Interpretability Of Machine Learning Algorithms In Cancer Diagnosis

ROHITH VADDEPALLY – W9529921
MSc Data Science With Advanced Practice
School of Computing, Engineering & Digital Technologies

Research Purpose

Introduction
In today's world, healthcare is one of the leading industries as it is most important one. No matter what happens this is only industry that should keep running. Now a days, Machine learning is included within the sector in various departments. One of the important department is Oncology. Oncology is the study of cancer. An oncologist is a doctor who treats cancer and provides medical care for a person diagnosed with cancer. So in this process different machine learning models are being used to detect the cancer in the patient. The purpose of this research is to investigate how the machine learning model comes to a specific decision that a patient has cancer or not?. This is where Interpretability in ML model comes into place. Interpretability means the ability to determine the cause and effect of a machine learning model.

Research Question
Is the machine learning algorithm's decision in the diagnosis of cancer in healthcare industry is interpretable to humans or not?

Research Objectives
Here are some of the research objectives:

- To investigate how far a Machine Learning algorithm or model is interpretable to the humans in the cancer diagnosis of the patient.
- This study also aims to find out the fairness of the algorithm in it's decision.
- To investigate how Explainability in ML model plays an important role in Cancer detection.
- Additionally, this research also aims to investigate different types of interpretable, explainable and black box models in the diagnosis of the cancer which play a vital role in model's decision.

Anticipated contribution to knowledge
Previously done literature review on this topic will be acting as base for the research. This will attempt to demonstrate case positives and negatives from the journals, conference papers, and other sources investigated. The most anticipated product will be the knowledge shared the article named "Interpretability of machine learning algorithms in cancer diagnosis".

Background Context

Importance of this research
Cancer has become one of the main cause of death in people all around the world. An interpretable method or model is necessary for diagnosing patients with cancer for a well performed treatment. A good and well performed treatment will always help to improve the patient's condition. As this is matter of people's lives, this research is very much useful and will be important to the mankind.

Main findings
As there are many cancer types and one is very much different from another, we cannot apply any one algorithm to all the cancer types.
For better interpretation and based on findings,

- IRFRE(Improved Random Forest-Based Rule Extraction Method) is used for Breast Cancer Diagnosis. Wang, S. *et al.* (2020)
- CFCMC (Cumulative Fuzzy Class Membership Criterion) is used for Colorectal Cancer Diagnosis. Sabol, P. *et al.* (2020)
- ISHAP (Improved Shapley Additive Explainations) for Lung Cancer Detection. He, W. *et al.* (2022) Etc.

As discussed with a diagram in element-1, the more accurate the models are, the less interpretable and explainable they are. Below a figure is provide how a interpretable methods provide value to end user.

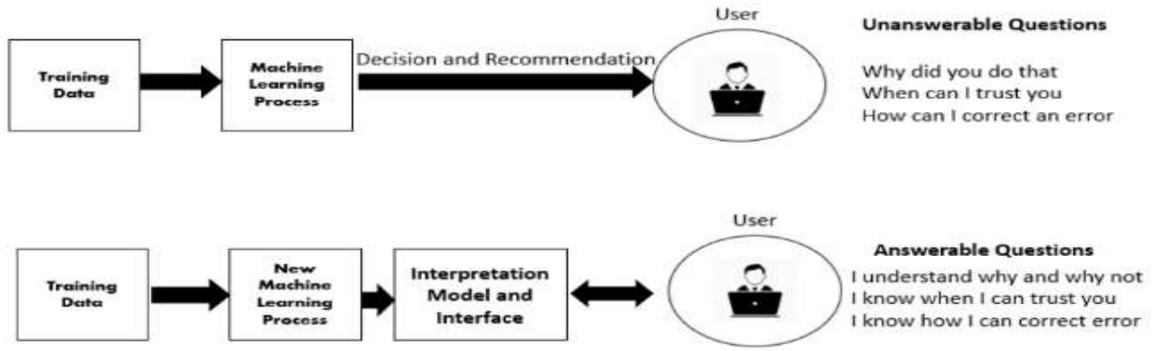


Fig 1: How interpretable model provide value to the end user. Gulum, Trombley and Kantardzic (2021)

For early breast cancer diagnosis, in this study I have come across using different interpretability techniques namely the Global Surrogate Model, Individual Conditional Expectation Plots and Shapley values.

Research Methodology

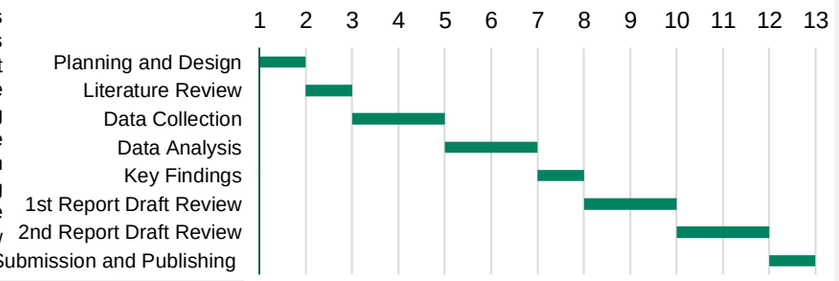
Research Strategy
A Secondary research method will be used for this as it will rely on peer-reviewed journal articles and conference papers. This Secondary research will be conducted based on the qualitative data or qualitative basis as the data that will be used in the research is the information gathered from various peer-reviewed journal articles, conference papers, case studies and record keeping.

Data Generation Methods
As a secondary qualitative research method will be used for this research, the data will be accessed and collected from various journals articles and conference papers which found in the element 1. If required, will gather additional data from surveys, online communities and groups on the internet.

Data Analysis Techniques
Inductive Analysis is used for this research as it systematically generate theory grounded in specific instances of empirical observation. Thomas (2003)

Schedule and Publishing

The total duration to complete this research project is 13 weeks. This is completely shown in the beside Gantt chart on a weekly basis, from the Planning, Literature review, Analyzing data and Reviewing and publishing. The journal article is aimed to be published in Science Direct which is right now serving many academic institutions. Science Direct always helps shed light on new research and innovative thinking.



Evaluation Methodology

This research's final product will be peer-reviewed by the module leader, module tutors, and fellow students.

The Evaluation methodology of this research will incorporate qualitative assessment, as mentioned in the research methodology, with data acquired from a variety of journals, conference papers, case studies, and records.

As this research revolves around various Interpretable machine learning algorithms evaluation will be done by Algorithmic impact assessment which is a case study generally done in the healthcare industry. This means learning about a difficult instance based on a thorough grasp of the instance gained by lengthy description and analysis of the instance in its entirety and context.

The Artifact in this research will be a good interpretable ML model with good reasonable accuracy results. This artifact will be helpful to the Cancer Patients, Oncologists, Physicians, clinics and hospitals in healthcare industries.

The assessment and evaluation of correct guidelines maintained in the development of this research module will also be included in this evaluation methodology.

Professional, Legal and Ethical Issues

During this research no patients and doctor's confidentiality will be violated. The case studies will be done according to the terms of Health Insurance Portability and Accountability Act (HIPAA).

Research will be overseen and done with research ethics and principles of Teesside University's University Research Ethics and Integrity Sub-Committee (UREISC).

This research will also be followed by six principles set by University Research Ethics Committee.

References

Wang, S. *et al.* (2020) 'An improved random forest-based rule extraction method for breast cancer diagnosis', *Applied Soft Computing,* 86, pp. 105941. doi: https://doi-org.ezproxy.tees.ac.uk/10.1016/j.asoc.2019.105941.

Sabol, P. *et al.* (2020) 'Explainable classifier for improving the accountability in decision-making for colorectal cancer diagnosis from histopathological images', *Journal of Biomedical Informatics,* 109, pp. 103523. doi: https://doi-org.ezproxy.tees.ac.uk/10.1016/j.jbi.2020.103523.

He, W. *et al.* (2022) 'An ISHAP-based interpretation-model-guided classification method for malignant pulmonary nodule', *Knowledge-Based Systems,* 237, pp. 107778. doi: https://doi-org.ezproxy.tees.ac.uk/10.1016/j.knosys.2021.107778.

Gulum, M.A., Trombley, C.M. and Kantardzic, M. (2021) 'A review of explainable deep learning cancer detection models in medical imaging', *Applied Sciences (Switzerland),* 11(10) doi: 10.3390/app11104573.

Thomas, D. (2003) 'A General Inductive Approach for Qualitative Data Analysis', *The American Journal of Evaluation,* 27.