A Study of Uncertainty in the LIDC

Jake Sauter, Victoria LaBarre,
Jacob Furst and Daniela Raicu

Overview:

- Study 1
  - Introduction
  - Motivation
  - Methodology
  - Results
  - Conclusion
- Study 2
  - Methodology
  - Results

Study 1

An Evaluation of Consensus Techniques for Diagnostic Interpretation

Introduction

- Lung cancer
- Computer-aided diagnosis (CAD)
- Variance in Ground Truth
- Mean consensus is the best technique to form ground truth with

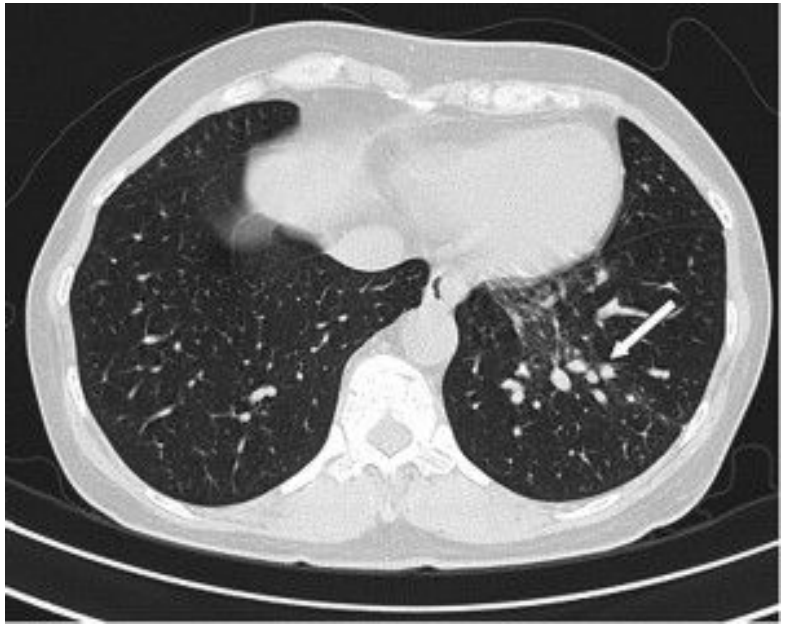


Figure 1: Lung Nodules in a Radiographic Image

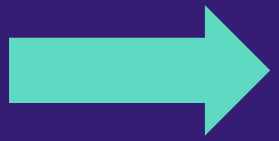
Literature Review:

- Lung Image Database Consortium: Developing a Resource for the Medical Imaging Research Community. Radiology
- Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity,Genetic Algorithm, and SVM

Motivation

- LIDC Data
- Consensus-based Label Extraction
  - Mean, Median, Mode
- Classification Models
- Hierarchy of Information Loss

PLV



- No loss of information
- Accuracy undefined

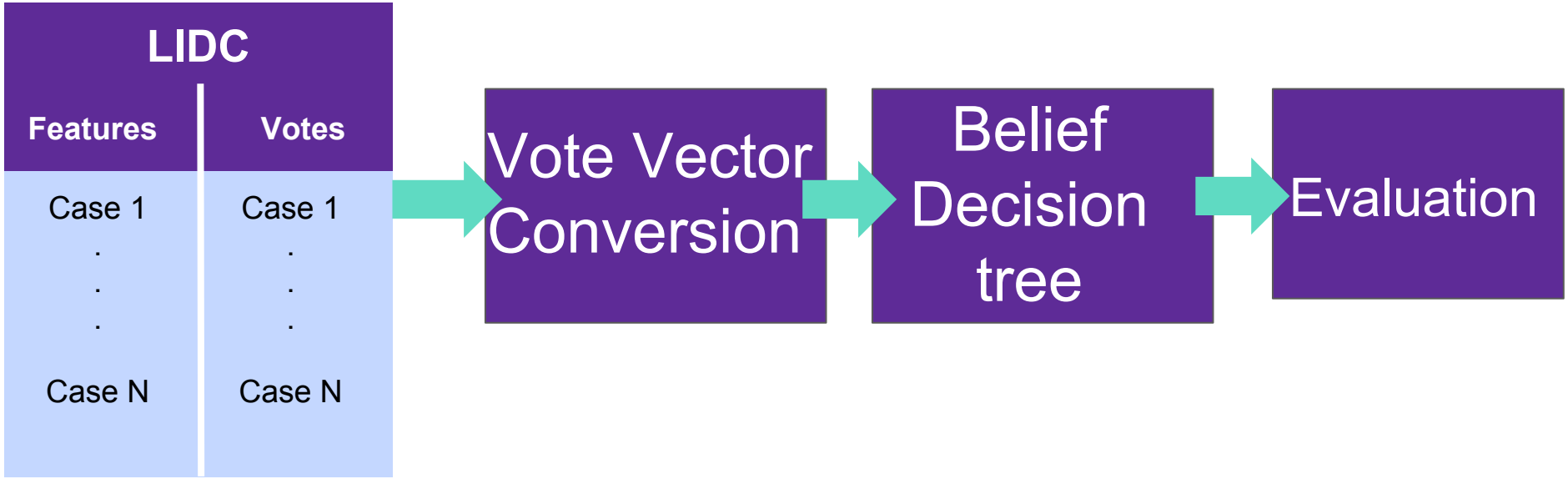
Label (1-5)

- Lose a little information
- Accuracy defined

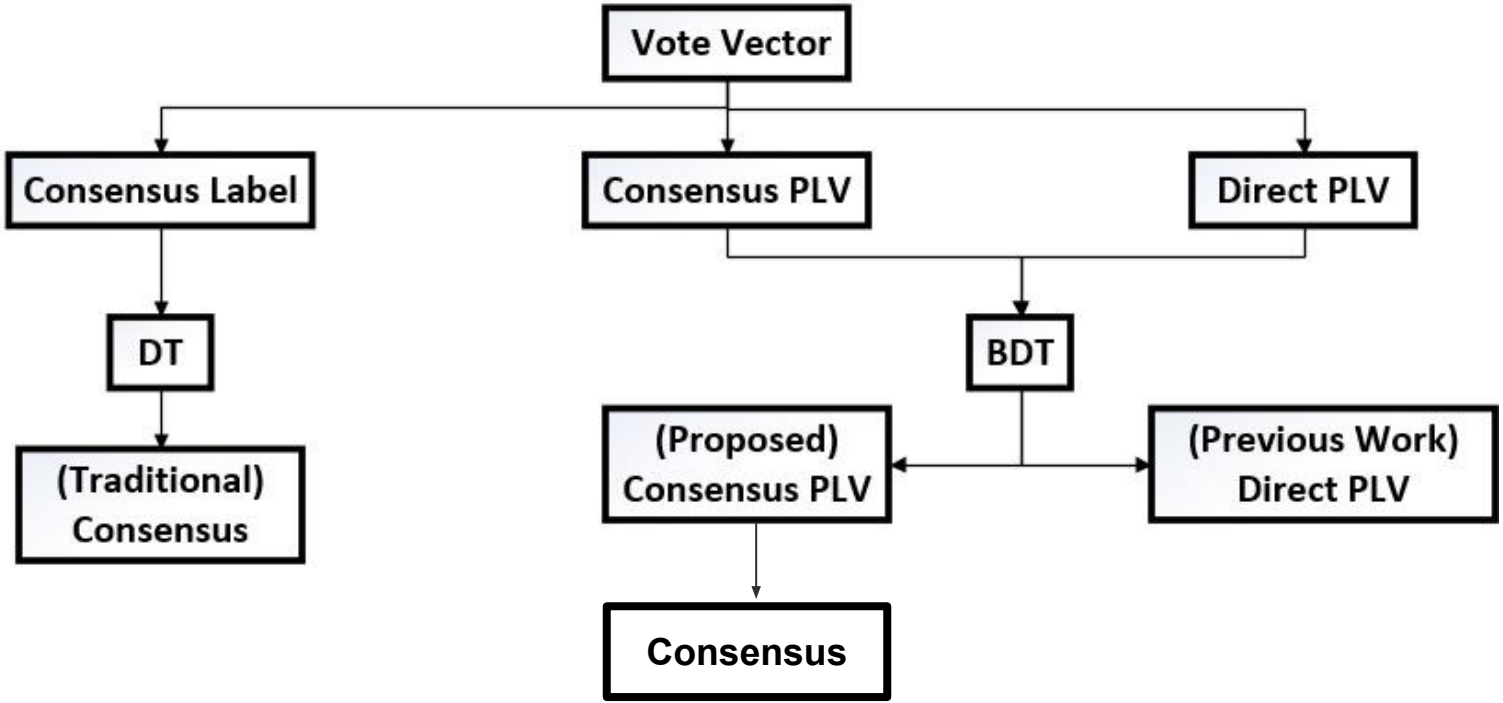
Binary

- Lose more information
- ROC, H-Measure, sensitivity, specificity...

Methodology



Methodology: Context



Methodology: Vote Vector Conversion

- Vote vector
  - Example: [2, 2, 3, 5]
- Converted using an consensus technique
  - Mean, median, mode
- Probabilistic label vector
  - Example of a Median PLV: [0, 0.5, 0.5, 0, 0]
  - Each index correlated to the probability of the nodule belonging to that malignancy class

Methodology: Vote Vector Conversion

- Mean
  - [2, 2, 3, 4] ➡ $\frac{2+2+3+4}{4}$ = 2.75 ➡ [0, .25, .75, 0, 0]
- Median
  - [2, 2, 3, 5] ➡ [~~2~~, 2, 3, ~~5~~] = 2.5 ➡ [0, 0.5, 0.5, 0, 0]
- Mode
  - [2, 2, 3, 5] ➡ [2, 2, 3, 5] = 2 ➡ [0, 1, 0, 0, 0]

Methodology: Label Extraction

- Mean

[.1,.1, .4, .2, .2] ➡ 1*.1+2*.1+3*.4+4*.2+5*.2 = 3.3 ➡ 3

- Median

[.1,.1, .4, .2, .2] ➡ RS: [.1, .2, .6, .8, 1] ➡ 3

- Mode

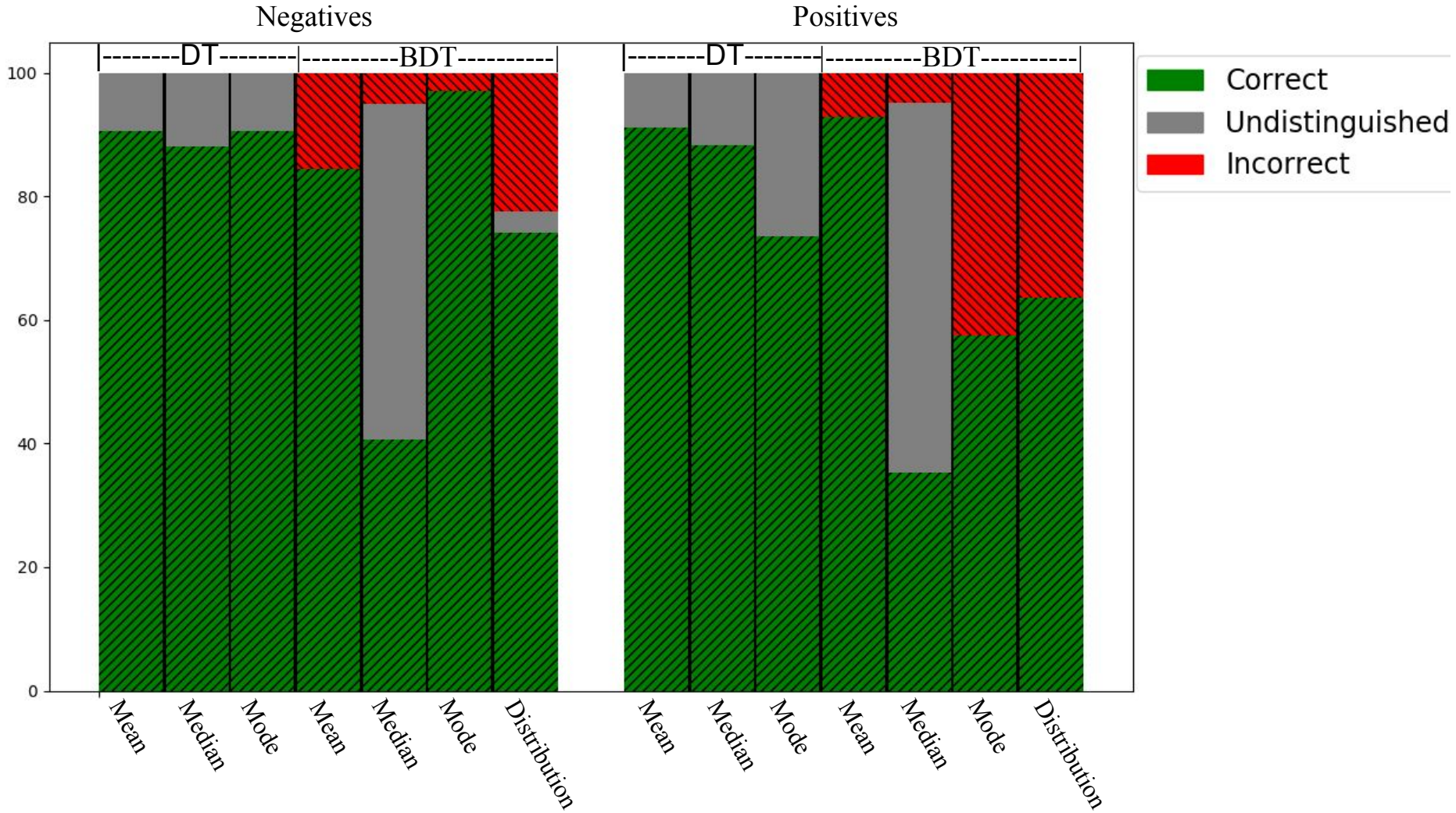
[.1,.1, .4, .2, .2] ➡ 3

Results:

- Compared consensus techniques
  - Mean, Median, Mode
- Compared forcing a label before/after
  - Multi Class Decision Tree (DT)
  - Belief Decision Tree (BDT)
- Compared consensus techniques to distribution(BDT)
  - Consensus provided better positive to negative response rate

Results: before looking at our chart

- DT- Decision Tree
- BDT- Belief decision Tree
- Mean, Median, Mode, Distribution
- Green- correct
- Red- incorrect
- Negatives
- Positives


Negatives
Positives
|-------DT-------|---------BDT---------|
|-------DT-------|---------BDT---------|
100
80
60
40
20
0
Correct
Undistinguished
Incorrect
Mean
Median
Mode
Mean
Median
Mode
Distribution
Mean
Median
Mode
Mean
Median
Mode
Distribution

Conclusions

- Consensus produced more robust results
- Consensus before model
  - better than consensus after model
- Mean is the best technique
  - Most positive responses to negative responses

Study 2

An Evaluation of Feature Characteristics of Atypical Nodules

Methodology: Support

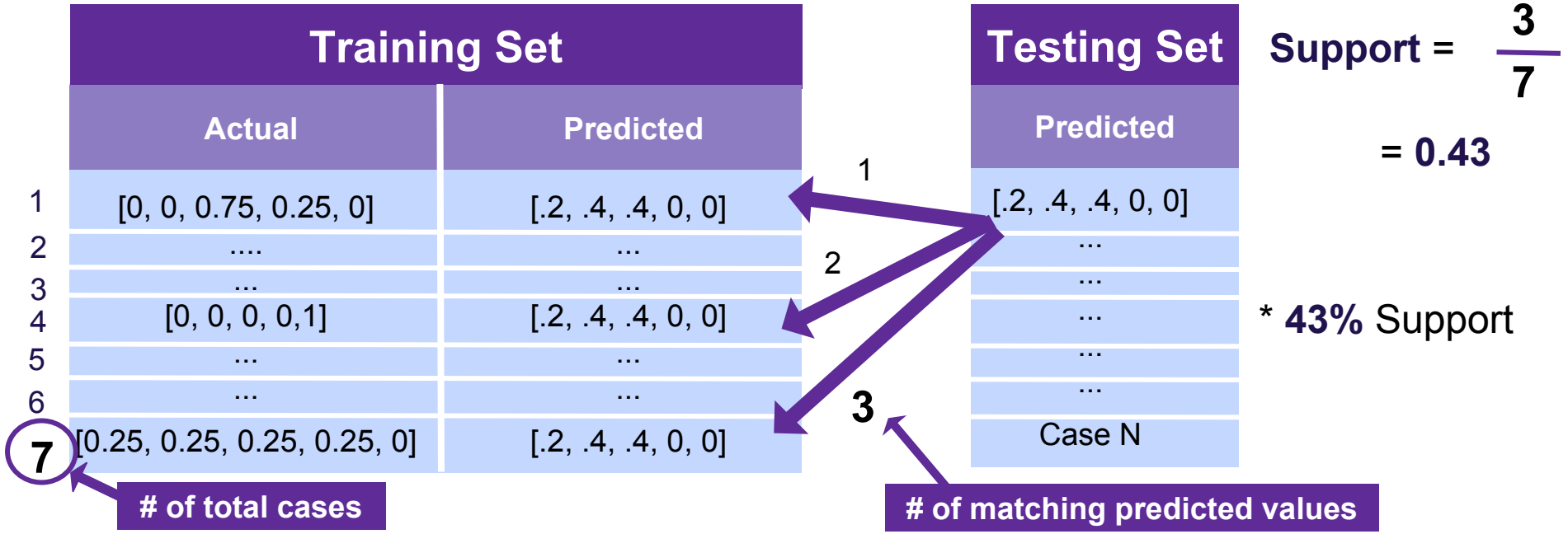
- Typical: There are “enough” similar nodules in the data
- **Atypical: the nodule is “significantly” different from the other nodules**

Methodology: Support

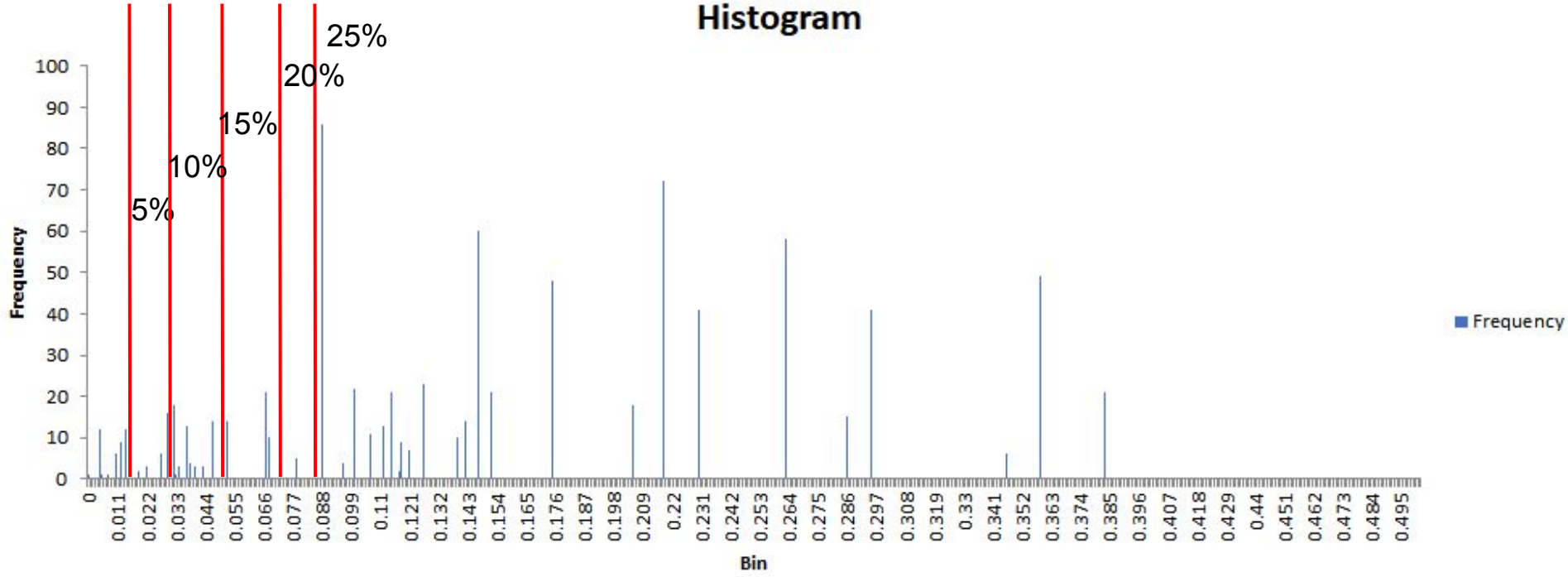
3 Strategies:

- **Based on labels predicted using BDT's**
- Based on image features
- Combinations the above strategies

Support


Training Set
Actual
Predicted
1
[0, 0, 0.75, 0.25, 0]
[.2, .4, .4, 0, 0]
2
....
...
3
...
...
4
[0, 0, 0, 0,1]
[.2, .4, .4, 0, 0]
5
...
...
6
...
...
7
[0.25, 0.25, 0.25, 0.25, 0]
[.2, .4, .4, 0, 0]
# of total cases
Testing Set
Predicted
[.2, .4, .4, 0, 0]
...
...
...
...
...
Case N
1
2
3
# of matching predicted values
Support = 3/7
= 0.43
* 43% Support

Results: Mean



Results: Discerning Features

Atypical nodules tend to be:

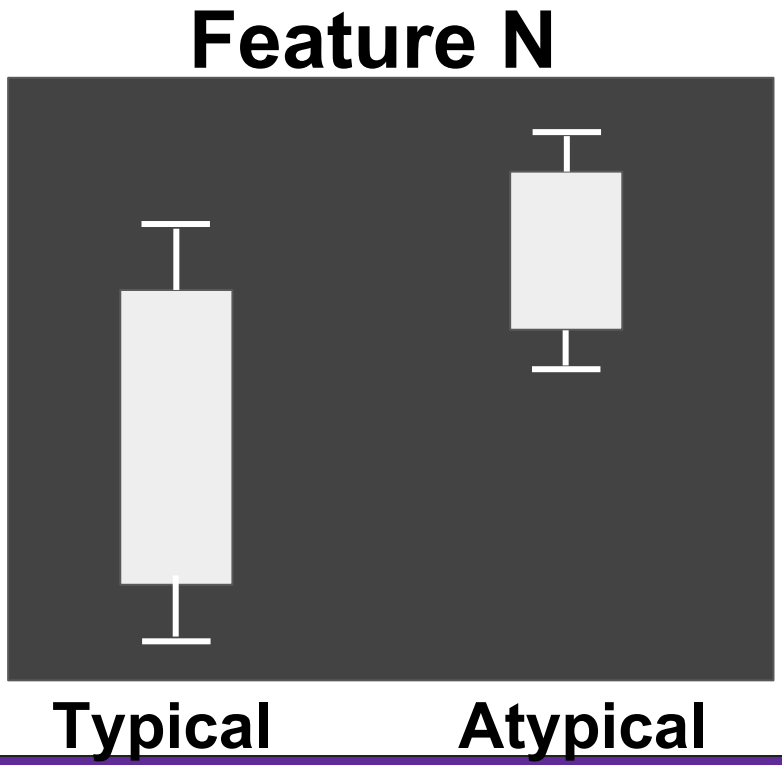
- Less solid
- Larger
- Less circular
- Less elongated
- Higher entropy

Results: Discerning Features

- Looked for a discerning feature of typicality
  - No one feature seemed to be able to do so
- Feature combinations necessary to determine atypical nodules
- Extracted the rules of the decisions of atypical cases
  - values of feature combinations


Questions?

Feature Boxplots



Results: extracted rules from atypical nodules

- Elongation < 1.230454272
- gaborSD_0_1 < 5782.78
- gabormean_3_1 > 59.0652
- MaxIntensityBG > 485.3564']

- gaborSD_0_0 < 1646.04
- MinorAxisLength >1.241783503
- Elongation < 1.390650159
- gaborSD_3_1 < 41.1318

- MinorAxisLength > 1.230454272
- SecondMoment < 0.170377481
- markov1 > 271.5484
- SDIntensity < 1412.0
- noduleID > 154.0