

Gaining Biological Intuition from Feature Importances

02-620 Machine Learning for Scientists Project

Siddharth Reed

Carnegie Mellon University

May 7th 2021

Introduction

Black Box Solvers

- ▶ Models that may provide accurate classification but

Black Box Solvers

- ▶ Models that may provide accurate classification but
- ▶ Difficult to “debug” classifications

Black Box Solvers

- ▶ Models that may provide accurate classification but
- ▶ Difficult to “debug” classifications
- ▶ can re-enforce existing biases/assumptions of researchers

Black Box Solvers

- ▶ Models that may provide accurate classification but
- ▶ Difficult to “debug” classifications
- ▶ can re-enforce existing biases/assumptions of researchers
- ▶ Examples

Black Box Solvers

- ▶ Models that may provide accurate classification but
- ▶ Difficult to “debug” classifications
- ▶ can re-enforce existing biases/assumptions of researchers
- ▶ Examples
 - ▶ Deep Neural Networks

Black Box Solvers

- ▶ Models that may provide accurate classification but
- ▶ Difficult to “debug” classifications
- ▶ can re-enforce existing biases/assumptions of researchers
- ▶ Examples
 - ▶ Deep Neural Networks
 - ▶ Convolutional NNs (image task)

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information
- ▶ Examples

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information
- ▶ Examples
 - ▶ Logistic Regression

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information
- ▶ Examples
 - ▶ Logistic Regression
 - ▶ Random Forest

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information
- ▶ Examples
 - ▶ Logistic Regression
 - ▶ Random Forest
 - ▶ SVM

Informative Models in Bioinformatics

- ▶ Want a biological intuition behind some classification (e.g. regular vs cancer cells)
- ▶ Helps researchers and doctors understand genetic drivers of disease
- ▶ Allows prediction vs. patient specific information
- ▶ Examples
 - ▶ Logistic Regression
 - ▶ Random Forest
 - ▶ SVM
 - ▶ Naive Bayes

Methods

Transcriptomic Data

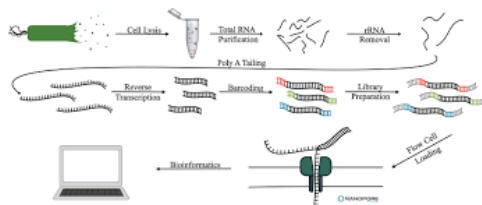


Figure 1: Transcriptomics Workflow

- ▶ use recount2 counts for TCGA (cancer,) and GTEx (normal, 373) lung samples

Transcriptomic Data

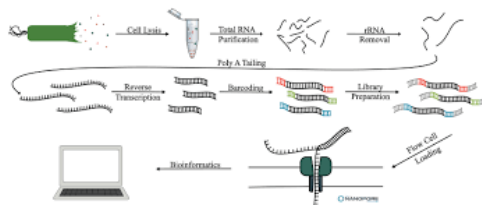


Figure 1: Transcriptomics Workflow

- ▶ use recount2 counts for TCGA (cancer,) and GTEx (normal, 373) lung samples
- ▶ **Features:** gene-level transcript counts for $\sim 20,000$ coding genes

Transcriptomic Data

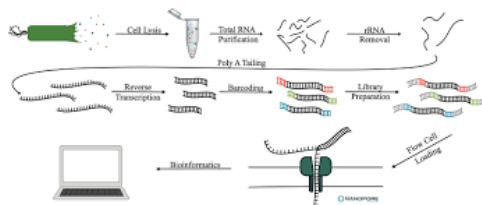


Figure 1: Transcriptomics Workflow

- ▶ use recount2 counts for TCGA (cancer,) and GTEx (normal, 373) lung samples
- ▶ **Features:** gene-level transcript counts for $\sim 20,000$ coding genes
- ▶ **Labels:** binary, cancer/not cancer

Classifier with Informative Models

- ▶ chose classifiers that

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values
 - ▶ relatively fast/simple

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values
 - ▶ relatively fast/simple
- ▶ Importance of feature (gene) g

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values
 - ▶ relatively fast/simple
- ▶ Importance of feature (gene) g
 - ▶ **Naive Bayes:** $P(g|y)$ for $y = 1$ (cancer)

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values
 - ▶ relatively fast/simple
- ▶ Importance of feature (gene) g
 - ▶ **Naive Bayes:** $P(g|y)$ for $y = 1$ (cancer)
 - ▶ **SVM:** contribution of g to separating $y = 0, y = 1$

Classifier with Informative Models

- ▶ chose classifiers that
 - ▶ provide feature importance values
 - ▶ relatively fast/simple
- ▶ Importance of feature (gene) g
 - ▶ **Naive Bayes:** $P(g|y)$ for $y = 1$ (cancer)
 - ▶ **SVM:** contribution of g to separating $y = 0, y = 1$
 - ▶ **Random Forest:** variability explained by g

Importance Analysis

- ▶ extract feature (gene) importances from each model in predicting cancer

Importance Analysis

- ▶ extract feature (gene) importances from each model in predicting cancer
- ▶ compare most important genes and GO terms found by each model

Importance Analysis

- ▶ extract feature (gene) importances from each model in predicting cancer
- ▶ compare most important genes and GO terms found by each model
- ▶ train classifier on chosen genes and compare to all genes

Results

Model Test accuracy

- ▶ prediction accuracy on the test set ($n=289$)

Model Test accuracy

- ▶ prediction accuracy on the test set ($n=289$)
- ▶ **SVM:** 0.986

Model Test accuracy

- ▶ prediction accuracy on the test set ($n=289$)
- ▶ **SVM:** 0.986
- ▶ **Naive Bayes:** 0.957

Model Test accuracy

- ▶ prediction accuracy on the test set ($n=289$)
- ▶ **SVM:** 0.986
- ▶ **Naive Bayes:** 0.957
- ▶ **Random Forest:** 0.993

Top 5 Most Important Genes

SVM

- ▶ Eukaryotic Translation Elongation FACTOR 1/2

Naive Bayes

Random Forest

Top 5 Most Important Genes

SVM

- ▶ Eukaryotic Translation Elongation FACTOR 1/2
- ▶ surfactant protein C, associated with lung disease

Naive Bayes

Random Forest

Top 5 Most Important Genes

SVM

- ▶ Eukaryotic Translation Elongation FACTOR 1/2
- ▶ surfactant protein C, associated with lung disease

Naive Bayes

- ▶ all mitochondrial proteins

Random Forest

Top 5 Most Important Genes

SVM

- ▶ Eukaryotic Translation Elongation FACTOR 1/2
- ▶ surfactant protein C, associated with lung disease

Naive Bayes

- ▶ all mitochondrial proteins

Random Forest

- ▶ surfactant protein 2

Top 5 Most Important Genes

SVM

- ▶ Eukaryotic Translation Elongation FACTOR 1/2
- ▶ surfactant protein C, associated with lung disease

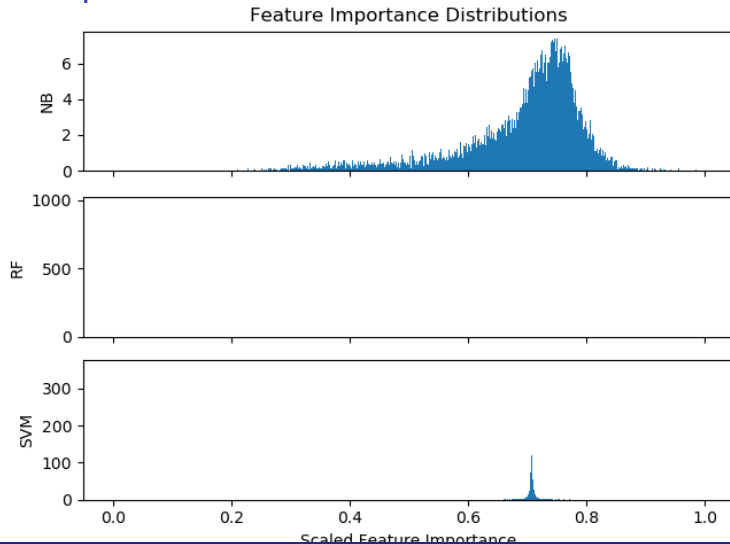
Naive Bayes

- ▶ all mitochondrial proteins

Random Forest

- ▶ surfactant protein 2
- ▶ **SLC11A1**, associated with tuberculosis

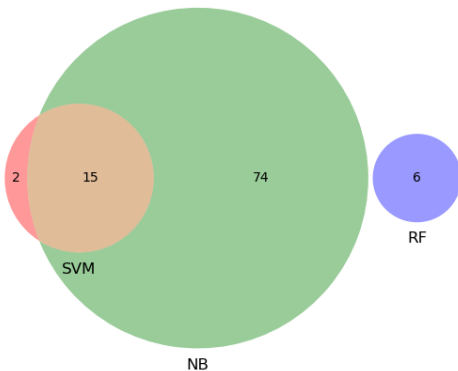
Gene Importance Distributions



GO Analysis of Important Genes

- ▶ Gene Ontology (GO) groups genes by function/mechanism

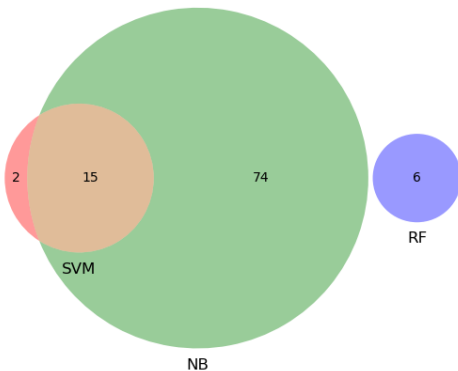
GO Term Overlaps



GO Analysis of Important Genes

- ▶ Gene Ontology (GO) groups genes by function/mechanism

GO Term Overlaps



- ▶ GO terms the most important genes across all methods?

Model Re-training

- ▶ compare model trained on top 10, 50, 250, 500 and all genes separately

Model Re-training

- ▶ compare model trained on top 10, 50, 250, 500 and all genes separately
- ▶ where do precision and recall start to drop off?

Conclusion

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things

Interpretable Deep Learning

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things
- ▶ Underlying structure of a model is important to which features are learned

Interpretable Deep Learning

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things
- ▶ Underlying structure of a model is important to which features are learned
- ▶ There may be equally useful orthogonal ways to approach treating lung cancer

Interpretable Deep Learning

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things
- ▶ Underlying structure of a model is important to which features are learned
- ▶ There may be equally useful orthogonal ways to approach treating lung cancer

Interpretable Deep Learning

- ▶ Seninge et al. propose a VAE Enchange by Gene Annotations

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things
- ▶ Underlying structure of a model is important to which features are learned
- ▶ There may be equally useful orthogonal ways to approach treating lung cancer

Interpretable Deep Learning

- ▶ Seninge et al. propose a VAE Enchange by Gene Annotations
- ▶ Rybakov et al. propose a similar method, like PCA along Reactome elements

Conclusion

- ▶ Even methods applied to the same problem with the same data can learn different things
- ▶ Underlying structure of a model is important to which features are learned
- ▶ There may be equally useful orthogonal ways to approach treating lung cancer

Interpretable Deep Learning

- ▶ Seninge et al. propose a VAE Enchange by Gene Annotations
- ▶ Rybakov et al. propose a similar method, like PCA along Reactome elements
- ▶ Associate latent variables with Biological Abstractions (gene sets)