*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10*
*DOI: https://doi.org/10.17762/ijritcc.v11i10.8512*
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

# A Machine Learning Pipeline and Application for Automatic Classification of Clinical Documents

**[1]Dr. G Uday Kiran, [2]Sneha Raga Soujanya, [3]Mrs. M Mounika, [4]Dr. Narasimha Chary CH**
[1]Associate Professor, Department of CSE(AI & ML), B V Raju Institute of Technology
udaykiran.goru@bvrit.ac.in
[2]assistant professor, Department of CSE, AVN college
snehatanta@gmail.com
[3]Assistant Professor, B V Raju Institute of Technology
mounika.m@bvrit.ac.in
[4]Associate Professor, Dept of CSE, Sri Indu college of engineering and technology (Autonomous) Sheriguda, Hyderabad, TS, INDIA- 501510
narasimhachary.dr@gmail.com

**Abstract---**Healthcare industry has many associated services including research on various trends or patterns in diseases and patients' life style. With the emergence of Artificial Intelligence (AI), it is made possible that problems in healthcare domain can be solved by using Machine Learning (ML) techniques. One such problem considered in this paper is known as clinical document classification. Existing methods in this area lack a systematic approach in filtering out false positives. In this paper we proposed a ML framework that considers pipelining of ML models at multiple levels. In the first level, clinical documents that do not have any content related to smoking are discarded. In the second level, the documents that talk about known smoking cases are retained. In the third level clinical document are classified into two categories such as currently smoking and past smokers. We proposed an algorithm known as Learning based Clinical Document Classification (LbCDC). This algorithm makes use of three models in pipeline in order to perform classification of clinical documents at multiple levels of granularity. Our experimental results revealed that the proposed system is efficient in clinical document classification.

**Keywords –** Clinical Document Classification, Machine Learning, Learning Based Framework, Exploratory Data Analysis

## I. INTRODUCTION

Clinical document classification is one of the research areas on healthcare domain. It has potential to improve operational efficiency and research pertaining to useful insights for decision m aking. Clinical document classification is useful at the level of healthcare units, state and central governments. It is also useful at global scale to understand disease and patients' life style patterns and help in making policy decisions. As the AI became popular ML models are widely used to solve real world problems. Clinical document classification is the problem that could be solved using ML models. However, existing methods in this area lack a systematic approach in filtering out false positives. There are many contributions found in the literature that are based on ML models.

Emilio *et al*. [3] used radiological reports for their empirical study. They used ML models to perform classification of such records useful to hospitals. Bevan *et al*. [6] proposed a hybrid ML model to detect cancer and come up with cancer statistics from death certificates. They used rule based phenomena for automatic detection of cancer and render its statistics. Zaid *et al*. [8] proposed a pipeline of deep learning models to detect cervical cancer from documents and classify the disease. Lin *et al*. [11] proposed a methodology based on deep learning for automatic analysis of heterogeneous medical data. Alzubaidi *et al*. [14] proposed an application based on deep learning tools to study medical images. Their study could help in discovering new knowledge from the medical images. Gupta *et al*. [18] used deep learning to make a histopathological model for image classification for distributed environments. From the literature, it is observed that pruning documents prior to classification is important. This insight is incorporated in the proposed system of this paper. Our contributions are as follows.

1. We proposed a ML framework that considers pipelining of ML models at multiple levels.
2. We proposed an algorithm known as Learning based Clinical Document Classification (LbCDC).
3. We built an application for evaluation of our system. Our experimental results revealed that the proposed system is efficient in clinical document classification.

The remainder of the paper is structured as follows. Section 2 reviews prior works on clinical document classification. Section 3 presents the proposed system. Section 4 presents results of our experiments. Section 5 concludes our work and throw light on future scope

## II. RELATED WORK

This section reviews prior works pertaining to clinical document classification based on machine learning techniques.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

Heath *et al*. [1] focused on automatic classification of health records in medical domain. Jahanzaib *et al*. [2] used ML models to classify electronic medical records in healthcare domain. It has modus operandi to fulfill this use case towards automation. Emilio *et al*. [3] used radiological reports for their empirical study. They used ML models to perform classification of such records useful to hospitals. Jonathan *et al*. [4] investigated on possibilities in healthcare domain with the emergence of AI based approaches. They found the utility of ML models for solving problems in medical domain. Eli *et al*. [5] proposed a deep learning model known as NiftyNet that is designed to analyze medical imagery. Bevan *et al*. [6] proposed a hybrid ML model to detect cancer and come up with cancer statistics from death certificates. They used rule based phenomena for automatic detection of cancer and render its statistics. Victor *et al*. [7] proposed a deep learning model that works in two stages for finding entities in medical documents and then find relationships among important medical texts. Zaid *et al*. [8] proposed a pipeline of deep learning models to detect cervical cancer from documents and classify the disease. Yunlu *et al*. [9] used deep learning based method for automatic classification of documents based on respiratory patterns in the patients' related documents. Jihad *et al*. [10] use clinical notes from medical domain and investigated on finding mental status of patients.

Lin *et al*. [11] proposed a methodology based on deep learning for automatic analysis of heterogeneous medical data. Jain *et al*. [12] explored different ML tools available for automatically finding useful information and solve problems in different domains. Pérez-García *et al*. [13] proposed a library for medical image analysis using Python language. It is developed for various ML approaches including pre-processing, sampling and classification. Alzubaidi *et al*. [14] proposed an application based on deep learning tools to study medical images. Their study could help in discovering new knowledge from the medical images. Tan *et al*. [15] proposed deep learning methodology for automatic monitoring of cardiovascular diseases in real time. Min *et al*. [16] used deep learning approach for automatic detection of presence of diseases in human samples. Pandey *et al*. [17] investigated on medical imaging and analysis using ML models. They could throw light on possible challenges and opportunities. Gupta *et al*. [18] used deep learning to make a histopathological model for image classification for distributed environments. Yasar *et al*. [19] investigated on deep learning models to understand medical documents and detect cases of Covid-19. Nathan *et al*. [20] studied hematological malignancies in medical field using ML models. Different methodologies are found in [21]- [30] for medical document analysis. From the literature, it is observed that pruning documents prior to classification is important. This insight is incorporated in the proposed system of this paper.

## III. PROPOSED SYSTEM

This section presents the proposed system that includes a ML framework, different ML models, dataset and performance evaluation procedure.

### *A. The Framework*

We proposed a framework, shown in Figure 1, for automatic classification of clinical documents. The framework is based on machine learning. The framework is designed to have multiple ML pipelines. Model 1 pipeline is designed to know whether given clinical document corpus has any information related to smoking habit of patients. This pipeline is meant for discriminating the clinical documents that have some text related to patient smoking. Thus in this model, the documents that have smoking information are retained and the rest are discarded. The retained documents are used in model 2 pipeline which is designed to classify clinical documents into smoked and never smoked categories. The smoked category of documents is retained while the other documents discarded. The smoked category of documents is considered in the model 3 pipeline where the system classifies clinical documents into two categories such as presently smoking and past smoker

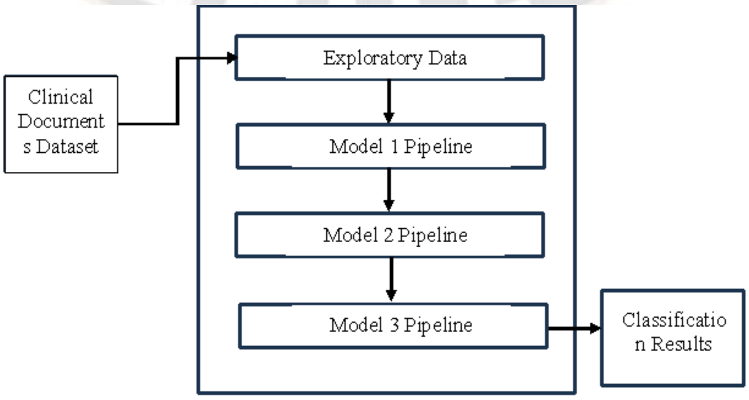



Figure 1: Overview of proposed ML based framework for clinical document classification

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

After completion of all pipelines, the final document classification results contain two categories such as presently smoking and past smokers. Different ML models are used in the proposed framework. In other words, each pipeline has set of ML models such Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF) and Gradient Boosting (GB). LR is a ML model that is based on supervised learning. Eq. 1 reflects logistic model obtained from logistic function.

$$Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \qquad (1)$$

The linier sum of Eq. 1 is then substitute with a function and the resultant expression is provided in Eq. 2.

$$f(z) = \frac{1}{1+e^{-z}} \qquad (2)$$

The expression is formulated as a mathematical model. The probability is modelled reflecting conditional probability as expressed in Eq. (3).

$$P(D = 1|X_1, X_2, \dots, X_k) = P(X) \qquad (3)$$

Finally, the logistic model is constructed and expressed as in Eq. 4. The logistic model is used for classification of medical documents in the proposed framework.

$$P(X) = \frac{1}{1+e^{-(\alpha+\sum \beta_i X_i)}} \qquad (4)$$

NB model is also widely used model for classification in real world applications. It is based on Bayes Theorem. In fact, it is a collection of algorithms meant for solving classification problems. The NB model is based on the expression in Eq. 5.

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)} \qquad (5)$$

In terms of input and output variables X and y, it is possible to rewrite the above expression as in Eq. 6.

$$P(y/X) = \frac{P(X/y) * P(y)}{P(X)} \qquad (6)$$

Because of the naive assumption that variables are independent given the class, we can rewrite $P(X/y)$ as follows:

$$P(X/y) = P(x_1, y) * P(x_2, y) * \dots * P(x_n, y) \qquad (7)$$

The goal of NB is to detect classes with highest probability. This model is widely used in classification tasks as it has potential to discriminate class labels. Thus the NB model finally focuses on maximum y value as expressed in Eq. 8.

$$y = argmax_y [P(y) * \prod_{i=1}^{n} P(x_i, y)] \qquad (8)$$

Random Forest (RF) is yet another model that is used extensively in many ML applications. It has many underlying decision trees that form an ensemble of models to improve classification performance. Gini index is often used in RF for branching nodes in different trees of RF as expressed in Eq. 9.

$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2 \qquad (9)$$

Probability and class are used to determine to know Gini of each branch. It is also possible to exploit entropy to determine branching as expressed in Eq. 10.

$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i) \qquad (10)$$

Random forest makes use of ensemble approach with voting in order to arrive at final classification results. Gradient boosting model is also based on decision tree. The tree can be modelled as in Eq. 11.

$$h_m(x) = \sum_{j=1}^{T} b_{jm} l_{Rjm}(x) \qquad (11)$$

Finally, Eq. 12 and Eq. 13 provide the modus operandi of the GB model used in the proposed framework.

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^{T} \alpha_{jm} l_{Rjm}(x) \qquad (12)$$

$$\alpha_{jm} = argmin_\alpha \sum_{x_i \epsilon R_{jm}} L(y_i, f_{m-1}(x_i) + \alpha) \qquad (13)$$

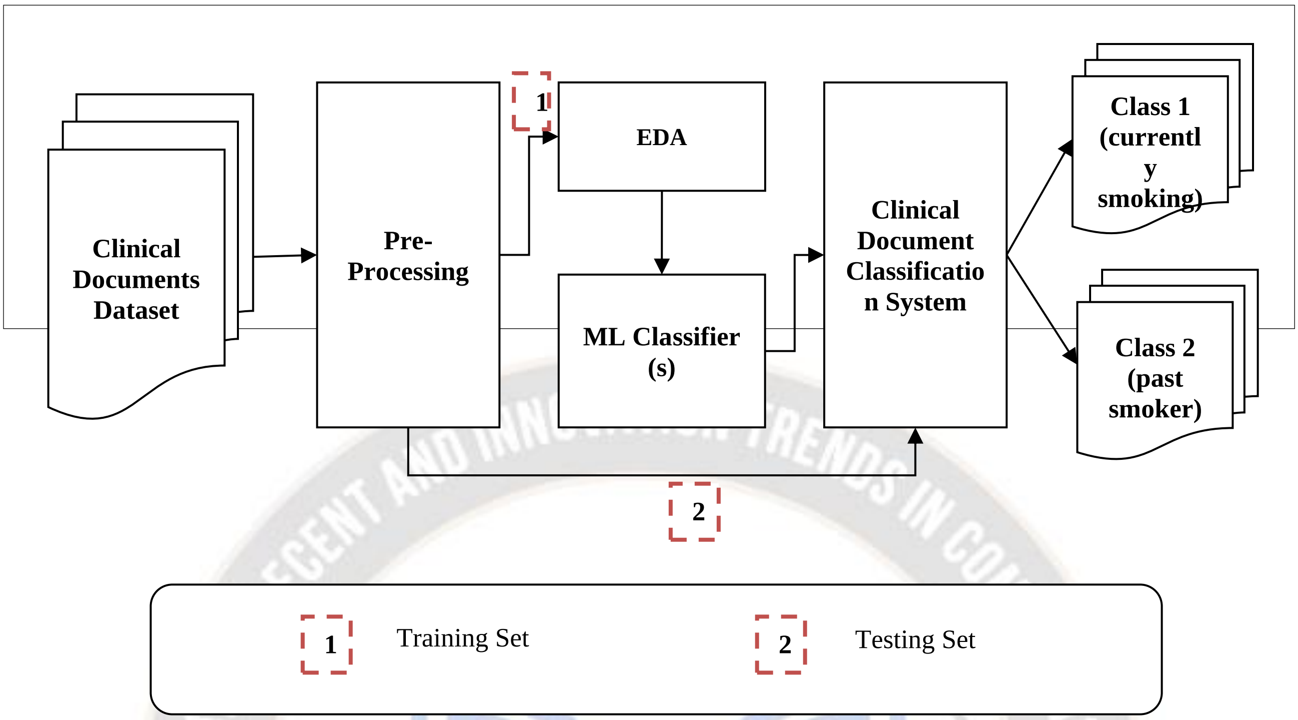
The GB model needs computation of pseudo residual, alpha and also regularization mechanism in order to get rid of overfitting.

| **Notation** | **Meaning** |
|---|---|
| $\alpha$ *and* $\beta_i$ | Parameters to be estimated |
| $P(X)$ | Denotes $P(D = 1|X_1, X_2, \dots, X_k)$ |
| $X$ | Set of variables |
| P(X) | Denotes probability dynamics |
| **Argmax** | Denotes argument with maximum value |
| $y_i$ | Testing data point |
| $f_i$ | Return value in decision tree |
| $p_i$ | Relative frequency value |
| Entropy | Measure used for branching of tree |
| Y | Actual value |
| Gamma | Value predicted |
| L | Denotes loss function |
| f(x) | Denotes a function |

Table 1: Shows notations used in this paper

The ML models in the proposed system are based on supervised learning and their approach in the classification of clinical documents is as in Figure 2.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*




**Algorithm:** Learning based Clinical Document Classification (LbCDC)

**Inputs:** Clinical documents dataset D, ML models M

**Output:** Clinical document classification results R

1. Begin
**1. Model-1 Pipeline**
2. (T1, T2)←PreProcess(D)
3. For each model m in M
4. Train m with T1
5. R←Classify T2 using m
6. Display R
7. D'←Assign docs with text on smoking
8. End For
**2. Model-2 Pipeline**
9. For each model m in M
10. Train m with m
11. R←Classify D' using m
12. Display R
13. Update D' with docs containing smoking positives
14. End For
**3. Model-3 Pipeline**
15. For each model m in M
16. Train m with m
17. R←Classify D' using m
18. Display R
19. End For
20. End

**FIGURE 2:** CLINICAL DOCUMENT CLASSIFICATION METHODOLOGY

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

As presented in Figure 2, each classification model used in the proposed framework (for all model pipelines), the modus operandi are provided. The given dataset is subjected to preprocessing and then divided into 80% training and 20% testing. Afterwards, the training data is subjected to EDA to determine any requirement for changes in the dataset such as treating missing values. Then the training data is used by ML model to gain knowledge from the training samples. Based on the knowledge gained, the model is further used to classify clinical documents.

*B. Dataset Details*

Clinical documents dataset [31], known as n2c2, is collected from https://portal.dbmi.hms.harvard.edu/. The dataset contains 502 patients' discharge summaries covering various details. The documents do have smoking status which is the main focus point in this research for document classification.

*C. Algorithm Design*

We proposed an algorithm known as Learning based Clinical Document Classification (LbCDC). This algorithm makes use of three models in pipeline in order to perform classification of clinical documents at multiple levels of granularity.

**Algorithm 1:** Learning based Clinical Document Classification (LbCDC)

As presented in Algorithm 1, it takes clinical documents dataset and ML models as input and performs classification in multi-level approach. There is pipelining of models in order to filter out documents in model 1 and model 2 and finally arrive at only documents that do have details of patients with smoking habit. Then in the third model pipeline, the documents are classified into two categories such as currently smoking and past smoker.

*D. Performance Evaluation*

The proposed algorithm is evaluated using different performance measures obtained through the cases provided in confusion matrix as shown in Figure 3.

In this paper, the clinical document classification performance of ML models is made by comparing prediction results with ground truth. True positive indicates a positive sample is correctly detected by the proposed algorithm. True negative indicates a negative sample is correctly detected by the algorithm. False positive indicates that a negative sample is detected as positive sample by the algorithm. In the same fashion, false negative indicates that a positive sample is detected as negative sample by the algorithm.

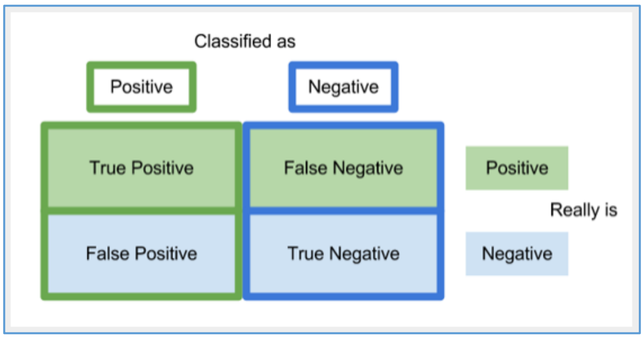



Figure 3: Confusion matrix

$$\text{Precision (p)} = \frac{TP}{TP+FP} \quad (14)$$

$$\text{Recall (r)} = \frac{TP}{TP+FN} \quad (15)$$

$$\text{F1-Score} = 2 * \frac{(p * r)}{(p+r)} \quad (16)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (17)$$

Based on confusion matrix cases, different metrics are derived and used for performance evaluation. Precision, recall and accuracy are used to measure capability of ML models. These metrics result in a value between 0 and 1 reflecting least performance to highest performance.

## IV. EXPERIMENTAL RESULTS

Experimental results are provided in this section for all the three model pipelines. For each underlying ML model, confusion matrix is provided as proof of performance followed by ROC curve and performance comparison of all models.

*A. Model-1 Pipeline*

This subsection presents experimental results in terms of confusion matrix, ROC curve and clinical document classification performance measures for model-1 pipeline.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

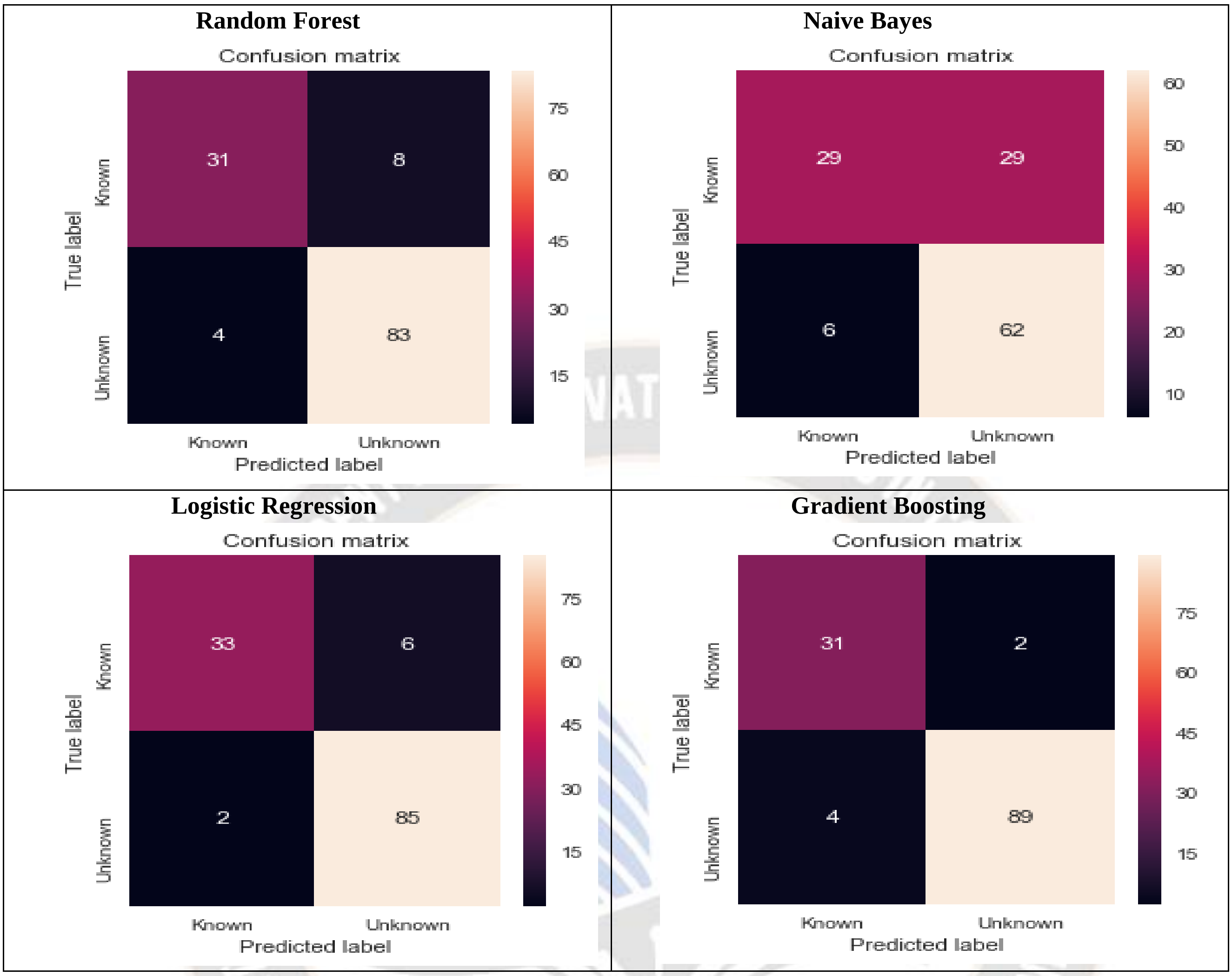



Figure 4: Model-1 pipeline confusion matrix

As presented in Figure 4, each model's confusion matrix is provided reflecting predicted and true label dynamics.

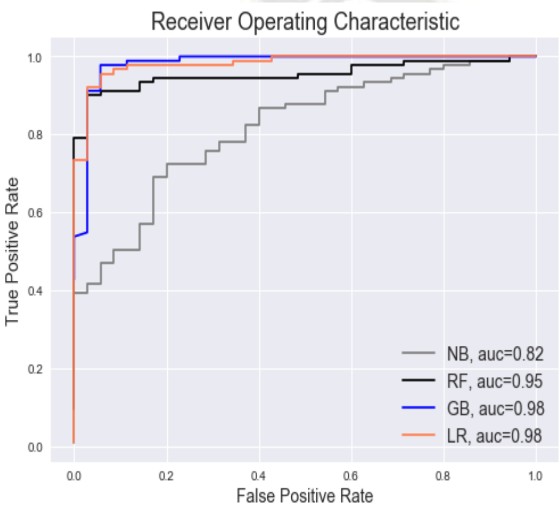



Figure 5: ROC curve for model-1 pipeline

As presented in Figure 5, ROC curve provides performance of different ML models in clinical document classification.

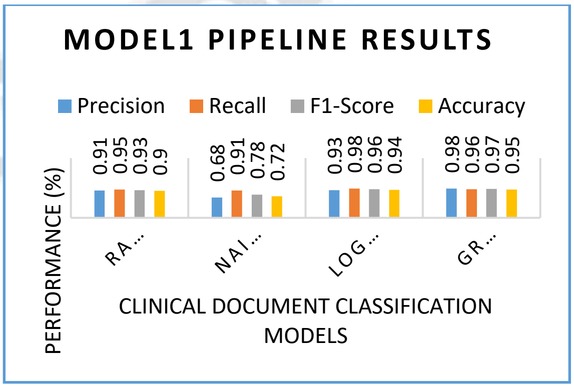



Figure 6: Shows performance of models in model-1 pipeline

As presented in Figure 6, each model is evaluated for the purpose of model-1 pipeline classification. Different performance metrics are used to evaluate the models. Higher in value for any measure indicates better performance. With

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

respect to accuracy, RF showed 90% accuracy, NB 72%, LR 94% and GB 95%. From the results, it is understood that GB showed highest performance in model-1 pipeline.

## *B. Model-2 Pipeline*

This subsection presents experimental results in terms of confusion matrix, ROC curve and clinical document classification performance measures for model-2 pipeline.

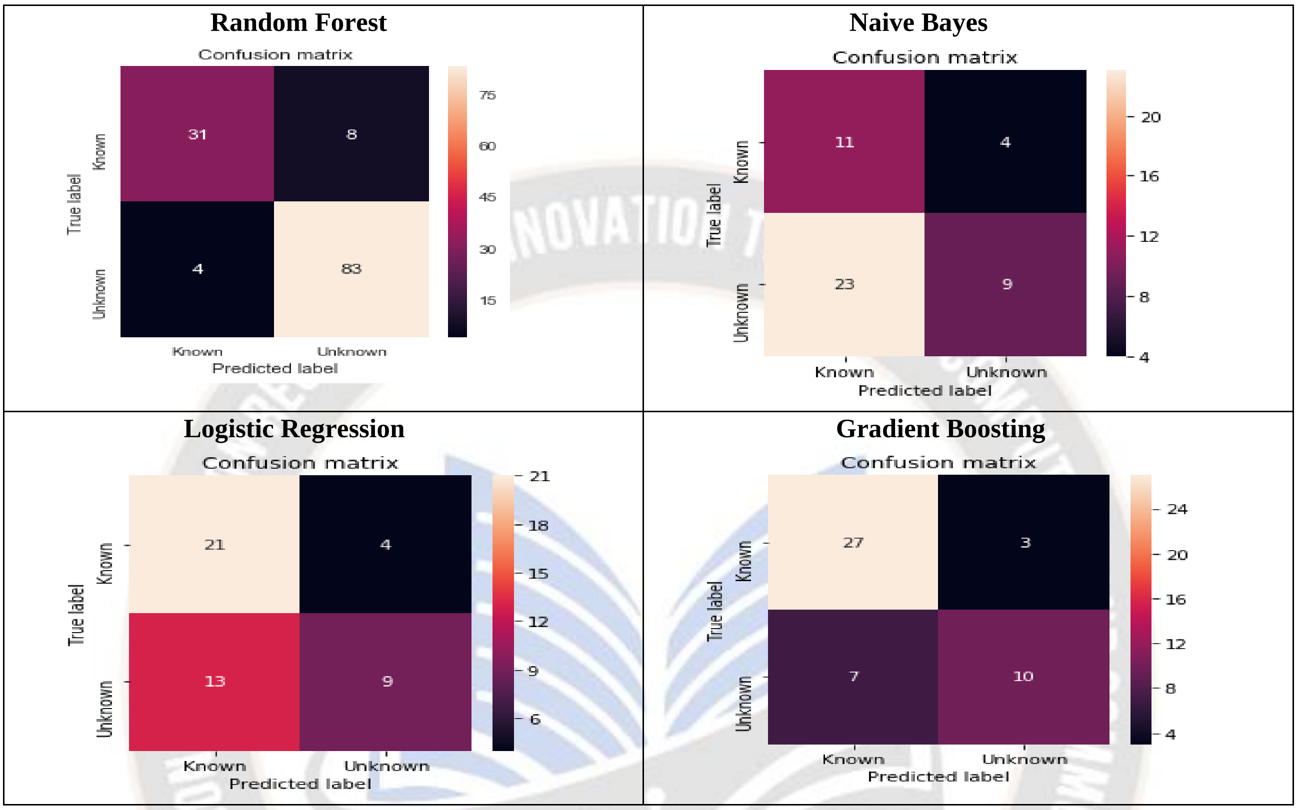



Figure 7: Model-2 pipeline confusion matrix

As presented in Figure 7, each model's confusion matrix is provided reflecting predicted and true label dynamics.

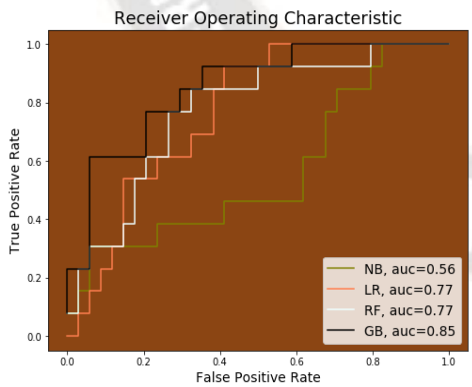



Figure 8: ROC curve for model-2 pipeline

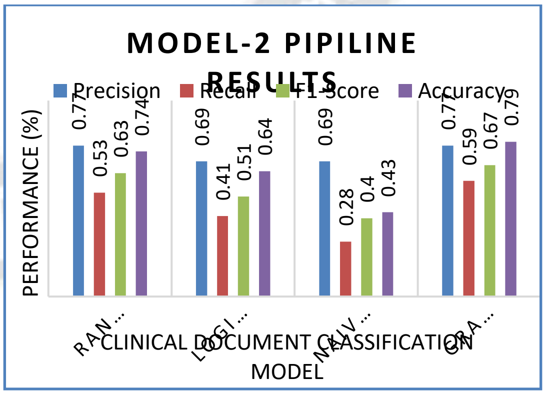



Figure 9: Shows performance of models in model-2 pipeline

As presented in Figure 8, ROC curve provides performance of different ML models in clinical document classification in model-2 pipeline.

As presented in Figure 9, each model is evaluated for the purpose of model-2 pipeline classification. Different performance metrics are used to evaluate the models. Higher in value for any measure indicates better performance. With respect to accuracy, RF showed 74% accuracy, NB 43%, LR

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 10***
***DOI: https://doi.org/10.17762/ijritcc.v11i10.8512***
***Article Received:** 29 July 2023 **Revised:** 26 September 2023 **Accepted:** 12 October 2023*

64% and GB 79%. From the results, it is understood that GB showed highest performance in model-2 pipeline.

### *C. Model-3 Pipeline*

This subsection presents experimental results in terms of confusion matrix, ROC curve and clinical document classification performance measures for model-3 pipeline.

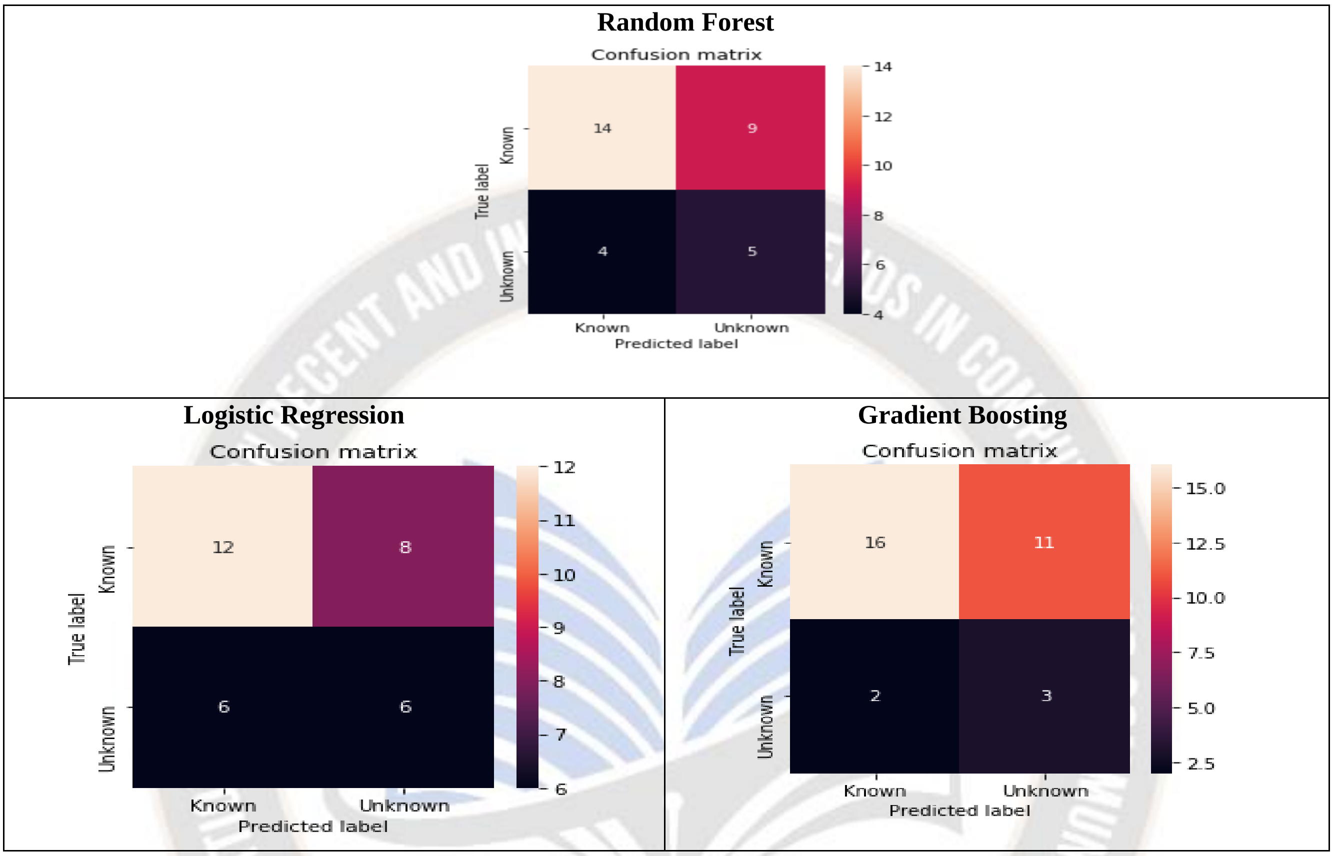



Figure 10: Model-3 pipeline confusion matrix

As presented in Figure 10, each model's confusion matrix is provided reflecting predicted and true label dynamics.

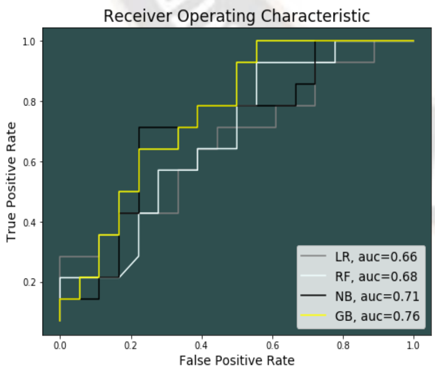



Figure 11: ROC curve for model-3 pipeline

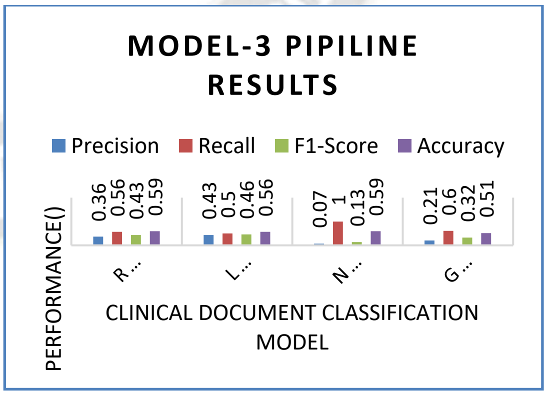



Figure 12: Shows performance of models in model-3 pipeline

As presented in Figure 11, ROC curve provides performance of different ML models in clinical document classification in model-3 pipeline.

As presented in Figure 12, each model is evaluated for the purpose of model-3 pipeline classification. Different performance metrics are used to evaluate the models. Higher in value for any measure indicates better performance. With respect to accuracy, RF showed 59% accuracy, NB 59%, LR

56% and GB 51%. From the results, it is understood that RF and NB showed highest performance in model-3 pipeline.

## V. CONCLUSION AND FUTURE WORK

In this paper we proposed a ML framework that considers pipelining of ML models at multiple levels. The framework is designed to have multiple ML pipelines. Model 1 pipeline is designed to know whether given clinical document corpus has any information related to smoking habit of patients. This pipeline is meant for discriminating the clinical documents that have some text related to patient smoking. Thus in this model, the documents that have smoking information are retained and the rest are discarded. The retained documents are used in model 2 pipeline which is designed to classify clinical documents into smoked and never smoked categories. The smoked category of documents is retained while the other documents discarded. The smoked category of documents is considered in the model 3 pipeline where the system classifies clinical documents into two categories such as presently smoking and past smoker. We proposed an algorithm known as Learning based Clinical Document Classification (LbCDC). Our experimental results revealed that the proposed system is efficient in clinical document classification. In future, we intend to improve our framework with deep learning techniques.

## REFERENCES

[1] Goodrum, Heath; Roberts, Kirk and Bernstam, Elmer V. (2020). Automatic classification of scanned electronic health record documents. International Journal of Medical Informatics, 144, 104302–. http://doi:10.1016/j.ijmedinf.2020.104302

[2] Latif, Jahanzaib; Xiao, Chuangbai; Tu, Shanshan; Rehman, Sadaqat Ur; Imran, Azhar and Bilal, Anas (2020). Implementation and Use of Disease Diagnosis Systems for Electronic Medical Records Based on Machine Learning: A Complete Review. IEEE Access, 1–1. http://doi:10.1109/ACCESS.2020.3016782

[3] Gerevini, Alfonso Emilio; Lavelli, Alberto; Maffi, Alessandro; Maroldi, Roberto; Minard, Anne-Lyse; Serina, Ivan and Squassina, Guido (2018). Automatic classification of radiological reports for clinical care. Artificial Intelligence in Medicine, S0933365717305912–. http://doi:10.1016/j.artmed.2018.05.006

[4] Waring, Jonathan; Lindvall, Charlotta and Umeton, Renato (2020). Automated Machine Learning: Review of the State-of-the-Art and Opportunities for Healthcare. Artificial Intelligence in Medicine, 101822–. http://doi:10.1016/j.artmed.2020.101822

[5] Gibson, Eli; Li, Wenqi; Sudre, Carole; Fidon, Lucas; Shakir, Dzhoshkun I.; Wang, Guotai; Eaton-Rosen, Zach; Gray, Robert; Doel, Tom; Hu, Yipeng; Whyntie, Tom; Nachev, Parashkev; Modat, Marc; Barratt, Dean C.; Ourselin, Sébastien; Cardoso, M. Jorge and Vercauteren, Tom (2018). NiftyNet: a deep-learning platform for medical imaging. Computer Methods and Programs in Biomedicine, S0169260717311823–. http://doi:10.1016/j.cmpb.2018.01.025

[6] Koopman, Bevan; Zuccon, Guido; Nguyen, Anthony; Bergheim, Anton and Grayson, Narelle (2018). Extracting cancer mortality statistics from death certificates: A hybrid machine learning and rule-based approach for common and rare cancers. Artificial Intelligence in Medicine, S0933365717301173–. http://doi:10.1016/j.artmed.2018.04.011

[7] Suárez-Paniagua, Víctor; Rivera Zavala, Renzo M.; Segura-Bedmar, Isabel and Martínez, Paloma (2019). A two-stage deep learning approach for extracting entities and relationships from medical texts. Journal of Biomedical Informatics, 99, 103285–. http://doi:10.1016/j.jbi.2019.103285

[8] Alyafeai, Zaid and Ghouti, Lahouari (2019). A Fully-Automated Deep Learning Pipeline for Cervical Cancer Classification. Expert Systems with Applications, 112951–. http://doi:10.1016/j.eswa.2019.112951

[9] Wang, Yunlu; Hu, Menghan; Zhou, Yuwen; Li, Qingli; Yao, Nan; Zhai, Guangtao; Zhang, Xiao-Ping and Yang, Xiaokang (2020). Unobtrusive and Automatic Classification of Multiple Peopleâ€™s Abnormal Respiratory Patterns in Real Time using Deep Neural Network and Depth Camera. IEEE Internet of Things Journal, 1–1. http://doi:10.1109/JIOT.2020.2991456

[10] Obeid, Jihad S.; Weeda, Erin R.; Matuskowitz, Andrew J.; Gagnon, Kevin; Crawford, Tami; Carr, Christine M. and Frey, Lewis J. (2019). Automated detection of altered mental status in emergency department clinical notes: a deep learning approach. BMC Medical Informatics and Decision Making, 19(1), 164–. http://doi:10.1186/s12911-019-0894-9

[11] Yue, Lin; Tian, Dongyuan; Chen, Weitong; Han, Xuming and Yin, Minghao (2020). Deep learning for heterogeneous medical data analysis. World Wide Web. http://doi:10.1007/s11280-019-00764-z

[12] Marshall, Iain J. and Wallace, Byron C. (2019). Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Systematic Reviews, 8(1). http://doi:10.1186/s13643-019-1074-9

[13] Fernando Pérez-García; Rachel Sparks and Sébastien Ourselin; (2021). TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning . Computer Methods and Programs in Biomedicine. http://doi:10.1016/j.cmpb.2021.106236

[14] Laith Alzubaidi; Mohammed A. Fadhel; Omran Al-Shamma; Jinglan Zhang; J. Santamaría and Ye Duan; (2021). Robust application of new deep learning tools: an experimental study in medical imaging . Multimedia Tools and Applications. http://doi:10.1007/s11042-021-10942-9

[15] Liang Tan; Keping Yu; Ali Kashif Bashir; Xiaofan Cheng; Fangpeng Ming; Liang Zhao and Xiaokang Zhou; (2021). Toward real-time and efficient cardiovascular monitoring for COVID-19 patients by 5G-enabled wearable medical

devices: a deep learning approach . Neural Computing and Applications. http://doi:10.1007/s00521-021-06219-9

[16] Li, Min; Fei, Zhihui; Zeng, Min; Wu, Fangxiang; Li, Yaohang; Pan, Yi and Wang, Jianxin (2018). Automated ICD-9 Coding via A Deep Learning Approach. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 1–1. http://doi:10.1109/TCBB.2018.2817488

[17] Babita Pandey; Devendra Kumar Pandey; Brijendra Pratap Mishra and Wasiur Rhmann; (2021). A comprehensive survey of deep learning in the field of medical imaging and medical natural language processing: Challenges and research directions . Journal of King Saud University - Computer and Information Sciences. http://doi:10.1016/j.jksuci.2021.01.007

[18] Koyel Datta Gupta; Deepak Kumar Sharma; Shakib Ahmed; Harsh Gupta; Deepak Gupta and Ching-Hsien Hsu; (2021). A Novel Lightweight Deep Learning-Based Histopathological Image Classification Model for IoMT . Neural Processing Letters. http://doi:10.1007/s11063-021-10555-1

[19] Yasar, Huseyin and Ceylan, Murat (2020). A novel comparative study for detection of Covid-19 on CT lung images using texture analysis, machine learning, and deep learning methods. Multimedia Tools and Applications. http://doi:10.1007/s11042-020-09894-3

[20] Radakovich, Nathan; Nagy, Matthew and Nazha, Aziz (2020). Machine learning in haematological malignancies. The Lancet Haematology, 7(7), e541–e550. http://doi:10.1016/S2352-3026(20)30121-6

[21] Arora, Ridhi; Rai, Prateek Kumar and Raman, Balasubramanian (2020). Deep featureâ€ "based automatic classification of mammograms. Medical & Biological Engineering & Computing. http://doi:10.1007/s11517-020-02150-8

[22] Vellido, Alfredo (2019). The importance of interpretability and visualization in machine learning for applications in medicine and health care. Neural Computing and Applications. http://doi:10.1007/s00521-019-04051-w

[23] L, Arokia Jesu Prabhu; Sengan, Sudhakar; G K, Kamalam; J, Vellingiri; Gopal, Jagadeesh; Velayutham, Priya and V, Subramaniyaswamy (2020). Medical information retrieval systems for e-Health care records using fuzzy based machine learning model. Microprocessors and Microsystems, 103344–. http://doi:10.1016/j.micpro.2020.103344

[24] Zabihollahy, Fatemeh; Schieda, Nicola; Krishna, Satheesh and Ukwatta, Eranga (2020).

[27] Jingyu Zhong, Chengxiu Zhang, Yangfan Hu, Jing Zhang, Yun Liu, Liping Si1, Yue Xing , Defang Ding, Jia Geng, Qiong Jiao, Huizhen Zhang, Guang Yang and Weiwu Yao (2022). Automated prediction of the neoadjuvant chemotherapy response in osteosarcoma with deep learning and an MRI-based radiomics nomogram. *Springer*, pp.1-11. https://doi.org/10.1007/s00330-022-08735-1

[28] CHERUBIN MUGISHA AND INCHEON PAIK. (2022). Comparison of Neural Language Modeling Pipelines for Outcome Prediction From Unstructured Medical Text Notes. *IEEE*. 10, pp.16489-16498. http://DOI:10.1109/ACCESS.2022.3148279

[29] Narmin Ghaffari Laleh, Hannah Sophie Muti, Chiara Maria Lavinia Loeffler, Amelie Echlea , Oliver Lester Saldanha , Faisal Mahmood , Ming Y. Lu , Christian Trautwein , Rupert Langer, Bastian Dislich, Roman D. Buelow, Heike Irmgard Grabsch, Hermann Brenner, Jenny Chang-Claude, Elizabeth Alwers, Titus J. Brinker, Firas Khader, Daniel Truhnn, Nadine T. Gaisa, Peter Boor, Michael Hoffmeister, Volkmar Schulz, Jakob Nikolas Kather. (2022). Benchmarking weakly-supervised deep learning pipelines for whole slide classification in computational pathology. *Elsevier*., pp.1-15. https://doi.org/10.1016/j.media.2022.102474

[30] Stefan Grafberger, Paul Groth, Julia Stoyanovich and Sebastian Schelter. (2022). Data distribution debugging in machine learning pipelines. *Springer*, pp.1-24. https://doi.org/10.1007/s00778-021-00726-w

[31] Clinical documents dataset. Retrieved from https://portal.dbmi.hms.harvard.edu/