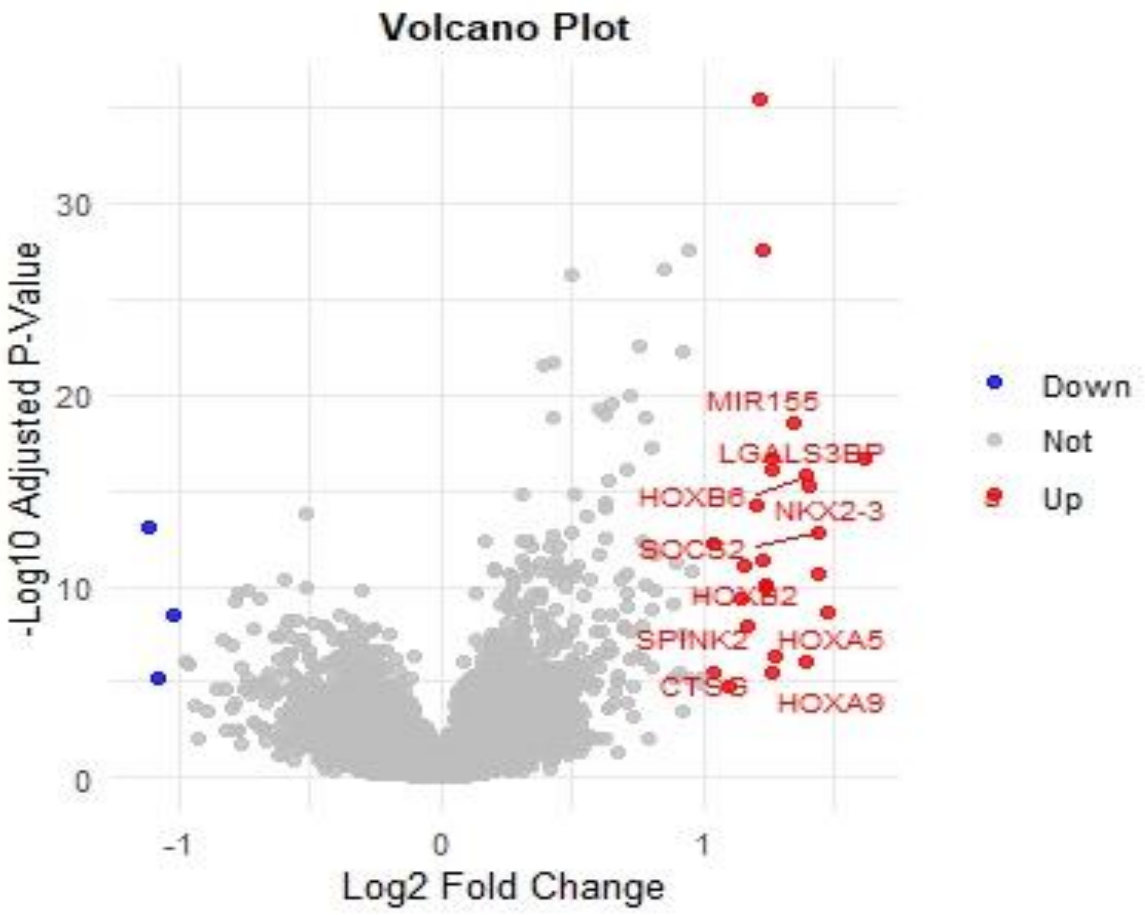


Figure 1: Volcano Plot of Differentially Expressed Genes (DEGs) Between FLT3-ITD-Positive and FLT3-ITD-Negative AML Samples. The volcano plot visualises DEGs identified by limma. Each point represents a gene; the **x-axis** shows the $\log_2$ fold change between FLT3-ITD-positive and -negative samples, while the **y-axis** shows $-\log_{10}$ of the FDR-adjusted *p*-value. Genes significantly upregulated (logFC > 1, adjusted $p < 0.05$) are coloured **red**, and significantly downregulated genes (logFC < –1, adjusted $p < 0.05$) are in **blue**. **Grey points** denote non-significant genes. The **top 10 DEGs** based on absolute $\log_2$ fold change are labelled, illustrating distinct transcriptomic alterations associated with FLT3-ITD mutation status.

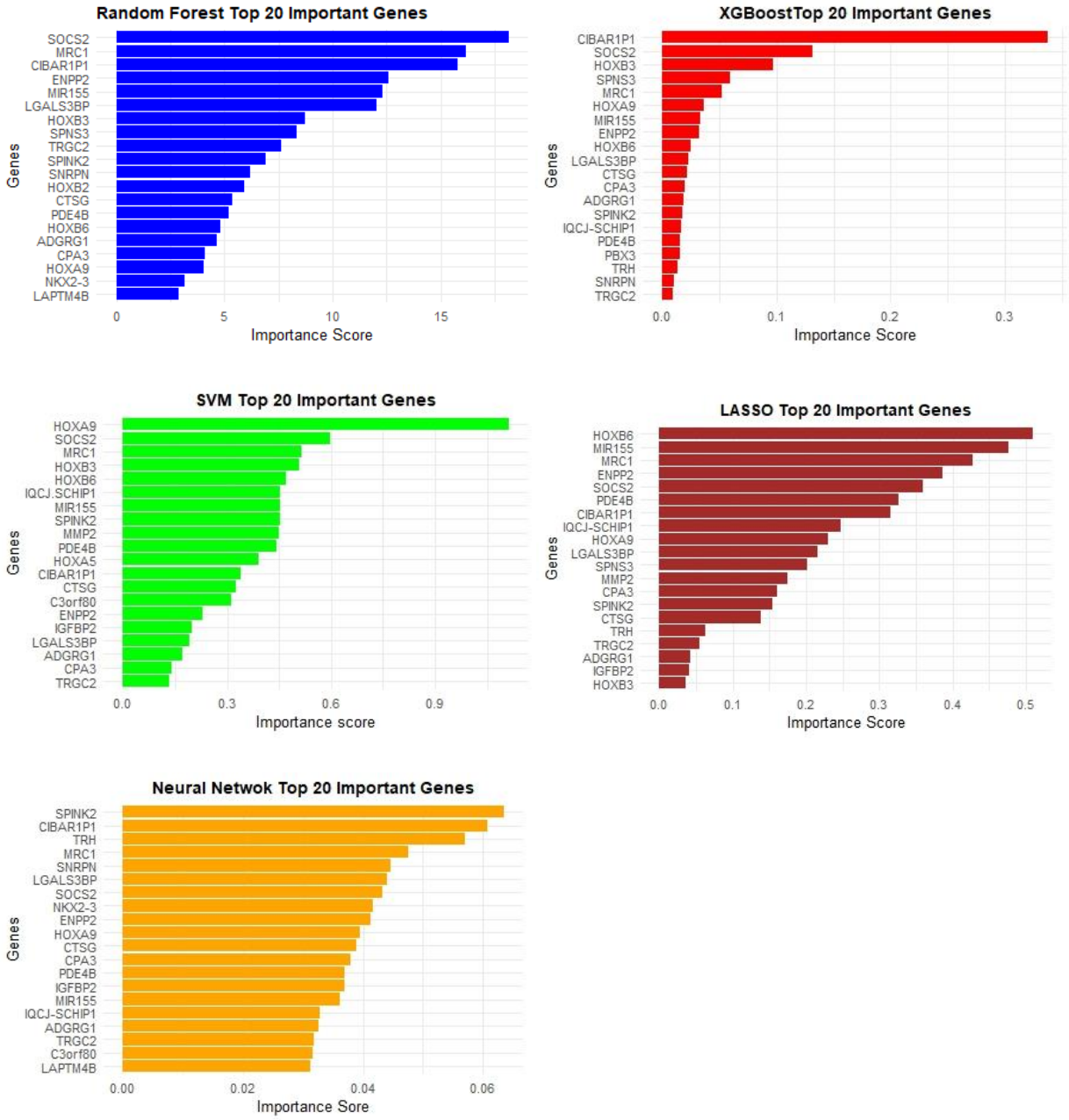


Figure 2. Horizontal Bar Charts of Top 20 Important Genes Identified Each of the Five ML Model Classifier. The bar chart ranks genes by importance score on the **x-axis** while the **y-axis** lists gene names. **Upregulated genes**, such as **SOCS2** and **MRC1**, were among the highest ranked in each model. Based on the trained ML model, the chart highlights the top 20 genes most informative for distinguishing FLT3-ITD-positive from negative AML samples.

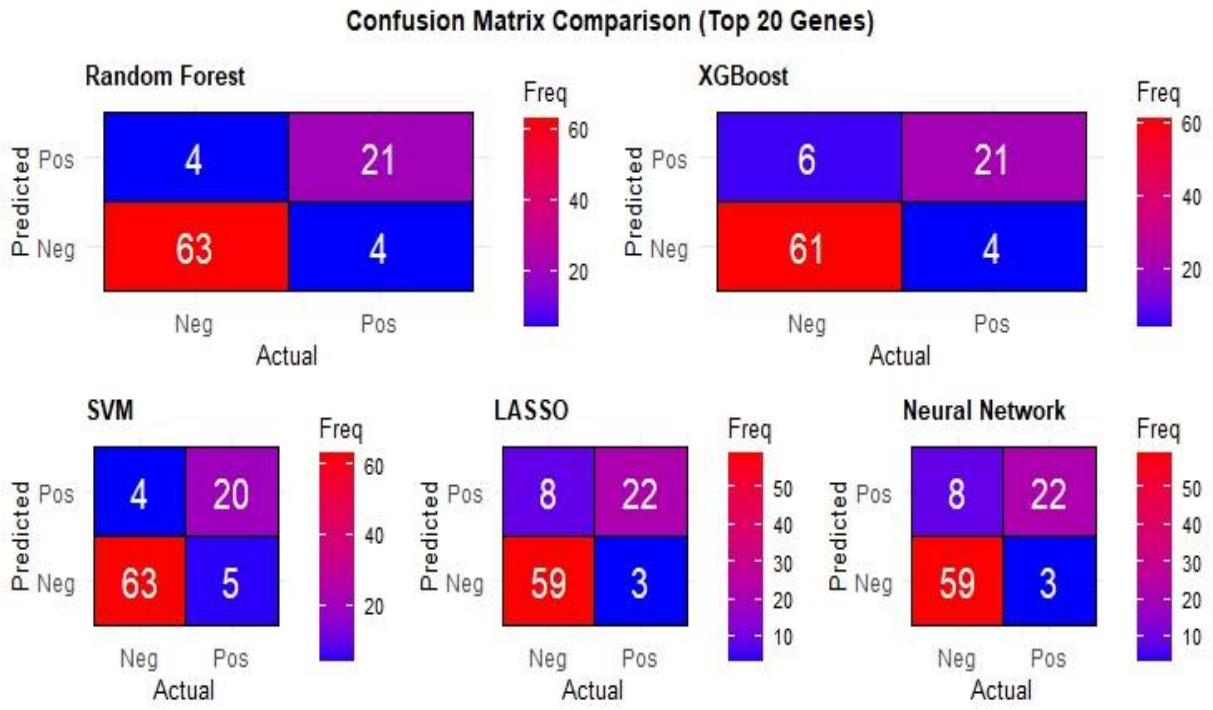


Figure 3. Confusion Matrices for Machine Learning Classifiers Retrained on Top 20 Genes. Each confusion matrix shows the classification performance of the ML model (RF, XGB, SVM, LASSO and NN) on the independent test set, following retraining on the model's top 20 ranked genes. The **x-axis** denotes actual FLT3-ITD status (Negative or Positive), while the **y-axis** shows predicted labels. Colour intensity reflects frequency **(red = high, blue = low)**. The matrices visually capture variation in classifier sensitivity and specificity across different ML strategies.

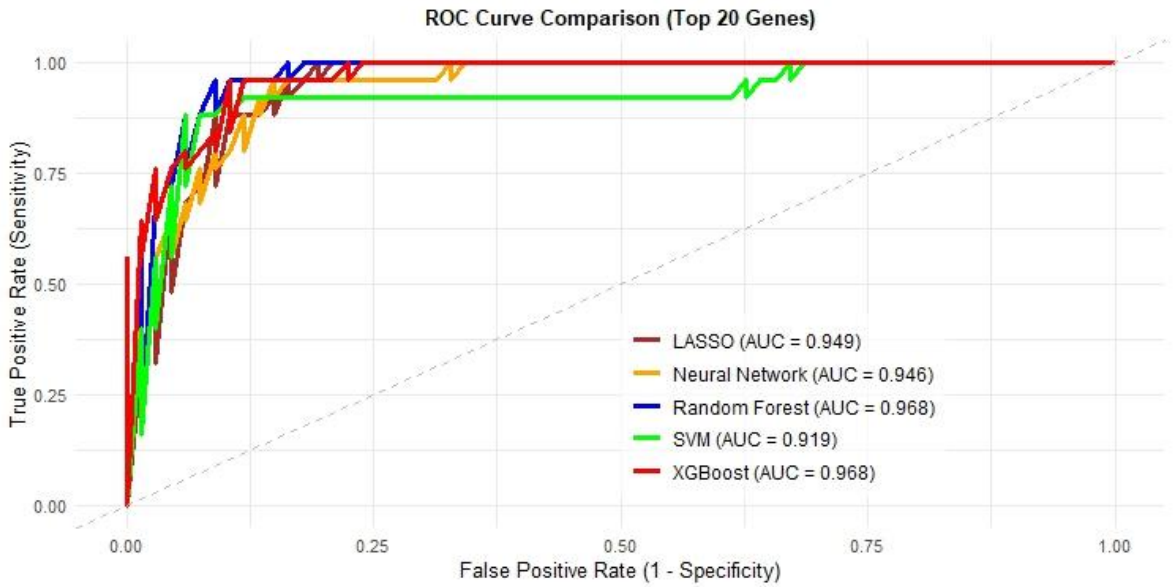


Figure 4. Overlay of ROC curves for all five models. Curves closer to the top-left corner indicate stronger discriminatory ability. All models exceeded an **AUC of 0.91,** confirming high predictive accuracy when using the top 20 genes selected per algorithm.

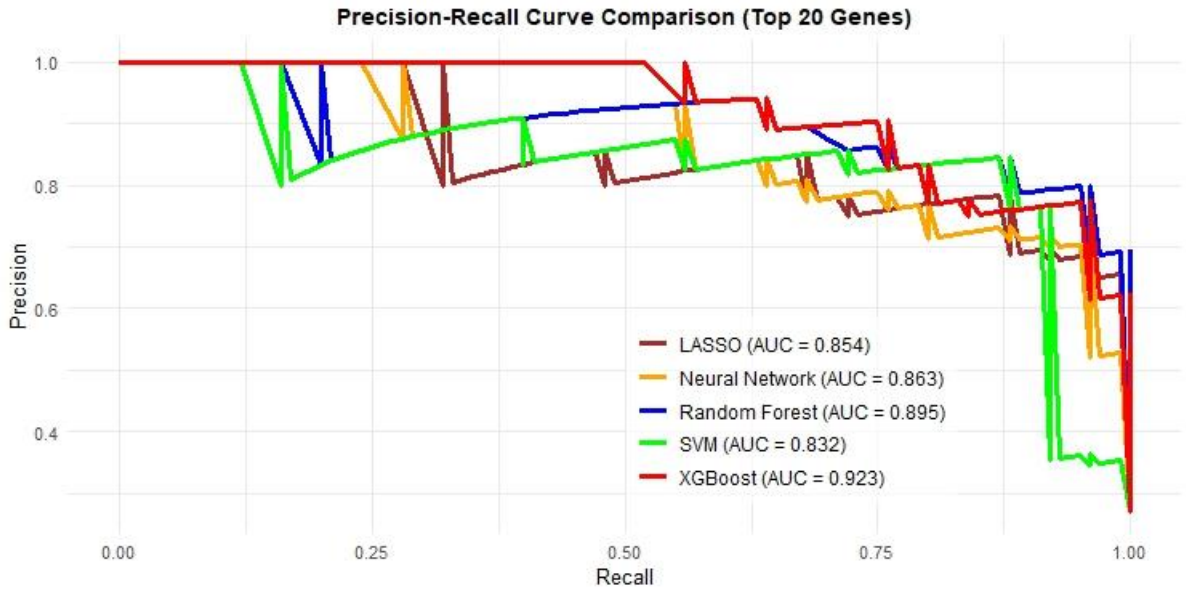


Figure 5. Combined PR curve overlay for all five models. Curves closer to the top-right reflect better performance in retaining precision while maximising recall. This comparison emphasises ensemble methods' robustness and predictive signatures' stability under data imbalance.

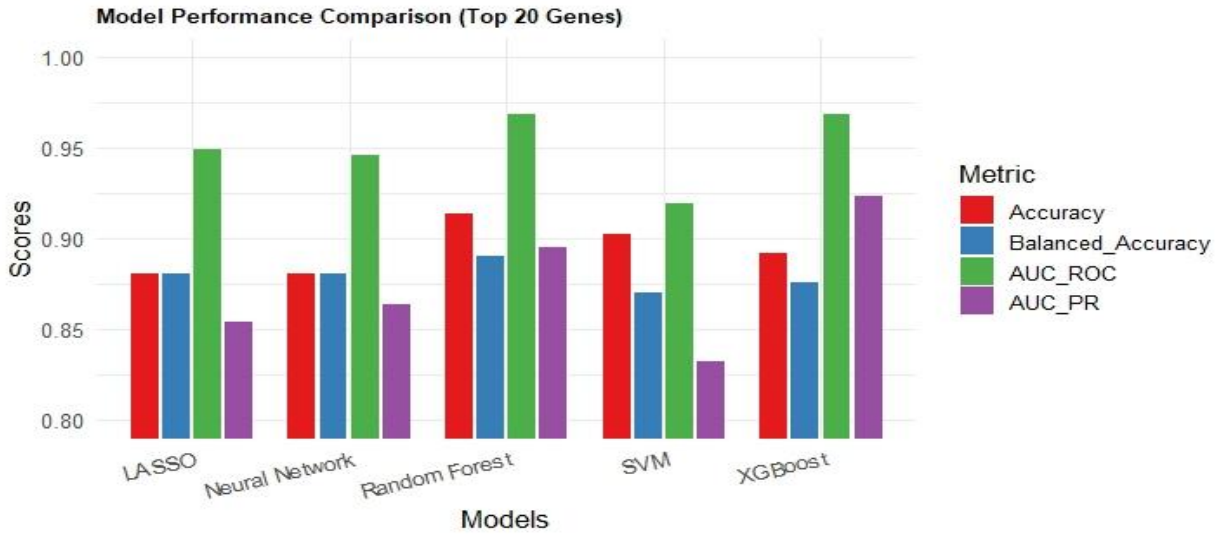


Figure 6. Model Performance Comparison Based on the Top 20 Genes Selected From Each Algorithm. Performance was evaluated using Accuracy, Balanced Accuracy, ROC AUC and PR AUC. **XGB** and **RF** exhibited the highest overall performance.

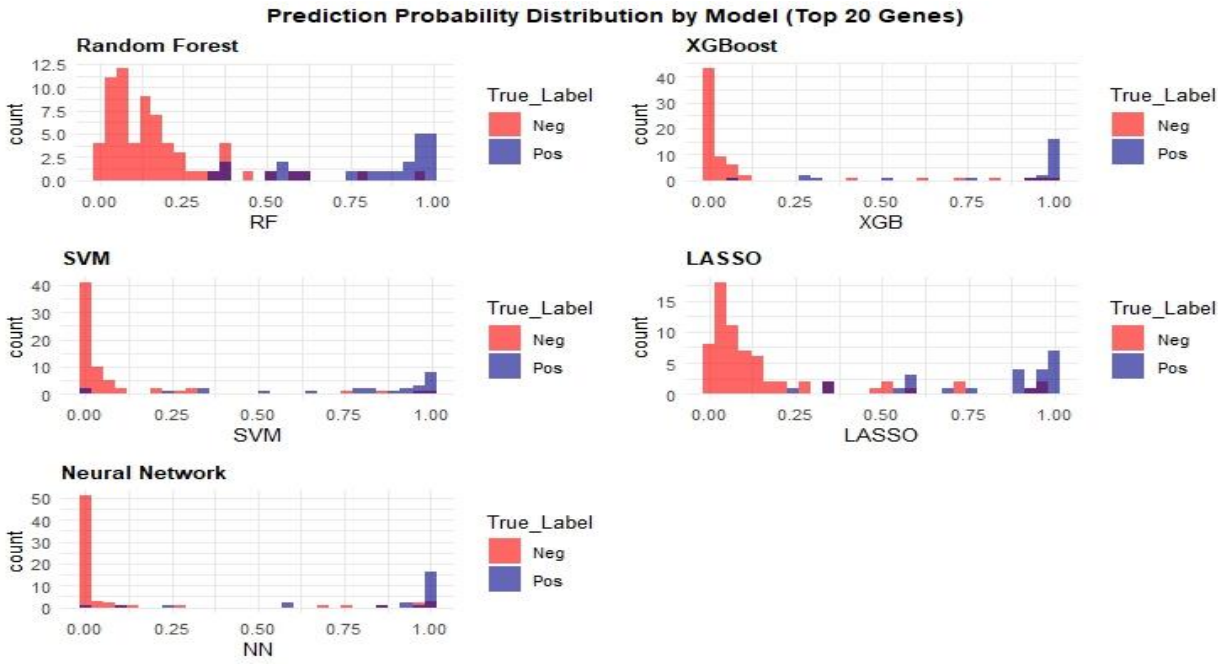


Figure 7. PR Distribution for ML Models Trained on Top 20 Genes. Overlaid histograms show the predicted probability outputs for each ML model, stratified by true FLT3-ITD class labels. **Red bars** represent FLT3-ITD-negative, and **blue bars** indicate positives. Sharp, bimodal distributions suggest stronger model confidence and class separation.

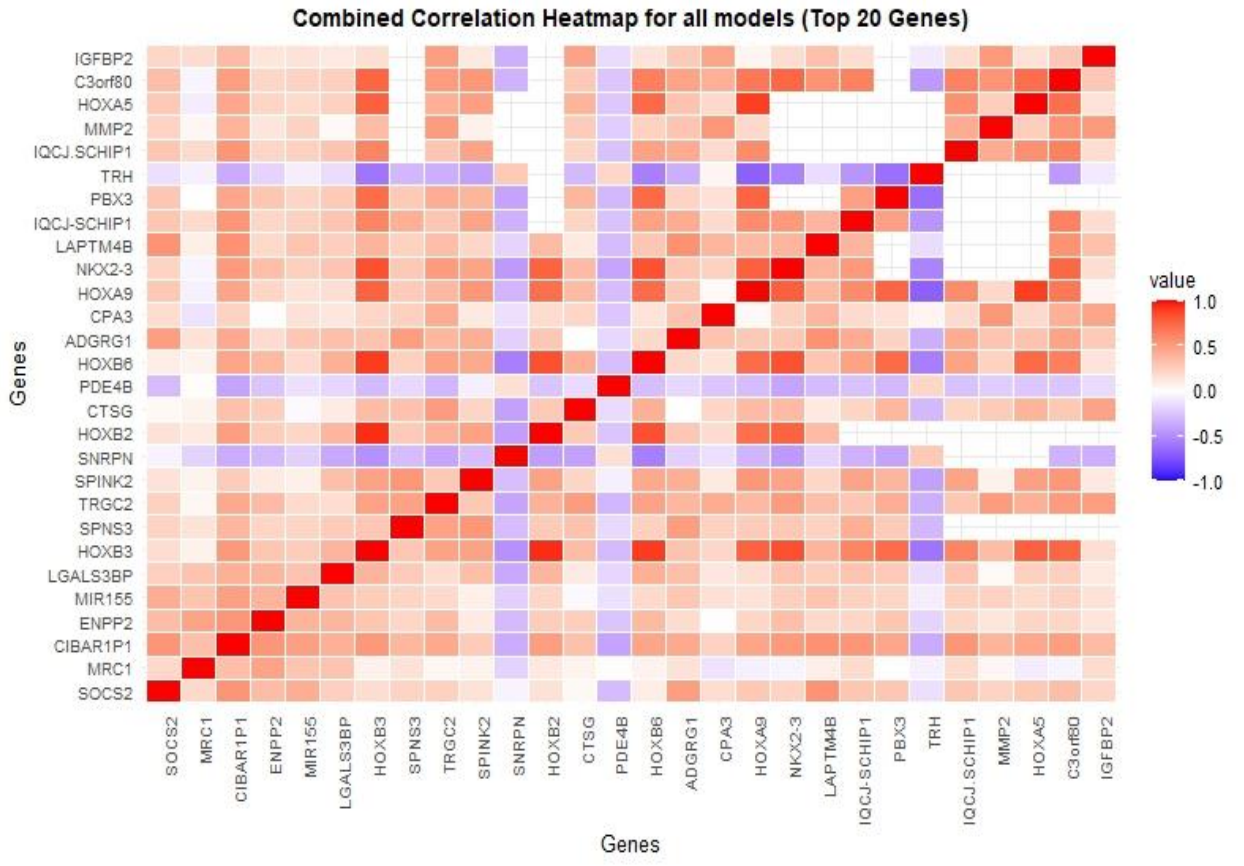


Figure 8. Combined Correlation Heatmap of Top 20 Genes Selected Across All ML Models. The heatmap visualises the averaged Pearson correlation coefficients between gene expression levels across all top 20 genes selected by all model classifiers. The gradient scale from **-1.0 (blue)** to **+1.0 (red)** represents perfect negative to perfect positive correlation, respectively. Diagonal values represent self-correlations **(value = 1.0)**, while off-diagonal cells indicate the degree of association between different gene pairs. Warmer colours signify stronger positive relationships; cooler colours denote weaker or negative correlations.

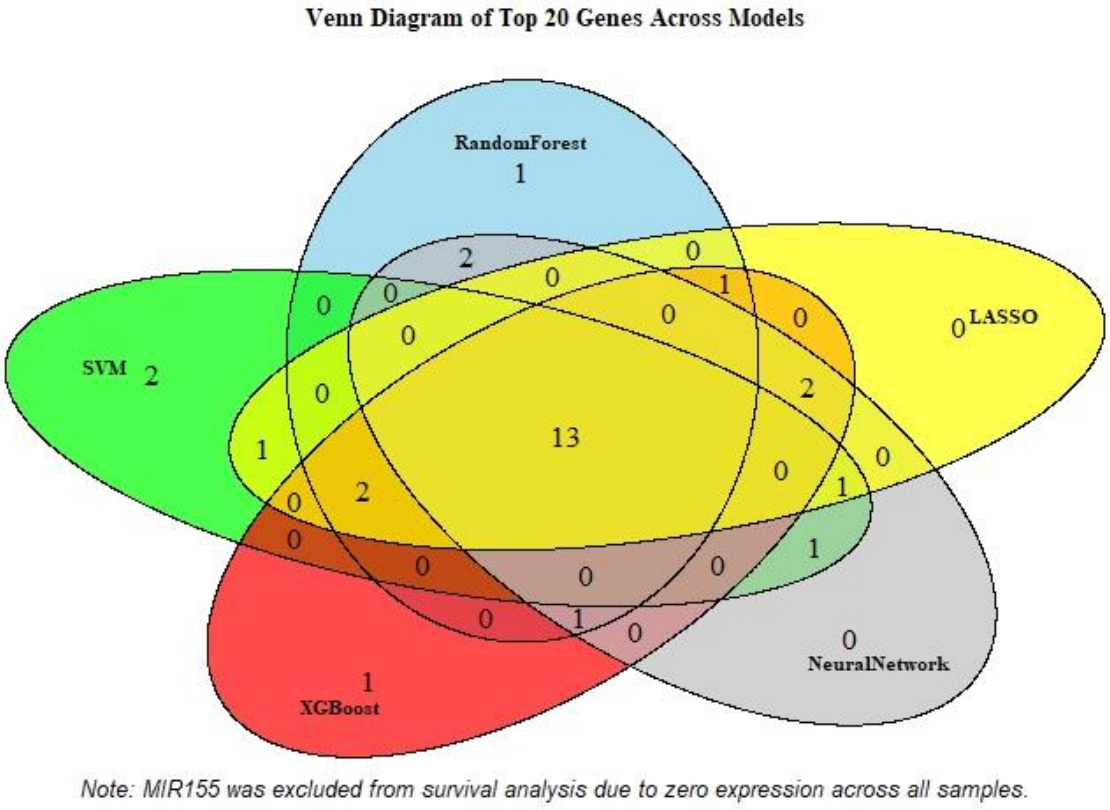
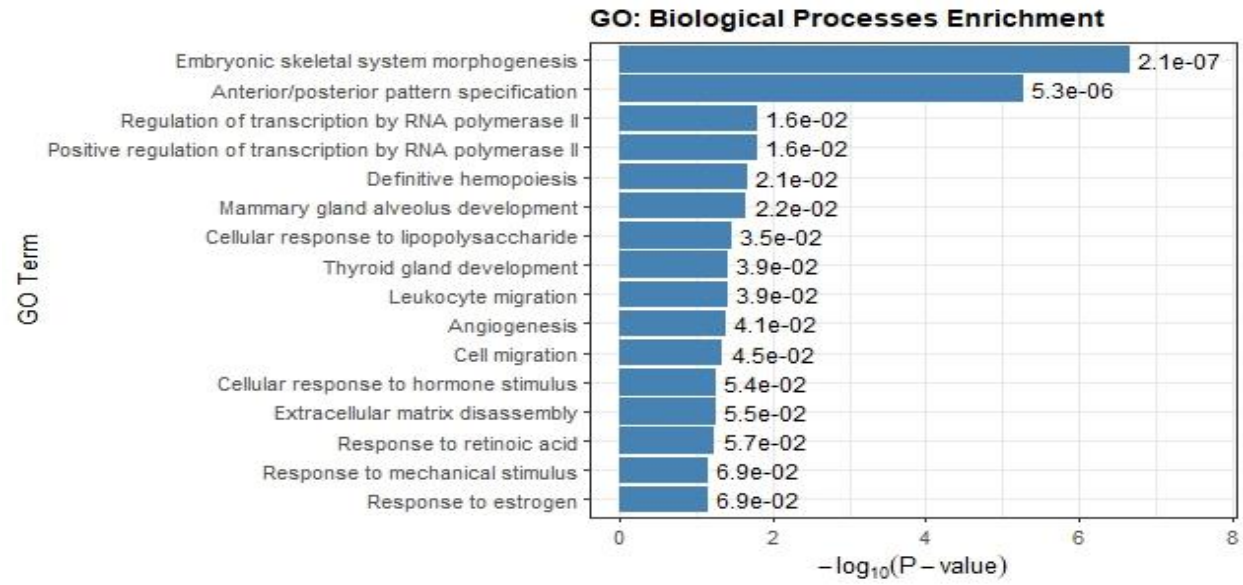
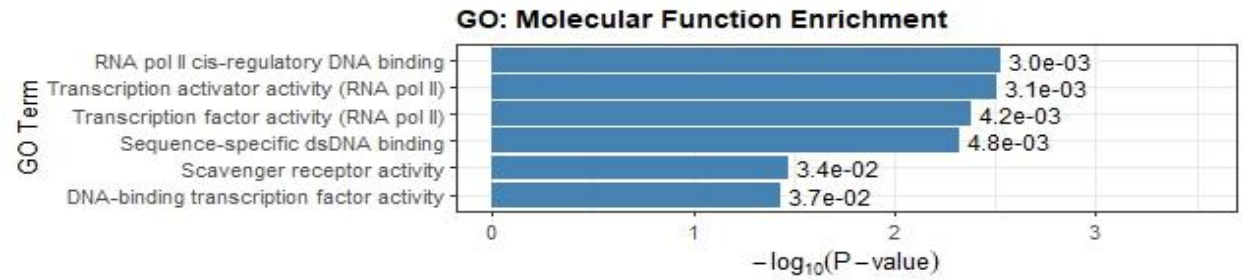


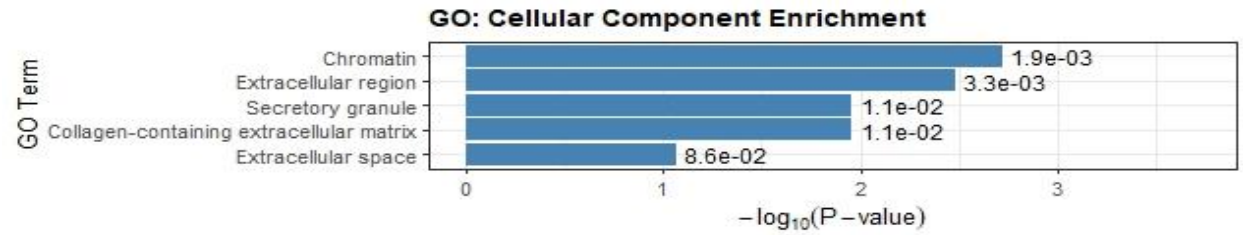
Figure 9. Venn Diagram of Top 20 Genes Across ML Classifiers. The diagram illustrates the overlap among the top 20 ranked genes identified by five machine learning classifiers: RF (blue), XGB (red), SVM (green), LASSO (yellow) and NN (grey). **13 genes** were consistently selected across all five models and were used in survival analysis and risk stratification. In addition to this core intersection, each classifier (RF, XGB, SVM, LASSO and NN) contributed **7 genes,** with varying degrees of overlap shared by two to four models. RF, XGB, and SVM contributed **1, 1** and **2 unique genes**, respectively, that were not selected by any other.



A



B



C

Figure 10. The Gene Ontology for the Top 20 Consistently Selected Genes Across All Models.

A. Biological processes horizontal bar chart displaying **$-\log_{10}$(p-value)**.

B. Molecular Function horizontal bar chart displaying **$-\log_{10}$(p-value)**.

C. Cellular Component horizontal bar chart displaying **$-\log_{10}$(p-value)**.

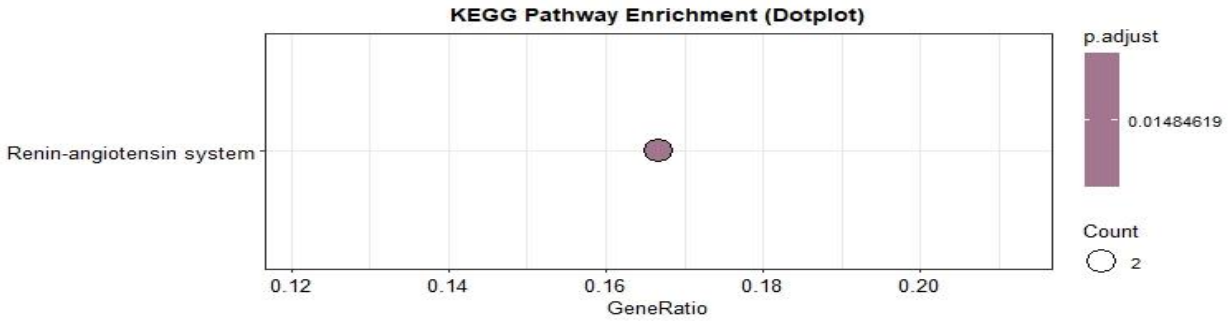


Figure 11. Dot plot of KEGG enrichment analysis showing the Renin–Angiotensin System pathway as significantly enriched (adjusted p = 0.0149).

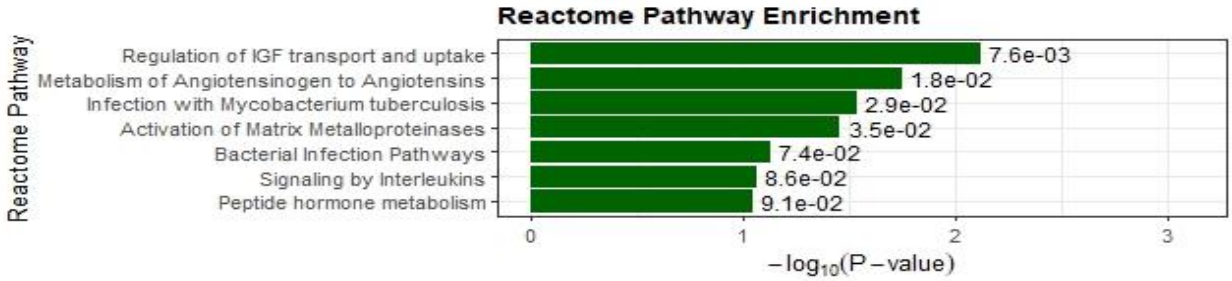


Figure 12. Horizontal bar chart displaying **$-\log_{10}$(p-value)** of the Reactome Pathway Enrichment for the Top 20 Consistently Selected Genes Across All Models.

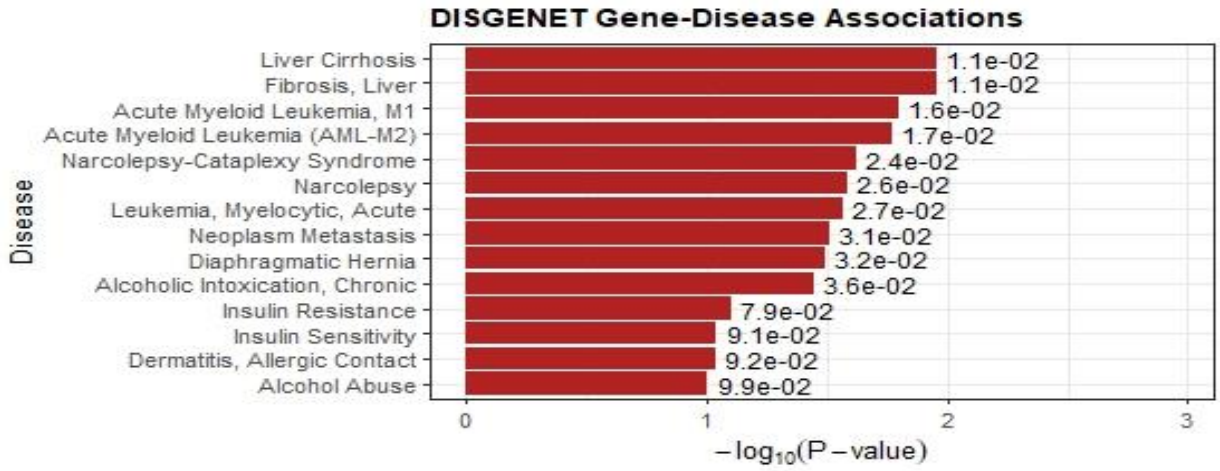


Figure 13. Horizontal bar chart displaying **$-\log_{10}$(p-value)** of the DISGENET Gene–Disease Enrichment Analysis for Top 20 Consistently Selected Genes.

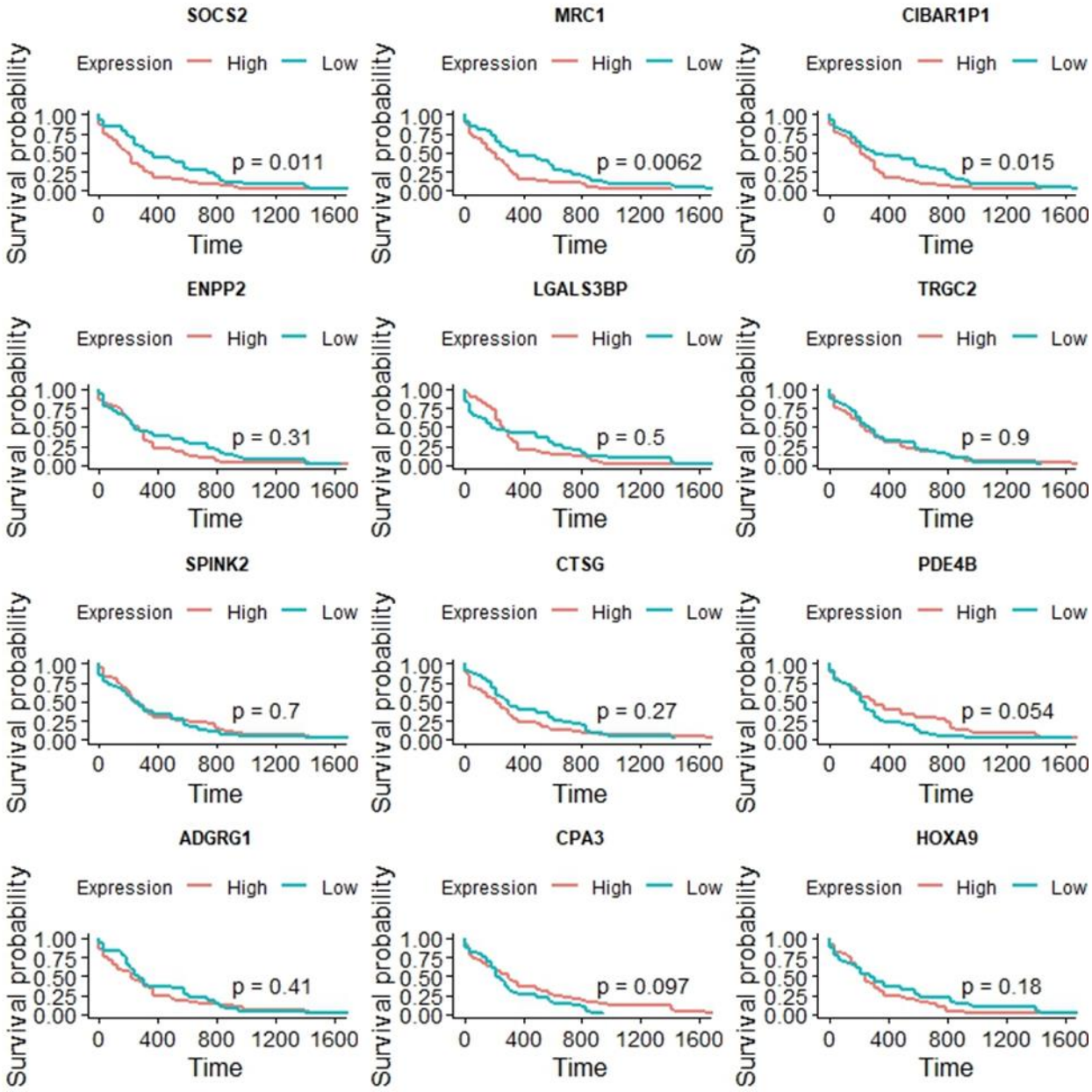


Figure 14. Kaplan–Meier Survival Analysis of Consistently Selected Genes. Survival curves for 12 genes were consistently selected across all ML classifiers. Significant differences were observed for MRC1, SOCS2 and CIBAR1P1. PDE4B and CPA3 showed borderline significance. MIR155 was excluded due to zero expression.

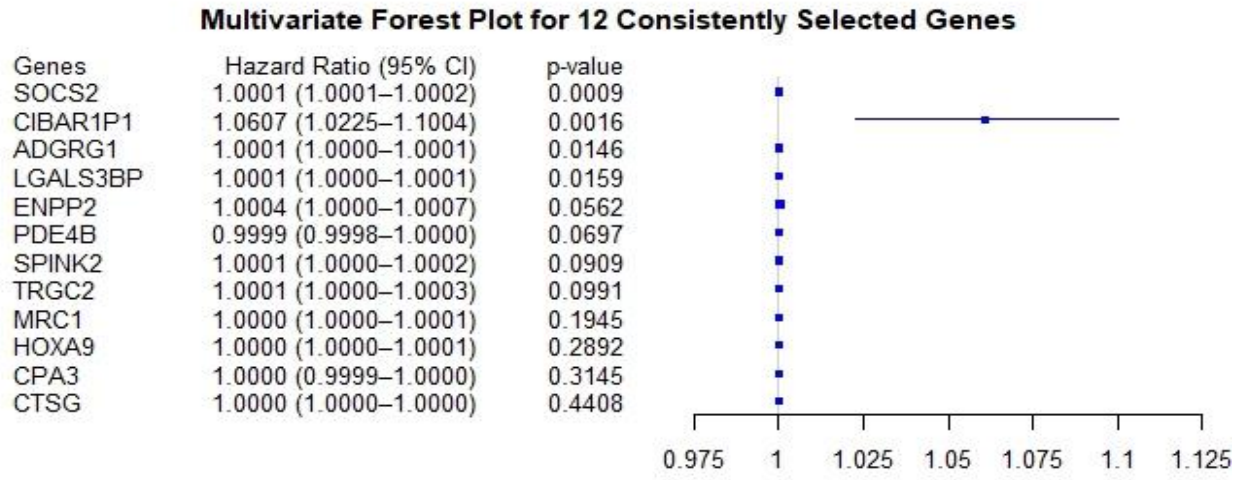


Figure 15. Multivariate Cox Regression Forest Plot for 12 Consistently Selected Genes. Forest plot showing HRs, 95% CIs and p-values for 12 genes after adjustment for age and gender. SOCS2, CIBAR1P1, ADGRG1 and LGALS3BP remained statistically significant. CIBAR1P1 had the highest HR (1.0607), with a CI not overlapping 1. Genes with borderline significance included ENPP2, PDE4B, SPINK2, and TRGC2. All remaining genes had HRs near 1 and non-significant p-values.

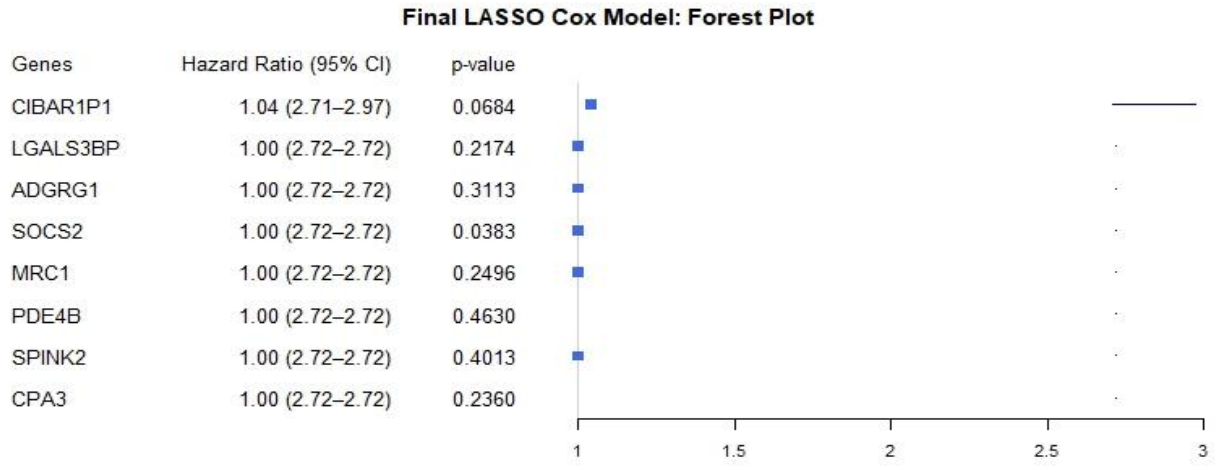


Figure 16. Final LASSO Cox Model (Forest Plot). Forest plot showing HRs, 95% CIs and p-values for the eight genes retained in the final LASSO Cox model. SOCS2 was statistically significant (p = 0.0383), while CIBAR1P1 showed borderline significance (p = 0.0684). The other six genes exhibited neutral effects with HR = 1.00 and wide, non-informative CIs.

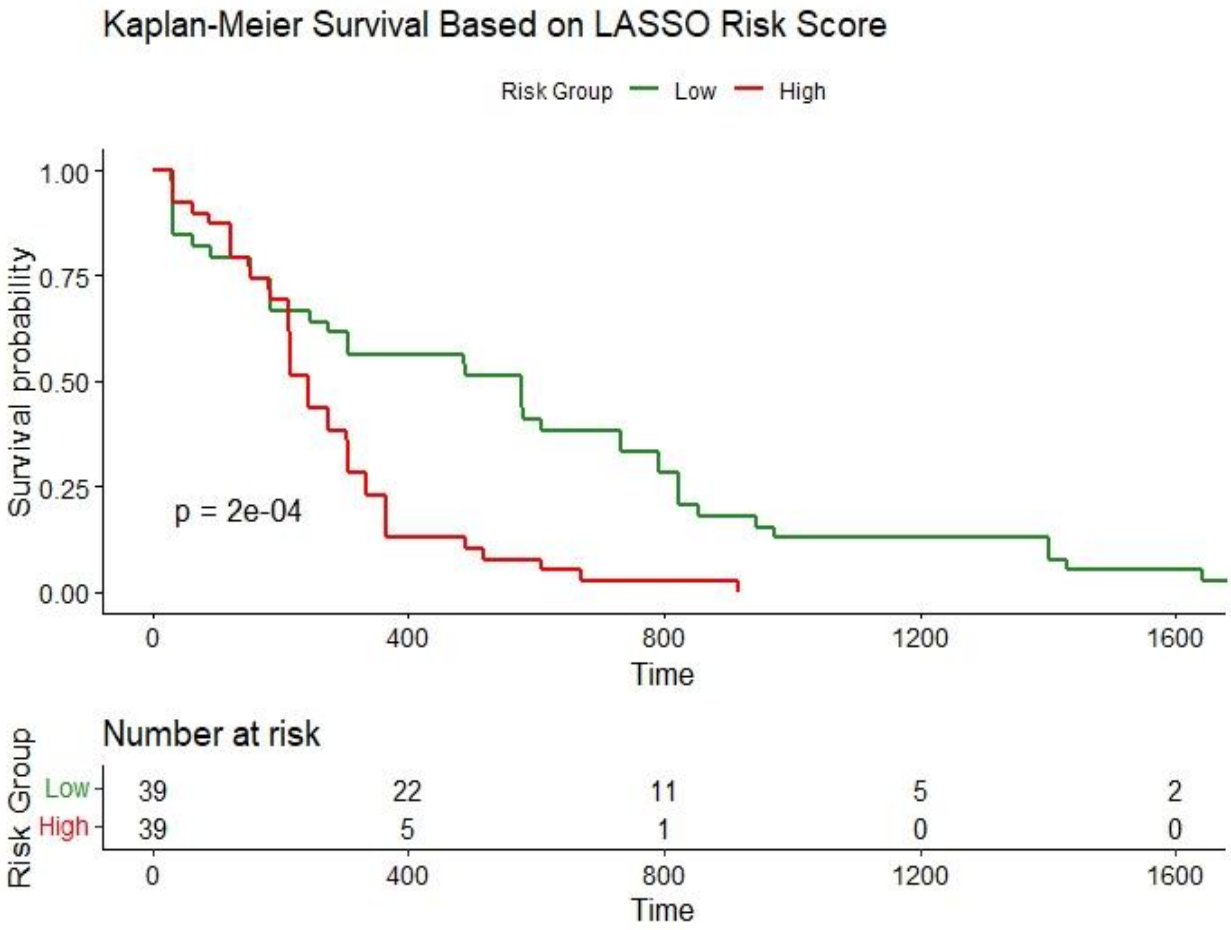


Figure 17. Kaplan–Meier Survival Analysis Based on LASSO Risk Score. Kaplan–Meier plot showing survival curves for high-risk **(red)** and low-risk **(green)** groups stratified by LASSO-derived risk scores. The survival curves diverge significantly **(p = 2e–04)**, indicating effective risk separation. The number of patients at risk over time is displayed below the plot. The low-risk group showed prolonged survival compared to the high-risk group, supporting the prognostic strength of the LASSO-selected multigene panel.