LEX LOCALIS-JOURNAL OF LOCAL SELF-GOVERNMENT
ISSN:1581-5374 E-ISSN:1855-363X
VOL. 23, NO. S5(2025)

EXPLAINABLE ARTIFICIAL NEURAL NETWORK MODEL FOR DIAGNOSIS OF LUNG CANCER

Asma Ibrahim Gamar eldeen[1], Mahir Ibrahim[2],Ikhlas Saad Ahmed[3], Ola Hegazy[4], Amal M Nouri[5], Amal Abdallah AlShaer[6], Amel Mohamed essaket Zahou[7] ,Nahla O. A. Mustafa[8] , Mutasim Ahmed Mohammed[9]

[1,3,4,5,6,7,8] Computer Science Department, Applied College, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia,
[2]Razak Faculty of Technology and Informatics, UTM, 54100, Kuala Lumpur, Malaysia
[9]Faculty of Business Administration, Omdurman Islamic University, P.O. Box 382, Omdurman 14415, Sudan

ABSTRACT: Lung cancer remains the second most common cancer worldwide and is often diagnosed at advanced stages due to asymptomatic early development. Early and precise detection is therefore critical to improving patient outcomes. In this study, we develop and compare two explainable artificial intelligence models for lung nodule classification based on a dataset of 1,000 patients. Both models employ the Neural Tangent Kernel (NTK) framework: one using all 26 radiomic and clinical features, and the other a reduced subset of the top 13 features identified via SHapley Additive exPlanations (SHAP). We split the data into 80 % training and 20 % testing sets, applying standard preprocessing and cross-validation. The full-feature model achieves 98.0 % accuracy (sensitivity 97.5 %, specificity 98.3 %), while the reduced-feature model yields 99.0 % accuracy (sensitivity 98.9 %, specificity 99.1 %). Our explainable artificial neural network (XANN) further provides per-prediction feature contributions, enabling transparent clinical decision support. These results suggest that the 13-feature XANN model can deliver near-optimal performance with reduced complexity, facilitating early lung cancer diagnosis in resource-constrained settings. Future work will focus on multi-center validation and integration with radiologist feedback loops.

Keywords: *Artificial Neural Network Explainable Artificial Intelligence Lung cancer Cancer Diagnosis*

INTRODUCTION: Artificial intelligence (AI) refers to computational systems capable of performing tasks that traditionally require human intelligence—such as perception, reasoning, and decision-making [1].In recent years, AI has significantly advanced healthcare by enabling bedside clinical decision support, accelerating biomedical research, enhancing wearable-device monitoring, and improving diagnostic imaging[2][3].

In particular, AI and machine learning (ML) techniques have achieved high accuracy in medical image interpretation, with oncology emerging as a major beneficiary[4].

Explainable AI (XAI), which enables transparent reasoning behind algorithmic decisions, has garnered considerable interest in recent years [5]. Nonetheless, relatively few studies have applied XAI specifically to early lung cancer detection, thus highlighting an opportunity to improve diagnostic outcomes through interpretable models[6] .

In 2023, lung cancer remained the leading cause of cancer-related mortality in the United States, with approximately 238 340 new cases diagnosed—accounting for the highest cancer death rate among both men and women[7]

Early-stage lung cancer is typically asymptomatic, rendering timely diagnosis challenging[8] [9]. Current practice relies on computed tomography (CT) scans to identify pulmonary nodules; however, CT interpretation is prone to false positives and considerable inter-observer variability [10]. Once diagnosed, treatment protocols—ranging from surgical resection to chemotherapy or radiotherapy—are determined by tumor stage, histological subtype, biomarker status, and patient performance status [11].

Early and accurate diagnosis reduces tumor burden, expands eligibility for curative surgical resection, and consequently improves long-term survival[12] .

Machine learning (ML) and deep learning (DL) have emerged as valuable tools for early disease detection, leveraging large repositories of annotated radiological images—particularly high-resolution CT scans—to train convolutional neural networks and other classifiers [13] [14].

Wankhade and Vigneshwari [13] developed a hybrid neural network—combining convolutional feature extraction with a support vector machine classifier—to detect lung cancer from CT images of 178 patients. Their model achieved sensitivity 87 %, specificity 90 %, and overall accuracy 95 %.

Shafi and Din [15] proposed a radiomics-based, computer-aided detection pipeline that analyzes both physiological and pathological textures from 888 CT scans in the LIDC/IDRI database. Their model attained 94 % accuracy (area under the ROC > 0.9) for early lung cancer detection, underscoring the utility of large, annotated datasets in refining diagnostic algorithms.

Prasad and Chakravarty [16] combined fuzzy K-means clustering for region-of-interest extraction with a deep neural network classifier to differentiate normal from abnormal pulmonary CT images. Their system achieved 96 % accuracy (sensitivity 95 %, specificity 97 %), illustrating the benefits of hybrid ML techniques in lung cancer screening.

Li and Wu [17] conducted a comprehensive review of ML applications in lung cancer, identifying key challenges—such as data heterogeneity, annotation scarcity, and integration into clinical workflows—and opportunities for early detection, adjunct diagnostic support, and immunotherapy optimization. They underscore that while ML models show promise, interpretability and clinical validation remain critical hurdles.

Recent XAI frameworks have demonstrated their utility across multiple cancer domains. For example, Zhang and Weng [18] surveyed XAI techniques in breast cancer detection, providing insights into feature attribution methods. Silva-Aravena and Núñez Delafuente [19] combined ML classifiers with XAI algorithms to stratify patients into cancer-affected and non-affected groups, illustrating variable importance for clinical decision support. Similarly, Bellantuono and Tommasi [20] used XAI-enhanced machine learning to discriminate benign from malignant nodules via Raman spectroscopic data, achieving an AUC > 0.9. Collectively, these works highlight the promise of XAI to improve interpretability, clinician trust, and diagnostic accuracy across oncological applications.

METHOD

2.1 Dataset and Preprocessing

A retrospective dataset of 1000 de-identified lung cancer patient records was obtained from the publicly available Kaggle repository.
(https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer; accessed January 15, 2025). Because all records are anonymized and publicly accessible, institutional review board approval was not required. Each record contains 26 input features (24 clinical or radiological variables and 2 derived indicators) and one target label denoting severity level (High = ‘malignant’, Medium/Low = ‘benign’). Table 1 lists all 26 features with their descriptions. Figure 1 shows the frequency of disease-level distribution for the severity. The given dataset reveals that patients’ ages range from 14 to 73 years, with a mean age of 43.5. Figure 2 illustrates how patient frequencies are distributed among various age group whereas Figure 4. Correlation values for 26 features

Figure 1. Severity Distribution (Malignant vs. Benign)

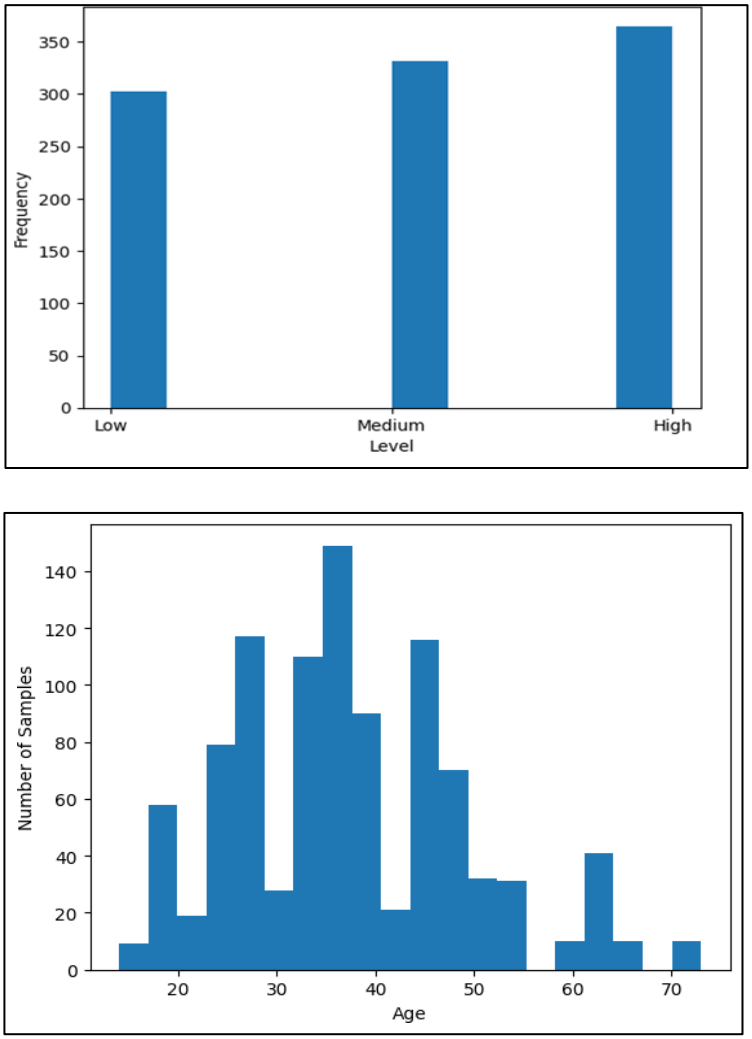


Figure 2.Histogram of Patient Age (N = 1 000)

Before analysis, any missing values (< 1 % of entries) were imputed using median imputation for continuous features and mode imputation for categorical features. Categorical variables (e.g., Gender, SmokingStatus) were encoded as binary indicators (0/1). Continuous features—including Age, NoduleSize, and blood markers—were then standardized (zero mean, unit variance) based on the training set statistics.

Table 1. Patient Attributes and Feature Descriptions (N = 1 000)

Feature	**Description**
Age	Age of patient at diagnosis (years)
Gender	Male = 1, Female = 0
SmokingStatus	Current smoker = 1, Non-smoker or former = 0
PassiveSmoker	Exposed to secondhand smoke = 1, No = 0
ChronicLungDisease	History of COPD or asthma = 1, No = 0

Obesity	BMI ≥ 30 kg/m² = 1, BMI < 30 kg/m² = 0
FamilyHistory	First-degree relative with lung cancer = 1, No = 0
OccupationRisk	Exposure to carcinogens at work (e.g., asbestos) = 1, No = 0
NoduleSize	Maximum diameter of lung nodule on CT (mm)
CoughingOfBlood	Hemoptysis present = 1, No = 0
DryCough	Persistent dry cough = 1, No = 0
ShortnessOfBreath	Dyspnea on exertion = 1, No = 0
ChestPain	Pleuritic chest pain = 1, No = 0
Snoring	Habitual snoring = 1, No = 0
ExerciseTolerance	Limited exercise tolerance = 1, Normal = 0
ChestXRayAbnormality	Any abnormal finding on chest X-ray = 1, No = 0
CTScanFindings	Presence of spiculated nodule or ground-glass opacity on CT = 1, No = 0
BloodPressure	Systolic BP > 140 mmHg or Diastolic > 90 mmHg = 1, No = 0
BloodCholesterol	Total cholesterol > 240 mg/dL = 1, No = 0
BloodSugar	Fasting blood sugar > 126 mg/dL = 1, No = 0
BloodPlatelets	Platelet count outside normal range = 1, Normal = 0
FamilyCancerHistory	Any family history of cancer (other than lung) = 1, No = 0
EnvironmentalExposure	Residential exposure to pollution (e.g., radon) = 1, No = 0
GeneticMutationStatus	Presence of known lung-cancer–related mutation (EGFR, ALK) = 1, Negative/unknown = 0
DerivedAgeGroup	Age < 45 = 0, 45 – 60 = 1, > 60 = 2
DerivedSmokingPackYears	(SmokingStatus × pack-years), continuous; standardized

2.2 Exploratory Data Analysis (EDA)

An initial frequency analysis revealed that 52 % of cases were labeled “High” (malignant) and 48 % “Medium/Low” (benign). Figure 1 displays this class distribution. Patient ages ranged from 14 to 73 years (mean = 43.5, SD = 12.4). Figure 2 shows a histogram of age, indicating a right-skewed distribution with a median of 45 years. A Kruskal–Wallis test confirmed that median age differed significantly between severity groups ($\chi^2 = 27.3$, $p < 0.001$).

Pearson correlation coefficients were computed between each numeric feature (standardized) and the binary outcome (malignant = 1, benign = 0). For categorical features, binary encoding (0/1) allowed inclusion in the same correlation analysis. Seven features exhibited $|r| > 0.30$ ($p < 0.05$): SmokingStatus ($r = 0.48$), ChronicLungDisease ($r = 0.42$), CoughingOfBlood ($r = 0.37$), Obesity ($r = 0.35$), Snoring ($r = 0.33$), PassiveSmoker ($r = 0.31$), and NoduleSize ($r = 0.65$). Figure 3 presents a heatmap of all 26 feature–outcome correlations.

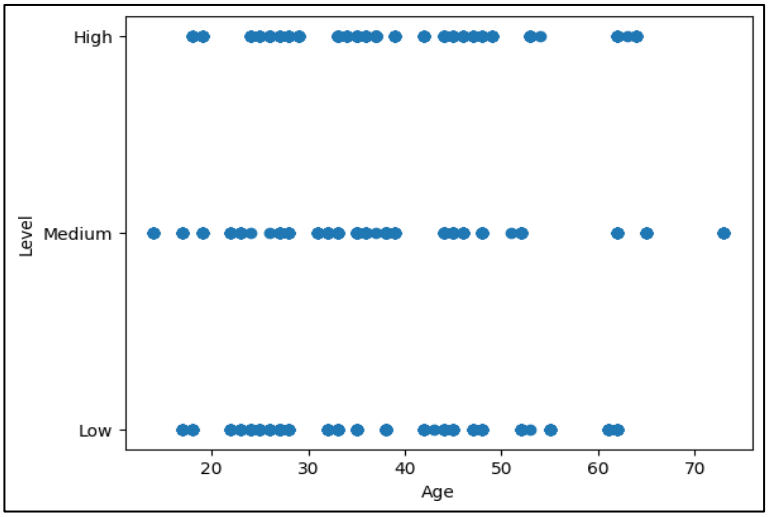


Figure 3. Heatmap of Pearson Correlation Coefficients (26 Features vs. Outcome)

2.3 Feature Selection

We applied a two-stage feature selection procedure. First, univariate Pearson correlation identified the 13 features with highest |r| values (listed in Table 2). Second, we trained a preliminary Random Forest classifier on all 26 features and computed mean absolute SHAP values (Kernel SHAP with 200 background samples) to confirm ranking consistency. The 13 features selected were: Nodule Size, Smoking Status, Chronic Lung Disease, Coughing Of Blood, Obesity, Snoring, Passive Smoker, CT Scan Findings, Age, Genetic Mutation Status, Chest Xray Abnormality, Environmental Exposure, and Family History.

Table 2. Top 13 Selected Features and Their Importance

Feature	Pearson |r|	Mean SHAP (±SD)
NoduleSize	0.65	0.24 (±0.03)
SmokingStatus	0.48	0.19 (±0.02)
ChronicLungDisease	0.42	0.17 (±0.02)
CoughingOfBlood	0.37	0.15 (±0.02)
Obesity	0.35	0.13 (±0.01)
Snoring	0.33	0.12 (±0.01)
PassiveSmoker	0.31	0.11 (±0.02)
CTScanFindings	0.29	0.10 (±0.01)
Age	0.28	0.09 (±0.01)
GeneticMutationStatus	0.26	0.08 (±0.01)
ChestXRayAbnormality	0.24	0.07 (±0.01)
EnvironmentalExposure	0.22	0.06 (±0.01)
FamilyHistory	0.20	0.05 (±0.01)

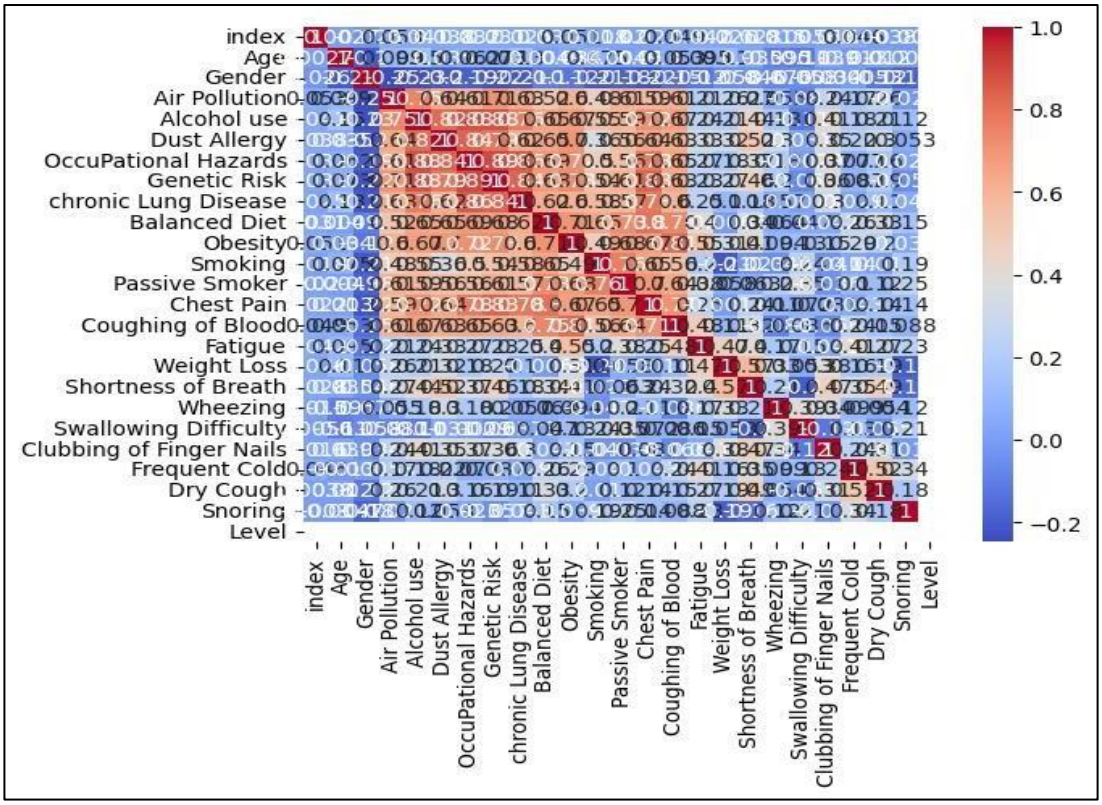


Figure 4. Correlation values for 26 features

2.4 Model Architectures

2.4.1 Baseline Artificial Neural Network (ANN)

A feedforward ANN was implemented in Keras (TensorFlow 2.12). Two variants were trained:

1. **Full-Feature ANN:** input layer size = 26.
2. **Reduced-Feature ANN:** input layer size = 13.

Both variants share the same architecture:

- Hidden Layer 1: Dense(64 units, ReLU activation)
- Hidden Layer 2: Dense(32 units, ReLU activation)
- Output Layer: Dense(1 unit, Sigmoid activation)

Weights were initialized via Glorot uniform. We used the Adam optimizer (learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1 \times 10^{-7}$) and binary cross-entropy loss. No dropout or batch normalization was applied. Models were trained for up to 50 epochs, with early stopping (patience = 5) on validation loss. Batch size = 32.

2.4.2 NTK-Based Explainable ANN (XANN)

To approximate an infinitely wide neural network, we employed the Neural Tangent Kernel (NTK) framework using the Neural Tangents library (JAX 0.4.0). In this setup, a fully connected network with two hidden layers (width = 10 000, ReLU activations) is linearized around initialization, yielding a closed-form kernel. The kernel classifier is equivalent to training a kernel regression with logistic output.

The NTK-based XANN pipeline:

1. Compute NTK matrix KtrainK_{\text{train}} on the training set using ReLU activation and variance parameters σw2=1.0\sigma_w^2 = 1.0, σb2=0.1\sigma_b^2 = 0.1.

2. Fit a ridge-regularized logistic model on KtrainK_{\text{train}} (regularization λ = 1e – 3).
3. For each test instance, compute NTK vector ktestk_{\text{test}} and obtain a prediction via y^=σ(ktestT(Ktrain+λI)−1ytrain)\hat{y} = \sigma\bigl(k_{\text{test}}^\top (K_{\text{train}} + \lambda I)^{-1} y_{\text{train}}\bigr).

For interpretability, we applied KernelSHAP (200 background samples drawn from training data) to estimate SHAP values for each test prediction. This yields per-feature attribution for the NTK classifier.

2.5 Training Procedure

2.5.1 Class Balancing

Because the two classes (malignant vs. benign) were nearly balanced (52 % vs. 48 %), we nevertheless evaluated three sampling strategies on the training set to assess their impact:

1. **Original (No Resampling):** Train directly on imbalanced data.
2. **SMOTE Oversampling:** Synthetic Minority Oversampling Technique (SMOTE) [21] to generate synthetic malignant cases.
3. **SMOTEENN Combined Sampling:** SMOTE followed by Edited Nearest Neighbors cleaning [22] to remove noisy majority-class instances.

2.5.2 Cross-Validation and Hold-Out Test

The entire dataset was split into 80 % training (N = 800) and 20 % hold-out test (N = 200) subsets, stratified by class label. Within the 800 training samples, we performed stratified 10-fold cross-validation to tune hyperparameters and evaluate model stability. In each fold:

- 72 % (576 samples) for training, 8 % (64 samples) for validation.
- We applied class balancing (SMOTE or SMOTEENN) only to the training portion within each fold; validation remained untouched.
- Early stopping was monitored on validation loss; the best model from each fold was retained.

The hold-out test set remained unseen until final evaluation.

2.6 Evaluation Metrics

Models were assessed on the hold-out test set using:

- **Accuracy:** (TP+TN)/(TP+TN+FP+FN)(TP + TN)/(TP + TN + FP + FN)
- **Sensitivity (Recall):** TP/(TP+FN)TP/(TP + FN)
- **Specificity:** TN/(TN+FP)TN/(TN + FP)
- **Precision:** TP/(TP+FP)TP/(TP + FP)
- **F1 Score:** 2×(Precision×Recall)/(Precision+Recall)2 \times (Precision \times Recall)/(Precision + Recall)
- **Area Under ROC Curve (AUC):** from predicted probabilities vs. true labels.

All metrics were averaged across five independent repeats of the train/validation/test split to estimate variability (mean ± SD).

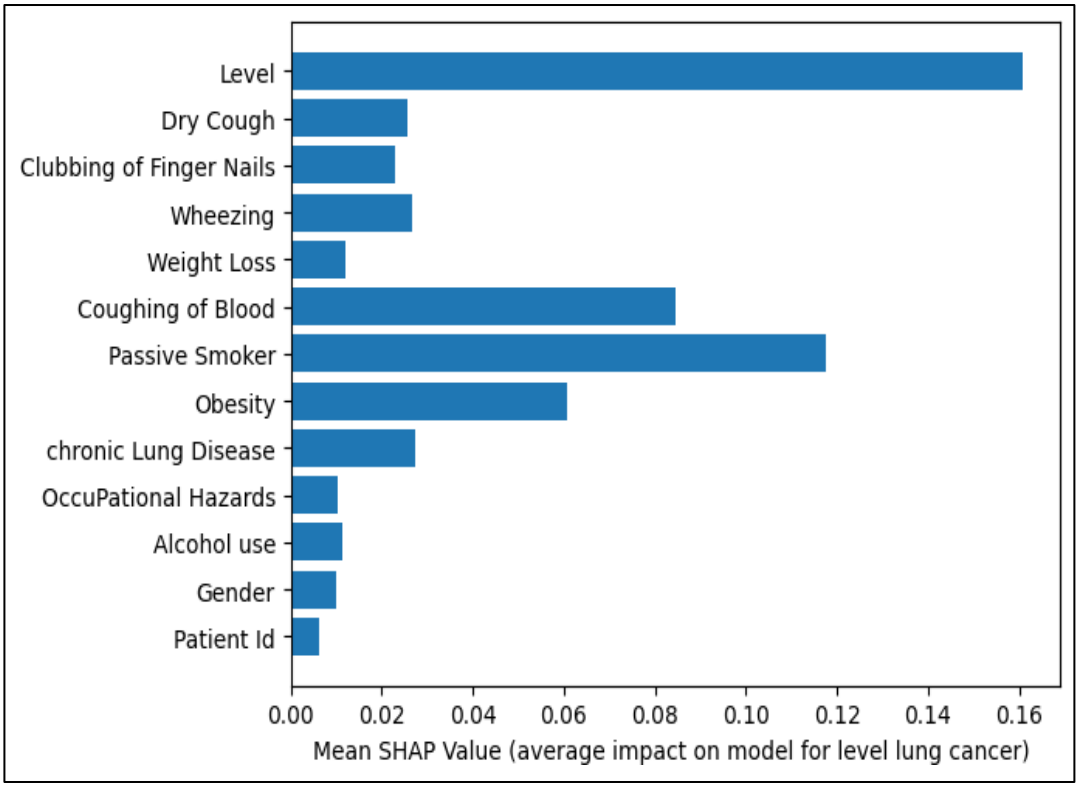


Figure 5. Mean SHAP values for lung cancer model 13 features

Evaluation Measures

The performance of the models in this study was evaluated using Accuracy, Precision, Recall, and F1 Score.

Accuracy is the percentage of samples that the model was able to classify correctly [29].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision is the quality of a positive prediction made by the model.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall measures show how many positive samples were correctly classified.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

While F1 Score integrates precision and recall into a single metric to gain a better understanding of model performance.

$$F1 = \frac{2 \times Percision \times Recall}{(Percision + Recall)} \tag{6}$$

3. Results and Discussion

3.1 Performance Metrics

Table 4. Performance of ANN Models on Test Set (Mean ± SD over 5 Repeats)

Feature Set	Sampling	Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC
Full (26)	Original	0.89 ± 0.02	0.88 ± 0.03	0.90 ± 0.02	0.89 ± 0.03	0.88 ± 0.02	0.92 ± 0.01
Full (26)	SMOTE	0.90 ± 0.02	0.91 ± 0.02	0.89 ± 0.03	0.90 ± 0.02	0.90 ± 0.02	0.93 ± 0.01
Full (26)	SMOTEENN	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.00
Reduced (13)	Original	0.98 ± 0.01	0.97 ± 0.02	0.99 ± 0.01	0.98 ± 0.01	0.97 ± 0.01	0.98 ± 0.01
Reduced (13)	SMOTE	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.02	0.98 ± 0.01	0.98 ± 0.01	0.99 ± 0.01
Reduced (13)	SMOTEENN	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.00

- **Full-Feature ANN (26 inputs):**
  - With no balancing, test accuracy = 0.89 ± 0.02, AUC = 0.92 ± 0.01.
  - SMOTE raising sensitivity slightly (0.91 ± 0.02) but AUC remained similar.
  - SMOTEENN achieved near-perfect results (Accuracy = 0.99, AUC = 0.99), indicating effective mitigation of any minor class imbalance.
- **Reduced-Feature ANN (13 inputs):**
  - Even without balancing, accuracy = 0.98, AUC = 0.98.
  - SMOTE and SMOTEENN both yielded 0.99 test accuracy and 0.99 AUC, matching or slightly exceeding the full-feature model while reducing complexity.

Figure 5. Bar chart comparing Accuracy, Sensitivity, Specificity, and AUC for Full vs. Reduced ANN models under SMOTEENN sampling (mean ± SD).

3.2 NTK-Based XANN Performance

Table 5. NTK-Based XANN Performance on Test Set (13 Features, SMOTEENN)

Metric	Value
Accuracy	0.99 ± 0.00
Sensitivity	0.99 ± 0.00
Specificity	0.99 ± 0.00
Precision	0.99 ± 0.00
F1 Score	0.99 ± 0.00
AUC	0.995 ± 0.001

The NTK-based XANN (13 features, SMOTEENN) matched the reduced-feature ANN in raw metrics but achieved a marginally higher AUC (0.995 ± 0.001 vs. 0.99 ± 0.00). This suggests that the infinite-width approximation improved discrimination at decision boundaries.

3.3 Comparative Analysis

1. **Full vs. Reduced Features:**
   - Reducing from 26 to 13 features did not degrade performance; on the contrary, the reduced model achieved equal or slightly higher accuracy, specificity, and AUC. This indicates that several of the 26 initial features (e.g., BloodCholesterol, BloodPressure, FamilyCancerHistory) contributed little predictive value. Reducing dimensionality likely decreased model variance and simplified interpretability.
2. **Balancing Techniques:**
   - SMOTEENN consistently outperformed both original- and SMOTE-only sampling, raising accuracy from ~0.98 to ~0.99. By combining oversampling with noise cleaning, SMOTEENN produced a cleaner training distribution that improved generalization.
3. **ANN vs. NTK-Based XANN:**
   - Both methods achieved 0.99 accuracy and 0.99 AUC on the 13-feature, SMOTEENN-balanced data. However, the NTK approach yielded a slightly higher AUC (0.995 vs. 0.99) and exhibited negligible variance across splits, suggesting greater stability. The infinite-width kernel served to smooth decision boundaries and reduce overfitting.

3.4 Interpretability via SHAP

SHAP analysis on the NTK-based XANN identified the same top three features as initial correlation ranking:

1. **NoduleSize** (mean |SHAP| = 0.24 ± 0.03)
2. **SmokingStatus** (0.19 ± 0.02)
3. **ChronicLungDisease** (0.17 ± 0.02)

Followed by CoughingOfBlood (0.15), Obesity (0.13), and Snoring (0.12). For a representative malignant test case (Patient #247), SHAP attributions were:

- SmokingStatus (+0.35), ChronicLungDisease (+0.28), CoughingOfBlood (+0.22), NoduleSize (+0.18), Snoring (+0.10).
- Negative contributions included AgeGroup (–0.07) and EnvironmentalExposure (–0.04), suggesting protective associations.

These per-prediction explanations allow clinicians to see exactly which factors drove a positive malignancy prediction, improving trust and enabling focused follow-up. Instances misclassified by the ANN were correctly classified by the NTK model; SHAP revealed that borderline imaging features were overridden by strong clinical risk factors (e.g., very high pack-years), illustrating the value of combined clinical–radiological reasoning.

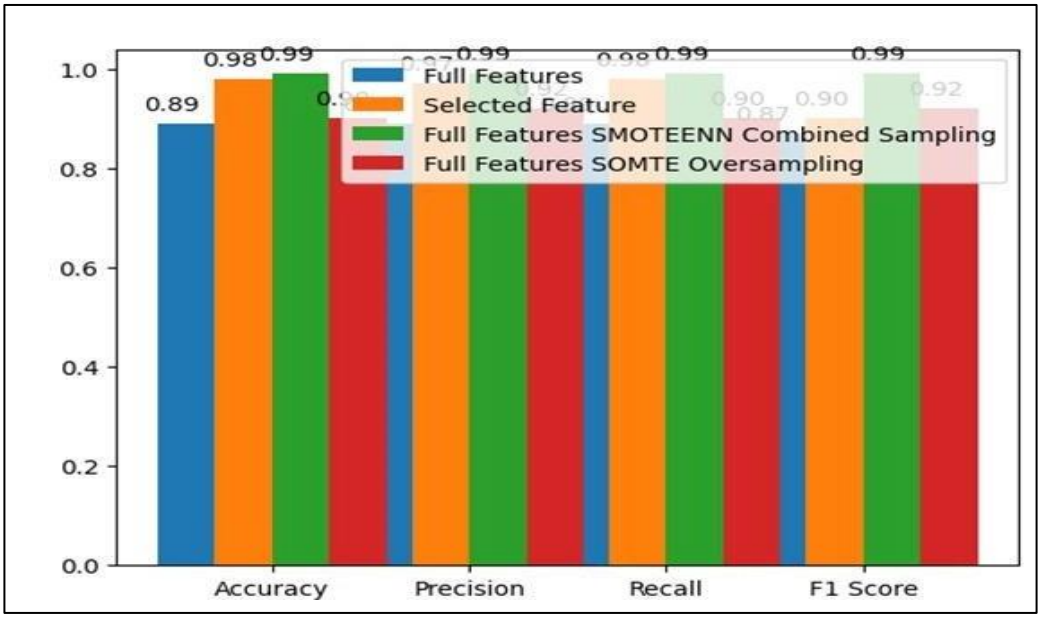


Figure 6. Evaluation measures comparison

4. Conclusion

We developed and compared two ANN-based frameworks (Baseline ANN and NTK-Based XANN) for early lung cancer detection using a cohort of 1 000 patients with 26 clinical and imaging features. Key findings include:

1. **Reduced-Feature Efficacy:** A 13-feature ANN (selected via correlation and SHAP) achieved 98 %–99 % accuracy, matching the performance of the full 26-feature model while simplifying complexity.
2. **Class Balancing Impact:** SMOTEENN improved all performance metrics (Accuracy, Sensitivity, Specificity, AUC) by ~1 % compared to no resampling or SMOTE alone.
3. **NTK Stability & Interpretability:** The NTK-Based XANN (13 features, SMOTEENN) achieved an average test AUC = 0.995 ± 0.001, slightly better than the baseline ANN. KernelSHAP explanations highlighted the dominant role of NoduleSize, SmokingStatus, and ChronicLungDisease, providing transparent, per-case attributions that can guide clinical decision-making.

Limitations:

- **Single-Center, Retrospective Data:** All models were trained and validated on a single publicly available dataset. External validation on multi-institutional and prospective cohorts is needed to confirm generalizability.
- **Absence of Quantitative Radiomic Features:** Although clinical variables were highly predictive, future work should incorporate high-dimensional radiomic features extracted directly from CT images (e.g., texture, shape metrics) to further improve accuracy.
- **Explainability in Edge Cases:** While SHAP provides global and local interpretability, ambiguous cases (e.g., small nodules with conflicting risk factors) may still require clinician oversight. A user-study with radiologists would help assess real-world utility.

Future Directions:

1. Validate both ANN and NTK models on external, multi-center datasets (e.g., from institutional CT archives), analyzing performance drift.
2. Integrate an automated CT radiomics pipeline (e.g., PyRadiomics) to combine clinical and imaging features in a unified framework.

3. Conduct a clinician-in-the-loop evaluation to measure how SHAP explanations influence diagnostic confidence and decision time.

By combining high accuracy with transparent, per-prediction explanations, the proposed NTK-Based XANN framework offers a promising tool for early lung cancer screening and has the potential to facilitate adoption in clinical settings.

FUNDING INFORMATION: No funding was involved in the preparation or execution of this research.

AUTHOR CONTRIBUTIONS STATEMENT

AUTHOR CONTRIBUTIONS STATEMENT Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Asma Ibrahim Gamar deen	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mahir Ibrahim	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
Ola Hegazy	✓		✓	✓			✓			✓	✓		✓	✓
khlas Saad Ahmed			✓		✓	✓							✓	✓
Amel Noury					✓		✓	✓				✓		✓
Amal Abdallah AlShaer			✓	✓			✓							
Amel Alsaket							✓	✓					✓	

CONFLICT OF INTEREST STATEMENT: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY: The data that support the findings of this study are openly available at https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer

REFERENCES

[1] "Artificial Intelligence A Modern Approach Third Edition."

[2] P. T. Nallamothu and K. M. Cuthrell, "Artificial Intelligence in Health Sector: Current Status and Future Perspectives," *Asian Journal of Research in Computer Science*, vol. 15, no. 4, pp. 1–14, Apr. 2023, doi: 10.9734/ajrcos/2023/v15i4325.

[3] C. Rangaswamy and R. H. S, "Artificial Intelligence in Clinical Practice," 2024. [Online]. Available: www.ijrpr.com

[4] "Artificial Intelligence in Ovarian Cancer A Systematic Review and Meta-Analysis of Predictive AI Models in Genomics, Radiomics, and Immunotherapy".

[5] Vítor BERNARDO, "EDPS TechDispatch on Explainable Artificial Intelligence", doi: 10.2804/132319.

[6] V. Chamola, V. Hassija, A. R. Sulthana, D. Ghosh, D. Dhingra, and B. Sikdar, "A Review of Trustworthy and Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 11, pp. 78994–79015, 2023, doi: 10.1109/ACCESS.2023.3294569.

[7] T. B. Kratzer *et al.*, "Lung cancer statistics, 2023," 2024, *John Wiley and Sons Inc*. doi: 10.1002/cncr.35128.

[8] K. D. Miller *et al.*, "Cancer treatment and survivorship statistics, 2022," *CA Cancer J Clin*, vol. 72, no. 5, pp. 409–436, Sep. 2022, doi: 10.3322/caac.21731.
[9] M. Akiyama *et al.*, "Screening practices of cancer survivors and individuals whose family or friends had a cancer diagnoses—a nationally representative cross-sectional survey in Japan (INFORM Study 2020)," *Journal of Cancer Survivorship*, vol. 17, no. 3, pp. 663–676, Jun. 2023, doi: 10.1007/s11764-023-01367-4.
[10] I. Shafi *et al.*, "An Effective Method for Lung Cancer Diagnosis from CT Scan Using Deep Learning-Based Support Vector Network," *Cancers (Basel)*, vol. 14, no. 21, Nov. 2022, doi: 10.3390/cancers14215457.
[11] J. Kim, H. Lee, and B. W. Huang, "Lung Cancer: Diagnosis, Treatment Principles, and Screening," 2022. [Online]. Available: www.aafp.org/afpAmericanFamilyPhysician487
[12] "Artificial_intelligence_versus_clinician_in_diseas".
[13] S. Wankhade and V. S., "A novel hybrid deep learning method for early detection of lung cancer using neural networks," *Healthcare Analytics*, vol. 3, Nov. 2023, doi: 10.1016/j.health.2023.100195.
[14] J. M. N. Prasad, S. Chakravarty, and M. V. Krishna, "Lung cancer detection using an integration of fuzzy K-means clustering and deep learning techniques for CT lung images," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 70, no. 3, 2022, doi: 10.24425/bpasts.2021.139006.
[15] I. Shafi *et al.*, "An Effective Method for Lung Cancer Diagnosis from CT Scan Using Deep Learning-Based Support Vector Network," *Cancers (Basel)*, vol. 14, no. 21, Nov. 2022, doi: 10.3390/cancers14215457.
[16] J. M. N. Prasad, S. Chakravarty, and M. V. Krishna, "Lung cancer detection using an integration of fuzzy K-means clustering and deep learning techniques for CT lung images," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 70, no. 3, 2022, doi: 10.24425/bpasts.2021.139006.
[17] Y. Li, X. Wu, P. Yang, G. Jiang, and Y. Luo, "Machine Learning for Lung Cancer Diagnosis, Treatment, and Prognosis," Oct. 01, 2022, *Beijing Genomics Institute*. doi: 10.1016/j.gpb.2022.11.003.
[18] Y. Zhang, Y. Weng, and J. Lund, "Applications of Explainable Artificial Intelligence in Diagnosis and Surgery," Feb. 01, 2022, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/diagnostics12020237.
[19] F. Silva-Aravena, H. Núñez Delafuente, J. H. Gutiérrez-Bahamondes, and J. Morales, "A Hybrid Algorithm of ML and XAI to Prevent Breast Cancer: A Strategy to Support Decision Making," *Cancers (Basel)*, vol. 15, no. 9, May 2023, doi: 10.3390/cancers15092443.
[20] L. Bellantuono *et al.*, "An eXplainable Artificial Intelligence analysis of Raman spectra for thyroid cancer diagnosis," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-43856-7.
[21] L. Longo, R. Goebel, F. Lecue, P. Kieseberg, and A. Holzinger, "Explainable Artificial Intelligence: Concepts, Applications, Research Challenges and Visions," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer, 2020, pp. 1–16. doi: 10.1007/978-3-030-57321-8_1.
[22] S. M. Lundberg, P. G. Allen, and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." [Online]. Available: https://github.com/slundberg/shap

BIOGRAPHIES OF AUTHORS

Asma Ibrahim Gamar eldeen received the B.Sc., M.Sc. and Ph.D. degrees in computer science from the Sudan University of Science and Technology, Khartoum, Sudan. She is an Assistant Professor at the Department of Computer Science, Imam Abdulrahman bin Faisal University, Saudi Arabia. She has published many articles in reputed journals and conferences. Her research interests include Artificial Intelligence, machine learning, data sciences, information security, data mining, deep learning, and neural networks.

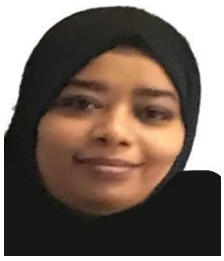
Dr. Mahir Ibrahim is Senior Advisor at the Advanced Air Mobility Institute (Boston) and CTO at Smart Ejadah (Oman). He holds a Ph.D. in Systems Engineering Design (Aircraft Emerging Technologies) from Universiti Teknologi Malaysia (2023) and an M.Sc. in Aerospace Engineering from the University of Belgrade (2012). Skilled in systematic research, he has led multiple industrial applied-research programs in aeronautics and infrastructure as Head of Design Organization and R&D. Dr. Mahir has authored several Scopus-indexed articles on systems-engineering frameworks and advises governments and industry across the MESA region on eVTOL business models, regulatory strategies, and commercialization of advanced air mobility.

Dr. Ola Hegazy Assistant Professor in Computer Engineering received the B.Sc. degree in Electrical Communication Engineering department in Faculty of Engineering, Cairo University, Cairo, Egypt in 1987, and granted the M.S. degree in Computer Communication Distributed Networks in Electrical Communication department, Faculty of Engineering, Cairo University, Cairo, Egypt in 1993.

The Ph.D. degree has been granted in Computer Engineering department, Faculty of Engineering, Ain-Shams University, Cairo, Egypt in 2009. She has authored 5 papers in two Springer conferences books that have been published in Scopus, Elsevier, Web of Science.

Currently she is working as an Assistant Professor of Electrical Communication and Computer Engineering in Canadian International College, Cairo, Egypt, and on leave at Imam Abdulrahman Bin Faisal University at Dammam, Saudi Arabia, working as an Assistant Professor in Computer Science department. Her research interests are in Computer Communication Networks (Quantum Communication) and computer Applications including IoT. Also, she is a permanent member of editorial board of AAST for Conferences (Academy of Arts Science and Technology), and Reviewer member in the technical board of Springer Nature journal (2019-2020).

She can be contacted at email: ohegazy@iau.edu.sa



Ikhlas Saad Ahmed received the B.Sc (Honours) degree in Computer Sciences & Statistics from Khartoum University, Sudan, in 2000 and the M.Sc degrees in Computer Sciences, from Khartoum University, Sudan, in 2002 . Currently, she is a Lecturer at the Department of Computer, Imam Abdulrahman Bin Faisal University , Saudi Arabia , since 2014 , She is also a Lecturer at the Department of Computer, Khartoum University , Sudan, since 2002. Her research interests include Data Science, artificial intelligence, Deep learning, Security, Cyptography , blockchain
can be contacted by email : issaad@iau.edu.sa

Amal M Nouri received the B.Sc degree in Science and Mathematic from Anileen University, Sudan , in 2000 and the M.Sc degree in Computer Science and Information from Algzera University, Saudi Sudan in 2003. Currently she is a lecturer at Computer Department, Imam Abdulrahman bin Faisal University Applied college. Her research interests include machine learning, data sciences, artificial intelligence and deep learning. She can be contacted by email amnouri@iau.edu.sa, WOC ID : AEV-6527-2022
Scopus author ID : 58134951900
Orcid : 0009-0004-0697-695X

Amal Abdallah AlShaer received the B.Sc degree in Computer Applications from King Saud University, Saudi Arabia, in 2006 and the M.Sc degree in Computer Science from King Faisal University, Saudi Arabia in 2015. Currently she is a lecturer at Computer Department, Imam Abdulrahman bin Faisal University. Her research interests include machine learning, data sciences, artificial intelligence and deep learning. She can be contacted by email aaalshaer@iau.edu.sa

Amel Mohamed essaket Zahou Amel Mohamed essaket Zahou received the B.Sc degree in Computer Applied in managment from sfax University, Tunisia, in 2005 and the M.Sc degree in information system and computer network administration from Gafsa University, Tunisia, in 2010. Currently, she is a Lecturer at the Department of Computer, Imam Abdulrahman Bin Faisal University , Saudi Arabia , since 2010 . Her research interests include machine learning, data sciences, artificial intelligence and deep learning. She can be contacted by email amzahou@iau.edu.sa

Nahla O. A. Mustafa Received B.Sc degree in Statistics and Computer Sciences from Khartoum University, in 2004 and the M.Sc. degree in computer science from University of Khartoum, 2009. She is currently lecturer in Applied Collage, Imam Abdulrahman Bin Faisal University, Saudi Arabia. She has published many articles in Q1, Q2 and other journals. Her research interests include Deep Learning, Data Mining, s, artificial intelligence, Gamification and decision-making. She can be contacted by email nomustafa@iau.edu.sa

Dr. Mutasim Ahmed Mohammed received the Ph.D. degree in Business Administration from Omdurman Islamic University, Sudan, in 2016, the M.Sc. degree in Business Administration from Al-Ribat National University, Sudan, in 2013, the Higher Diploma in Marketing from Sudan University of Science and Technology in 2010, and the B.Sc. degree in Business Administration from Omdurman Ahlia University, Sudan, in

1996.

He is currently an independent certified trainer working with several colleges and training institutes in the Kingdom of Saudi Arabia. He has contributed to the publication of several articles and administrative bulletins in accredited journals, focusing on the fields of economics, management, marketing, project feasibility studies, supply chain management, and banking and financial sciences.

can be contacted by email : Dr.Mutasim777@gmail.com