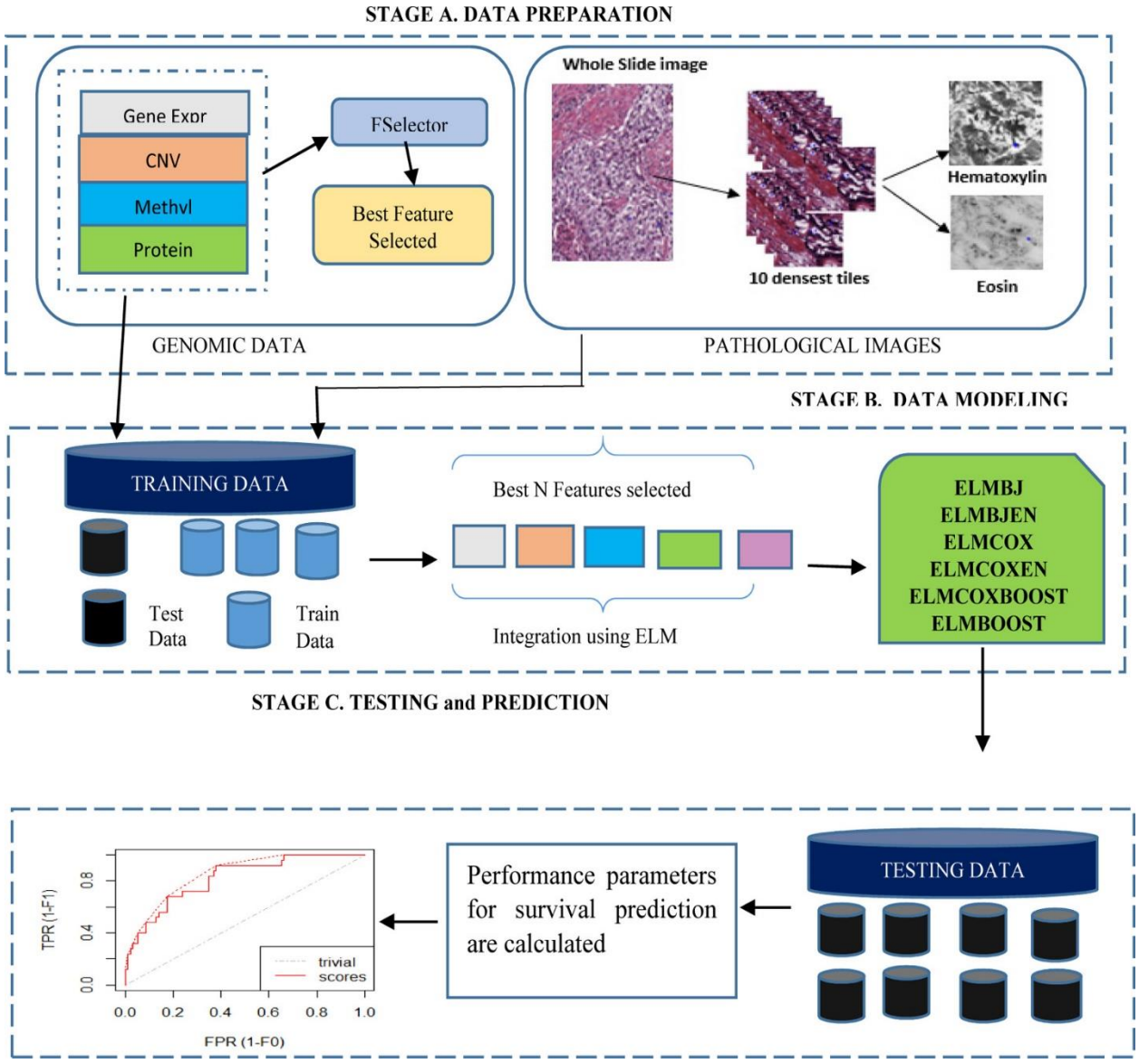
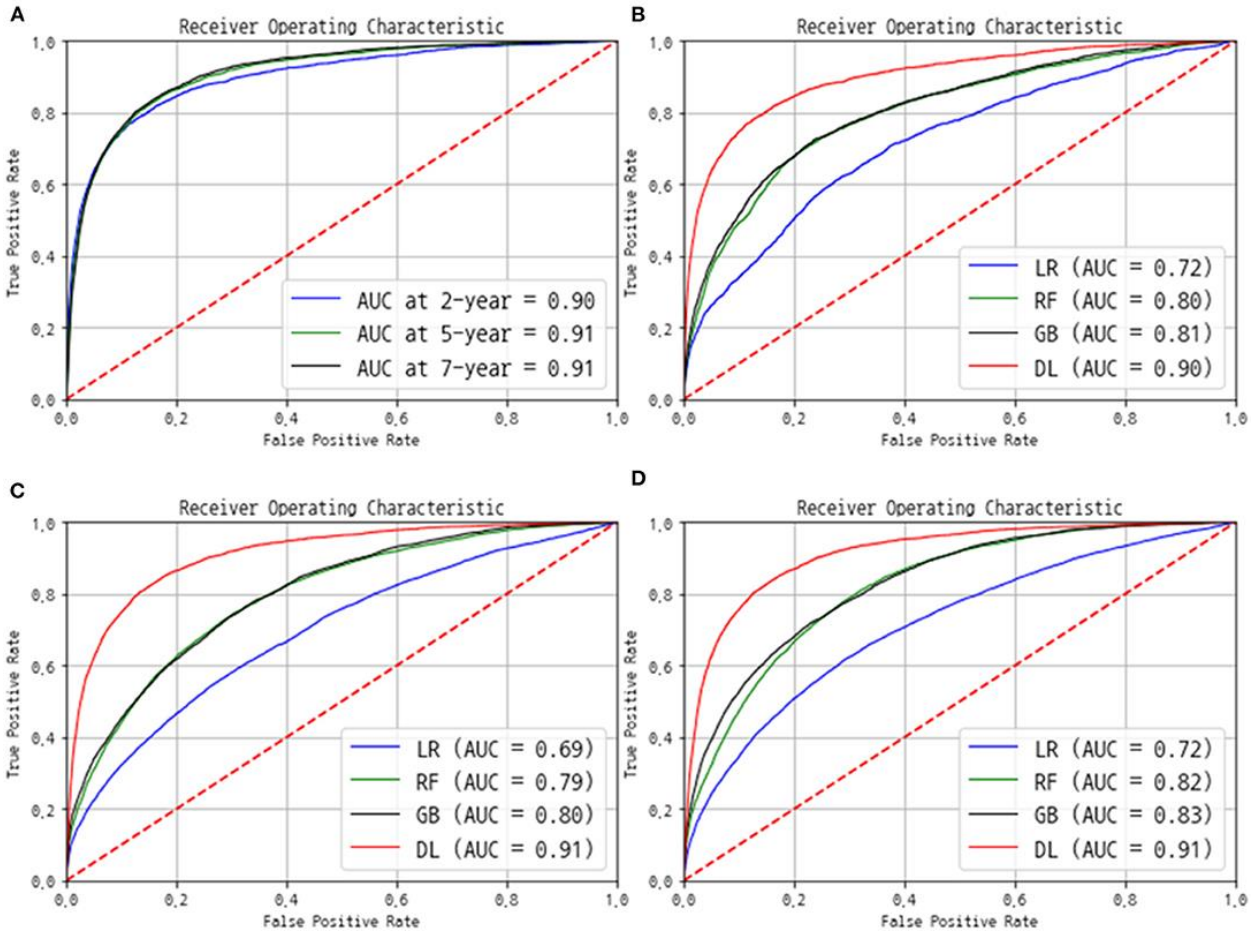
Model Performance Metrics

Team ID	NM2023TMID18418
Project name	CancerVision: Advanced Breast Cancer Prediction with Deep Learning





To learn the various measurement period data, we needed to find a fixed value for the window size (*k*) during the training stage.

Longer periods are better, but a look-back period of 12 months was the optimal window size for our training data set, according to the results of grid-search algorithms.

The AUC and C-index were used to evaluate the performance. The C-index eventually reached 0.92 for the training data set and 0.89 for the validation and test data sets.

The AUC value was 0.90 at the 2-year point, 0.91 at the 5-year point, and 0.91 at the 7-year point . We also compared our model with the performances of three other machine learning prediction models .

The logistic regression model produced AUC in the range of 0.69–0.72, and the random forest and gradient boosting methods, which are the most popular ensemble models, showed similar AUC values in the range of 0.80–0.83.

The deep learning-based final model only exceeded an AUC of 0.90, outperforming the existing machine learning-based models.

We also evaluated the model from a clinical point of view with pathologic T stage, pathologic N stage, subtypes according to ER/PR and HER2 status, EGFR status, and CK5/6 status by comparing the predicted recurrence proportion with the actual recurrence proportion .

The MAE and wMAE of each group showed great results of as little as 3.5%. The model errors for pathologic T stage and N stage features were similar to each other but differed from those for the other pathologic features.

The subtypes had similar error values of around 2.5%.

Performance of the final model in terms of model error (predicted–observed) over each of the 2/5/7 years. Patients were grouped by breast cancer features.

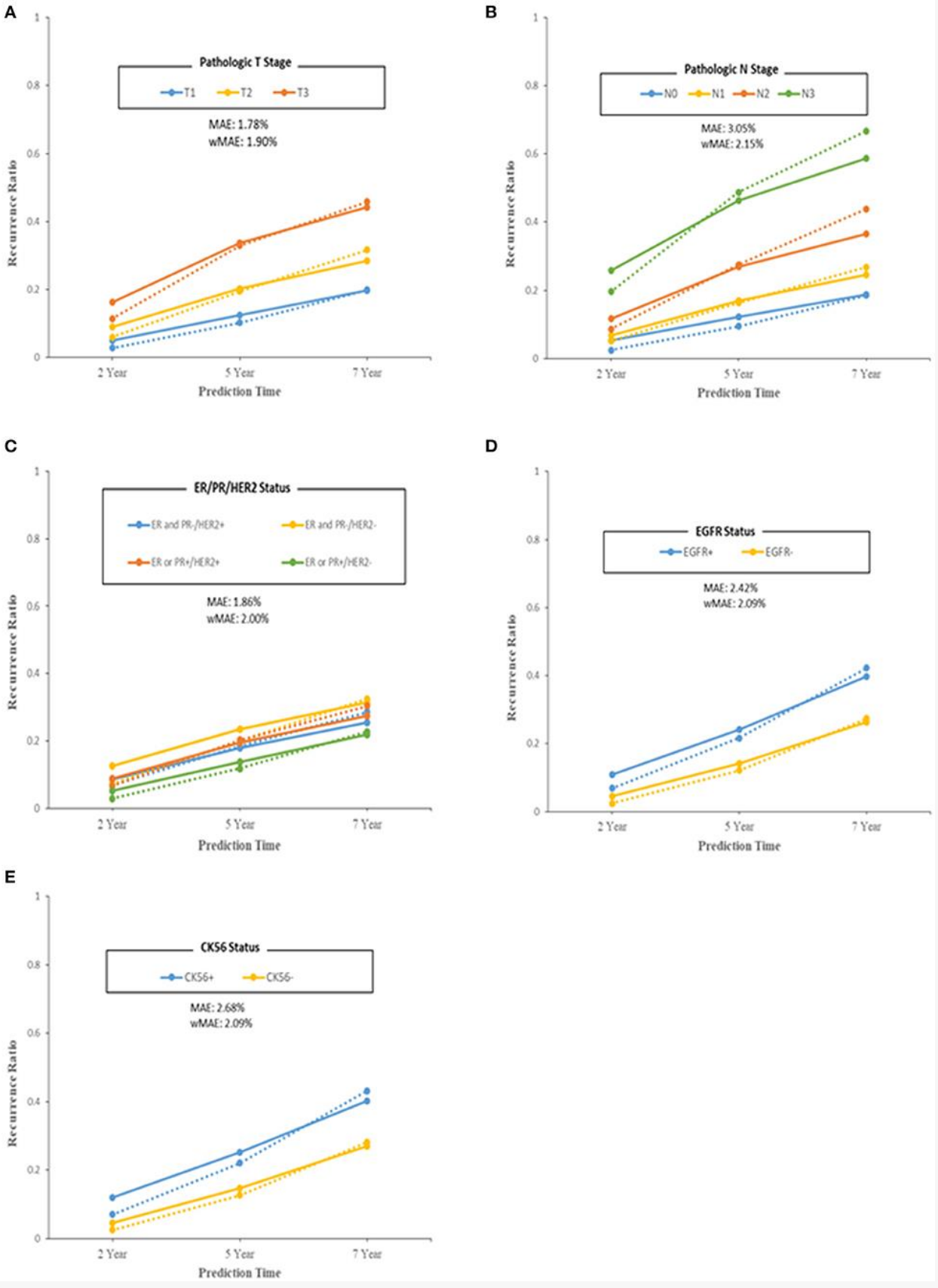
(A) Pathologic T stage.

(B) Pathologic N stage.

(C) ER/PR/HER2 status.

(D) EGFR status.

(E) CK56 status.


A
Pathologic T Stage
T1
T2
T3
MAE: 1.78%
wMAE: 1.90%
Recurrence Ratio
Prediction Time
2 Year
5 Year
7 Year
B
Pathologic N Stage
N0
N1
N2
N3
MAE: 3.05%
wMAE: 2.15%
C
ER/PR/HER2 Status
ER and PR-/HER2+
ER and PR-/HER2-
ER or PR+/HER2+
ER or PR+/HER2-
MAE: 1.86%
wMAE: 2.00%
D
EGFR Status
EGFR+
EGFR-
MAE: 2.42%
wMAE: 2.09%
E
CK5/6 Status
CK5/6+
CK5/6-
MAE: 2.68%
wMAE: 2.09%

For model specificity and accuracy, Median Absolute Error (MAE) and weighted Mean Absolute Error (wMAE) were calculated at each group or bin. Solid lines indicate observed recurrence proportion, and dashed lines indicate predicted recurrence proportion