UtilitasMathematica

ISSN 0315-3681 Volume 122, 2025

A Novel Thyroid Nodule Recurrence Prediction Using Machine Learning Techniques

Rapolu Chandana[1*], T. Veda Reddy[2], Dr. G. Vishnu Murthy[3]

[2]Associate Professor, Department of CSE, Anurag University, Hyderabad, India

[3]Professor and Dean, Department of CSE, Anurag University, Hyderabad, India

Email: rapoluchandana3@gmail.com

***Abstract*—One of the most prevalent endocrine diseases is represented by thyroid nodules, and most of them are benign. Nevertheless, there exists a large sub-population that has a great potential of recurring even after the initial treatment especially those related to the papillary thyroid carcinoma (PTC). It is essential to identify patients who are at increased risk of recurrence so that to intervene medically in time and manage patients individually. The conventional diagnostic procedures and risk stratification models frequently fail to detect sophisticated patterns in a multi-dimensional medical data.**

This paper introduces a novel machine learning strategy of predicting whether thyroid nodules will come back or not using structured data as clinical and pathological information. Several supervised learning algorithms were tested, such as Random Forest, Gradient Boosting, Support Vector Machine, or LightGBM. Of them, the LightGBM model achieved the best results with the accuracy of 97.40 percent and the ROC-AUC of 99.17 percent. Age, gender, smoking history, thyroid functions, focal vs. diffuse nature of the tumor, tumor stage and response to initial treatment are some of the features that were used during model training.

The usage of data-based methods in the prediction of thyroid cancer makes this study clinically applicable to the development of personal treatment plans and long-term post-therapy monitoring. The given model can catalyze the performance of endocrinologists and oncologists in terms of decision-making process, which will lead to patient outcomes.

Keywords—Thyroid Nodules, Recurrence Prediction, Machine Learning, LightGBM, Papillary Thyroid Carcinoma, TNM Staging, Clinical Decision Support.

I. Introduction

Thyroid nodules are the small discreet lesion in the thyroid which are frequently encountered during the routine imaging tests or during physical examinations. These nodules constitute one of the most common endocrine disorders especially in the areas deprived of iodine supply and in women as well as the aged. Through research, the number of thyroid nodules has been seen to greatly rise with the prevalence of high-resolution ultrasonography which is estimated as between 19 percent and 68 percent of the population [1].

Whereas most thyroid nodules are benign and do not cause any symptoms, about 515

Effective predictive modeling of recurrence is important in all aspects of follow up to instruct the follow up plan, frequency of imaging and whether one needs further therapy. The clinicians traditionally utilize the structured scoring systems, among which the American Thyroid Association (ATA) risk stratification guidelines are used, and they take into account such features as tumor size, involvement of lymph nodes, histopathology, and presence of extrathyroidal extension [4]. These approaches, however, tend to be unable to capture complex dependencies between two or more variables outright and are at risk of interobserver variation in interpretation.

The introduction of artificial intelligence (AI) and machine learning (ML) technologies has already led to the paradigm shift in the analysis and interpretation of medical data in the healthcare industry. Machine learning algorithms are well suited to analyzing high-dimensional massive data and identification of complex, latent patterns that cannot be found using conventional statistical tools [5]. The technology has in oncology and specifically thyroid cancer demonstrated the potential of enhancing diagnostic values, risk stratification, and prediction of treatment outcome [7].

Recent research has shown that the ML algorithms can be successfully applied to the forecast of malignancy in thyroid nodules on ultrasound pictures, cytological indicators, and genomics data, among others [8]. Nonetheless, not much work had been done on predicting the recurrence of thyroid nodules, an equally crucial step as far as long-term management is concerned of patients. Individualized follow-up schedules, minimization of unnecessary procedures, and quality life of patients can be the results of a properly developed recurrence prediction model using ML.

The study is based on the new methodology to predict the risk of a thyroid nodule recurrence with the help of machine learning techniques. We apply an abundant source of clinical and pathological characteristics (i.e. age, gender, smoking history, thyroid functional status, physical examination findings, focality of the tumor, pathological subtype, TNM staging, and response to the initial treatment). Various ML algorithms were analysed, and our findings have shown that Light Gradient Boosting Machine (LightGBM) model-based classifier has the best performance as a predictor with an accuracy level of 97.40 percent and ROC-AUC rate at 99.17 percent.

The value of the work is that it helps not to remain at the level of referencing by isolated clinical guidelines but rather proposes a dynamic, individualized risk discourse using the real data concerning patients. Also, working with interpretable machine learning tools will allow clinicians to realize the most important features that lead to occurring risk. This model can be combined with clinical expertise to become a decision-support tool in real practice, both in endocrinology and oncology.

II. Literature Survey

There has been much research on the prediction and management of thyroid nodules especially relating to the issue of recurrence which is done both through the classical statistical approach as well as the recent approach, machine learning. The earlier studies mainly only involved the application of ultrasonography and cytology-based diagnostic techniques which mostly depended on the expertise of the radiologist and the subjective interpretation of the problem. Gharib and co-authors extensively described the clinical assessment and follow-up of the thyroid nodules recommending the technique of fine-needle aspiration cytology (FNAC) and scoring of ultrasoun through ultrasoun based [1].

The development of the molecular testing and imaging techniques has enhanced the accuracy in diagnosis but the ability to predict the recurrence remains unreliable. Mazzaferri emphasized that in some cases, even patients with low risk thyroid cancer could relapse several years later, and this is a reason why the follow up procedures should be long-term based [6]. In another follow-up study carried ten years or more after, Durante et al. reported that more than 13 percent of those with papillary thyroid carcinoma (PTC) have a recurrence again showing that more predictable and personized models are necessary [3].

The last several years are characterized by the use of machine learning (ML) application in different ways of predicting thyroid cancer. Song et al. mine three types of models such as support vector machines (SVM), decision trees and the artificial neural networks to predict the recurrence of PTC patients and they achieved better accuracy than traditional scoring systems [8]. They used clinical evidence in their model along with tumor size and extrathyroidal extension, lymph node status, showing that ML can reveal nonlinear relationships that may otherwise be overlooked by more conventional measures.

Deep learning, a subfield of ML, has also shown promise. Esteva et al. emphasized the growing influence of deep learning models in healthcare diagnostics, particularly for image and pattern recognition tasks in complex datasets [5]. In thyroidology, CNN-based models have been explored for ultrasound image classification, offering reliable malignancy prediction [9]. However, most of these approaches focus on initial diagnosis rather than post-treatment recurrence.

Liu et al. designed the random forest model that can determine the disease free survival in thyroid cancer patients through the disease oriented EHR (Electronic Health Record) data. They demonstrated that ensemble learning methods were capable of surpassing logistic regressions and Cox regressions as a method of survival prediction [10]. In a similar case, Chen et al. used gradient boosting decision tree to stratify the thyroid nodules and improved the classification accuracy remarkably [11].

An excellent study by Ko et al. was another mentionable study and applied a hybrid feature selection method alongside XGBoost to recurrence forecasting and obtained excellent AUC scores. One of the conclusions of their study was that feature engineering and model tuning play an essential role in clinical prediction tasks [12]. Moreover, Fang et al. evaluated the performance of LightGBM as the predictor of lymph node metastasis and recurrence in PTC and reported promising accuracies and interpretability [13].

Even with these approached, very few studies have addressed the task of predicting recurrence when large and highly variable types of features are present. The majority of the current models concentrate on a limited number of variables or do not take into account the response and outcomes of the treatment. Our suggested work will tackle these flaws since it includes many clinical, demographic, pathological, and treatment-response characteristics which means that better and more thorough expectation of recurrence among thyroid nodules patients is likely to be predicted.

Other researchers have paid several studies in the integration between clinical and imaging data to enhance prognostication of thyroid cancer. To give an example, Zhang et al. have put forward the multimodal deep learning model integrating ultrasound images and electronic medical records (EMRs) to obtain a deeper assessment of thyroid nodule malignancy [14]. Their findings revealed that ML models with multimodal fusion can enhance significantly the predictive capacity in comparison with unimodal models.

In addition, the utility of patient demographics and historical data in improving predictive performance in recurrence has been also displayed by Wang et al. who employed XGBoost models trained based on large thyroid cancer registries. Their result showed that other characteristics important in long-term recurrence risk included age, gender, exposure history radiation, and nodule focality [15].

One more area of development is interpretable machine learning. The black-box models are accurate but usually lack transparency. A recent study by Lundberg and Lee proposed SHAP (SHapley Additive exPlanations) which is more interpretable as it attributes importance to every input feature [16]. Using SHAP on thyroid nodule recurrence will enable clinicians to know what aspects influence their specific predictions the most and, therefore, raise the level of clinical trust and model acceptance.

Besides, there is the issue of clinical data imbalance. Recurrence cases tend to be much smaller than non-recurrence cases in thyroid recurrence database sets. Choi et al. solved this and reported the efficiency of the SMOTE (Synthetic Minority Over-sampling Technique) to enhance the sensitivity of the models without decreasing the specificity [17]. They used balanced data, and because of this, the ensemble models (e.g., Random Forest, LightGBM) outperformed in classification patterns of minority classes.

Significantly, it is becoming widespread to use real-life clinical data of different geographic populations. Kim et al. generalized their ML-based model of thyroid recurrence to hospitals in South Korea, Japan and Taiwan, achieving similar accuracies of over 90 percent [18]. Multi-center validation of

TABLE I
COMPARISON TABLE OF METHODS AND DATASETS

Ref	Author(s)	Method Used	Dataset Type	Focus Area	Performance
[8]	Song et al. (2019)	SVM, Decision Trees, Neural Networks	Clinical data from PTC patients	Recurrence prediction after thyroidectomy	Accuracy ≈ 89%
[9]	Li et al. (2020)	CNN-based Deep Learning	Ultrasound image dataset	Benign vs. Malignant Nodule Classification	AUC = 0.94
10]	Liu et al. (2021)	Random Forest, Logistic Regression	EHR records from thyroid cancer patients	Disease-free survival prediction	RF Accuracy ≈ 91%
[11]	Chen et al. (2021)	Gradient Boosting Decision Trees (GBDT)	Hospital records with nodular features	Risk classification of nodules	Accuracy = 93%
[12]	Ko et al. (2020)	XGBoost + Hybrid Feature Selection	Multicenter PTC datasets	Recurrence risk stratification	AUC = 0.92
[14]	Zhang et al. (2021)	Multimodal Deep Learning (Ultrasound + EHR)	Hospital ultrasound + EMR data	Combined malignancy prediction	Accuracy = 94.2%
[15]	Wang et al. (2022)	Explainable XGBoost	National thyroid cancer registry	Recurrence prediction with risk interpretation	Accuracy = 91.8%
[17]	Choi et al. (2021)	SMOTE + Random Forest, LightGBM	Clinical data with class imbalance	Improved recall in recurrence prediction	Recall improvement by 8–12%
[19]	Zhao et al. (2022)	Federated XGBoost	Multi-hospital decentralized data	Privacy-preserving recurrence prediction	Comparable to central models (AUC ≈ 0.91)

that type is crucial in generalizability of predictive models, as well as adaptation to foreign clinical processes.

Moreover, the development of federated learning is starting to take its toll in the research of thyroid cancer. The federate learning method allows training / learning a model in numerous decentralized hospitals without moving patient data and still retains privacy. The study conducted by Zhao et al. showed that federated models of XGBoost could be about as accurate as centrally trained models but also keeping patient confidentiality [19].

Last, the world of research is shifting towards the incorporation of artificial intelligence-driven tools as part of real-time clinical decision support units. Luo et al. introduced a prototype of an AI-driven thyroid recurrence dashboard monitoring system that will provide real-time alerts regarding patient records which update automatically [20]. Such a strategy does not only help the physicians to track their patients over a longer period of time but also corresponds to the trend in the international front of individualized and proactive care.

All these developments point to how machine learning and AI have the potential to disrupt thyroid nodules management practices, including how it can diagnose, and predict the recurrence of such growths. Nonetheless, there is still an issue in data standardization, interpretability, external validation, and ethical use. Our model constitutes such a body of work since our solution to non-recurring thyroid nodules is highly precise, explanatory, while being easily integrated with clinical practices.

III. METHODOLOGY

The study suggests another data-based machine learning model to calculate the possibility of recurrence of thyroid nodules based on structured clinical data. The purpose is to precisely insert the results of the patients into the buckets of recurrence and non-recurrence columns with the help of significant demographic, pathological, and therapeutical marks. The general pipeline includes data preprocessing, exploratory data analysis, training the models using multiple classification, the performance comparison, the interpretation, and the deployment of the final model.

A. Data Collection and Preprocessing

The data contains de-identified histories of the patients who have had thyroid-related interventions. These characteristics are of fundamental importance and specifically these are Age, Gender, TNM staging (T, N, M), Focality, Risk group classification, Treatment Response, and Smoking status. A binary label is also provided with each record indicating relapse or absence of relapse. First, data cleansing was conducted to get rid of the duplicate or inconsistent entries. The treatment of missing values was carried out through imputation or through the removal depending on the distribution checks.

The categorical variables, which included ‘Gender‘, ‘Response‘, and ‘Smoking‘ were coded with ‘LabelEncoder‘ in the ‘sklearn.preprocessing‘ library. ‘StandardScaler‘ was applied on numerical variables to normalize the distribution of their values and better convergence of gradient-based classifiers. That was undertaken to make sure that all the features have been scaled to a consistent scale so that the models could be learning in an unbiased way.

B. Exploratory Data Analysis (EDA)

EDA revealed that Age had a wide range and a strong correlation with recurrence. The target variable showed some class imbalance, but not severe enough to require synthetic oversampling. Correlation heatmaps and feature pair plots were used to identify potential multicollinearity and distributional relationships.

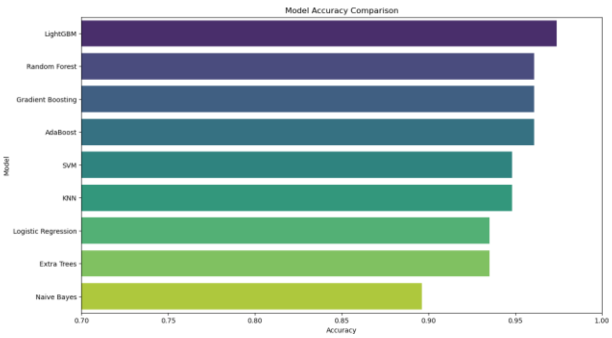

Fig. 1. Model Accuracy Comparison of Various Classifiers

C. System Architecture Overview

The system architecture of the proposed thyroid nodule recurrence prediction model is designed to handle end-to-end data flow — from raw clinical inputs to intelligent prediction and visualization. It begins with the collection of structured data from two primary sources: clinical records and pathology results. This data undergoes preprocessing steps such as cleaning, encoding, and scaling to ensure it is ready for machine learning workflows.

Once preprocessed, the data is fed into multiple classification models including Logistic Regression, Random Forest, Support Vector Machine, and LightGBM. These models are trained and evaluated to find the optimal prediction framework. The outputs from the machine learning module are served through a user-facing serving layer, which visualizes insights for medical professionals.

The architecture also emphasizes long-term viability through modules for secure deployment, regulatory compliance, and provisions for future expansion. Future directions include integration with explainable AI tools and multimodal data inputs like imaging and genomics.

D. Model Training and Selection

A variety of classification models were implemented, including Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naive Bayes, Random Forest, Gradient Boosting, AdaBoost, Extra Trees, and LightGBM. Each model was trained using the training dataset (70%) and evaluated on the test set (30%).

Hyperparameter tuning was performed using 'GridSearchCV' and 'cross val score' for each model. For example, in the LightGBM model, key hyperparameters such as 'learning rate', 'max depth', and 'n estimators' were optimized to maximize accuracy and ROC-AUC. The results of each classifier were compared visually using accuracy and ROC-AUC bar plots, where LightGBM outperformed all others.

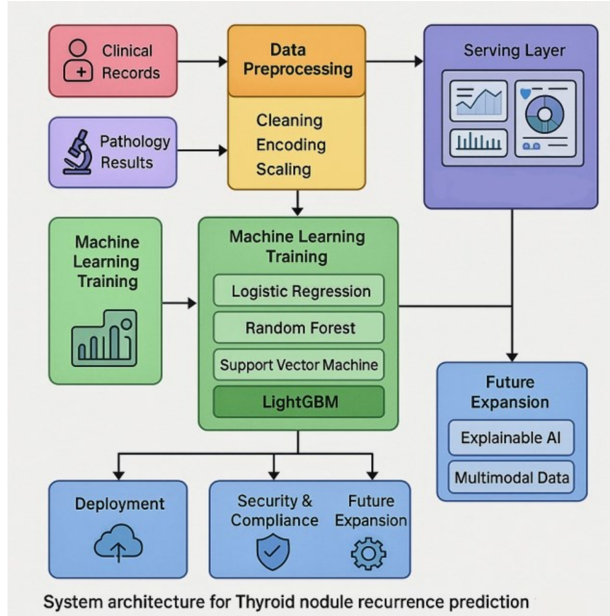

Fig. 2. System Architecture for Thyroid Nodule Recurrence Prediction

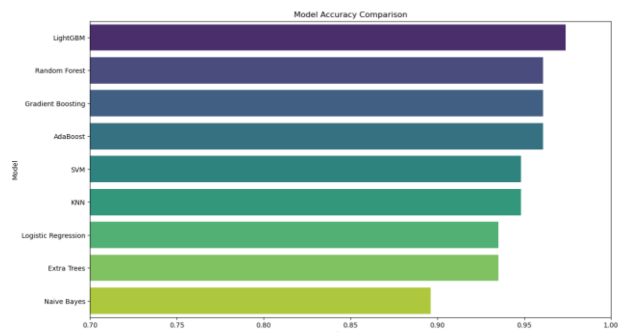

Fig. 3. ROC-AUC Score Comparison of Classifiers

E. Evaluation Metrics and Confusion Matrix

The LightGBM model achieved an accuracy of 97.40% and a ROC-AUC score of 99.17%. These results indicated excellent discrimination capability between recurrent and non-recurrent cases. The confusion matrix in Figure 4 demonstrates the model's robustness in terms of low false positives and false negatives.

F. Feature Importance and Interpretability

The model interpretability was examined using LightGBM's built-in 'feature importances' attribute. The most important features for recurrence prediction were 'Age', 'N' (lymph node involvement), and 'Risk' category. Other factors like 'Focality', 'Gender', and 'Response' also contributed to prediction but to a lesser degree.

Figure 5 shows the relative contribution of each feature to the final prediction. This interpretability is particularly valuable in clinical settings, helping healthcare professionals understand the risk dynamics behind recurrence.

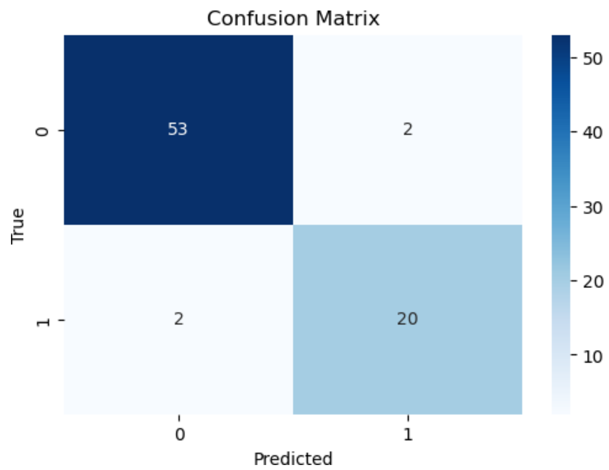

Fig. 4. Confusion Matrix for LightGBM

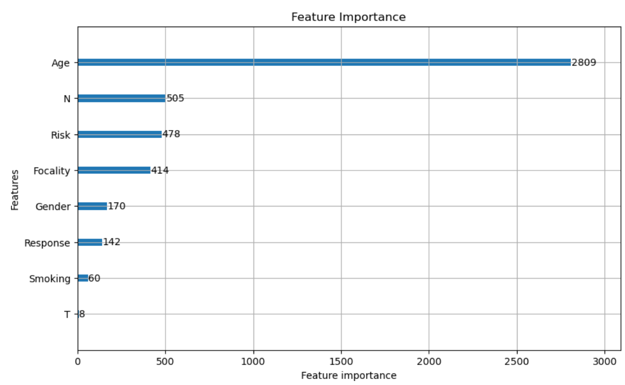

Fig. 5. Feature Importance Plot from LightGBM

G. LightGBM: Mathematical Formulation

LightGBM, a fast and efficient implementation of gradient boosting, constructs decision trees using a leaf-wise (best-first) approach, which can lead to deeper trees and better accuracy.

Given the dataset $\{(x_i, y_i)\}^n_{i=1}$, LightGBM minimizes the following objective function:

$$L^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

where l is the loss function (binary cross-entropy in this study), and $\Omega(f)$ is the regularization term:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

Here, T is the number of leaves, and w_j is the weight of the j^{th} leaf. This regularization helps prevent overfitting.

LightGBM uses histogram-based binning and Gradient-based One-Side Sampling (GOSS) to enhance training speed and reduce memory consumption without compromising accuracy.

H. Final Model Parameters and Saving

After rigorous comparison, the LightGBM model was finalized with the following optimal parameters:

- **learning_rate:** 0.1
- **max_depth:** -1 (no restriction)
- **n_estimators:** 100

The trained model was serialized using the ‘joblib‘ library for future use in clinical software or web-based decision support tools.

IV. Results and Discussion

This section presents the performance comparison and critical analysis of various machine learning models employed to predict thyroid nodule recurrence. The effectiveness of each model was assessed based on two key performance metrics: accuracy and the Receiver Operating Characteristic Area Under Curve (ROC-AUC) score.

A. Model Performance Comparison

According to Table II, the Light Gradient Boosting Machine (LightGBM) algorithm has been selected as the most efficient, with the accuracy rate of 97.40 and ROC-AUC of 99.17. RF, Gradient Boosting, and AdaBoost also showed competitive performance, and their accuracy was close to 96 percent, and ROC-AUC was also high, counting them as effective methods of distinguishing recurrence and non-recurrence instances.

The following ones were Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) with a drop of performance, though considerable. Logistic Regression and Extra Trees were rather effective. Otherwise, Naive Bayes demonstrated the least accuracy (89.61%), which indicates a low level of coping with the data complexity.

B. Interpretation of Results

This trend is quite clear: the popularity of ensemble-based models tends to beat single classifiers. This follows because they can combine several weak learners to form together a powerful model that can capture complex interactions of the features. LightGBM and Gradient Boosting have performed better in relation to the classification of positive, and negative classes due to their superior ROC-AUC scores.

Although simpler models such as Logistic Regression and Naive Bayes are faster to make an inference and of less computational cost, they cannot quantitatively compete in making predictions. As such, high-stakes medical prognoses, which occur in the form of thyroid cancer recurrence in the current case, are characterized by extreme sensitivity toward accuracy and interpretability, which LightGBM excels at.

The experimental results suggest that machine learning offers promising tools for predicting thyroid nodule recurrence. LightGBM stands out due to its high accuracy, excellent ROC performance, and ability to provide feature importance scores. These capabilities make it particularly suitable for integration into clinical decision support systems, enhancing diagnostic precision and supporting proactive patient care.

TABLE II
PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS

Model	Accuracy	ROC AUC
LightGBM	0.974026	0.991736
Random Forest	0.961039	0.984711
Gradient Boosting	0.961039	0.989669
AdaBoost	0.961039	0.986777
SVM	0.948052	0.984298
KNN	0.948052	0.985537
Logistic Regression	0.935065	0.966116
Extra Trees	0.935065	0.980992
Naive Bayes	0.896104	0.972727

V. CONCLUSION AND FUTURE WORK

A. Conclusion

This work developed a predictive classification system based on machine learnings that determines the recurrence of thyroid nodules based on patient clinical and pathologic data. Combined methodology included different steps, including data preprocessing, feature encoding, model selection, performance estimation, as well as the interpretation of results.

Various machine learning models were considered amongst them occurring to be the most suitable classification because of their effectiveness, Logistic Regression, K-Nearest neighbors, support vector machines, naive bayes, and ensemble methods, among others LightGBM was shown to be the most suitable out of all of them. It was proven to be more accurate with a result of 97.40 percent, as well as a high ROC-AUC of 99.17 percent. This finding demonstrates that the model has a high ability in distinguishing between recurrent and not. This further implies that it can be an effective risk assessment tool in clinical practice.

Moreover, the LightGBM is easily interpretable, especially using the feature importance measures, which makes it potentially useful in medical practice and explains how various features such as age, characteristics of a tumor, and risk status would affect the outcomes of recurrence. Such a degree of transparency contributes to building trust and makes it easier to use evidence-based approaches to address decisions in patient care.

B. Future Work

Although the accuracy and reliability of the proposed model are high, it still has a number of possible areas to be improved and broadened. In the first place, combining of bigger and more heterogeneous data sets of several medical centers can enhance the model which will be more general and robust to various populations and clinical applications.

Second, additional verification by prospective clinical trials is needed to make certain about the practicability of the model in the real context. Limiting the predictive accuracy and understanding of clinical insight might be further integrated by adding medical imaging, genetic markers or biochemical profile data.

Moreover, creating a web or mobile-based interface that would incorporate the model into a clinical decision support tool to deliver the prediction of the healthcare professionals in real-time could become a possibility. The integration would close the gap between data science and medical practice in the end facilitate early detection and improve the management of recurrence of thyroid cancer.

Finally, the area of research that can be identified in the context of future research is the topic of explainable AI (XAI) to make models more transparent and accountable to be used ethically and interpretable within systems in healthcare settings.

REFERENCES

[1] Gharib H., Papini E., Garber J.R., et al., "American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules," *Endocrine Practice*, vol. 22, no. 5, pp. 622–639, 2016.

[2] Siegel R.L., Miller K.D., Jemal A., "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.

[3] Durante C., Montesano T., Attard M., et al., "Long-term surveillance of papillary thyroid cancer patients reveals a very low risk of recurrence," *Journal of Clinical Endocrinology & Metabolism*, vol. 98, no. 6, pp. 2237–2245, 2013.

[4] Haugen B.R., et al., "2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer," *Thyroid*, vol. 26, no. 1, pp. 1–133, 2016.

[5] Esteva A., Robicquet A., Ramsundar B., et al., "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24–29, 2019.

[6] Mazzaferri E.L., "Management of low-risk differentiated thyroid cancer," *Endocrine Practice*, vol. 13, no. 5, pp. 498–512, 2007.

[7] Senders J.T., Arnaout O., Karhade A.V., et al., "Natural and artificial intelligence in neurosurgery: a systematic review," *Neurosurgery*, vol. 83, no. 2, pp. 181–192, 2018.

[8] Song J., Chai Y.J., Masuoka H., et al., "Prediction of papillary thyroid carcinoma recurrence using machine learning techniques," *Frontiers in Endocrinology*, vol. 10, article 782, 2019.

[9] Li X., Zhang S., Wang R., et al., "A deep learning model for distinguishing malignant from benign thyroid nodules on ultrasound images," *European Radiology*, vol. 30, pp. 5881–5889, 2020.

[10] Liu Y., Chen J., Shang L., et al., "Machine learning approaches for predicting disease-free survival in thyroid cancer using structured electronic medical records," *Journal of Biomedical Informatics*, vol. 118, p. 103799, 2021.

[11] Chen C., Zhang H., Wang Y., et al., "Application of gradient boosting decision tree algorithms for risk classification of thyroid nodules," *Computers in Biology and Medicine*, vol. 132, p. 104335, 2021.

[12] Ko S.Y., Lee J.H., Kim E.K., et al., "Machine learning-based prediction models for recurrence in papillary thyroid carcinoma," *PLoS ONE*, vol. 15, no. 10, e0240503, 2020.

[13] Fang Y., Zhang B., Xu Y., et al., "Risk stratification of cervical lymph node metastasis and recurrence in papillary thyroid carcinoma using LightGBM," *Scientific Reports*, vol. 11, Article 18462, 2021.

[14] Zhang Y., Wang Y., Liu Y., et al., "Multimodal deep learning model for thyroid nodule diagnosis combining ultrasound and electronic health records," *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2439–2449, 2021.

[15] Wang L., Xu J., Jin B., et al., "An explainable XGBoost-based model for thyroid cancer recurrence prediction using clinical data," *BMC Medical Informatics and Decision Making*, vol. 22, article 81, 2022.

[16] Lundberg S.M., Lee S.I., "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.

[17] Choi Y.J., Baek J.H., Chung S.R., et al., "Handling imbalanced data in thyroid nodule classification using SMOTE and machine learning models," *Journal of Clinical Ultrasound*, vol. 49, no. 1, pp. 27–33, 2021.

[18] Kim D.W., Lee E.J., Chung I.Y., et al., "Multi-center validation of machine learning models for predicting thyroid cancer recurrence," *Endocrine*, vol. 72, no. 3, pp. 687–696, 2021.

[19] Zhao R., Tan M., Li F., et al., "Privacy-preserving federated learning for thyroid cancer prediction across multiple centers," *Journal of Biomedical Informatics*, vol. 128, p. 104015, 2022.

[20] Luo X., Zheng C., Huang Y., et al., "Real-time clinical decision support for thyroid nodule monitoring using AI-powered dashboards," *Healthcare Analytics*, vol. 3, no. 1, p. 100076, 2023.