1.INTRODUCTION

1.1 Overview

Lung cancer survival rate is very limited post-surgery irrespective of if it is small cell or non-small cell. A lot of work has been carried out by employing machine learning in life expectancy prediction post thoracic surgery for patients with lung cancer. Many machine learning models like multi-layer perceptron (MLP), SVM, naïve Bayes, decision tree, random forest, logistic regression have been applied for post thoracic surgery life expectancy prediction based on data sets from UCI. Also, work has been carried out towards attribute ranking and selection in performing better in improving prediction accuracy with machine learning algorithms. So accordingly, the authors, here, have developed a deep neural network-based approach in prediction of post thoracic life expectancy which is the most advanced form of neural networks. This is based on dataset obtained from Wroclaw Thoracic Surgery Centre machine learning repository which contained 470 instances. On comparing the accuracy, the results indicate that the deep neural network can be efficiently used for predicting the life expectancy.

1.2 Purpose

In order to improve quality initiatives, healthcare administration, and consumer education, it is critical to track health outcomes. The data obtained from patients who had large lung resections for primary lung cancer is referred to as thoracic surgery. Attribute ranking and selection are critical components of successful health outcome prediction when using machine learning algorithms. Researchers used several procedures, such as early-stage examinations, to determine the type of cancer before symptoms appeared. The most relevant attributes are identified using attribute ranking and selection, and the duplicated and unnecessary attributes are removed from the dataset. The goal of our study is to look at patient mortality over the course of a year after surgery. More precisely, we\'re looking into the patients\' underlying health issues, which could be a powerful predictor of surgical-related mortality.

2.LITERATURE SURVEY

2.1 Existing Problem

It has always been a difficult task to accurately predict the life expectancy post an operation. The prediction relies upon on numerous fitness elements of which a few have a far important function in comparison to the opposite elements. A famous approach used within side the beyond became to investigate the CT test photos of the lungs and expect primarily based totally at the everyday check-ups. The thirty-day mortality price is one statistic that has been used to estimate mortality charges within side the beyond. This statistic, however, won't be absolutely correct due to the fact many sufferers die or come to be very frail right now after this time period, requiring them to be transferred to any other organization earlier than passing death. As a result, a large number of these deaths go unreported.

2.2 Proposed Solution

In order to improve quality initiatives, healthcare administration, and consumer education, it is critical to track health outcomes. Data from patients who have undergone extensive lung resection for primary lung cancer is called thoracic surgery. Attribute ranking and selection are important components for correctly predicting health outcomes when using machine learning algorithms. Researchers have used several techniques, such as early screening, to determine the type of cancer before symptoms appear. The most relevant attributes / characteristics are identified using attribute ranking and selection, and duplicate unwanted attributes / characteristics are removed from the dataset.

The goal of our study is to look at patient mortality over the course of a year after surgery. More precisely, we're looking into the patients' underlying health issues, which could be a powerful predictor of surgical-related mortality.

3. Theoretical Analysis

3.1 Block Diagram

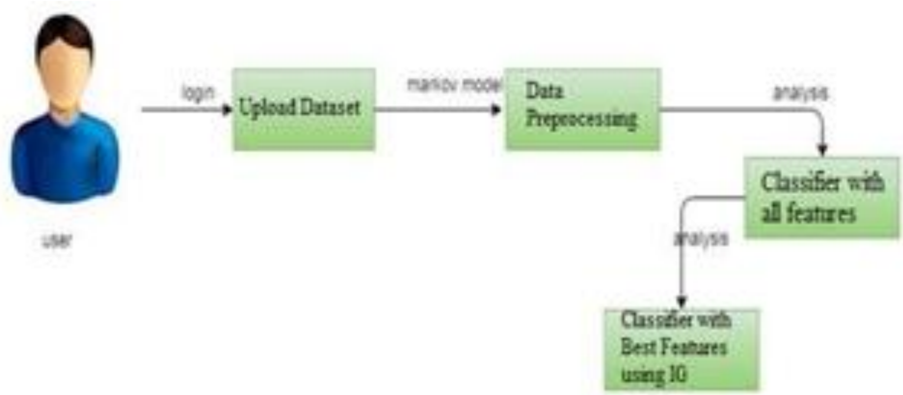
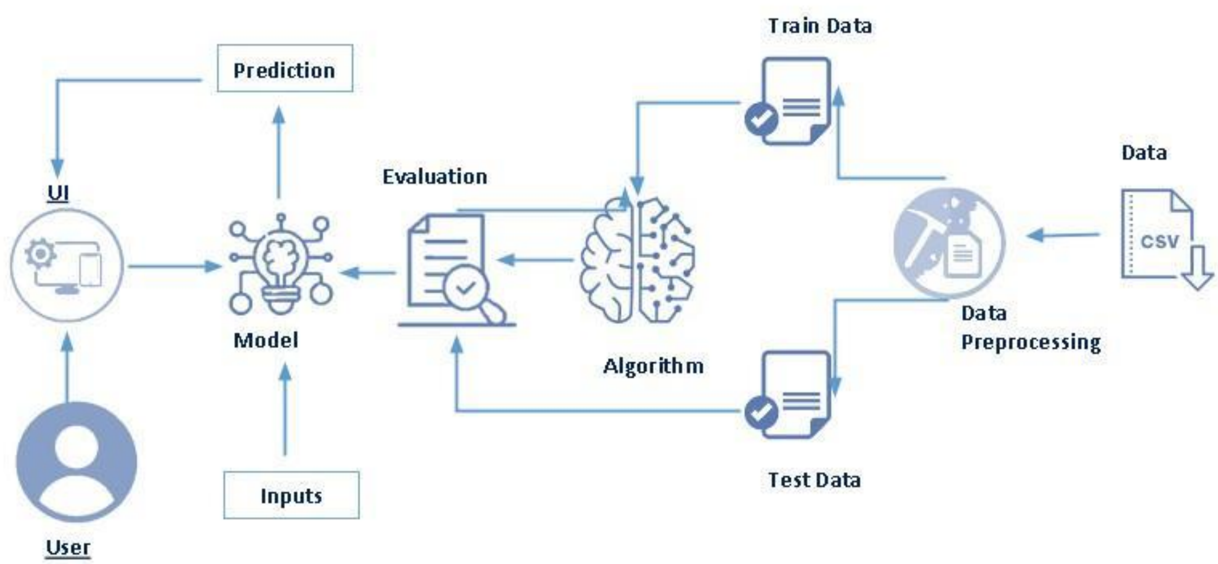


Figure 1: Block diagram

3.2 Hardware/Software designing



5. Flowchart

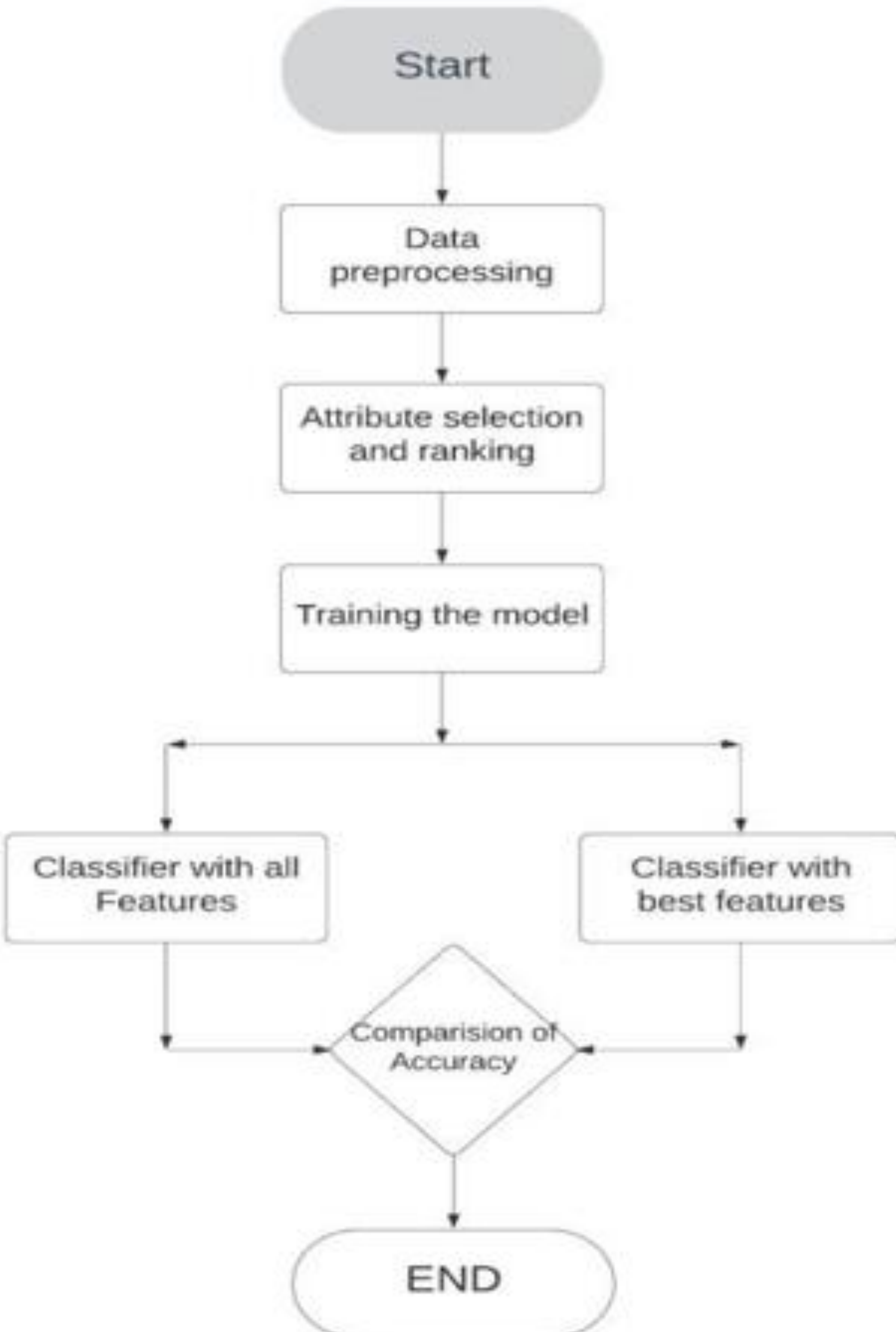
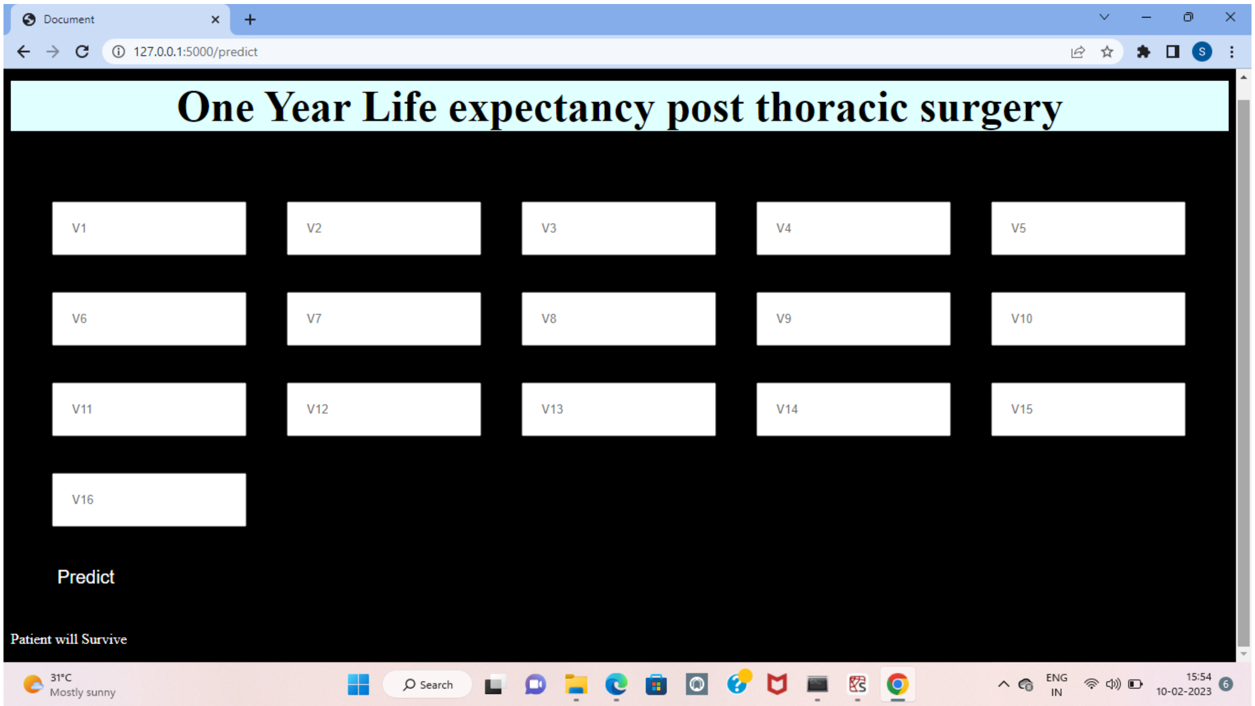


Figure 2: Flowchart

6. Result



7. Advantage and Disadvantage

Benefits includes enhancing predictability that facilitates the derivation of more appropriate decisions. A study conducted by Chong et al. revealed this benefit since the random forest classifier (RFC) ML model applied was cost-efficient (19). It also prompted the prediction of lymph node metastasis during or after surgery for early-stage adenocarcinomas with a sensitivity and specificity of 87.5 and 82.2 percent, respectively (19). In a similar perspective, Wu et al. indicated that the RFC machine learning algorithm was the best predictive model and facilitated the recognition of lymph node metastasis among patients with early T-stage non-small cell lung cancer before undergoing an operation (20). Another study by Liu et al. revealed that machine learning algorithms had the potential to assist in personalizing treatment decisions (21). This conclusion arose from the excellent performance of the nomogram comprising CT-derived radiomic features and risk factors developed by the authors in predicting post-surgical progression-free survival of patients with lung adenocarcinoma (21).

Additionally, Liang et al. associated machine learning algorithms with improved lung cancer screening and better assessment of treatment efficiency after the ML-assisted deep methylation sequencing detected tumor-derived signals among patients with surgery-resectable lung cancer (22). These arguments imply that the advantages derived from application of ML in lung surgery ensure accurate predictability and opportunities for enhancing patient outcomes.

On the contrary, the main shortcoming of machine learning in lung surgery is the lack of sufficient research on using them appropriately. For instance, a study by Feng et al. revealed that most ML techniques require adopting different protocols and settings that require complex analysis in literature to avoid any errors (23). The second limitation of ML algorithms in lung surgery is the tendency of most outcome prediction researchers to adopt a small dataset that increases the chances of biased study results (24). Therefore, it remains vital to deal with the general challenges facing ML implementation to ensure the successful integration of the concept into lung surgery. This factor affected the derived results since the researchers could have excluded articles in other languages that provided information relevant to the study. The second limitation of the study is a wide literature gap in the field. This conclusion arises from the derivation of only nineteen articles to include in the research. The final limitation stems from adopting the narrative review approach since the technique led to broad findings, making it hard to draw essential conclusions.

8. Applications

- Rapid detection and screening of post-surgical complications
-
- Rapidly diagnose or screen cardiac abnormalities
- Objective and consistent aortic diameter measurement
- Objectively and rapidly measure cardiac function in real time
- Screening high-risk populations and basic research
- Stratifying patient with type A aortic dissection for closer monitoring and follow-up
- Stratifying patients for closer monitoring and follow-up
- Rapid and continuous surgical skill evaluation without need for manual annotation
- Real-time technique feedback in the operating room
- Inform surgical risks of high-risk patients
- Risk stratify post-operative patients for closer follow-up
- Increase the scale and speed of basic research projects by eliminating manual diameter measurement

9.CONCLUSION

In this study, the quality of the three classification methods/ algorithms have been tested to improve the prognosis for the life of patients with thoracic thoracic cancer after thoracic cancer surgery. Three ways to train the machine before and after using the quality level options are compared with their improved versions. The results show that boosting is not always the best solution, where the level of responsibility and choice can make better at improving predictive accuracy. Other qualifications and machine learning strategies can be introduced in the future work to find the best performance of the data forecast model. The results indicate that decision tree, logistic regression and KNN give us better performance than other data mining algorithms, we also have monitored and mentioned the performance of each algorithm. From this study we deducted that FEV1, Dyspnoea before surgery, Size of the tumour, Weakness before surgery and performance status are the attributes/features that are mainly responsible for the survival of patients after the thoracic surgery. This study asserts the use of data mining algorithms in medical field as the results are better than the existing system which are confirmed by statistical analysis.

10.FUTURE SCOPE

This study does have few limitations. The results/outputs which were obtained are particular to a nation or an organization which collected the data set. Results obtained may be time-limited (2007–11). The dataset used as part of this project has very less records and may impede the accuracy of few algorithms that are used. In any case, this dataset can serve as a starting point to raise a better understanding of thoracic surgery patients. These tests can be further extended.

This analysis only makes and uses three data mining methods. Therefore, some more machine learning methods could be used to get more knowledge about the data-set as a future work.

11.BIBLIOGRAPHY

[1] V. Sindhu, S. A. S. Prabha, S. Veni and M. Hemalatha. (2014), “Thoracic surgery analysis using data mining techniques”, International Journal of Computer Technology & Applications, Vol. 5 pp.578-586. [2] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis and Dimitrios I. Fotiadisa. (2015), “Machine learning applications in cancer prognosis and prediction”, Computational and Structural Biotechnology Journal, Vol. 13, pp.8-17. [3] Kwetishe Joro Danjuma. (2015), “Performance evaluation of machine learning algorithms in post-operative life expectancy in the lung cancer patients”, IJCSI International Journal of Computer Science Issues, Vol. 12, No. 2, pp.189-199. [4] Joseph A. Cruz, David S. Wishart. (2006), “Applications of machine learning in cancer prediction and prognosis”, Cancer Informatics, Vol. 2, pp.59-77 2006. [5] Mehdi Naseriparsa, Amir-Masoud Bidgoli and Touraj Varaee. (2013), “A hybrid attribute/feature selection method to improve performance of a group of classification algorithms”, International Journal of Computer Applications, Vol. 69, No. 17, pp.28-35. [6] Desuky, A.S. and El Bakrawy, L.M., 2016. Improved prediction of post- operative life expectancy after Thoracic Surgery. Advances in Systems Science and Applications, 16(2), pp.70-80.