



پیش‌بینی متاستاز سرطان پستان با استفاده از شبکه‌های عصبی گراف: یک

مطالعه جامع

امیرحسین رحمانی^{*۱}

۱. دانشجوی کارشناسی، گروه مهندسی پزشکی، دانشکده فنی و مهندسی، موسسه آموزش عالی روزبهان، ساری، ایران.
ایمیل نویسنده مسئول: Amir.h.rahmani9007@gmail.com

چکیده

پیش‌بینی زود هنگام متاستاز سرطان پستان برای مداخله به موقع و بهبود نتایج بیمار بسیار مهم است. در این مطالعه، ما مدل‌های شبکه عصبی گراف (GNN) را برای پیش‌بینی متاستاز در سرطان پستان با ادغام داده‌های مولکولی و بالینی با ابعاد بالا بررسی می‌کنیم. ما از دو گروه بزرگ سرطان پستان METABRIC - (حدود ۲۰۰۰ تومور) و TCGA-BRCA - (۱۰۸۲ تومور) - که پروفایل‌های چند آمیکس (بیان ژن، تعداد کپی و غیره) و وضعیت متاستاز را ارائه می‌دهند، استفاده می‌کنیم. پس از پیش‌پردازش (نرمال‌سازی، انتخاب ویژگی و ساخت نمودارهای شباهت بیمار)، چندین نوع GNN (شبکه کانولوشن گراف، شبکه توجه گراف و GraphSAGE) را برای طبقه‌بندی بیماران به‌عنوان متاستاتیک در مقابل غیرمتاستاتیک آموزش می‌دهیم. مدل‌ها با استفاده از اعتبارسنجی متقابل ۵ گانه با معیارهایی از جمله دقت، AUC، حساسیت و ویژگی ارزیابی می‌شوند. بهترین مدل GNN ما به AUC حدود ۰/۹۵ و دقت حدود ۰/۹۳ دست می‌یابد که از روش‌های پایه (جنگل تصادفی، SVM، رگرسیون لجستیک) که AUC حدود ۰/۸۵-۰/۹۰ را ارائه می‌دهند، بهتر عمل می‌کند. تجسم نتایج از طریق منحنی‌های ROC، قدرت تمایز برتر GNN‌ها را تأیید می‌کند. ما در مورد چگونگی ثبت روابط پیچیده بین بیمار و تعاملات بین ژن‌ها توسط رویکرد مبتنی بر نمودار، که منجر به پیش‌بینی بهبود یافته می‌شود، بحث می‌کنیم. ما همچنین، قابلیت تفسیر و محدودیت‌های مدل (به‌عنوان مثال، نیاز به ساخت نمودار، تنظیم هاپر پارامتر) را تجزیه و تحلیل می‌کنیم. این مطالعه جامع نشان می‌دهد که GNN‌ها مزایای قابل توجهی برای پیش‌بینی خطر متاستاز در سرطان پستان ارائه می‌دهند و زمینه را برای مدل‌های پیش‌آگهی مبتنی بر نمودار چندوجهی آینده فراهم می‌کنند.

کلمات کلیدی: سرطان سینه؛ پیش‌بینی متاستاز؛ شبکه‌های عصبی گراف (GNN)؛ یادگیری ماشین؛ مولتی اومیکس؛ تحلیل بقاء



Prediction of Breast Cancer Metastasis Using Graph Neural Networks: A Comprehensive Study

Amir Hossein Rahmani^{1,*}

1. Bachelor Student, Department of Biomedical Engineering, Technical and Engineering Faculty, Rouzbahan Institute of Higher of Education, Sari , Iran

* Corresponding author e-mail: Amir.h.rahmani9007@gmail.com

Abstract

Early prediction of breast cancer metastasis is crucial for timely intervention and improved patient outcomes. In this study, we investigate Graph Neural Network (GNN) models for predicting metastasis in breast cancer by integrating high-dimensional molecular and clinical data. We leverage two large-scale breast cancer cohorts – METABRIC (~2,000 tumors) and TCGA-BRCA (1,082 tumors) – which provide multi-omics profiles (gene expression, copy number, etc.) and metastasis status. After preprocessing (normalization, feature selection, and construction of patient similarity graphs), we train several GNN variants (Graph Convolutional Network, Graph Attention Network, and GraphSAGE) to classify patients as metastatic vs. non-metastatic. Models are evaluated using 5-fold cross-validation with metrics including accuracy, AUC, sensitivity, and specificity. Our best GNN model achieves an AUC of ~0.95 and accuracy ~93%, outperforming baseline methods (random forest, SVM, logistic regression) which yield AUCs around 0.85–0.90. Visualization of results via ROC curves confirms superior discriminative power for GNNs. We discuss how the graph-based approach captures complex inter-patient relationships and interactions between genes, leading to improved prediction. We also analyze model interpretability and limitations (e.g., need for graph construction, hyperparameter tuning). This comprehensive study demonstrates that GNNs offer significant advantages for metastasis risk prediction in breast cancer, laying the groundwork for future multimodal graph-based prognostic models.