Reproducing *Data Mining Applied to Acoustic Bird Species Recognition* paper by Vilches et. al

Talha Ansari

April 30th, 2014

Columbia Engineering
The Fu Foundation School of Engineering and Applied Science

Abstract

In this work, we explore the application of data mining techniques to the problem of acoustic recognition of bird species. Most bird song analysis tools produce a large amount of spectral and temporal attributes from the acoustic signal. The identification of distinctive features has become critical in resource constrained applications such as habitat monitoring by sensor networks. Reducing computational requirements makes affordable to run a classifier on devices with power consumption constraints, such as nodes in a sensor network. Experimental results demonstrate that considerable dimensionality reduction can be achieved without significant loss in classification efficiency.

Agenda

- Background of the paper
- Reproduction
  - Data acquisition
  - Data processing
  - Feature extraction
  - Classification
  - Results
- Conclusions

Background

- *Data*: Cornell Lab of Ornithology, Macaulay Library
- *Feature Extraction* using Sound Ruler
- *Classification* methods: Decision Trees (ID3 and J4.8), Naive-Bayes
- *Dimensionality* reduction using the decision trees
- *Results*: High successful classification rate, even after dimensionality reduction

Reproduction - Steps

1. Data Acquisition
2. Data processing
3. Feature Extraction
4. Classification
5. Results

1) Data acquisition…

- Dataset bought from the Macaulay Library - total of 400 files
- *Problem*: No information on which exact files used by the authors
- *Solution*: About 228 files shortlisted by looking at the file size and quality. Files above 10MB ignored.

Specie	Original	Reproduced
Taraba Major	49	68
Cercomacra tyrannia	79	79
Thamnophilus Doliatus	76	81
Total	204	228

1) …Data acquisition

Metadata for sound files provided

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Format	Catalog	Scientific	Common	Male	Female	Adult	Immatur	Call	Song	Mechanic	Behavior	Environm
2	Audio	165049	Campylo	Rufous-naped Wren; White-tipped Dove; Barred Antshrike									Y
3	Audio	62018	Cercoma	Dusky A	Y	Y	Y			Y		song	
4	Audio	62017	Cercoma	Dusky A	Y	Y	Y			Y		song	
5	Audio	62013	Cercoma	Dusky A	Y	Y	Y		Y	Y		call;song	
6	Audio	62008	Cercoma	Dusky Antbird									
7	Audio	29130	Thamnop	Barred Antshrike			Y			Y		song	
8	Audio	7228	Taraba m	Great Antshrike					Y			call	
9	Audio	7053	Cercoma	Dusky Antbird									
10	Audio	134643	Cercoma	Dusky A	Y	Y	Y		Y	Y		call;duet;song	
11	Audio	164013	Thamnop	Barred A	Y	Y	Y			Y		duet;song	
12	Audio	128084	Taraba m	Great Antshrike						Y		song	
13	Audio	127527	Thamnop	Barred Antshrike									
14	Audio	117018	Cercoma	Dusky A	Y	Y	Y			Y		duet;song	
15	Audio	88840	Taraba m	Great Antshrike			Y			Y		song	
16	Audio	88703	Thamnop	Barred A	Y	Y	Y			Y		song	
17	Audio	88217	Cercoma	Dusky A	Y		Y			Y		song	
18	Audio	72418	Cercoma	Dusky A	Y	Y	Y		Y	Y		call;duet;song	

2) Data processing…

Original

1. Sound Ruler software used
2. Sounds passed through LP and HP filters (specific for each specie)
3. Each sound file divided into 'pulses'

Number of pulses/segments for each class

Specie	original	reproduced
Taraba Major	21,360	1400
Cercomacra tyrannia	5,373	1757
Thamnophilus Doliatus	911	1533

Reproduction

1. Sound Ruler found to be outdated, could not be used
2. Filters created and implemented in Matlab
3. Sound files divided into 'segments' using *smooth()*, applying a threshold, and looking at the differential (code shown)

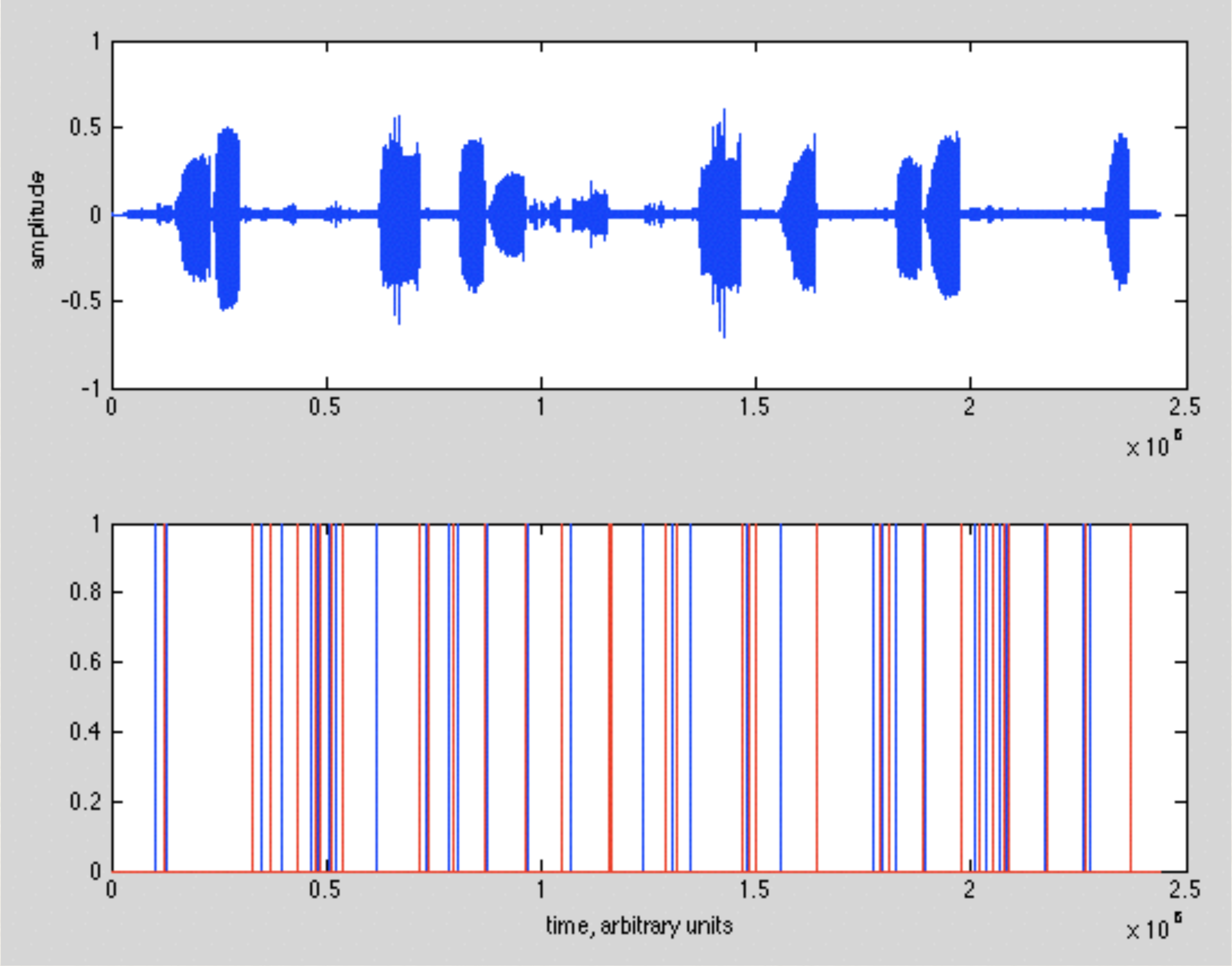
	Taraba Major	Cercomacra Tyrannina	Thamnophilus Doliatus
Low-pass filter	3597 Hz	4200 Hz	3597 Hz
High-pass filter	517 Hz	920 Hz	686 Hz

Table 1. Low-pass and High-pass filters per species

2) …Data processing (the code)

```
function X = gen_features2(audioarray, fs, savepath)
% Write feature files of the input sound data - according to the
setting, .mat
% example usage:
% gen_features(y, '1.mat');
% if no segment detected, return 0 - not sure how often it happens
% Created by Talha Ansari 2014, talhajansari@hotmail.com

head;
plot = false;
y = audioarray;

% Do some smoothing
yabs = abs(y);
ys = smooth(yabs,5000);
yss = smooth(ys,5000);
yssd = diff(yss,1);
yssds = smooth(yssd, 5000);

% Indentify the segments of the audiofile by applying a threshold
filter
on = [];
off = [];
t_on = false;
t_off = true;
cnt = 0;
frame = 5000;
for i=6:length(yssds)-5
    if yss(i)>=0.005 && mean(yssds(i-5:i+5)>0) && t_on==false &&
t_off==true
        on = [on, i];
        t_on = true;
        t_off = false;
    elseif yss(i)<=0.005 && mean(yssds(i-5:i+5)<0) && t_on==true &&
t_off==false && cnt>=frame
        off = [off, i];
        t_on = false;
        t_off = true;
        cnt = 0;
    end
    if t_on==true
        cnt = cnt +1;
    end
end
% if the last 'off' hasnt been found yet
if numel(off) < numel(on)
    if (length(y)-on(end))>=frame % if enough frame size
        off = [off, length(y)]; %if on has not ended, then off is
the last index of the audio
    else % remove the last on (frame) altogether
        on = on(1:end-1);
    end
end

% Break the audio into segments
Y = cell(length(on),1);
X = cell(length(on),1);
for k = 1:length(on)
    Y{k} = y(on(k):off(k));
    X{k} = melcepst(Y{k}, fs);
end

save(savepath, 'X');
```

2) …Data processing (the plot)



3) Feature extraction

Original

1. Using Sound Ruler (assumed)
2. 72 features - no information given on the type of features, or how they were obtained

Reproduction

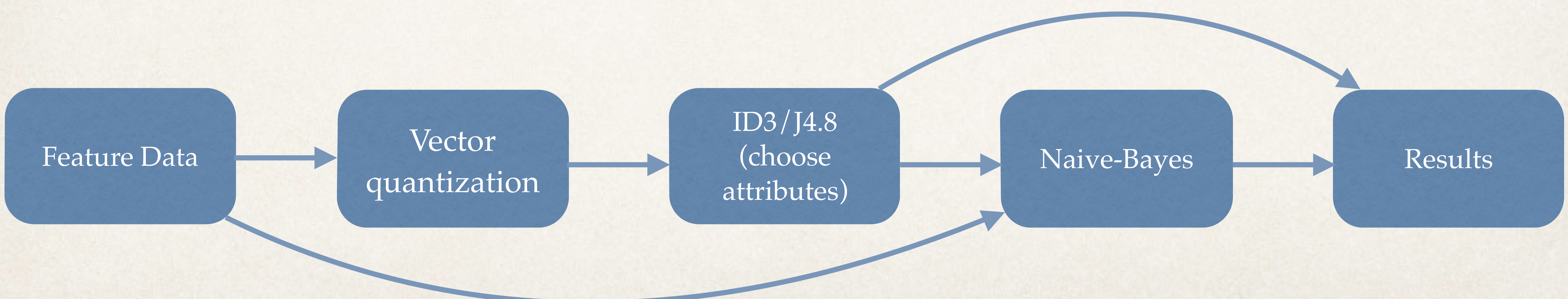
1. Sound Ruler was not available: Matlab code used to extract features. Mike Brooke's *Voicebox* library (http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html)
2. 55 features - based on MFCC (12th order)

Features = 12 mean(MFCCs), 12 min(MFCCs), 12 max(MFCCs),…
…12 var(MFCCs), a few others

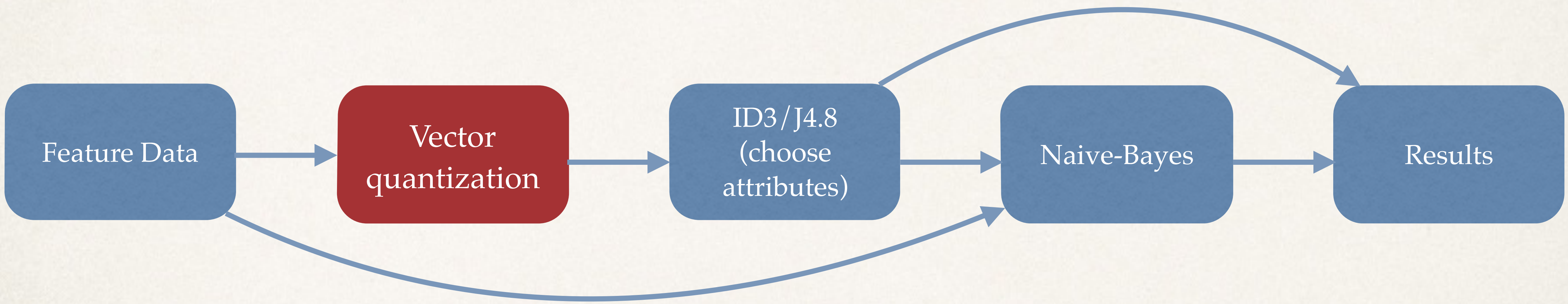
4) Classification…

Original & Reproduction

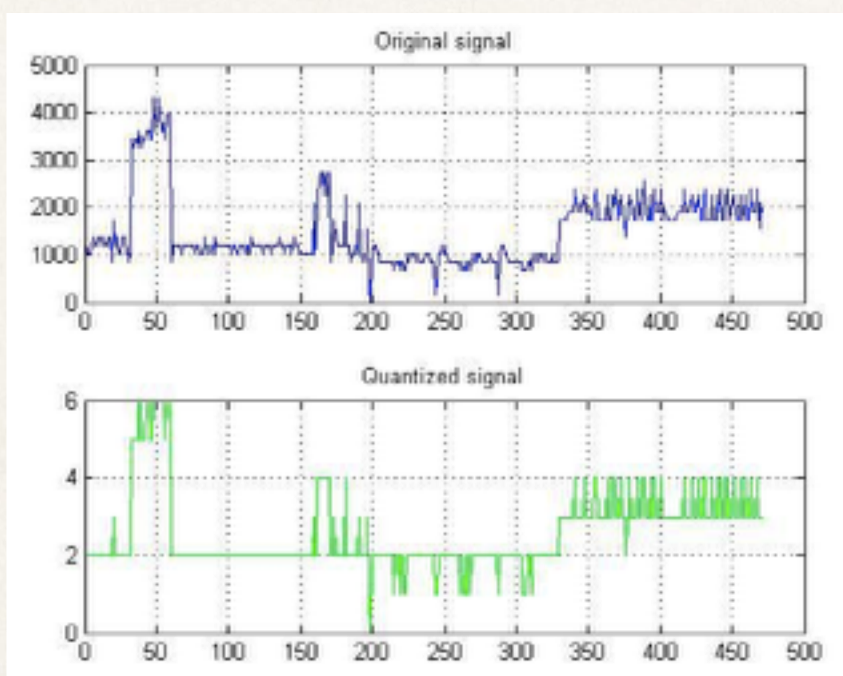
1. Weka Software used for almost everything (other than vector quantization)
2. Cross-validation of 70% - 30%
3. Decision Tree (ID3 and J4.8) used to select the most important features
4. Naive Bayes performed on the reduced feature set
5. Performance compared to the performance of Naive-Bayes on the full feature set



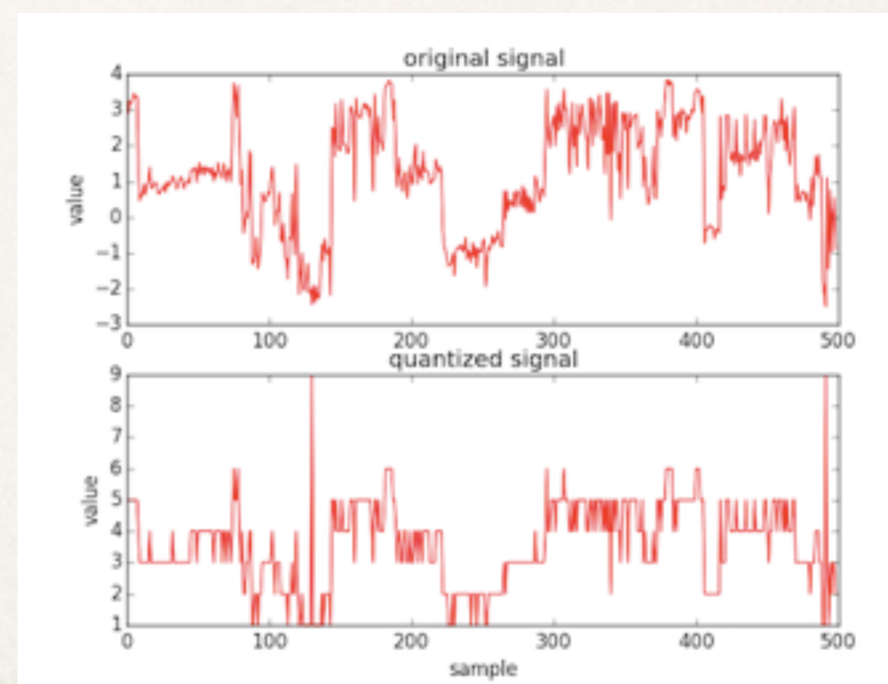
4) …Classification (vector quantization)



Original



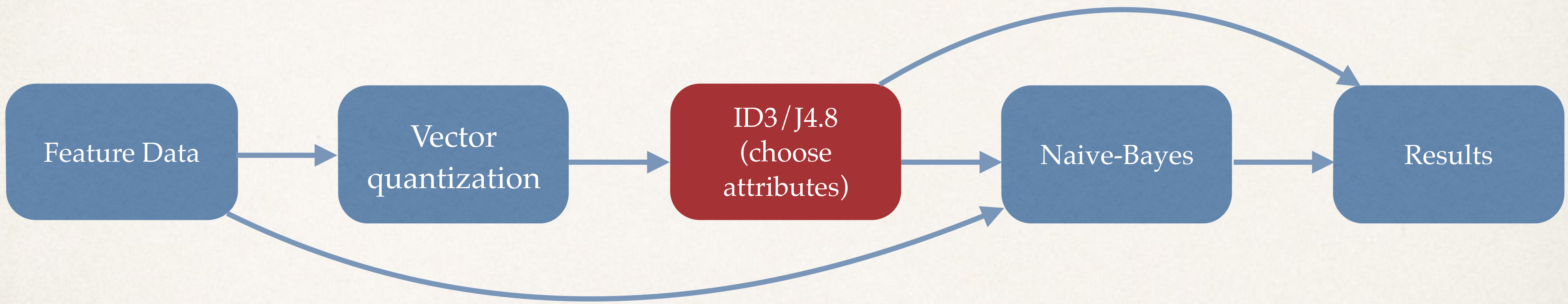
Reproduction



Link to Github

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

4) …Classification (decision trees)



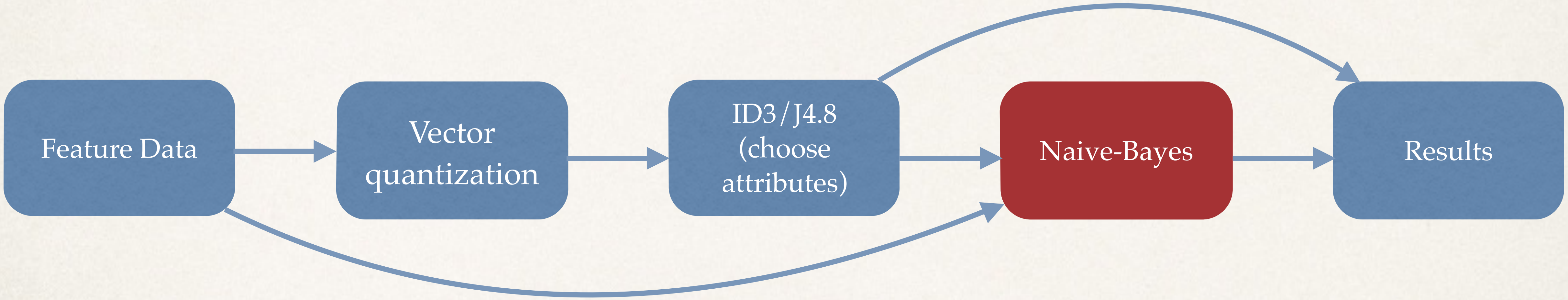
$$Entropy(p1, p2, \ldots pn) = \sum_{i=1}^{n} -pi \log_2 pi$$

$$Gain(P, xP) = Entropy(P) - Entropy(x \mid P)$$

<u>Reproduction</u>

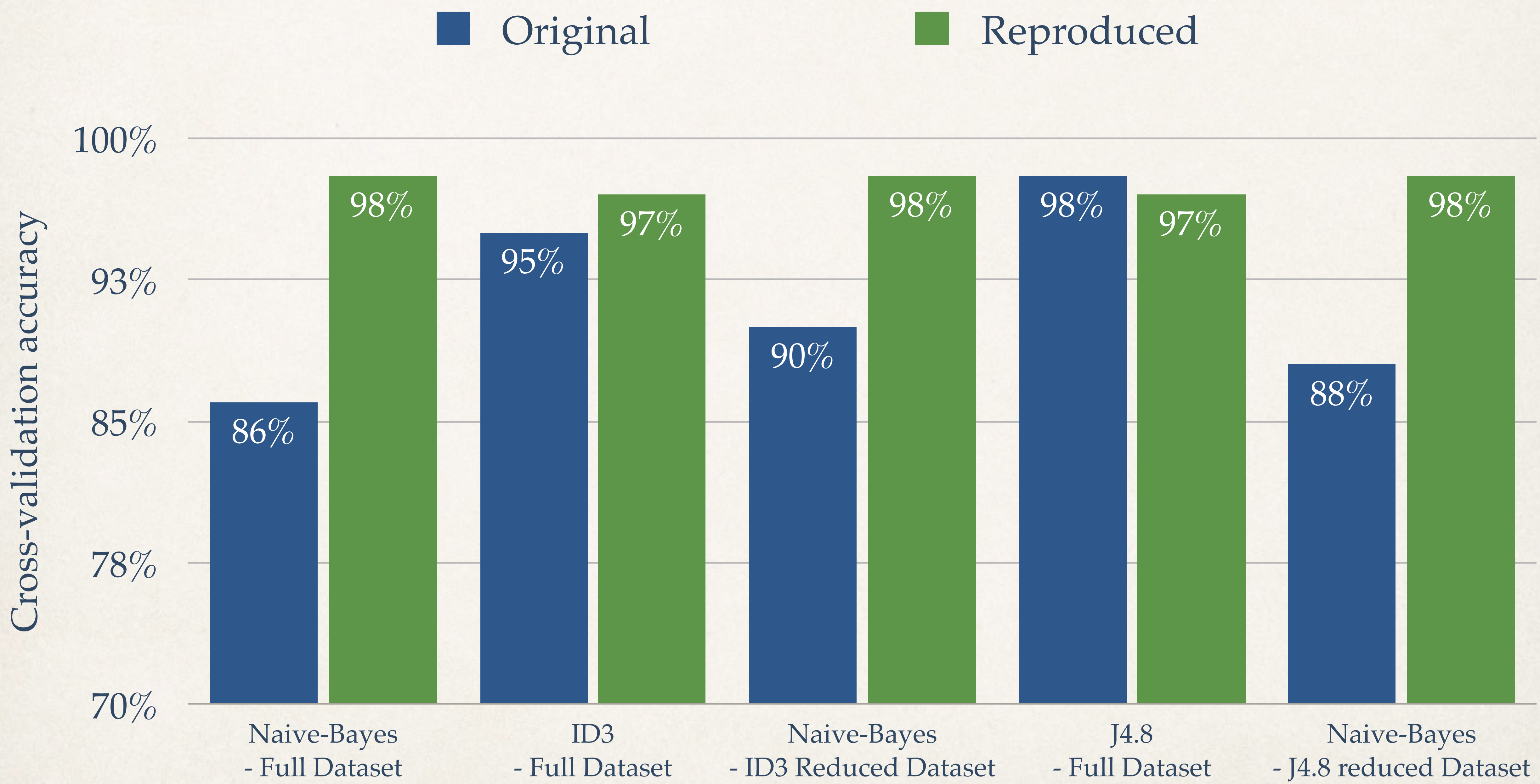
- ID3 -> 26 out of 55 features used
- J4.8 -> 18 out of 55 features used (stronger post prune)

4) …Classification (Naive-Bayes)

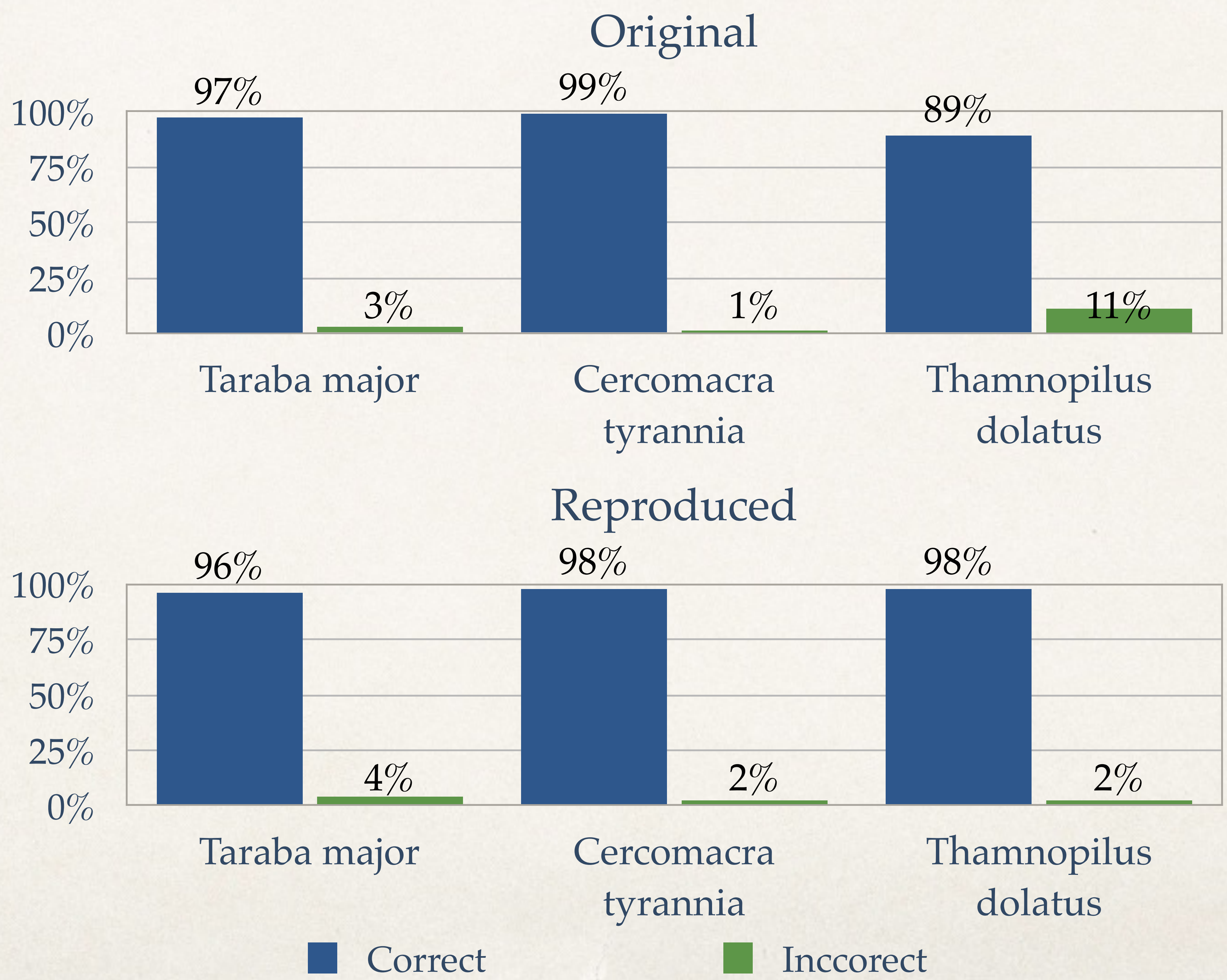


- Naive-Bayes uses Bayesian rule, to calculate posteriori probabilities from a priori probabilities
- Carried out on the full dataset and the reduced datasets from ID3 and J4.8

Results


Original
Reproduced
Cross-validation accuracy
100%
93%
85%
78%
70%
86%
98%
95%
97%
90%
98%
98%
97%
88%
98%
Naive-Bayes
- Full Dataset
ID3
- Full Dataset
Naive-Bayes
- ID3 Reduced Dataset
J4.8
- Full Dataset
Naive-Bayes
- J4.8 reduced Dataset

J4.8 accuracy compared



Conclusions

- A lot of crucial information not provided by authors, forcing us to make assumptions
- Outdated software used - so had to write alternative code
- Data skewed towards one specie - *Taraba Major* - over 75% of samples
- Results were, arguably, replicated. Results better in reproduction perhaps because of i) more sound files, ii) better segmentation and/or iii) better features

Github repository (by tonight)

- https://github.com/talhajansari/columbia_e6891
- *README.md* has instructions on how to use the reproduction package
- External packages also included
- Data provided through a separate downloadable link

References

[1] Vilches, Erika; Escobar, Ivan A.; Vallejo, E E; & Taylor, C E. (2006). Data Mining Applied to Acoustic Bird Species Recognition. Center for Embedded Network Sensing, 3, 400 - 403. doi: 10.1109/ICPR.2006.426. UCLA: Center for Embedded Network Sensing.

Thank you!

Questions?