IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Cervical Cancer Prediction Using Ensemble Algorithm In Health Care

Mr T.MANJUNATHA REDDY (M.C.A). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Mr. V.V. NAGENDRA KUMAR (MTech, Ph.D). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Abstract

Health professionals identify cervical cancer as a potentially fatal condition. Patients' valuable lives are at risk due to the difficult late diagnosis and treatment. The formal screening for disease identification suffers in both industrialized and developing countries because of high medical costs, a lack of healthcare facilities, social norms, and the late onset of symptoms. Early diagnosis of various different diseases, including cervical cancer, is possible because to machine intelligence. It is also cost-effective and computationally cheap. Modern, time-consuming medical treatments are not necessary for the patients, and machine intelligence can help with early cervical cancer diagnosis. The reliance on a single classifier's prediction accuracy is the issue with the present machine classification approaches for illness identification. Due to bias, over-fitting, improper treatment of noisy data, and outliers, the use of a single classification method does not guarantee the best prediction. In order to provide an appropriate diagnosis that addresses the patient's symptoms or problems, this research study presents an ensemble classification method based on majority voting. In the study, a variety of classifiers, including Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression (LR) classifiers, are put to the test. The study shows an increase in prediction accuracy of 94%, which is significantly higher than the prediction accuracies of individual classification methods tested on the same benchmarked datasets. As a result, the suggested paradigm grants health professionals a second opinion to aid in the early diagnosis and prompt treatment of diseases.

KEYWORDS: Machine Learning Algorithms, Classifiers, Cervical Cancer, Ensemble Classification.

1. INTRODUCTION

1.1 Introduction

Today's deadliest diseases include cancer [1]. The disease exhibits specific aberrant behavior in the afflicted cells. Normal tissues suffer damage from cancerous cells, which alters how well they normally function [2]. The likelihood of cancer spreading to different bodily areas is likewise very high. Consequently, in many situations, the inability to identify cancer at an early stage can result in mortality. Early cancer detection often increases the chance of survival, depending on the kind of cancer [3]. One of the most prevalent cancers in women is cervical cancer, which is primarily brought on by HPV. Papillomavirus in humans. Numerous studies have demonstrated that cervical cancer therapy and recovery can be considerably impacted by early identification [4], [5]. Pap smear (Papanicolaou test) [6], HPV DNA genotyping [7], HCII (Hybrid Capture II), hybrid capture [8], [9], and Southern blot hybridization assay [10] are frequently employed methods for the identification of cervical cancer. The most recent innovation in this area is the biosensor, which has enormous potential because to its low cost, quick findings, and simplicity of use [11], [12]. Each of these

methods has its drawbacks. The primary drawbacks of the pap smear test include its high false negative rate [13], low sensitivity, and expensive price [13]. The southern blot hybridization assay procedure requires a lot of time and is not very sensitive [14]. Genotyping of HPV DNA needs a lengthy It takes a long time to complete the test and is rather expensive [15]. One of the most cutting-edge methods for the early identification of cervical cancer is HCII hybrid capture. Only 13 HPV strains can be detected, which is its biggest drawback [16]. Several operational limitations, including a lack of electrode reusability, operational instability, and a short lifetime, affect the performance of biomedical sensors [17]. These shortcomings of the current approaches highlight the demand for more effective methods for cervical cancer early detection.

2. Literature Survey

• [1] X. Li, and J. Liu, "Application of cyclodextrins in cancer treatment," J. Inclusion Phenomena Macrocyclic Chem., vol. 89, nos. 3–4, pp. 229–246, Dec. 2017, doi: 10.1007/s10847-017-0752-2:, One of the leading causes of death is cancer. Chemotherapy is a common form of treatment that employs a mixture of medications to either kill cancer cells or inhibit their growth. The majority of cytotoxic chemotherapy medicines, however, are water insoluble, which makes formulation challenging. Utilizing cyclodextrins (CDs), which are frequently used to increase the solubility, bioavailability, stability, and safety of medicinal molecules by generating non-covalent inclusion complexes, is one promising tactic. This review's goal is to clarify how CDs are used in various cancer therapy methods. The fact that CDs have been found to have anticancer effect both in vitro and in vivo is very intriguing. It is addressed how CDs might be used as an anticancer agent and how they might stop the proliferation of cancer cells. Future potential uses are discussed along with CDs/anti-neoplastic-drug complexes that have better solubility, more stability, and higher anti-cancer action. This paper goes into detail on how CDs are used and their benefits in various drug delivery systems such liposomes, conjugates, nanoparticles, and siRNA carriers for the treatment of cancer..

[2] L. Schwartz, C. Supuran, and K. Alfarouk, "The warburg effect and the hallmarks of cancer," Anti-Cancer Agents Med. Chem., vol. 17, no. 2, pp. 164–170, Jan. 2017, doi: 10.2174/1871520616666161031143301: Whether cancer is one disease or a group of quite different diseases is a topic of ongoing dispute. The purpose of this study is to make a convincing case that the Warburg's effect may be the cause of the majority, if not all, of the characteristics of cancer. A reduction in ATP concentration results from the metabolic impairment of oxidative phosphorylation. Massive glucose absorption, anaerobic glycolysis, and an over-regulation of the Pentose Phosphate Pathway, which results in greater biosynthesis and enhanced cell division and local pressure, are all used to make up for the decreased energy production. The fractal structure of the tumor is caused by this increased pressure, which also causes the fibroblasts to secrete collagen and is essential for the progression of metastatic disease. the enormous extrusion The Pentose Phosphate Pathway causes an increase in biosynthesis, which in turn causes an increase in cell division and regional pressure. The fractal structure of the tumor is caused by this increased pressure, which also causes the fibroblasts to secrete collagen and is essential for the progression of metastatic disease. The extracellular acidity and immune system activation are both influenced by the large extrusion of lactic acid. Due to increased intracellular alkalosis and the contribution of carbonic acid to extracellular acidosis, caused by the decreased oxidative phosphorylation, CO2 levels both within and outside of the cell are impaired. This is mediated by at least two carbonic anhydrase isoforms that are related with cancer.

3. OVERVIEW OF THESYSTEM

3.1 Existing System

According to statistics, cervical cancer is the third most fatal disease affecting women globally. The main cause of this illness is HPV, which infects a location and causes a tumor there. Health professionals advise early detection and vaccination as treatments for the illness. Nearly 99.7% of cases of cervical cancer are brought on by HPV infection. To obtain the specifics of our health status, this screening and immunization process requires a lot of time.

3.1.1 Disadvantages of Existing System

- Less feature compatibility
- Low accuracy.

3.2 Proposed System

In the suggested approach, an ensemble classification method based on majority voting is used to make a precise diagnosis that takes into account the patient's symptoms or circumstances. In the study, a variety of classifiers, including Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression (LR) classifiers, are put to the test. The study shows an increase in prediction accuracy of 94%, which is significantly higher than the prediction accuracies of individual classification methods tested on the same benchmarked datasets.

3.3 Methodology

In this project work, I used five modules and each module has own functions, such as:

1. System Module
2. User Module

1. System:

1. Pre-processing:

In this step preprocessing i.e. data cleaning and data filling.

1.2 Training:

Use the pre-processed training dataset is used to train our model using our machine learning algorithms.

1.3 Generate accuracy

System generates accuracy for our model and dataset. This tells us how much efficiently model is working.

1.3 Generates results:

The results will be displayed are which disease.

2. User:

2.1 Data collection

The user has to upload an image which needs to be classified.

2.2 Model building

User builds the models to fit our data for prediction of cancer disease.

2.3 View Accuracy

Users view the generated accuracy from the system.

2.4 View Results

Users can view the generated classification from the user.

4 Architecture

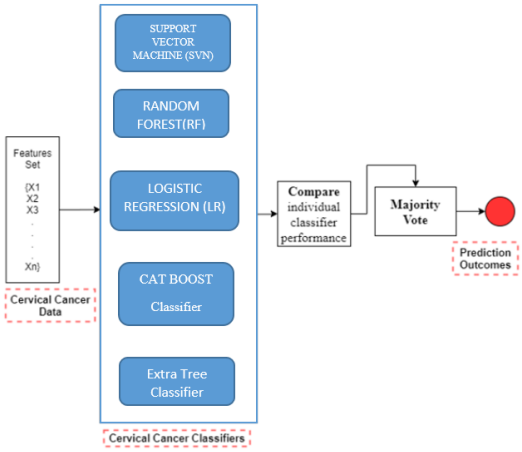
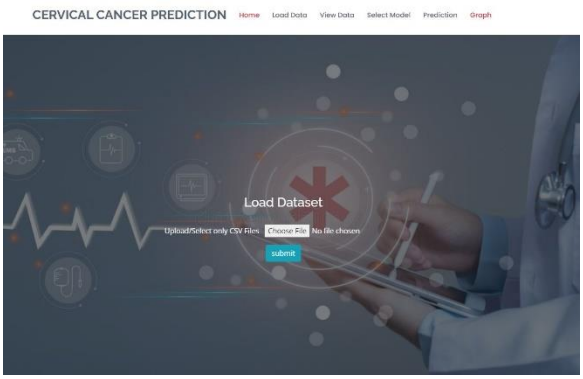


Fig 1: Frame work of proposed method

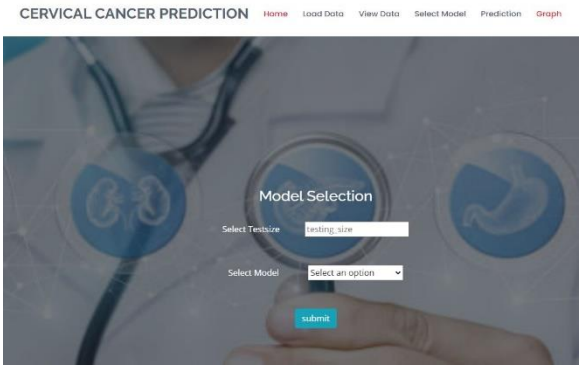
5 RESULTS SCREEN SHOTS

Home Page:

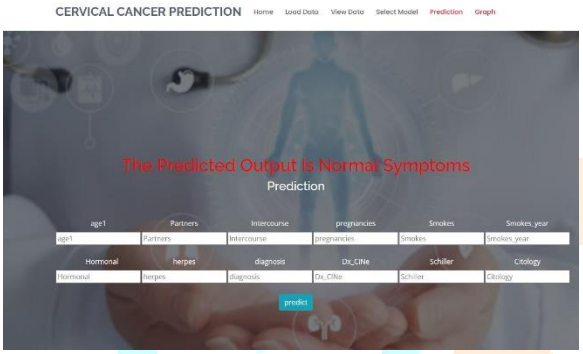
Upload Data:



Choose options:



Predict Result:



7. CONCLUSION

✓ To save the subjects' priceless lives, the health 4.0 requirements need an accurate and reliable cervical cancer detection. Despite the difficulties and problems, machine intelligence-based solutions now available are regarded as reliable. However, the effectiveness of various categorization systems for the detection of cervical cancer is also a problem. Different classifiers produce distinctly different results when used to the same datasets because the specific classification algorithms are sensitive to the nature of the data. Based on a major voting procedure to accept the best classification outcomes for cervical cancer prediction, this research study proposed an ensemble classification approach for cervical cancer diagnosis. Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression were among the many classifiers used in the study. classifiers. When compared to other classifiers, the suggested ensemble classifier performed better than individual classifiers, achieving the greatest accuracy at 94%. Health professionals can use the findings of this research to give cervical cancer patients a knowledgeable and confident second opinion to treat the condition more effectively.

8. References

[1] N. Qiu, X. Li, and J. Liu, ‘‘Application of cyclodextrins in cancer treatment,’’ J. Inclusion Phenomena Macrocyclic Chem., vol. 89, nos. 3–4, pp. 229–246, Dec. 2017, doi: 10.1007/s10847-017-0752-2.

[2] L. Schwartz, C. Supuran, and K. Alfarouk, ‘‘The warburg effect and the hallmarks of cancer,’’ Anti-Cancer Agents Med. Chem., vol. 17, no. 2, pp. 164–170, Jan. 2017, doi: 10.2174/1871520616666161031143301.

[3] E. M. Hassan and M. C. DeRosa, ‘‘Recent advances in cancer early detection and diagnosis: Role of nucleic acid based aptasensors,’’ TrAC Trends Anal. Chem., vol. 124, Mar. 2020, Art. no. 115806, doi: 10.1016/j.trac.2020.115806.

[4] N. A. Parmin, U. Hashim, S. C. B. Gopinath, S. Nadzirah, Z. Rejali, A. Afzan, and M. N. A. Uda, ‘‘Human papillomavirus e6 biosensing: Current progression on early detection strategies for cervical cancer,’’ Int. J. Biol. Macromolecules, vol. 126, pp. 877–890, Apr. 2019, doi: 10.1016/j.ijbiomac.2018.12.235.

[5] T. A. Kessler, ‘‘Cervical cancer: Prevention and early detection,’’ Seminars Oncol. Nursing, vol. 33, no. 2, pp. 172–183, May 2017, doi: 10.1016/j.soncn.2017.02.005.

[6] N. Wentzensen and M. von Knebel Doeberitz, ‘‘Biomarkers in cervical cancer screening,’’ Disease Markers, vol. 23, no. 4, pp. 315–330, 2007, doi: 10.1155/2007/678793.

[7] C. J. Meijer, P. J. Snijders, and P. E. Castle, ‘‘Clinical utility of HPV genotyping,’’ Gynecol. Oncol., vol. 103, no. 1, pp. 12–17, Oct. 2006, doi: 10.1016/j.ygyno.2006.07.031.

[8] H. N. Luu, K. R. Dahlstrom, P. D. Mullen, H. M. Vonville, and M. E. Scheurer, ‘‘Comparison of the accuracy of hybrid capture II and polymerase chain reaction in detecting clinically important cervical dysplasia: A systematic review and meta-analysis,’’ Cancer Med., vol. 2, no. 3, pp. 367–390, Jun. 2013, doi: 10.1002/cam4.83.

[9] H. N. Luu, K. Adler-Storthz, L. M. Dillon, M. Follen, and M. E. Scheurer, ‘‘Comparing the performance of

hybrid capture II and polymerase chain reaction (PCR) for the identification of cervical dysplasia in the screening and diagnostic settings,'' Clin. Med. Insights, Oncol., vol. 7, Jan. 2013, Art. no. CMO.S12811, doi: 10.4137/CMO.S12811.