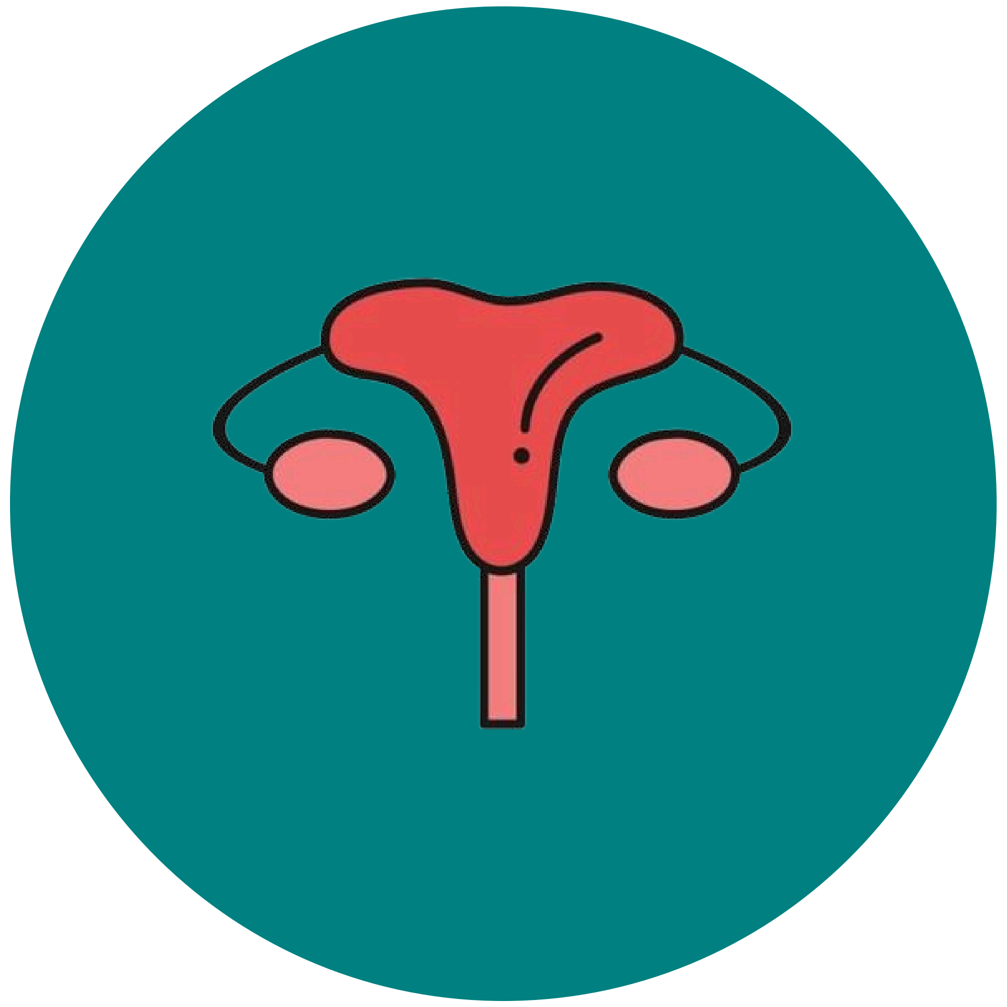
Multi-omic Inspired Approach for **Diagnosis** of **Ovarian Cancer**

Mishika Sood

Vardhaman Kalloli

Relevance of the Problem



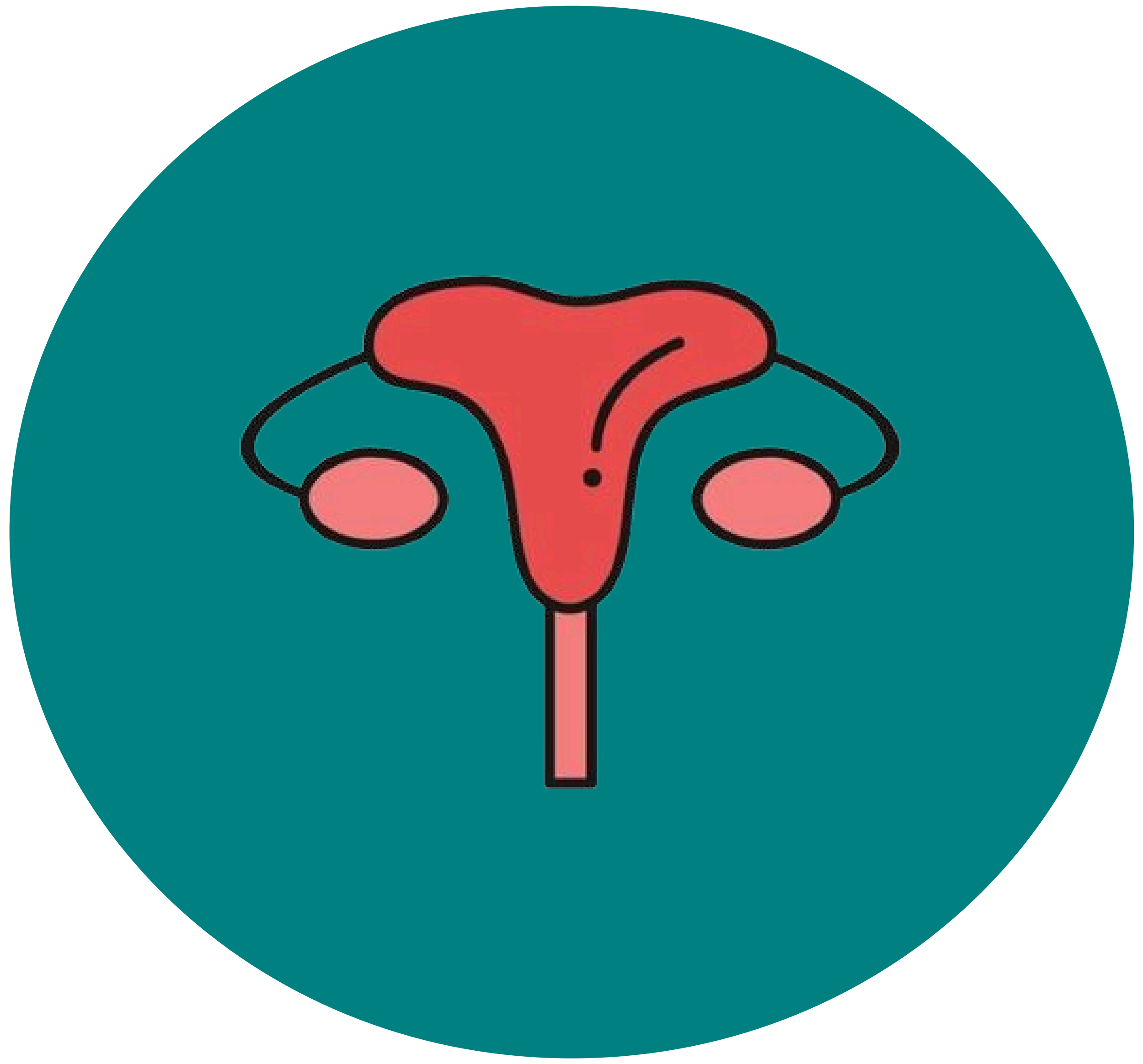
A woman's lifetime risk of ovarian cancer is **1 in 91**

Her chance of dying from it is **1 in 143**

In 2022, there were **206,956** OC deaths worldwide

American Cancer Society. (2024, January 19). Ovarian Cancer Statistics | How Common is Ovarian Cancer. Www.cancer.org. https://www.cancer.org/cancer/types/ovarian-cancer/about/key-statistics.html

Relevance of the Problem



In a study of **1,577** patients, doctors used a method called **frozen section analysis** to diagnose **Ovarian Tumours**

20% of the time, tumors were **under-diagnosed**

10.5% were **over-diagnosed**

Huang, Z., Li, L., Li, C., Ngaujah, S., Yao, S., Chu, R., Xie, L., Yang, X., Zhang, X., Liu, P., Jiang, J., Zhang, Y., Cui, B., Song, K., & Kong, B. (2018). Diagnostic accuracy of frozen section analysis of borderline ovarian tumors: a meta-analysis with emphasis on misdiagnosis factors. Journal of Cancer, 9(16), 2817–2824. https://doi.org/10.7150/jca.25883

Problem Statement

- Distinguishing benign from malignant ovarian tumors is difficult. Frozen section analysis shows only **69.2% accuracy** for borderline tumors, with **20.2% under-diagnosed** and **10.5% over-diagnosed**.

- **Missing** a **malignant tumor** can delay treatment and worsen outcomes. Over-diagnosis can lead to unnecessary surgeries, causing avoidable *risks and reduced quality of life*.

- Current tools like **imaging** and **CA-125 test** lack precision. Even with combined markers, the positive predictive value remains **low**, highlighting the need for **multi-model, data-driven** approaches.

Abramowicz, J. S., Condous, G., & Timmerman, D. (2018). Ovarian mass-differentiating benign from malignant. Why the International Ovarian Tumour Analysis rules should be implemented in Australasia. Australasian Journal of Ultrasound in Medicine, 21(3), 121–124. https://doi.org/10.1002/ajum.12108

Literature Review

Paper 1:

Context - This study aimed to develop a deep learning model using grayscale and color Doppler ultrasound images, comparing its performance with the O-RADS and expert assessments

Data - Analysis of 422 women with ovarian tumors, utilising grayscale and colour Doppler ultrasound images.

ML Approach - Two ResNet-based models—decision fusion and feature fusion

Gaps - Solely on imaging data, without incorporating other clinical information such as serum markers or patient history

Wang, Z., Luo, S., Chen, J., Jiao, Y., Cui, C., Shi, S., Yang, Y., Zhao, J., Jiang, Y., Zhang, Y., Xu, F., Xu, J., Lin, Q., & Dong, F. (2024). Multi-modality deep learning model reaches high prediction accuracy in the diagnosis of ovarian cancer. IScience, 27(4), 109403. https://doi.org/10.1016/j.isci.2024.109403

Paper 2:

Context - The study assessed a multi-modal deep learning model combining ultrasound, menopausal status, and serum indicators to classify ovarian tumors.

Data - 1,054 cases (699 benign, 355 malignant) using retrospective ultrasound and clinical data.

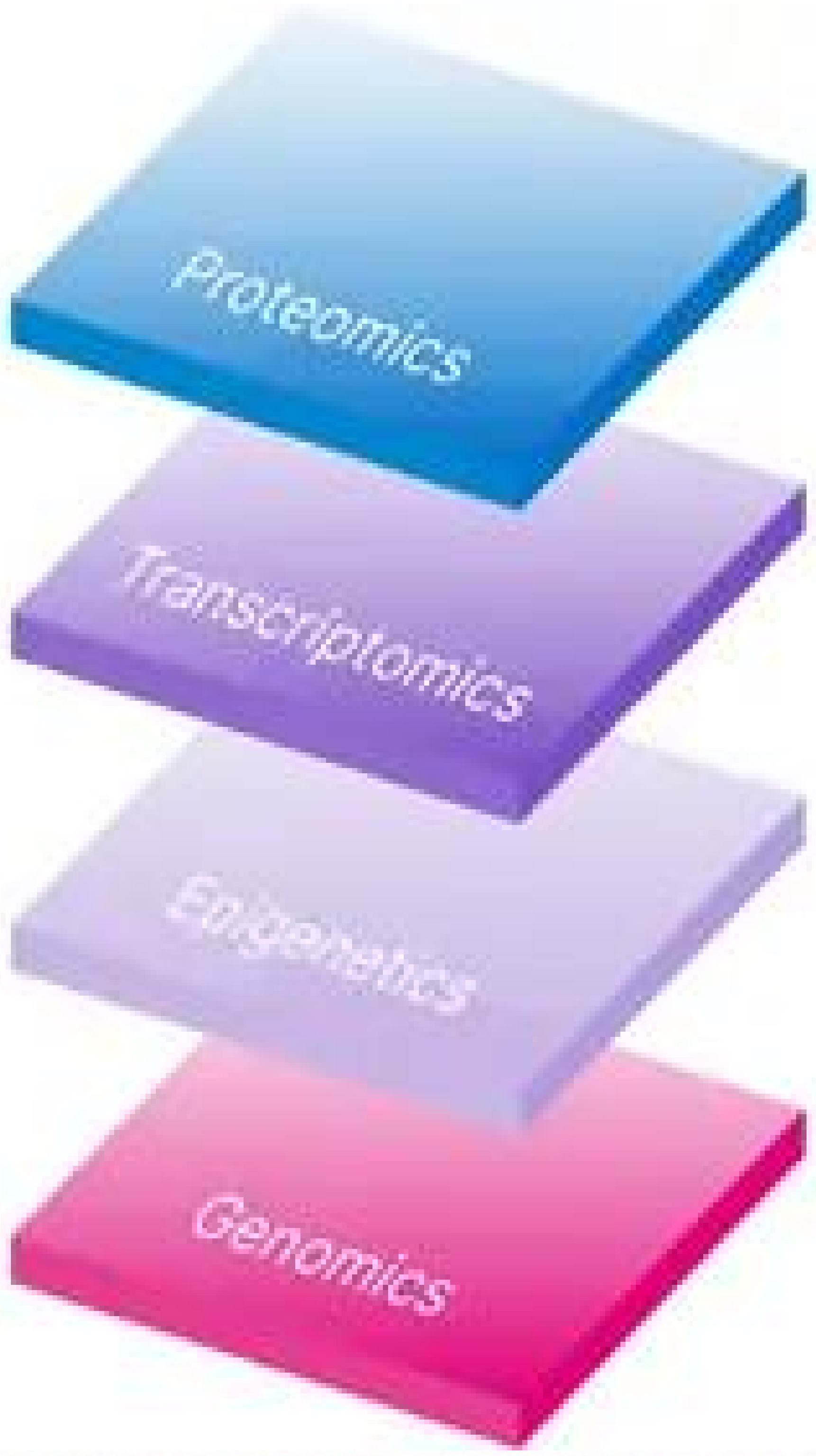
ML Approach - Three ResNet-50 models to evaluate how adding clinical data improves ovarian tumor classification.

Gaps - The study lacked advanced fusion techniques and broad validation across diverse clinical settings.

Chen, H., Yang, B.-W., Qian, L., Meng, Y.-S., Bai, X.-H., Hong, X.-W., He, X., Jiang, M.-J., Yuan, F., Du, Q.-W., & Feng, W.-W. (2022). Deep Learning Prediction of Ovarian Malignancy at US Compared with O-RADS and Expert Assessment. Radiology, 304(1), 106–113. https://doi.org/10.1148/radiol.211367

Rationale

Ovarian cancer detection is inherently multi-modal – it benefits from combining clinical, tumor markers and haematological features. A hybrid model can overcome the limitations of any single data source.



For example, a rise in CA-125 could trigger a closer look at haematological indices (like a high neutrophil-to-lymphocyte ratio) which together raise your suspicion.

By fusing modalities, the model can learn these conditional inferences. Research strongly supports multi-modal integration: integrated models using clinical data, tumor markers, and haematological data can outperform models based on any single type of data. Numerous reviews show that multi-omic approaches like this consistently outperform single-source models.

Hatamikia, S., Nougaret, S., Panico, C., Avesani, G., Nero, C., Boldrini, L., Sala, E., & Woitek, R. (2023). Ovarian cancer beyond imaging: integration of AI and multiomics biomarkers. European Radiology Experimental, 7(1). https://doi.org/10.1186/s41747-023-00364-7

USING MACHINE LEARNING TO PREDICT OVARIAN CANCER

Nature

The dataset comprises clinical, hematological, and tumor marker data from 349 Chinese female patients, aimed at differentiating between benign ovarian tumors and ovarian cancer.

Reason

It offers a diverse set of 49 features across multiple domains (tumor markers, clinical and hematological data), making it ideal for training and evaluating multi-model machine learning architectures.

Collection

Data were retrospectively collected from hospital records, encompassing demographics, blood tests, general chemistry, and tumor markers and shared under a CC BY-NC 3.0 license maintaining patient privacy and restricted to non-commercial usage.

Size & Features

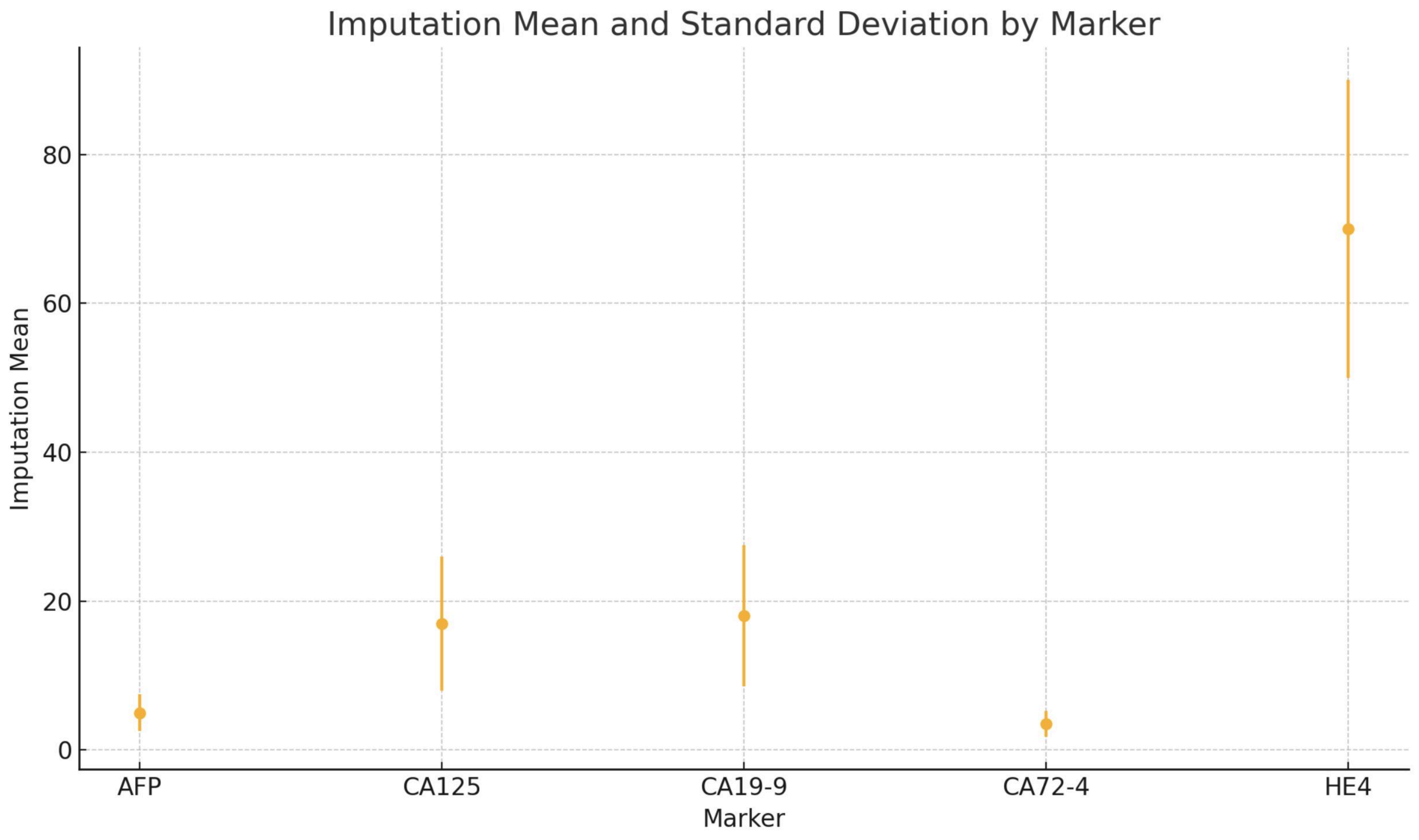
It includes 349 patient records with 49 variables, covering blood routine tests, general chemistry, and tumor markers.

Mi, Qi; Jiang, Jingting; Znati, Ty; Fan, Zhenjiang; Li, Jundong; Xu, Bin; Chen, Lujun; Zheng, Xiao; Lu, Mingyang (2020), "Data for: USING MACHINE LEARNING TO PREDICT OVARIAN CANCER", Mendeley Data, V11, doi: 10.17632/th7fztbrv9.11

Feature Preprocessing

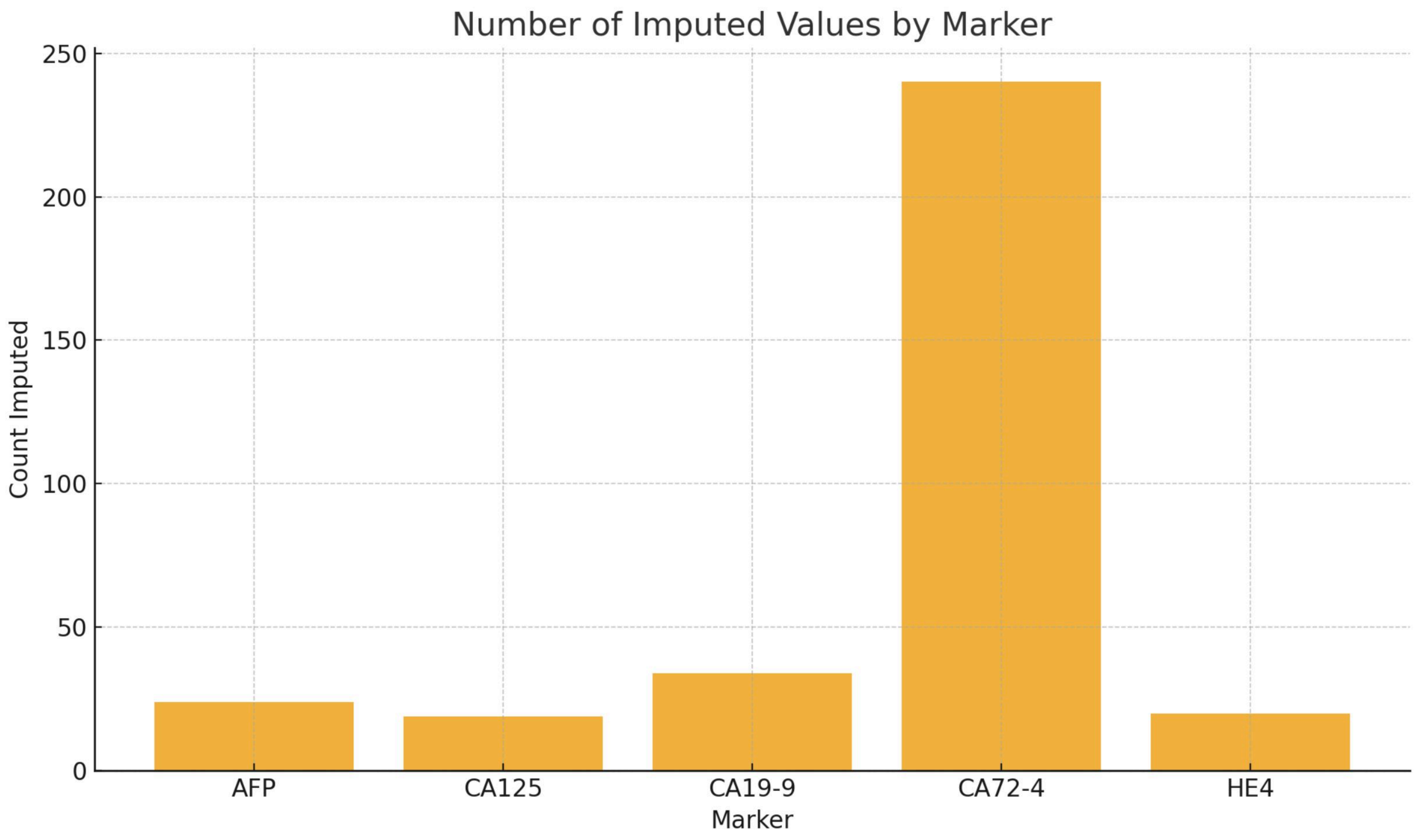
Missing-value handling

Missing entries in each column were imputed with the column mean



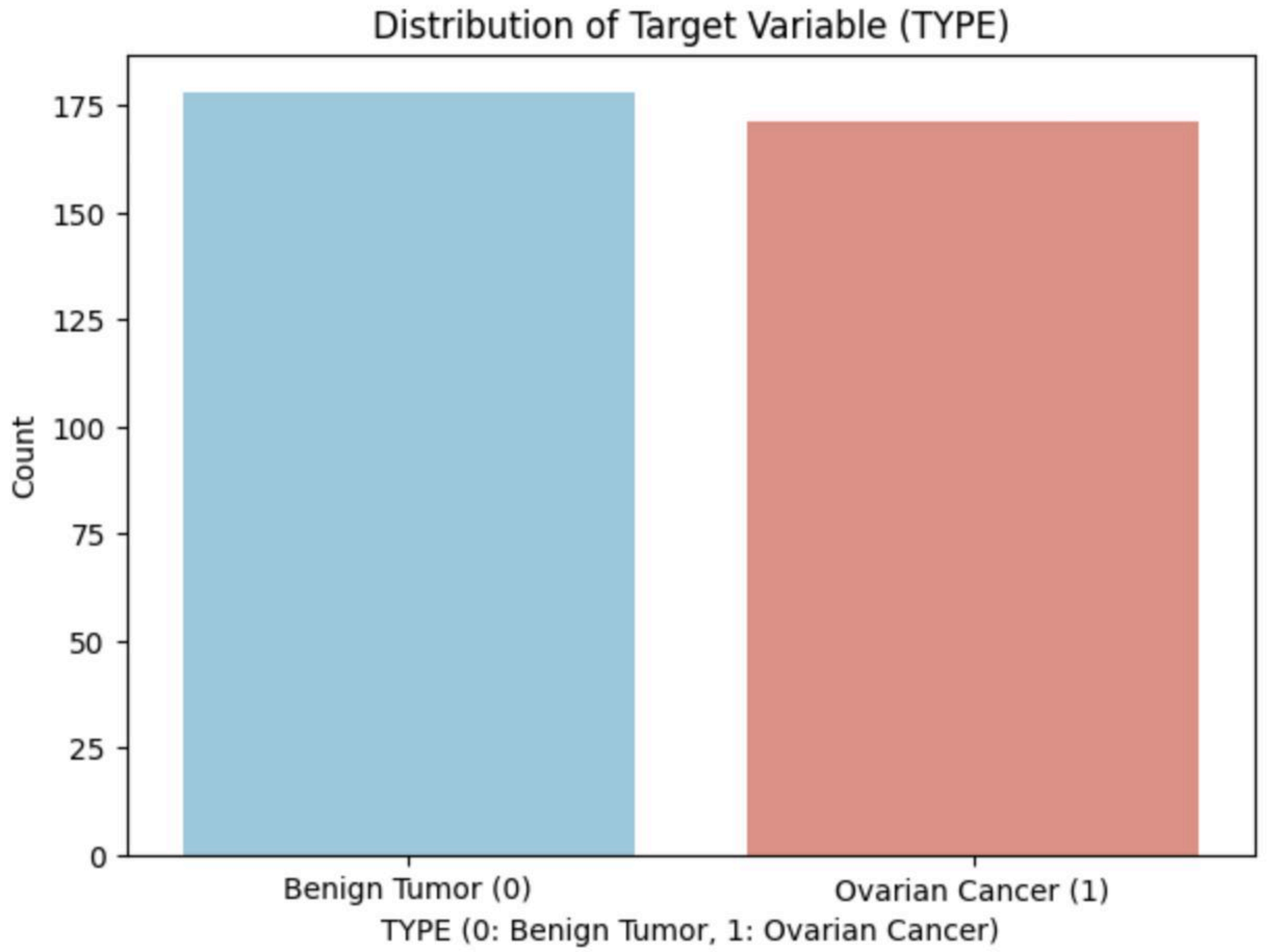
Scaling

In both modeling models we've normalized all numeric features with *MinMaxScaler*

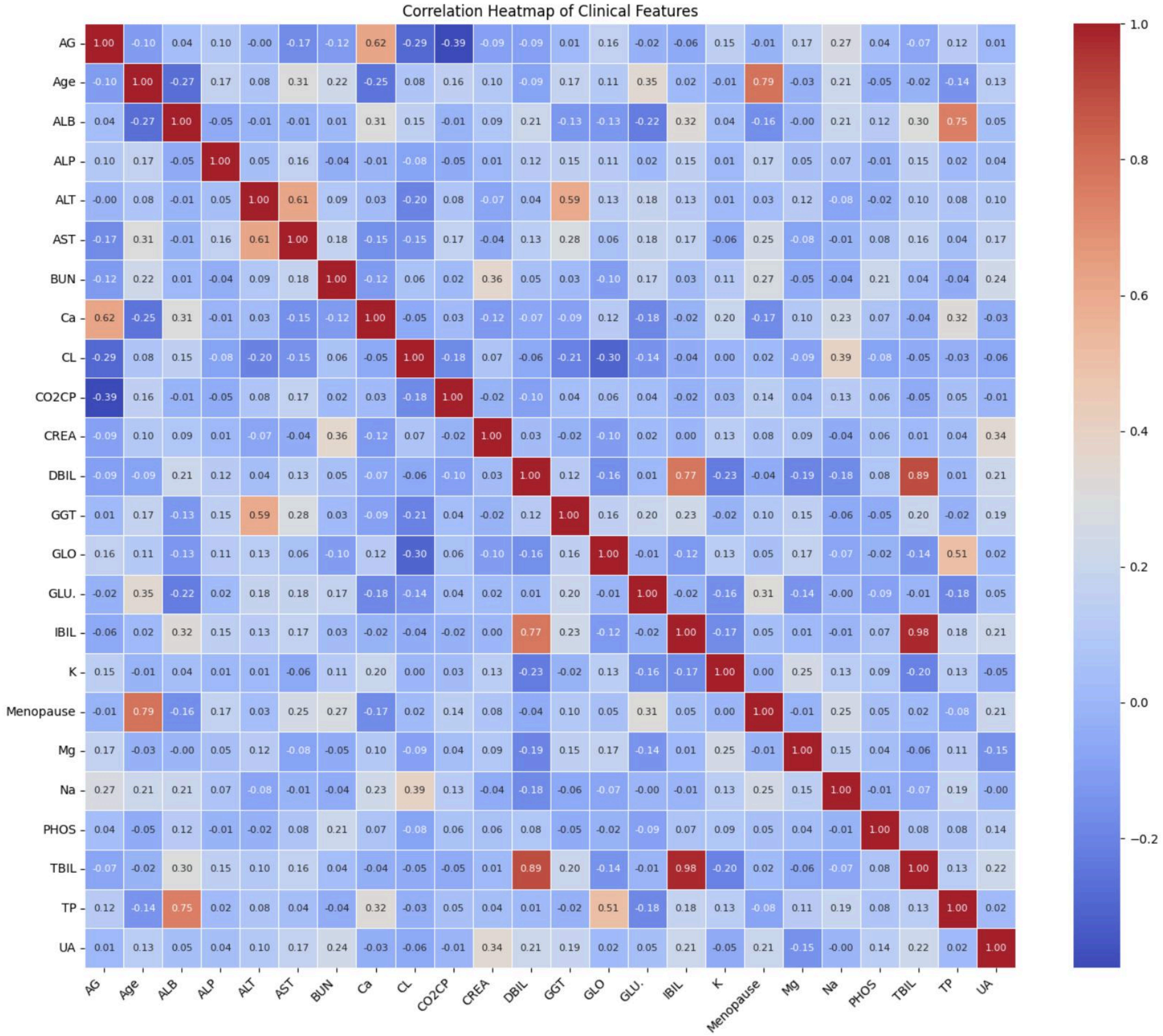


Data Visualisations

Target Variables Distribution

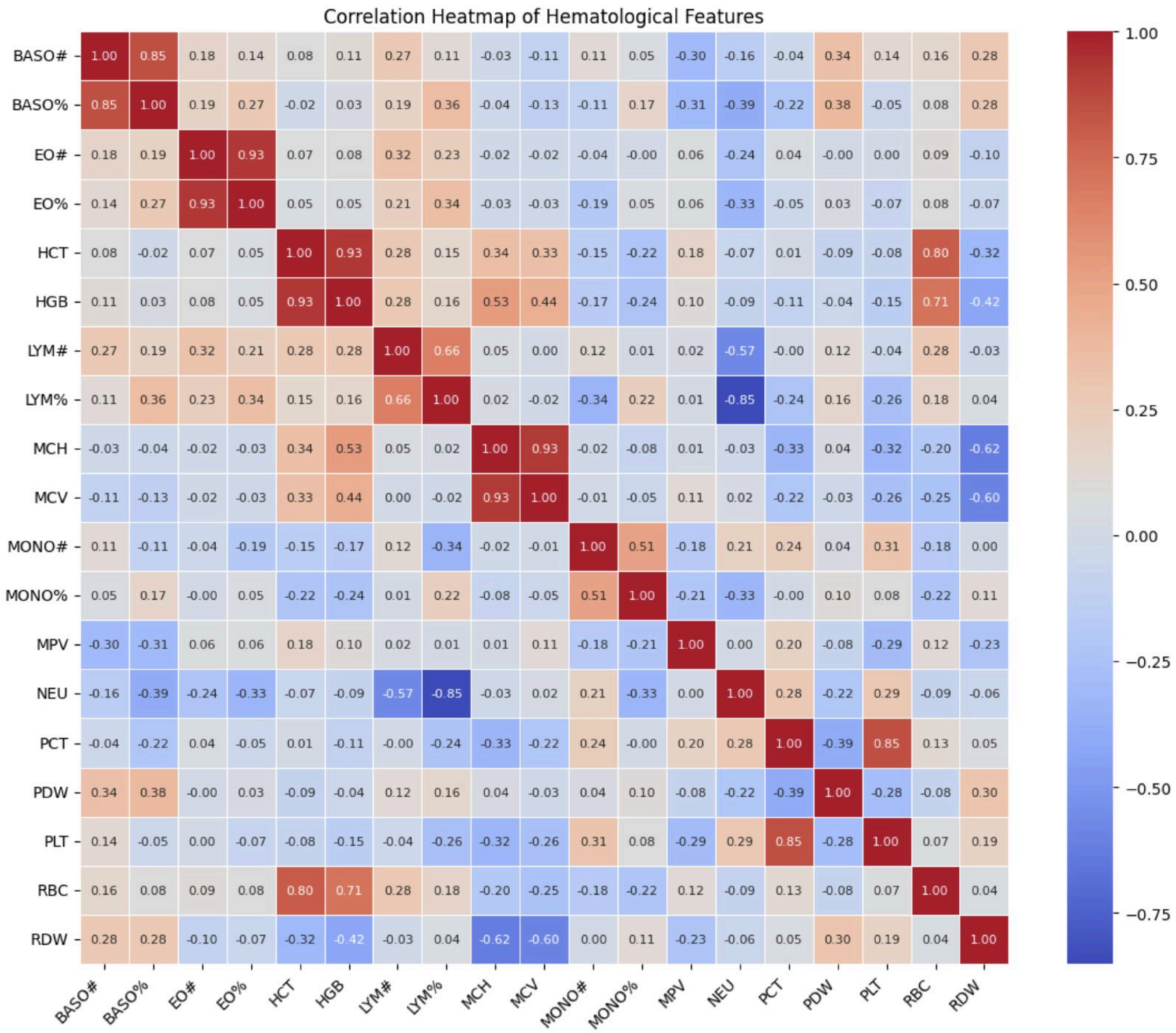


Clinical Features

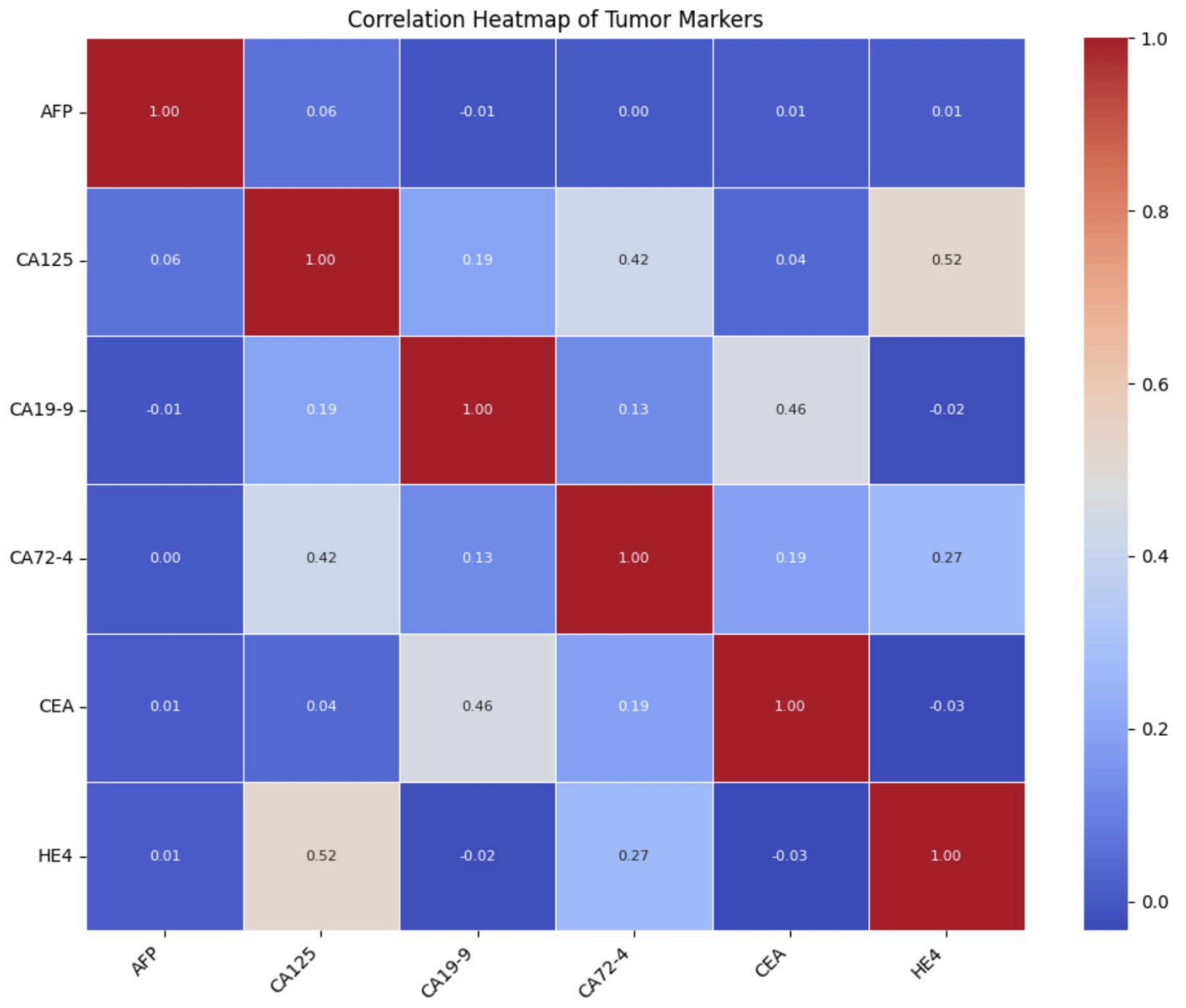


Data Visualisations

Haematological Features



Tumor Markers



Naive methods

Deep Learning (MLP)

- Two separate MLPs: Clinical + Hematological labs (43 features) & Tumor Markers (6 biomarkers)

Why:
- Captures complex, non-linear interactions and hierarchies in the data.
- Early-stopping, dropout, L2 regularization to combat overfitting on a small cohort.

How they work:
- Stacked dense layers (128→64→32) with ReLU activations
- Dropout + batch-norm + L2 shrinkage
- Backprop on binary cross-entropy (or focal loss with tumor markers)

Gradient Boosting Ensemble (XGBoost + CatBoost Soft Voting)

- Independently tuned XGBoost & CatBoost via randomized CV on balanced-accuracy, then average their predicted probabilities.

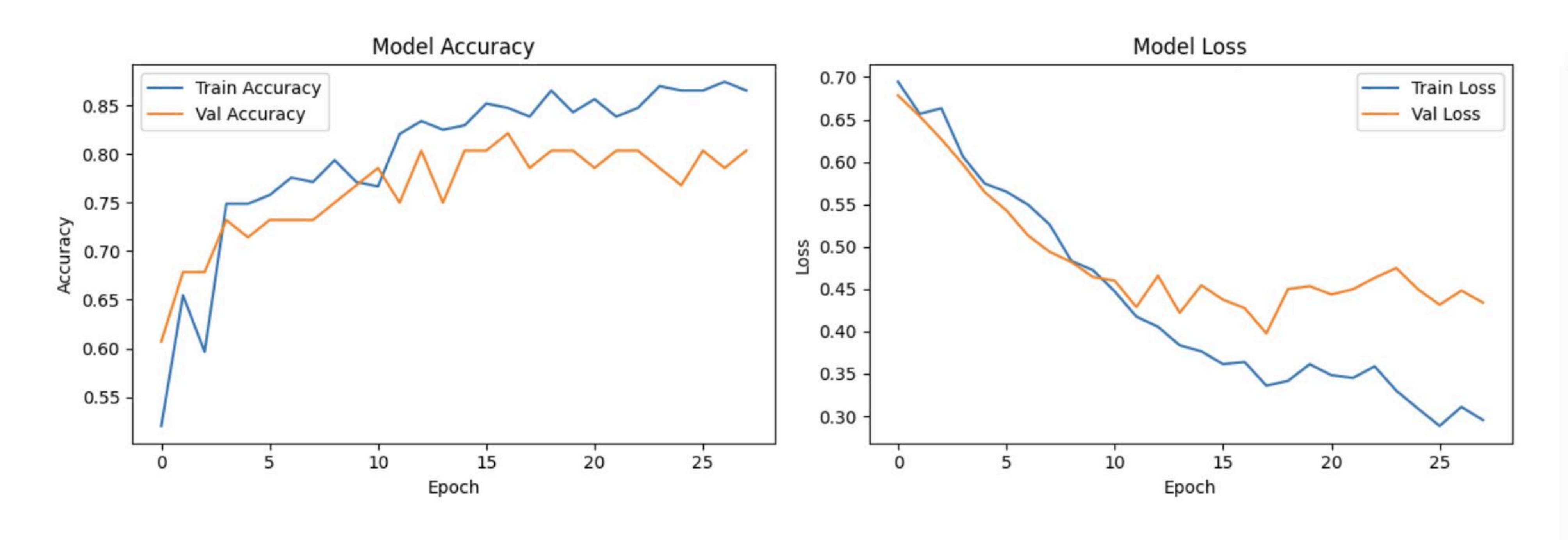
Why:
- Built-in regularization, missing-value handling, and class-imbalance controls (scale_pos_weight, per-leaf weights).
- Ensembling smooths out each model’s high-variance mistakes.

How they work:
- Greedy, additive tree construction on residual gradients
- Per-tree shrinkage & regularization (L1/L2)
- Each model outputs P(cancer) → final P = average → thresholded for label

Naive methods

Deep Learning (MLP)

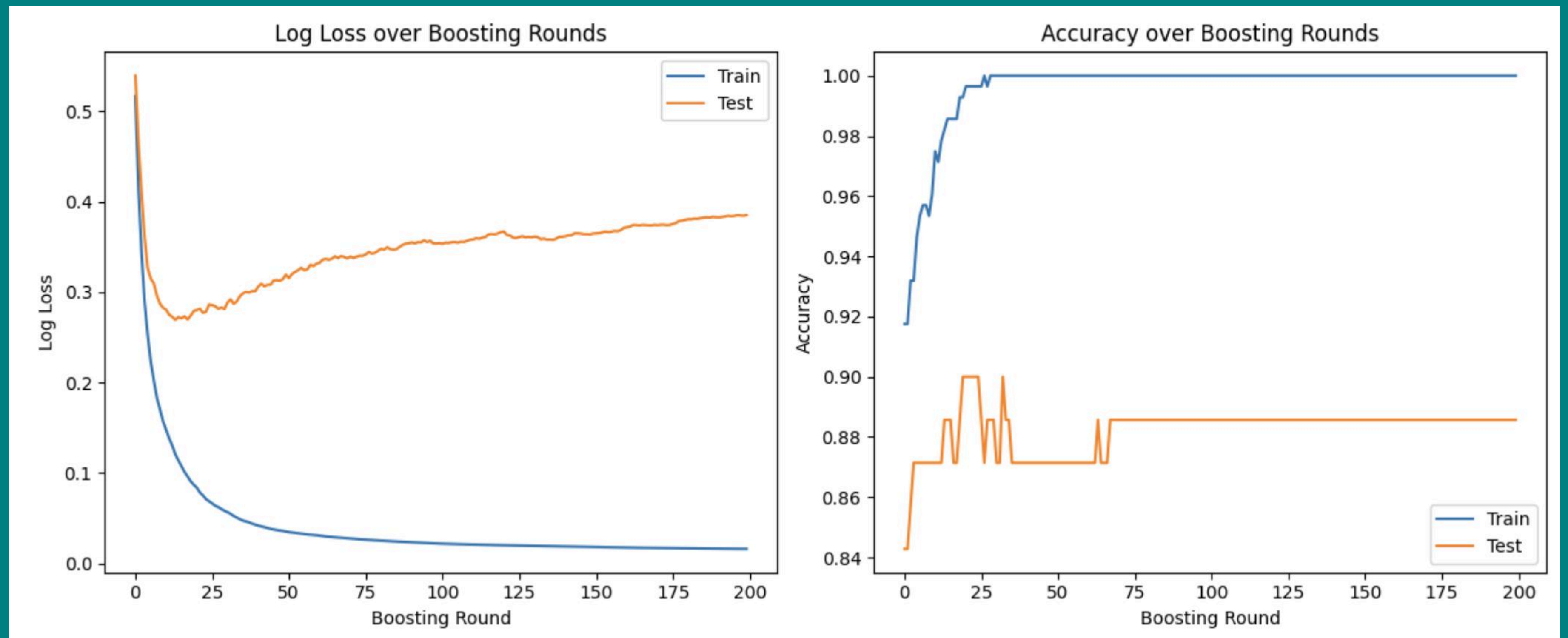
- Two separate MLPs: Clinical + Hematological labs (43 features) & Tumor Markers (6 biomarkers)



- Test Accuracy (Deep Learning): **84.29%**

Gradient Boosting Ensemble (XGBoost + CatBoost Soft Voting)

- Independently tuned XGBoost & CatBoost via randomized CV on balanced-accuracy, then average their predicted probabilities.



- Test Accuracy: **88.5%**

ML Methodology Used

Random Forest

A RandomForestClassifier (200 trees, class_weight='balanced')

Why

Simple, robust baseline.

Often surprisingly strong on small, noisy datasets—and in our case, it ultimately outperformed the more complex learners.

How they work

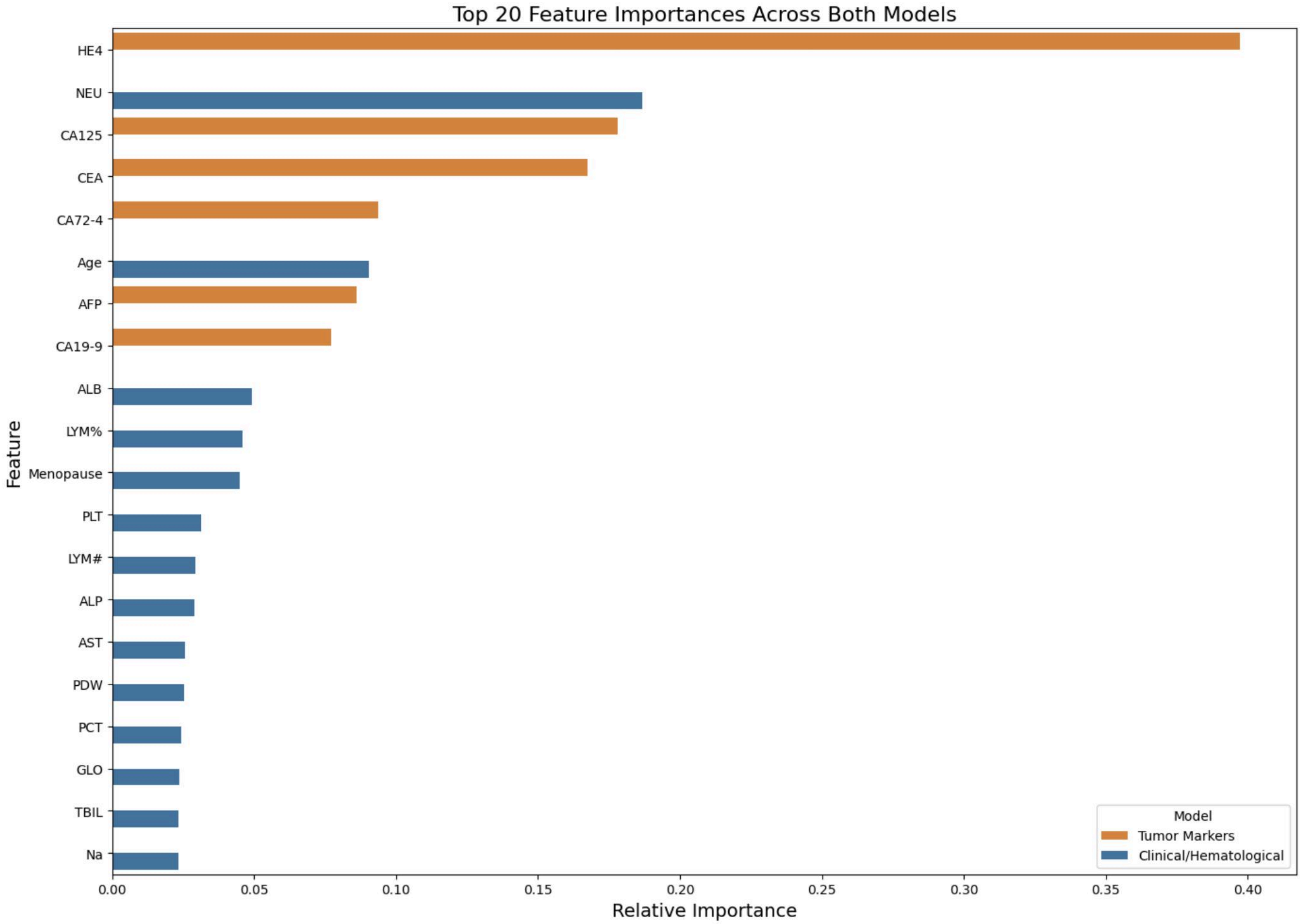
Trains many decision trees on bootstrapped samples with random feature subsets → majority vote

Test Accuracy: 90%

ML Methodology - Random Forest Fusion

We combined the **probability outputs** of the tumor-marker RF and clinical/hematology RF into a meta-learner, allowing **each model's strengths** to contribute where they're most informative.

By fusing across domains, we capture non-overlapping signals (e.g. biochemical vs. hematological), improving robustness **against noise** or **missing pattern**s in any single dataset.



UCSF Health. (2024, May 21). Taking charge: Who gets ovarian cancer? ucsfhealth.org. https://www.ucsfhealth.org/education/taking-charge-who-gets-ovarian-cancer

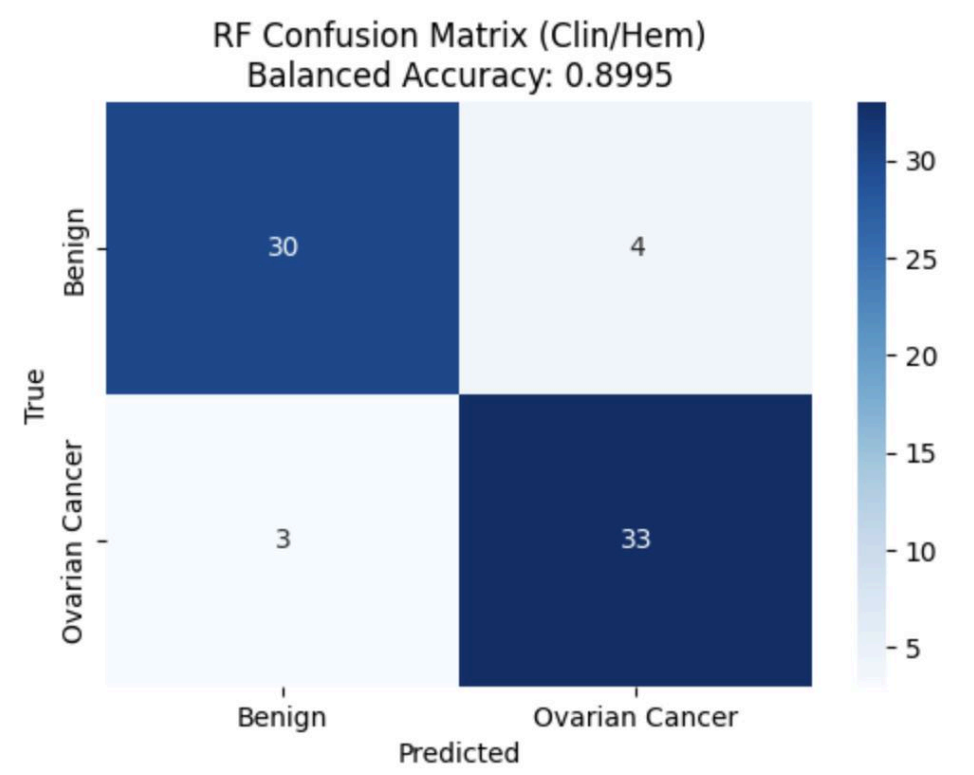
Performance Metrics

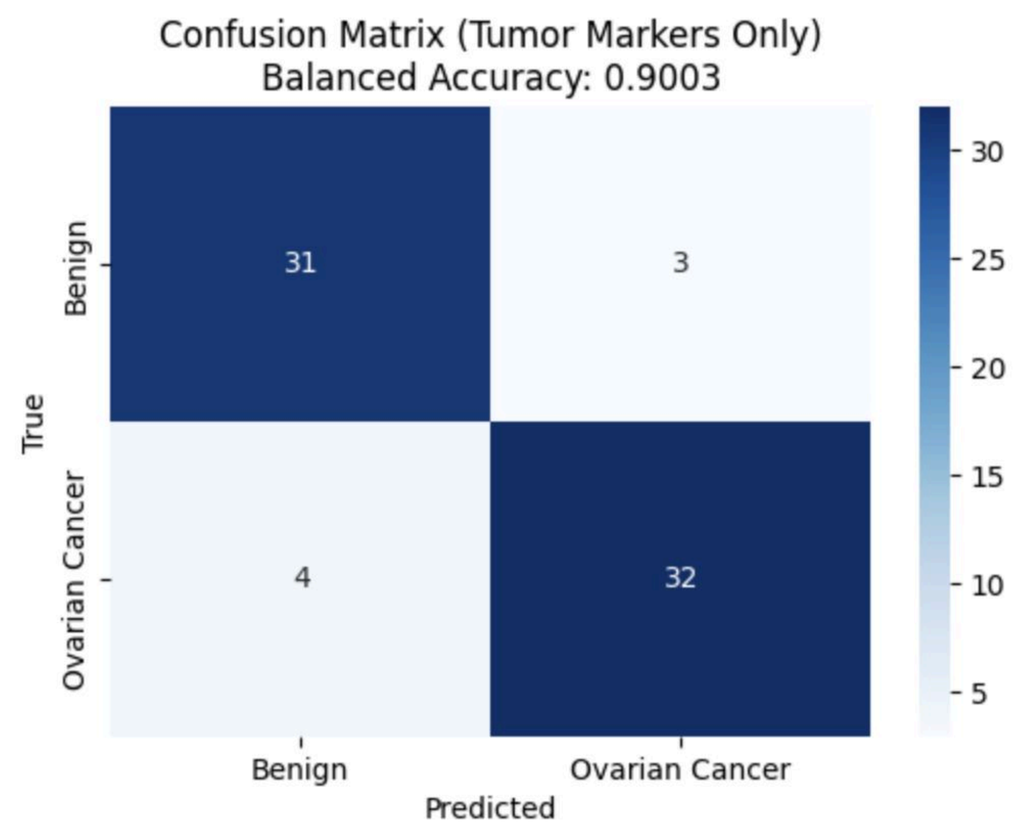
Accuracy = 94% - shows good amount cases correctly classified by the fused model.

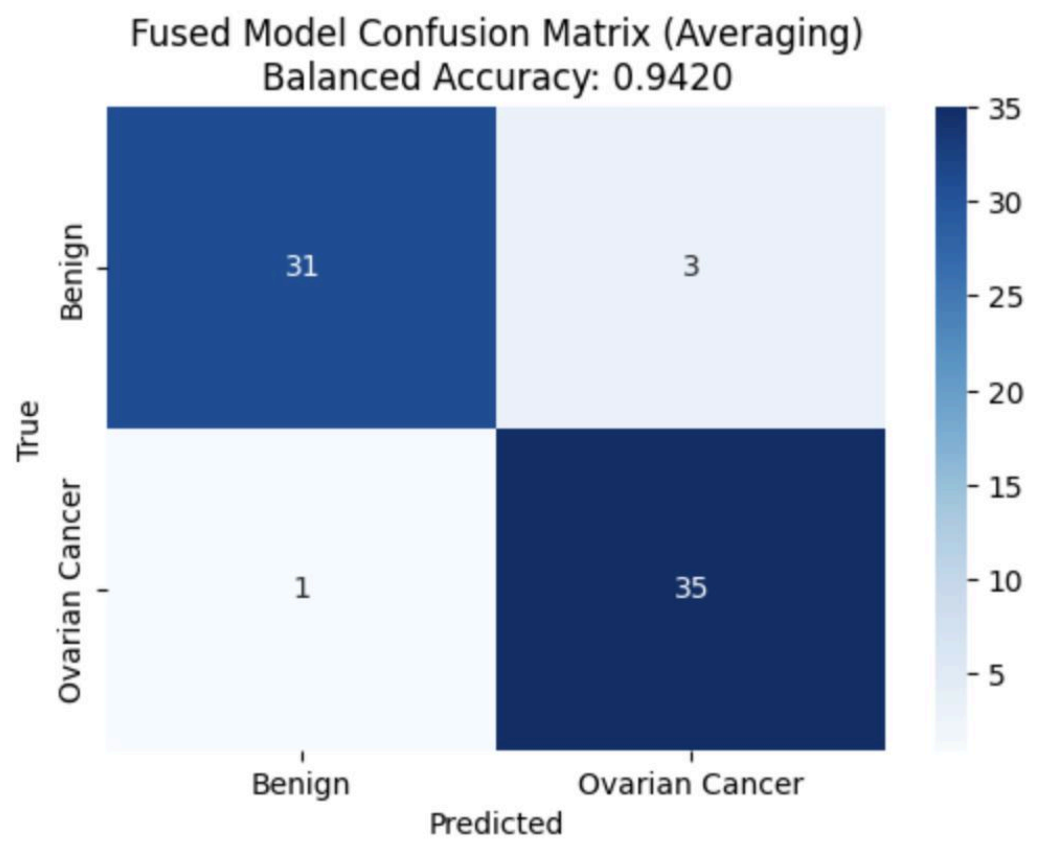
ROC AUC = 0.92 – shows excellent overall discrimination between benign and malignant cases.

Sensitivity = 0.89 & **Specificity = 0.87** – high true-positive rate for cancer detection, with few false alarms.

F1-Score = 0.88 – **balances precision (0.90)** and **recall (0.87)**, indicating reliable positive predictions.





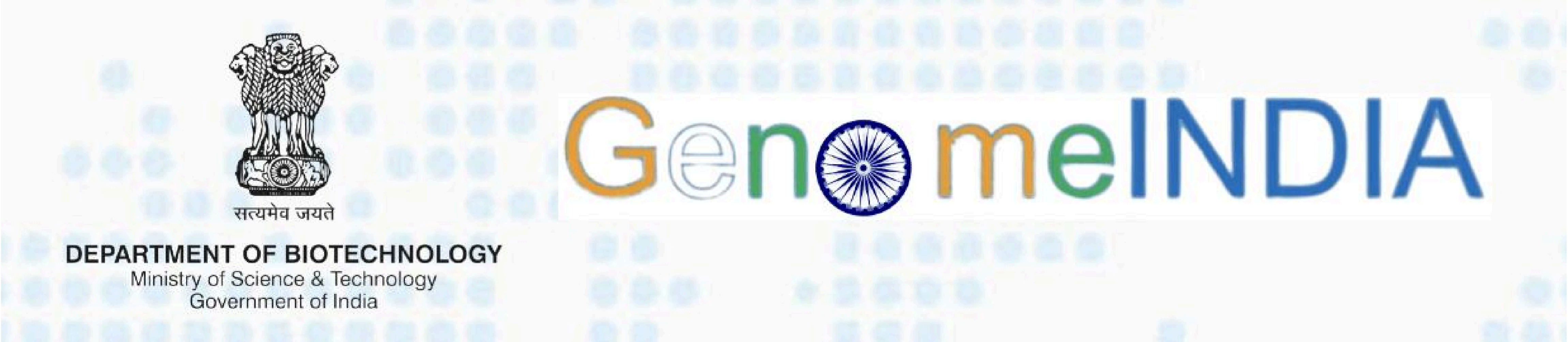


Challenges & How We Tackled Them

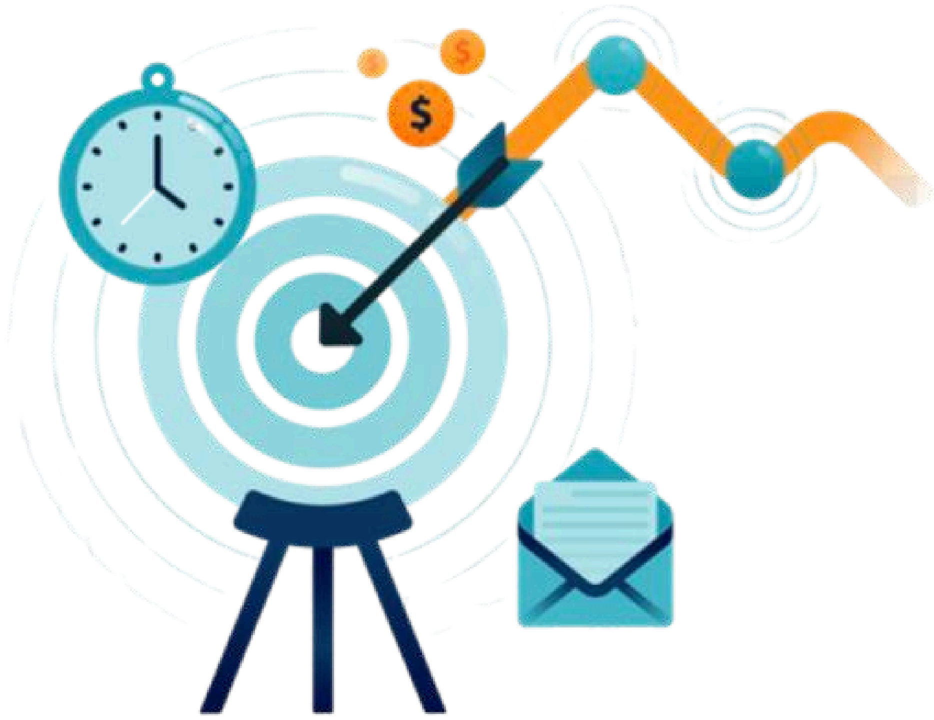
Challenge	Impact	Mitigation
Small, Noisy Dataset	Overfitting; models collapsed to chance (≈50 % balanced acc).	– Heavy regularization (dropout, L2, early-stop) – Imputation with clinical priors + mean fill
Class Imbalance (~52/48)	Tendency to predict majority class → low sensitivity.	– scale_pos_weight in boosting – Class weights in DL – Threshold optimization (0.1–0.9 scan)
Algorithmic Complexity & Tuning	8–10 hyperparams per boosting model → multi-hour grid search.	
Missing & Mixed-Type Data	Parsing errors, feature gaps.	– Regex-based string cleaning – Biomarker-specific normal imputation – Column-mean fill

Potential Applications and Impact

India Genome Project -
Enhanced risk prediction, tailored screening protocols, development of targeted therapies, public health benefits- tailored specifically to Indian population.



Clinical Decision Support - Enables general practitioners to rapidly distinguish benign from malignant ovarian cases using fused multi-domain data, improving diagnostic accuracy and patient referral decisions.



Screening Protocol Development - Refine screening guidelines and protocols for early diagnosis will lead to better quality of life due to reduced mental toll.

Plaksha Deployability

This solution enables general practitioners (GP) to rapidly distinguish between benign and malignant tumors, this can be implemented in our infirmary as an assistive technology to the GP to analyse the blood report and make classification.

Scaling Challenges

In order to use this in a fully fledged production environment, some potential challenges and factors to be considered are :

- **Data Integration:** Lab reports arrive in different formats/units across hospitals. We need robust parsers and unit-conversion routines.
- **Multi-Modal Fusion:** To boost accuracy, incorporate imaging or image derived features, then fuse those with lab-based predictions in a final ensemble.
- **Privacy and Compliance:** Handling real patient data and complying with local health-data regulations.

Thank You