Using MicroRNA and Deep Learning to Noninvasively Diagnose Gynecologic Conditions

Palak Yadav
Advisors: Dr. Jill Moore & Dr. Kevin Crowthers

Research Question

- How can miRNA classification be used as a noninvasive diagnostic candidate?
- What implication do miRNAs have about the pathology and comorbidities of gynecological conditions?

Objectives

- Identify differentially expressed and common miRNA for each target disease.
- Design a robust machine learning classification model.
- Model miRNA-mediated pathways to advance understanding of disease pathology and etiology. .

Background

80% of **ovarian cancer** cases are found at a late stage

Undiagnosed gynecologic conditions can have a great burden on families and individuals

Black women have a higher death rate from **breast cancer**

Endometriosis impacts 1 out of 10 women yet take 5-7 years to diagnose

Need for non-invasive and accessible diagnostic tools.

microRNAs:

- noncoding RNA
- control gene expression
- can be collected from blood, urine, and sailva

Main Takeways

Deep learning can be used for disease detection using microRNA expression.
Breast Cancer, Ovarian Cancer, and endometriosis have a shared pathology and onset of one disease can increase the risk of other.

Machine Leanrning Model Performance

Binary Classfication	Ovarian Cancer	Breast Cancer	Endometriosis
Logistic Regression	96%	93%	97%
Random Forest	95%	94%	98%
DNN	95%	94%	96.5%

Multi-Disease Classfication	Overall	Ovarian Cancer	Breast Cancer	Endometriosis
Random Forest	93%	97%	92%	97%
DNN	85.3%	81%	83%	89%

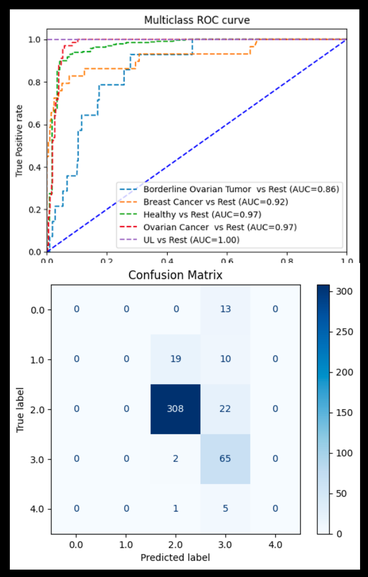


Fig. 1: Random Forest ROC Curve showing model performance for each set

Fig. 2: DNN Confusion Matrix showing the predicted and actual value for each disease

miRNA Panel

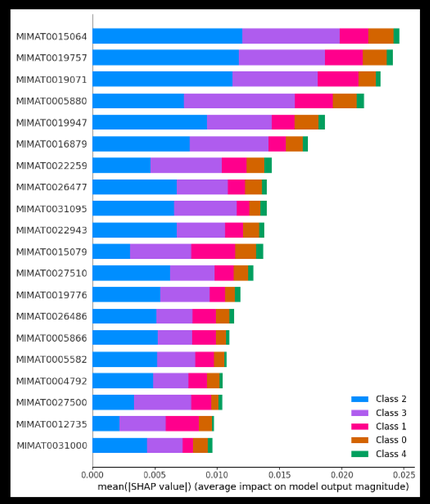


Fig. 3: Significant miRNA and their contribution in prediction in DNN

Significant miRNAs

- miR-let-7d
- miR-1307-3p
- miR-320a
- miR-6855-3p

Pathway Modeling

Targeted Pathways

- Pathways in Cancer
- Rap 1 Signaling
- Actin Cytoskeleton Regulation
- Ras Signaling

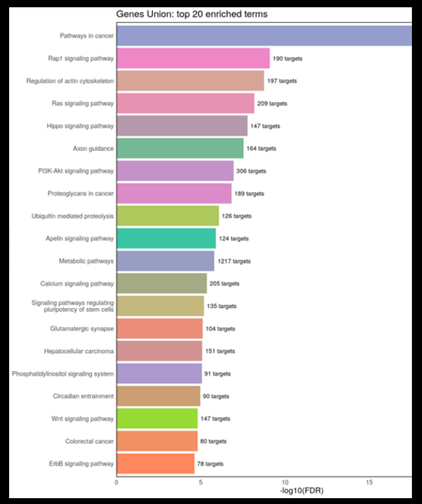
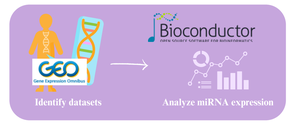


Fig. 4: Biological pathways targeted by significant miRNAs

Methodology



Develop a unique miRNA panel for each gynecological condition from public databases: Gene Expression Omnibus and the Cancer Atlas TCGA. Analyze miRNA expression using statistical analysis tools from the Bioconductor package to determine differential expression.

The results from the last stage were used to train machine-learning models to classify diseases based on miRNA expression profiles. The following models were designed: Binary Classification: Logistic Regression, DNN Multi-Disease: Random Forest, DNN

Use feature extraction techniques to identify significant miRNA and model biological pathways.

Analysis

- Greater than 90% accuracy for binary classification and 85% for deep learning model suggests that **miRNA expression can be used for disease prognosis**
- Feature extraction techniques applied to each model i**dentified unique and shared miRNA**
- Pathway modeling showed that pathways of cancers are targeted by miRNA found in benign and malignant female conditions and identified potential diagnostic and therapeutic targets:
  - **miR-Let-7d:** highly expressed across all three --> cell proliferation and carcinogenesis (Zhang, et. al, 2017).
  - **miR-320a**: migration and invasiveness in breast cancer
  - **miR-1307-3p**: linked to ovarian tumor chemoresistance and future malignancy of endometriosis

Future Work

- miRNA expression and deep learning can be harnessed for early disease detection, reduce waiting periods, and guide future therapeutic development beyond just gynecologic diseases.
- Collect larger miRNA datasets in a standard methodology from diverse populations
- Test miRNA targets identified in clinical settings to evaluate reliance as a diagnostic candidate
- Develop serum, saliva, and urine miRNA testing kits for under-resourced communities
- Test aberrant miRNA-mediated pathways as therapeutic targets