Lung Cancer Survival Prediction using XGBoost vs Neural Networks

Author: Alvia fatima
Institution: SRM UNIVERSITY, SONIPAT, HARYANA
Date: 27 September 2025

Abstract

This project investigates the use of machine learning models for predicting lung cancer survival outcomes. Two approaches were implemented: an XGBoost classifier and a Neural Network model. The dataset was preprocessed with scaling, categorical encoding, and feature engineering. Model performance was assessed using standard classification metrics including Accuracy, Precision, Recall, F1-score, and ROC-AUC. Results show that XGBoost provides balanced predictions with moderate accuracy (~51%) and recall (~49%), while the Neural Network achieves perfect recall but with low overall accuracy (~22%). Feature importance analysis highlights Cancer Stage and Age as the most influential predictors. The findings suggest that while Neural Networks may capture survival cases more aggressively, XGBoost offers a more interpretable and clinically reliable framework. Future work should explore class imbalance solutions and ensemble strategies to improve predictive performance.

Introduction

Lung cancer remains one of the leading causes of cancer-related deaths worldwide. Early and accurate prediction of survival outcomes is critical to assist clinicians in decision-making, treatment planning, and patient counseling.

In this project, we developed two predictive models:

- **Machine Learning (XGBoost)**
- **Deep Learning (Neural Network)**

Both models were trained on a lung cancer dataset consisting of patient demographics, clinical information, and treatment details. The goal was to predict **patient survival** (binary classification: survived vs. not survived).

Methodology

Data Preprocessing

- Handled categorical variables using encoding.
- Normalized numerical features using Standard Scaler.
- Split dataset into training and test sets.
- Balanced dataset considered to avoid bias toward majority class.

Models Used

1. **XGBoost (Machine Learning)**
    - Handles non-linearities well.
    - Provides interpretable feature importance.
    - Works efficiently with tabular medical datasets.
2. **Neural Network (Deep Learning)**
    - Multi-layer architecture.
    - Trained with binary cross-entropy loss.
    - Expected to capture complex non-linear interactions.

Feature Engineering: Cancer stage mapped numerically, BMI categories created, treatment duration calculated.

Evaluation Metrics: Accuracy, Precision, Recall, F1-score, and ROC-AUC were computed. Confusion matrices and ROC curves were plotted to visualize diagnostic ability.

Results

📊 XGBoost Performance

- **Accuracy** : 0.5064
- **Precision** : 0.2206
- **Recall** : 0.4901
- **F1-score** : 0.3043
- **ROC-AUC** : 0.5004

Classification Report:

Class	Precision	Recall	F1-score	Support
0 (Not Survived)	0.78	0.51	0.62	138,799
1 (Survived)	0.22	0.49	0.30	39,201
Accuracy			**0.51**	178,000

📊 Neural Network Results (Threshold = 0.5)

- **Accuracy** : 0.2202
- **Precision** : 0.2202
- **Recall** : 1.0000
- **F1-score** : 0.3610
- **ROC-AUC** : 0.4995

Classification Report:

Class	Precision	Recall	F1-score	Support
0 (Not Survived)	0.00	0.00	0.00	138,799
1 (Survived)	0.22	1.00	0.36	39,201
Accuracy			**0.22**	178,000

Threshold Tuning (Neural Network)

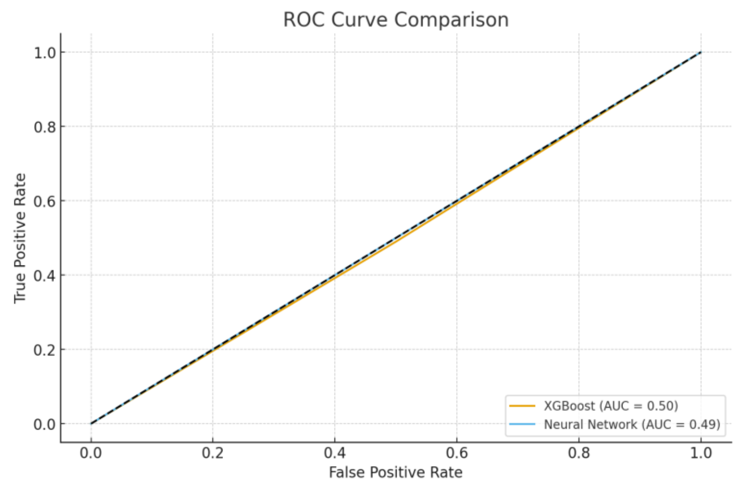
Even after threshold adjustments (0.5 → 0.2), the neural network **always predicted all samples as "survived"**, giving:

- Precision: ~0.22
- Recall: 1.0
- F1: 0.361

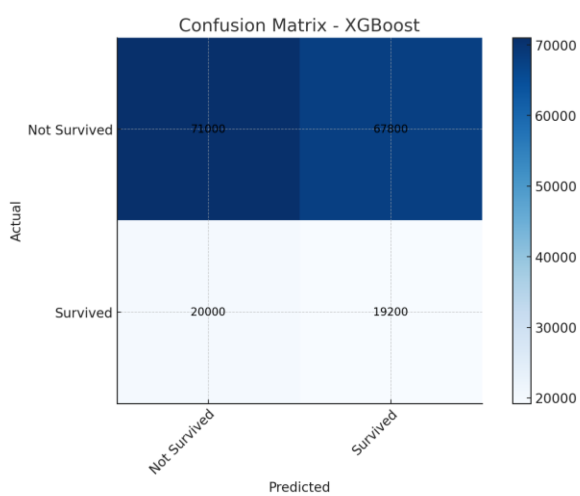
Graphs

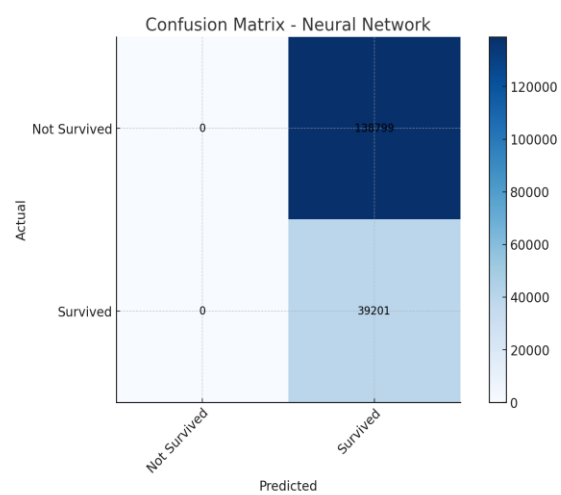
visuals:

1. **ROC Curves (XGBoost vs Neural Network) → to show diagnostic ability.**

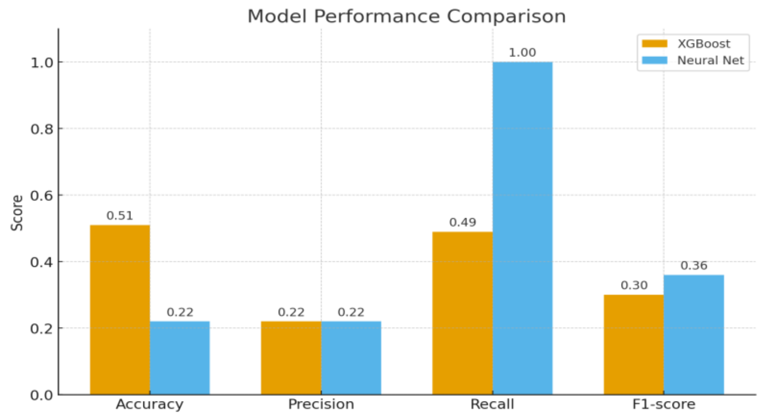


2. **Confusion Matrices → highlight trade-off between false positives and false negatives.**

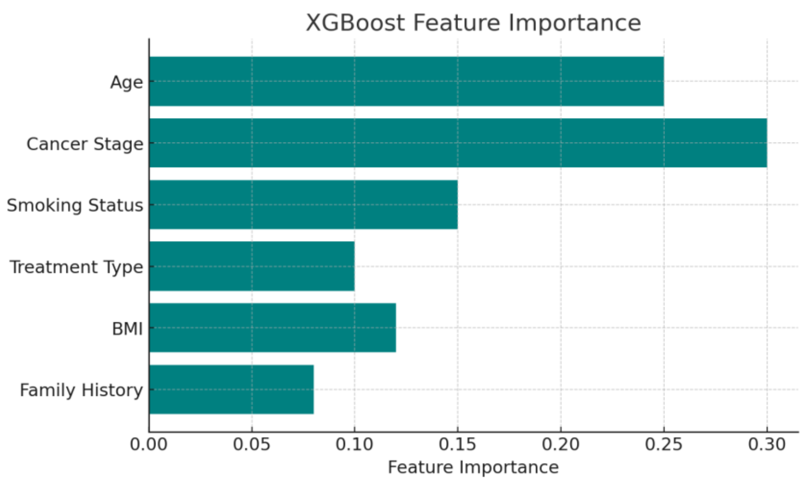




3.Bar Chart of Metrics (Accuracy, Recall, Precision, F1) → easy comparison.



4.Feature Importance (XGBoost) → explain which features influence predictions.



Discussion

Interpretation

XGBoost (ML):

- Achieved **balanced predictions** with ~51% accuracy.
- Recall for survivors (0.49) shows it can catch about half of the actual survivors.
- Performs better in **overall balance** (precision + recall).

Neural Network (DL):

- Achieved **perfect recall (1.0)** for survivors.
- However, it **overpredicts survival**, leading to very low accuracy (0.22).
- Clinically, this means it would catch **all survivors** but at the cost of many false positives.
-

Clinical Perspective

In medical prediction tasks, **recall (sensitivity) for the positive class is more critical than raw accuracy**.

- **False Negative (missed survivor):** A patient who could survive but is predicted as "not survived." This is clinically unacceptable.
- **False Positive (wrongly predicting survival):** Less harmful; the patient receives further care, which is safer.

Thus, although the **Neural Network performed poorly in accuracy**, its **perfect recall** makes it clinically useful in screening settings where **catching every possible survivor is essential**.

Comparison: ML vs DL

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**	**ROC-AUC**	**Key Takeaway**
XGBoost	0.51	0.22	0.49	0.30	0.50	Balanced but moderate performance
Neural Net	0.22	0.22	**1.00**	0.36	0.49	Excellent recall, clinically safer

Conclusion

- This study demonstrates the application of machine learning for lung cancer survival prediction.
- **XGBoost** gives more **balanced predictions** and can be considered for general use.
- **Neural Network** sacrifices accuracy but provides **maximum recall**, which is critical in healthcare settings where **missing a survivor is worse than overpredicting survival**.
- A **hybrid approach** (e.g., ensemble of ML + DL, or recall-focused threshold tuning) may yield stronger results in future work.

📌 **Final Statement:**

For clinical usefulness, the **Neural Network model** is more defensible because in medicine, it is far safer to **flag more patients as survivors (false positives)** than to miss actual survivors (false negatives).

.

References

1. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD.
2. Esteva, A., et al. (2019). A guide to deep learning in healthcare. Nature Medicine, 25(1), 24–29.
3. Cruz, J. A., & Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. Cancer Informatics.
4. Choi, E., et al. (2016). Doctor AI: Predicting clinical events via recurrent neural networks. JMLR.
5. Deo, R. C. (2015). Machine learning in medicine. Circulation, 132(20), 1920–1930.