Prediction of Metastatic Risk in Breast Cancer by the Expression of Mechanobiological Markers

Ksenia Maksimova[1], Margarita Pustovalova[1], Sergey Leonov[1], Yulia Merkher[2]

1 Institute of Future Biophysics, Moscow Center for Advanced Studies, 123592 Moscow, Russia
2 Department of Biomedical Engineering, Technion - Israel Institute of Technology, Haifa, 3200003, Israel

INTRODUCTION & AIM

Breast cancer (BC) is among the most common malignancies worldwide, with metastasis being the main cause of cancer-related death. Accurate prediction of distant metastasis risk is crucial for improving patient outcomes. While gene expression assays have improved prognostic stratification in select cancers, they often overlook dynamic cellular behaviors such as cytoskeletal remodeling and mechanotransduction. Based on our previous research[1,2], we selected a panel of actin cytoskeleton-related genes as potential biomarkers. This study aimed to develop machine learning models predicting distant metastases in primary BC using their expression profiles.

METHODS

- Genes: *TAGLN, ANXA6, ANXA2, ECM1, PFN1, MYH9, CFL1, EZR, ACTN4, GSN,* and *FSCN1*
- Target: Metastasis status by AJCC staging
- Dataset: TCGA Breast Invasive Carcinoma data (Firehose Legacy)

- ML Models:
  - k-NN
  - Logistic Regression
  - Decision Trees
  - Random Forest
  - XGBoost

Load expression data via cBioPortal TCGA Breast Invasive Carcinoma data

↓

Normalization & SMOTE Log2(TPM+1)

↓

Train/test splitting (80/20%) Per-gene standartization

↓

Training ML models with best hyperparameters

↓

Feature (Gene) importance analysis

RESULTS & DISCUSSION

- All models performed well - Accuracy ≥ 0.89
  - Exception: Logistic Regression with Accuracy = 0.65
- Best models:
  - Random forest - Accuracy = 0.983, Kappa = 0.967
  - XGBoost - Accuracy = 0.961, Kappa = 0.922

- Top predictors: *CFL1, ANXA2,* and *MYH9*
- *TAGLN, FSCN1,* and *ECM1:*
  - High importance in XGBoost
  - Were also important in our prior cell line studies

Model Comparison by Metrics

Model Comparison by Gene Importance
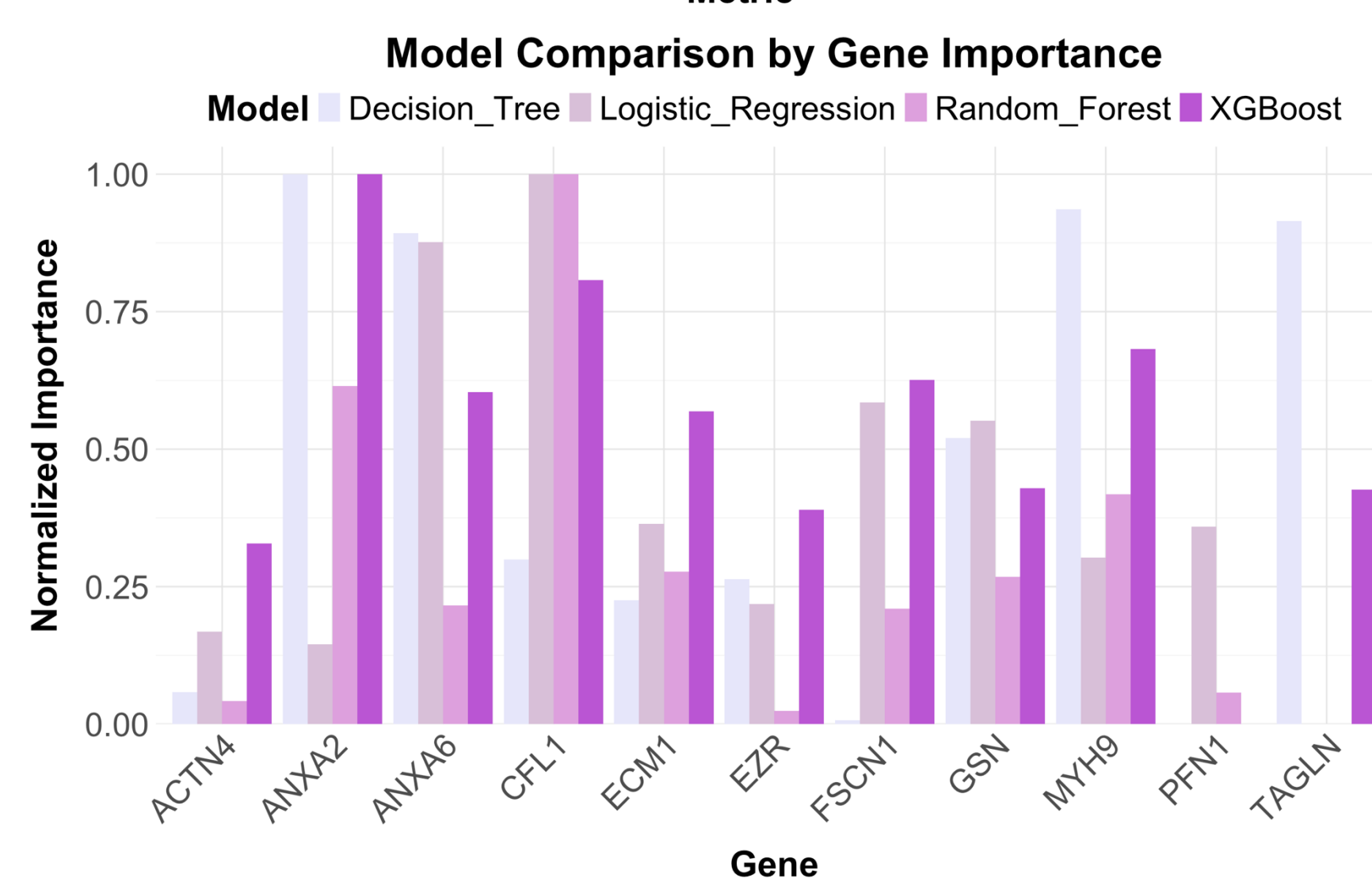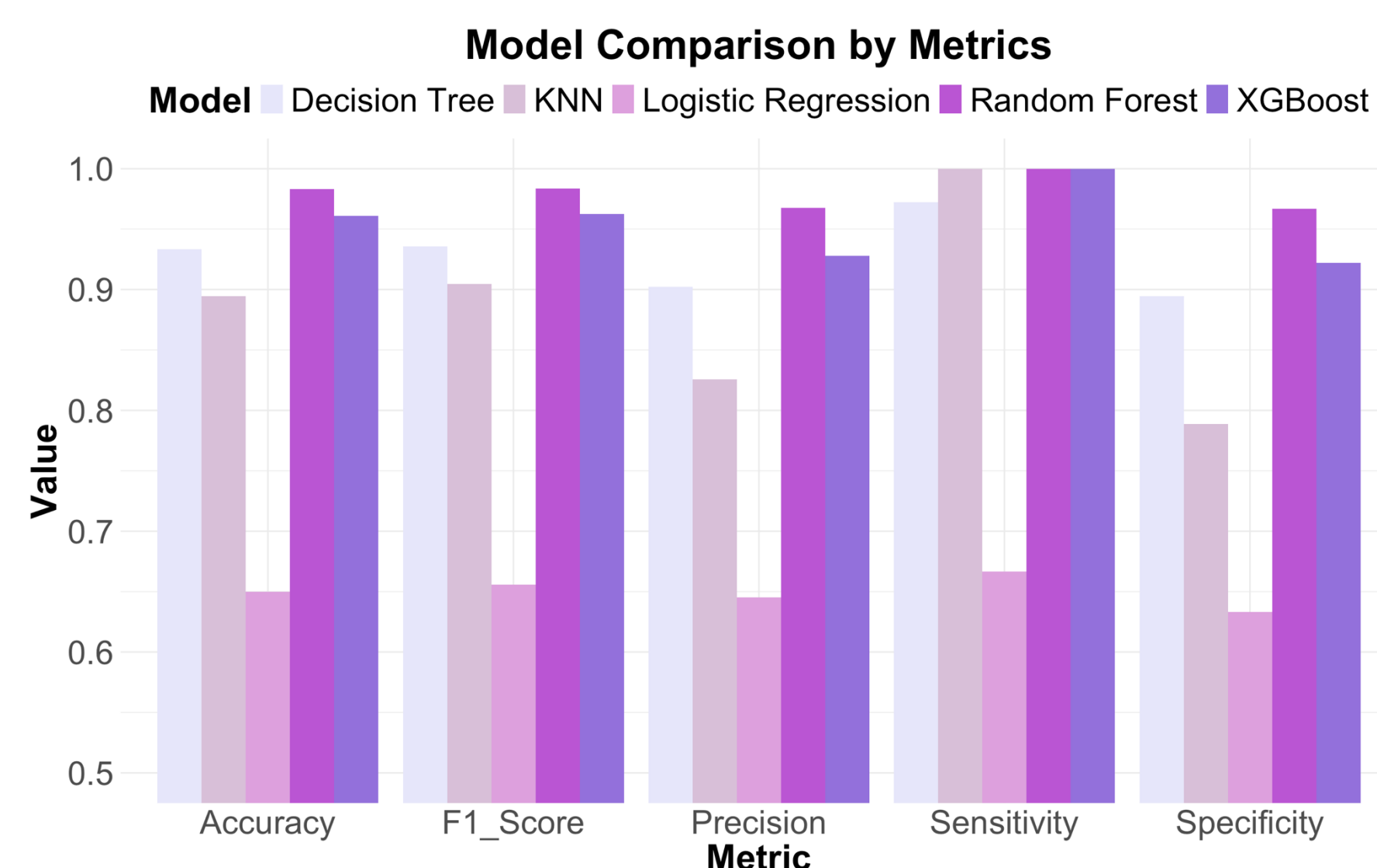
CONCLUSION

Expression of cytoskeleton-related genes demonstrates strong potential for predicting distant metastasis in BC and warrants further validation in other cancer types.

ACKNOWLEDGEMENT

REFERENCES AND CONTACT INFORMATION

1. Merkher, Y. et al, *Sci Rep* 13, 12289 (2023).
2. Leonov, S. et al, *Int. J. Mol. Sci.* 24, 4773, (2023).

Ksenia Maksimova: maksimova.kv@phystech.edu
Yulia Merkher: merkher@technion.ac.il