redlands.edu
W: https://roy-jafari.com
https://www.linkedin.com/in/roy-jafari-marandi-84077932/
https://www.researchgate.net/profile/Roy_Jafari
https://twitter.com/jafariroy

FURTHER READING

R Jafari-Marandi, Supervised or Unsupervised Learning? Investigating the role of pattern recognition assumptions in the success of binary predictive prescriptions, Neurocomputing, 2021, 434, 165–193, https://www.sciencedirect.com/science/article/abs/pii/S0925231220319639?via%3Dihub

H Long, JM Brooks, M Harvie, A Maxwell, DP French, How do women experience a false-positive test result from breast screening? A systematic review and thematic synthesis of qualitative studies, British Journal of Cancer, 2019, 121, 351–358.

R Jafari-Marandi, S Davarzani, M Soltanpour Gharibdousti, BK Smith, An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN learning and decision-making goals, Applied Soft Computing, 2018, 72, 108–120.

M Bond, R Garside, C Hyde, A crisis of visibility: The psychological consequences of false-positive screening mammograms, an interview study, British Journal of Health Psychology, 2015, 20, 792–806.