Photometric Detection of Quasars

ISRO - Funding
Santhom Computing Centre@
St. Thomas College in Kozhencheri

Sheelu Abraham

Ninan Sajeeth Philip

Understanding

Analysis

Archives

Data

Universe

Objective

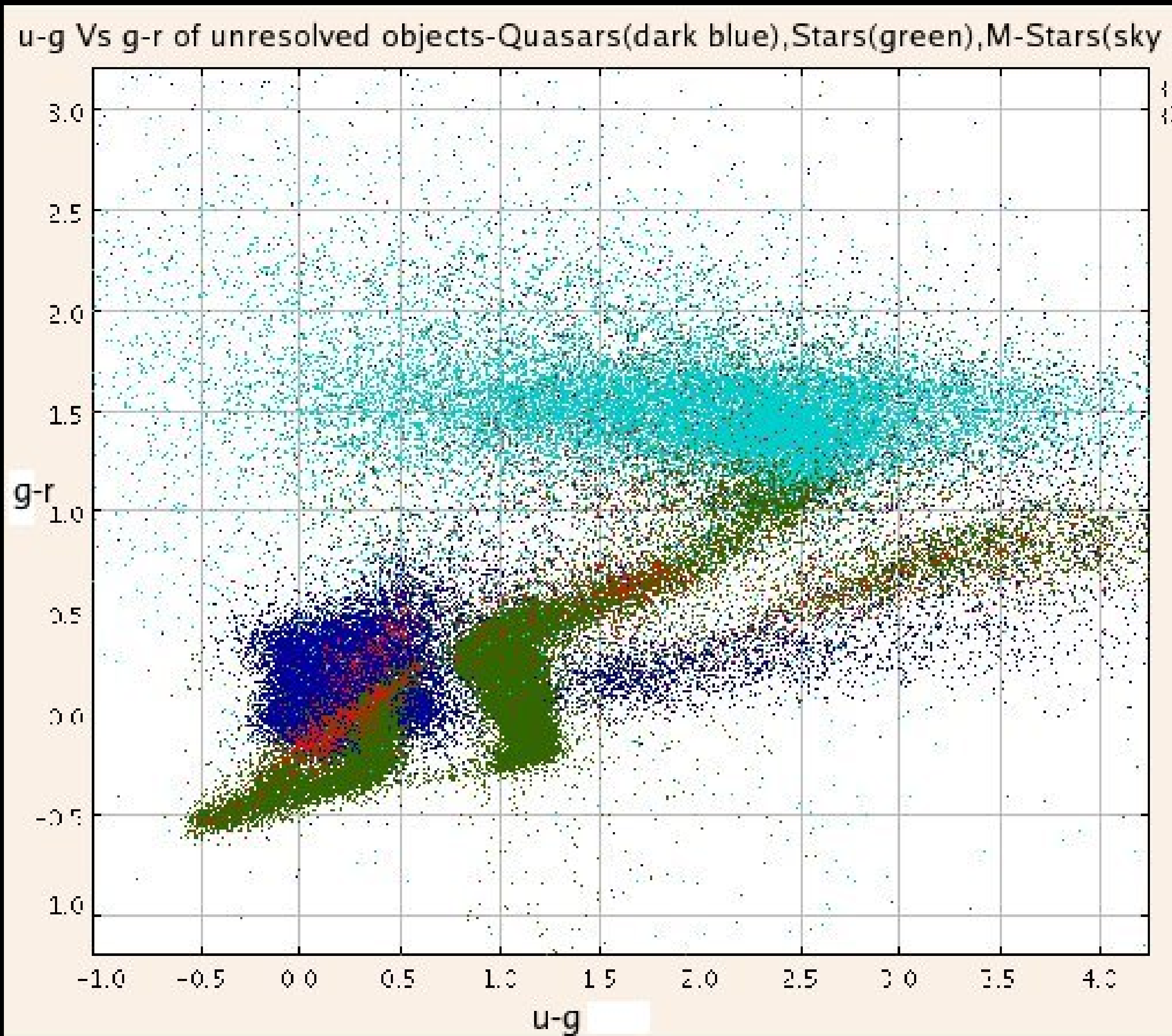
Illustration of Machine Learning

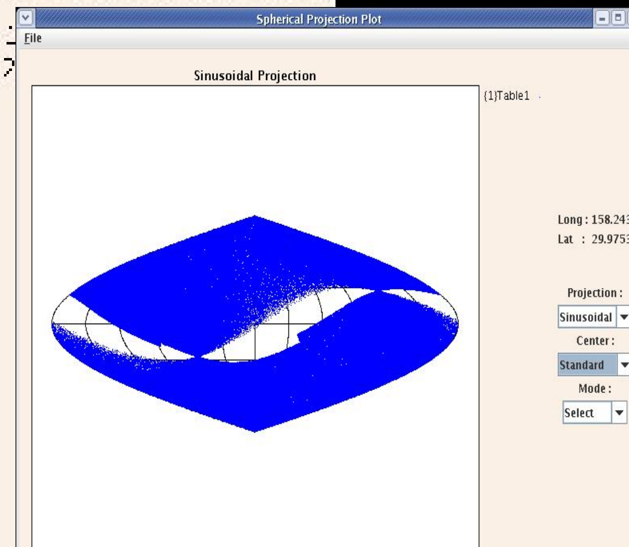
Applications in Astronomy

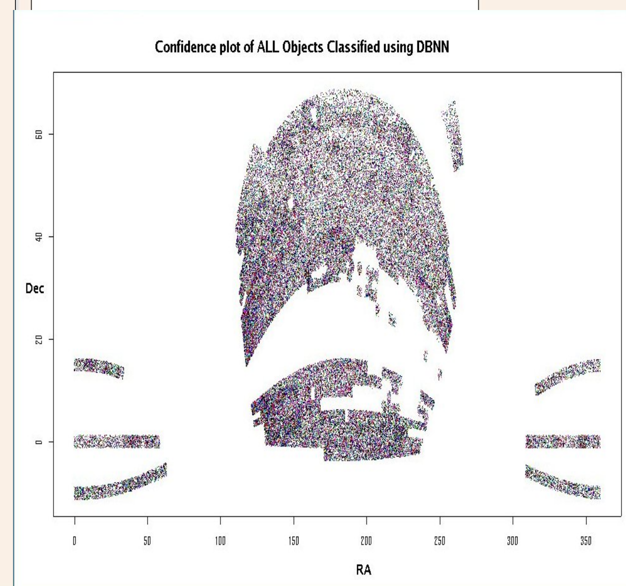
Using SDSS data Archive


清酒 千歳鶴

Quasar Colours







Difference Boosting
Neural Network

Object types are identified on the basis of subtle differences in features.

What is Bayes Rule?

Suggests to update your belief on the basis of evidences in 3 steps.

Prior P(A) : Initial belief

Evidence B : Make an Observation

P(A|B) : Update Prior with a posterior

Repeat until the confidence is acceptable

How to compute the Posterior Probability?

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

P(A) is prior belief that the outcome is A

P(A|B) is the updated belief about the outcome

P(B|A) is the likelihood for B to cause A

P(B): probability for B to occur.

Iterative learning

$$P(C_k \mid U) = \frac{\prod_m \hat{P}(U_m \cap C_k) W_m}{\sum_j \prod_m \hat{P}(U_m \cap C_j) W_m}.$$

$$\Delta W_m = \alpha \left[1 - \frac{P_k}{P_k^*} \right]$$

Where P_k stands for the posterior for the actual class and P_k* for that of the predicted class.

Start with a small and equal value for W_m initially and as each new data is seen, update W_m along the negative gradient of the probability estimates – which in effect is the difference of their posterior estimates.

Training data

Since the entire learning process depends upon the data, it is important that the data used for training has all the features to be expected in the problem domain.

Iterative identification of Principal Samples

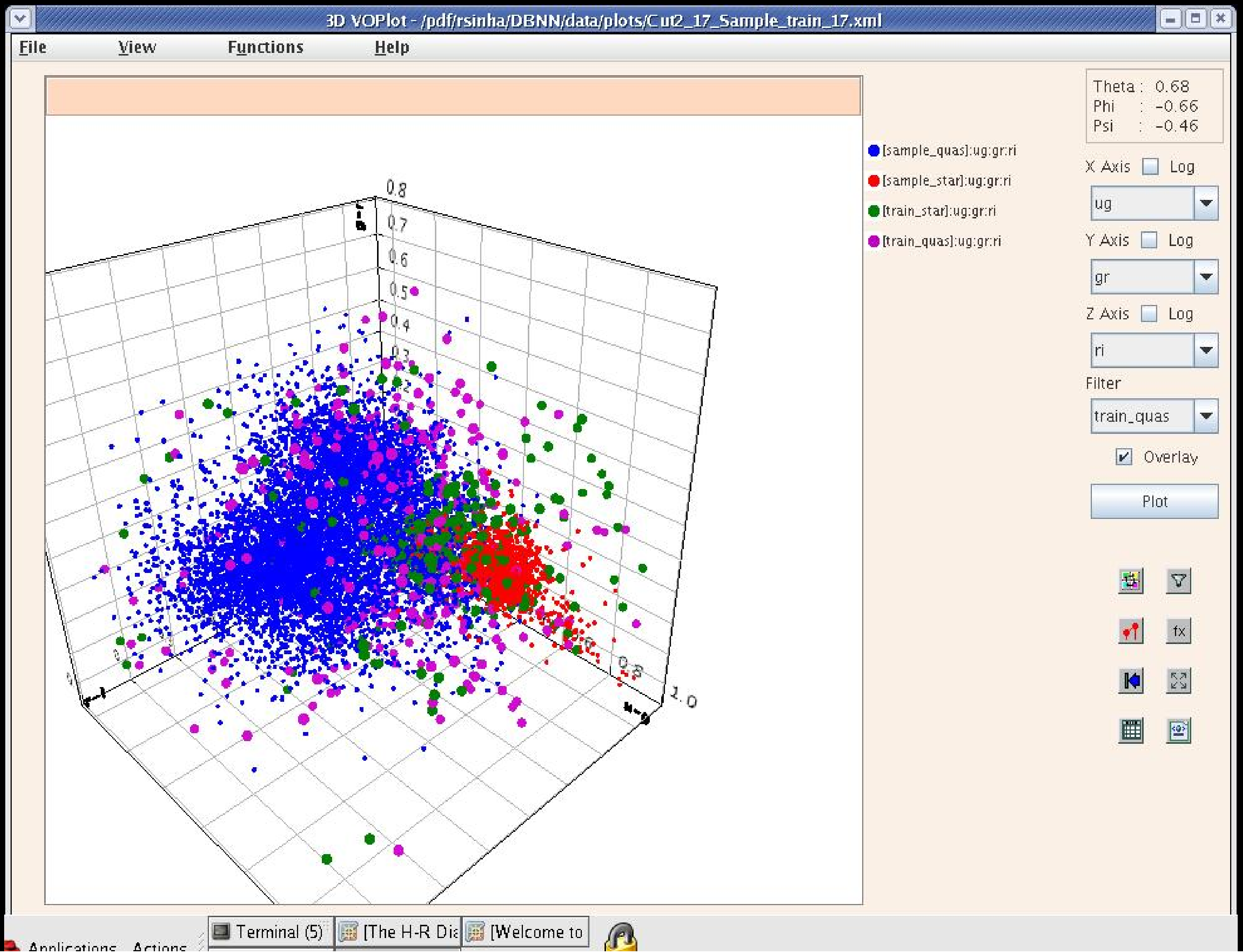
Randomly pick up just one example from each class and train the network on it.

Test the network on the remaining data

Look for the example that has failed with maximum posterior probability from each class and add it to the training data.

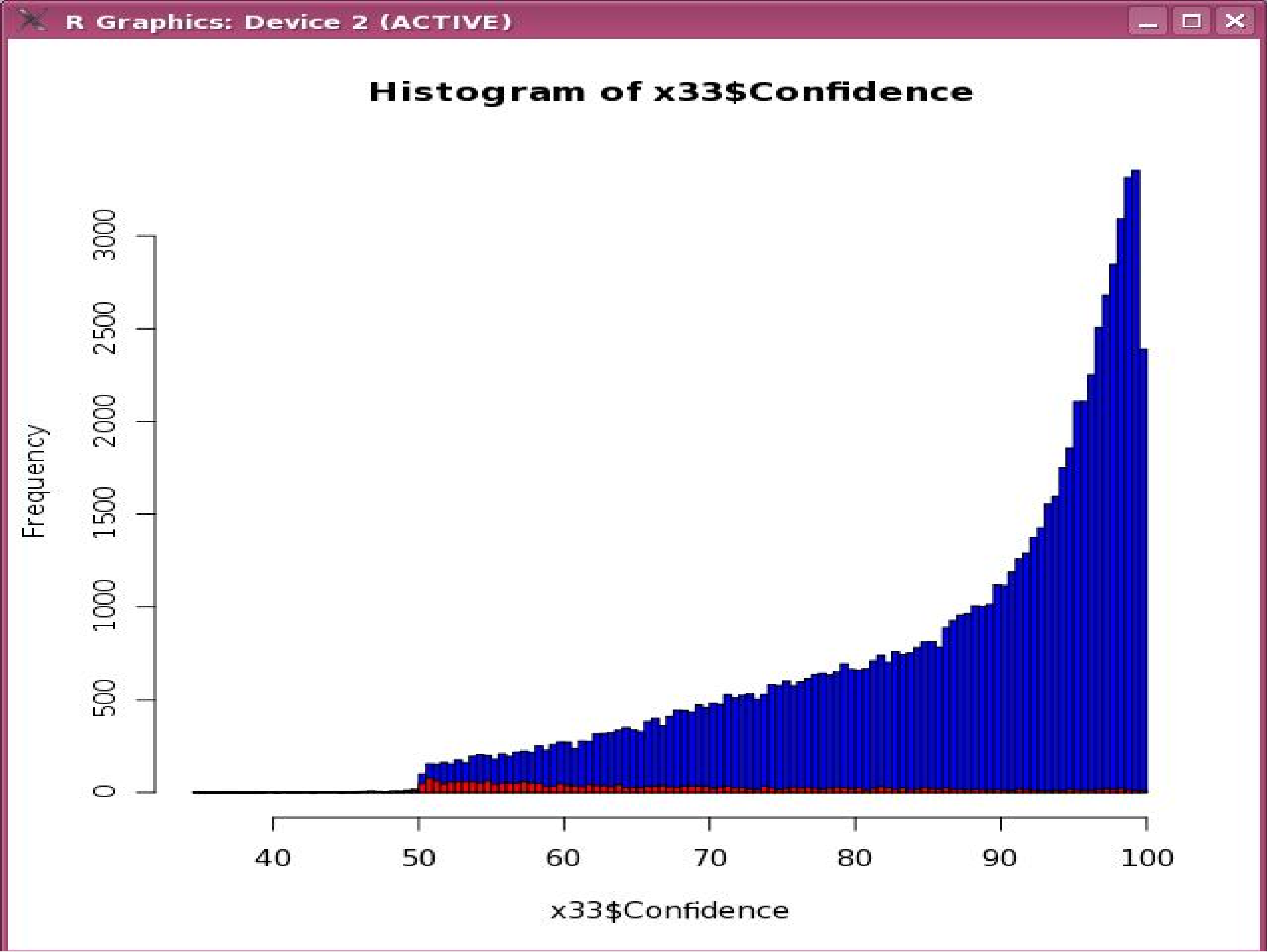
REPEAT THE PROCESS UNTIL ALL EXAMPLES ARE CORRECTLY CLASSIFIED.

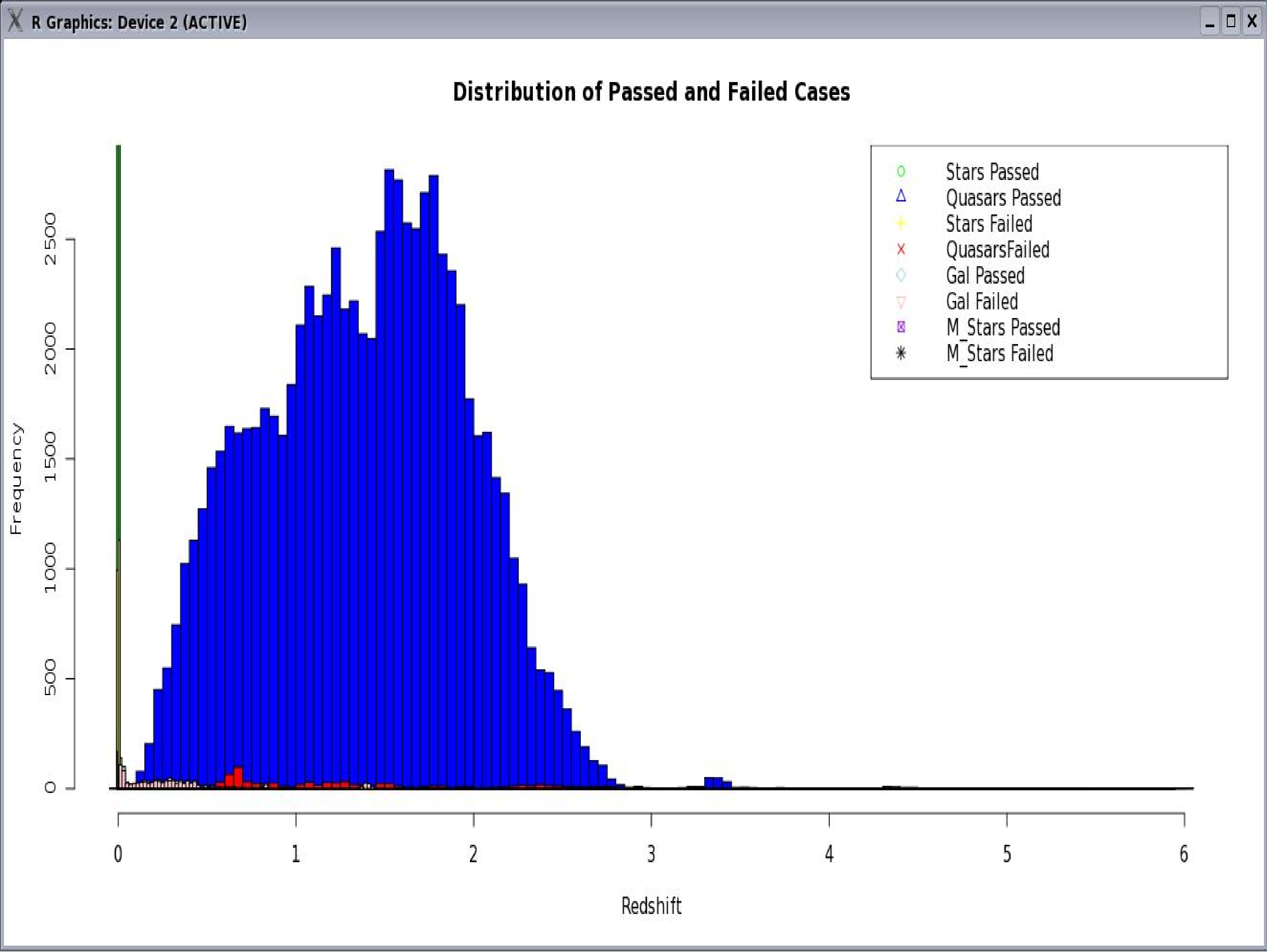
Optimal Data Selection

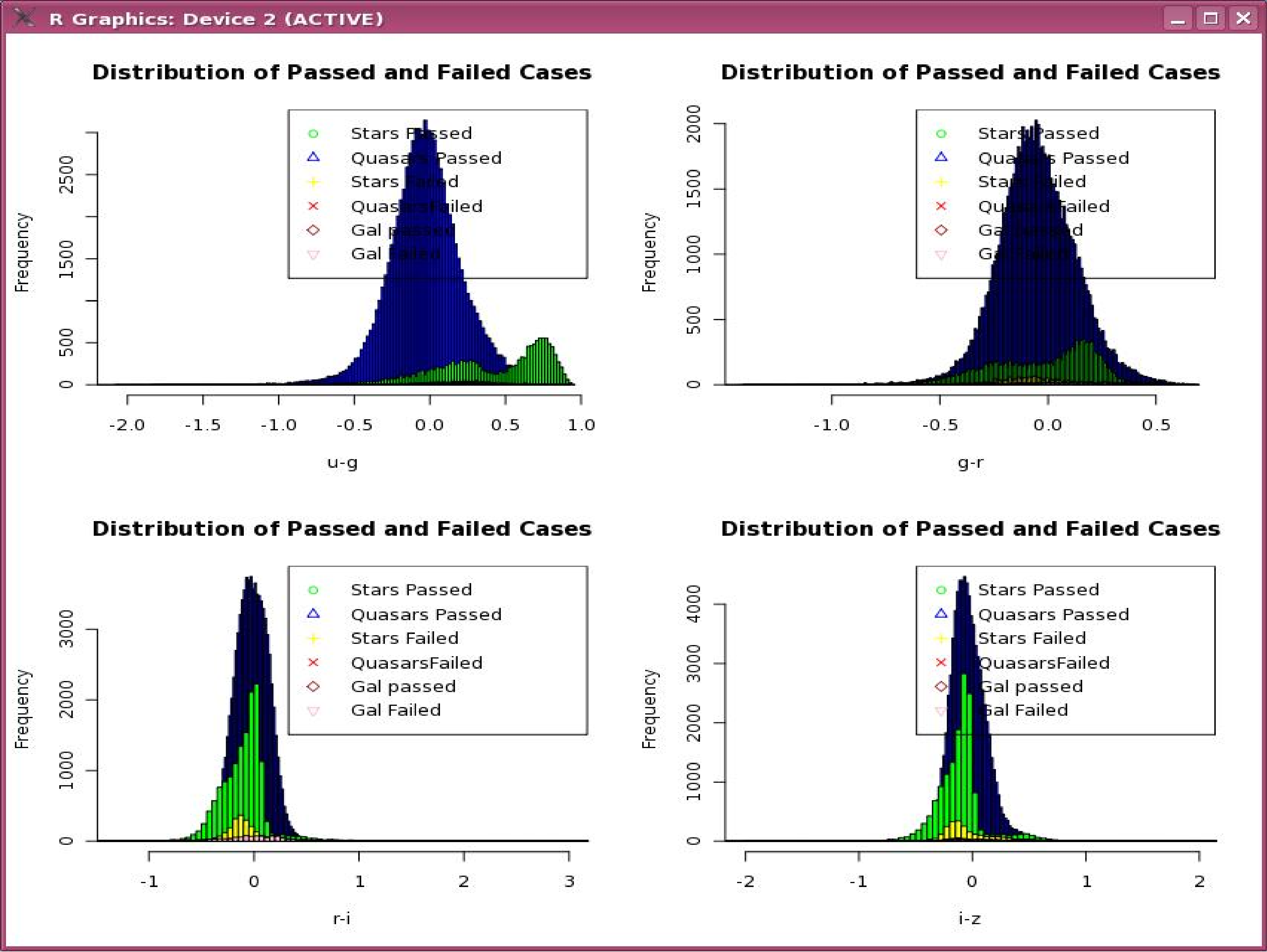


Classification Results

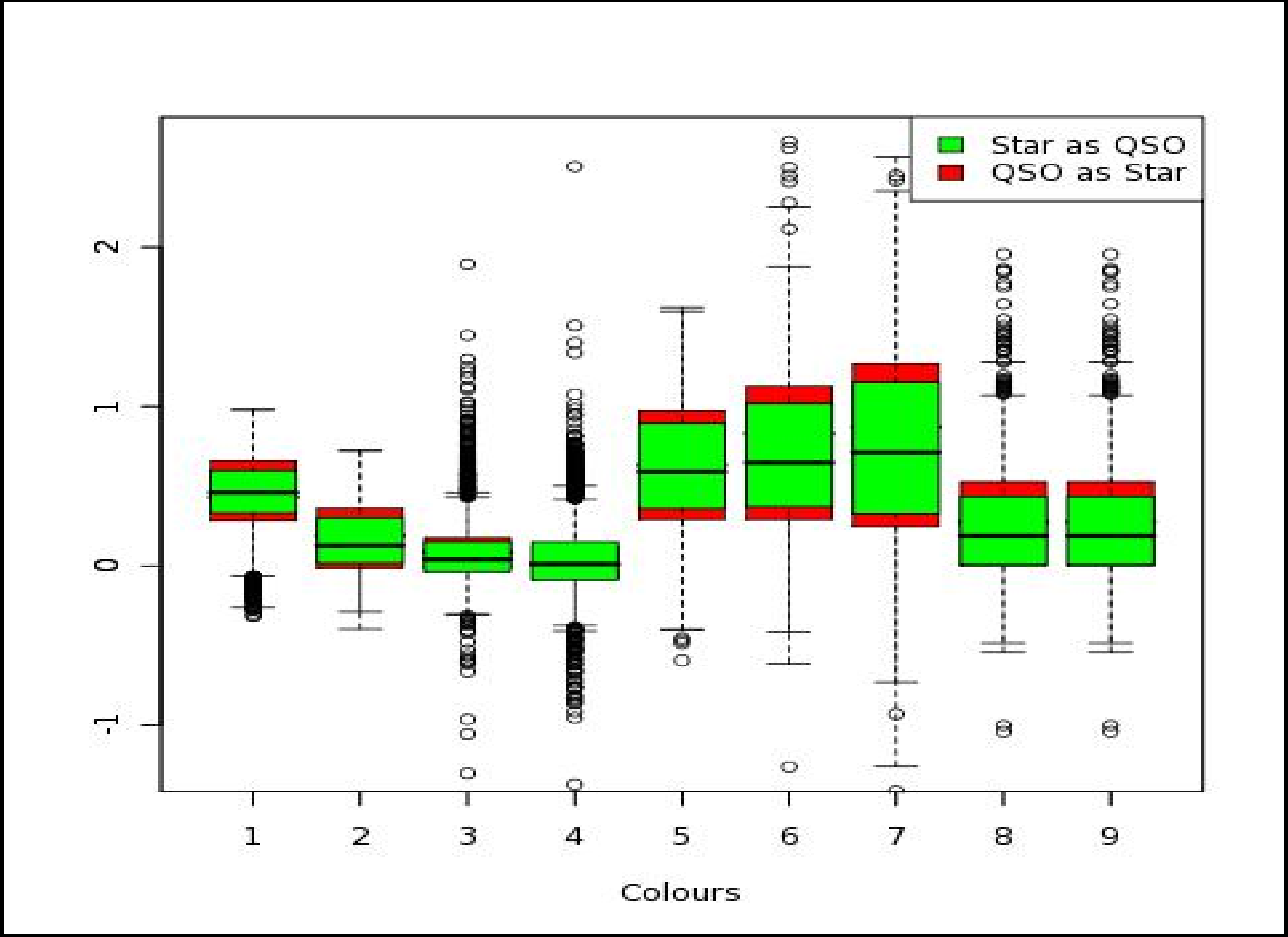
Name	SDSS	DBNN	%
Star	72449	68747	94.89
Galaxy	1868	238	12.74
Quasar	90363	88290	97.7
HizQSO	3702	3289	88.86
Star-Late	759	184	24.24
Total	169141	160748	95.03


R Graphics: Device 2 (ACTIVE)
Histogram of x33$Confidence
Frequency
0
500
1000
1500
2000
2500
3000
40
50
60
70
80
90
100
x33$Confidence


R Graphics: Device 2 (ACTIVE)
Distribution of Passed and Failed Cases
Stars Passed
Quasars Passed
Stars Failed
QuasarsFailed
Gal Passed
Gal Failed
M_Stars Passed
M_Stars Failed
Frequency
0
500
1000
1500
2000
2500
0
1
2
3
4
5
6
Redshift

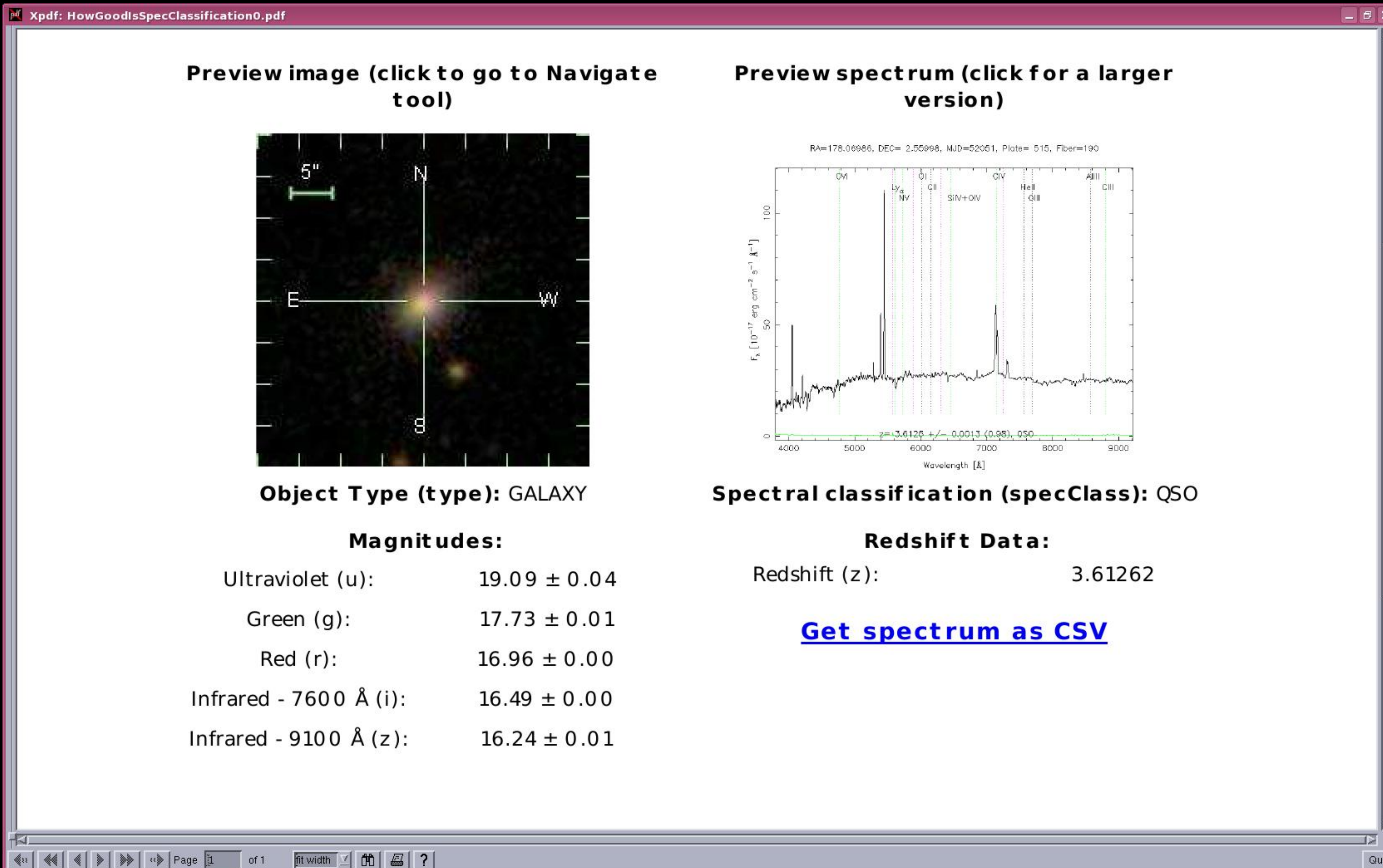

R Graphics: Device 2 (ACTIVE)
Distribution of Passed and Failed Cases
Stars Passed
Quasars Passed
Stars Failed
QuasarsFailed
Gal passed
Gal Failed
Frequency
u-g
g-r
r-i
i-z

What caused failures?



Colors of quasars and stars are indistinguishable at some redshift patches

What caused failures?



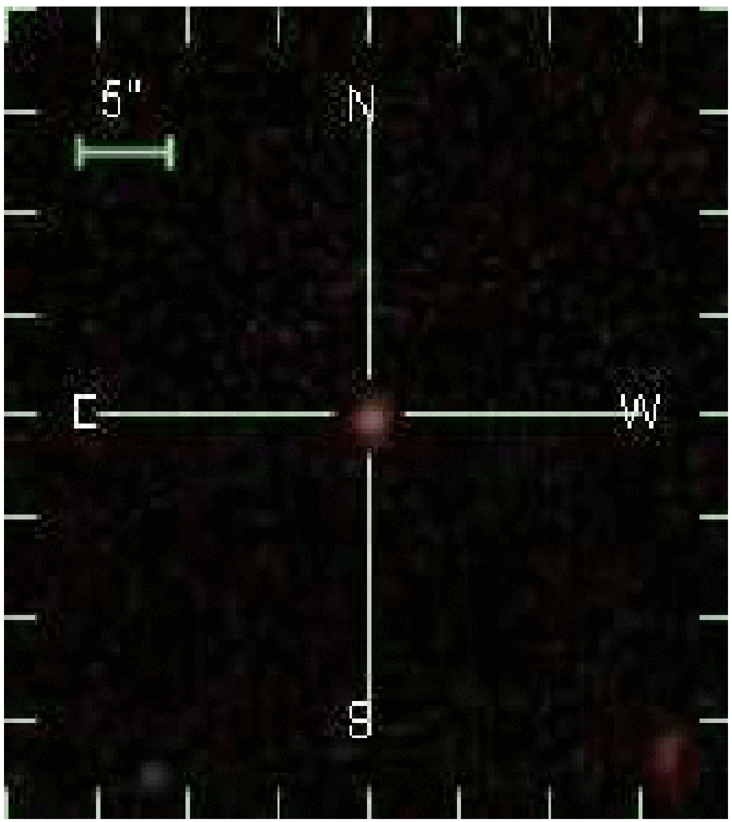
Incorrect Spectroscopic classifications

Object ID (objID):	Right ascension (ra):	Declination (dec):
587722982287802775	204.76416145	-0.71141558

Image Data
(What does it look like?)

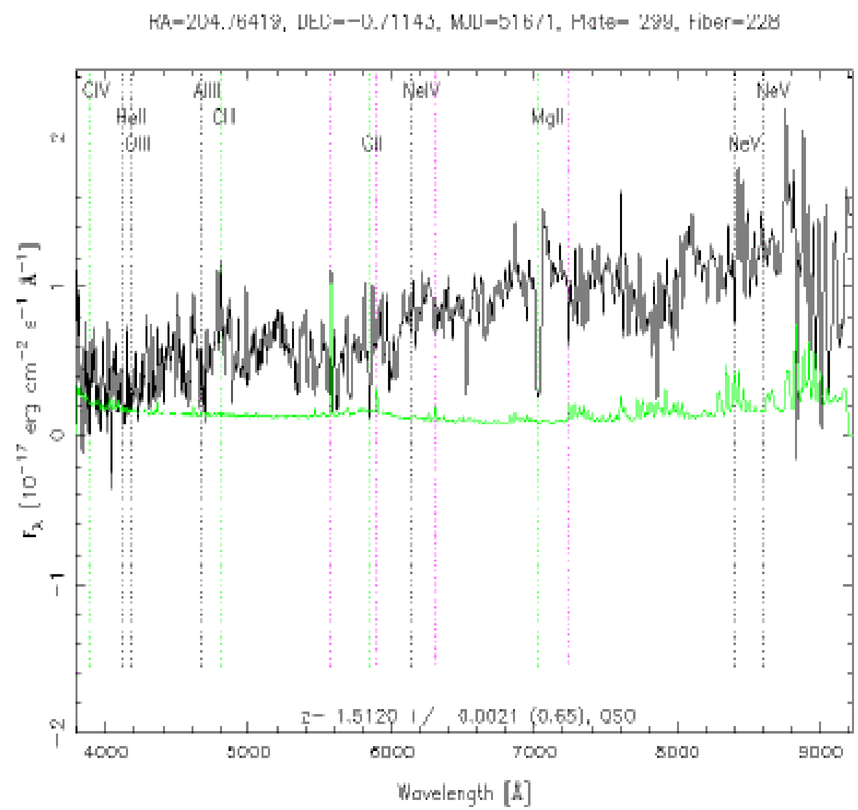
Spectrum Data
(What does its spectrum look like?)

iew image (click to go to Navigate tool)



Object Type (type): STAR

Preview spectrum (click for a larger version)



Spectral classification (specClass): QSO

What caused failures?

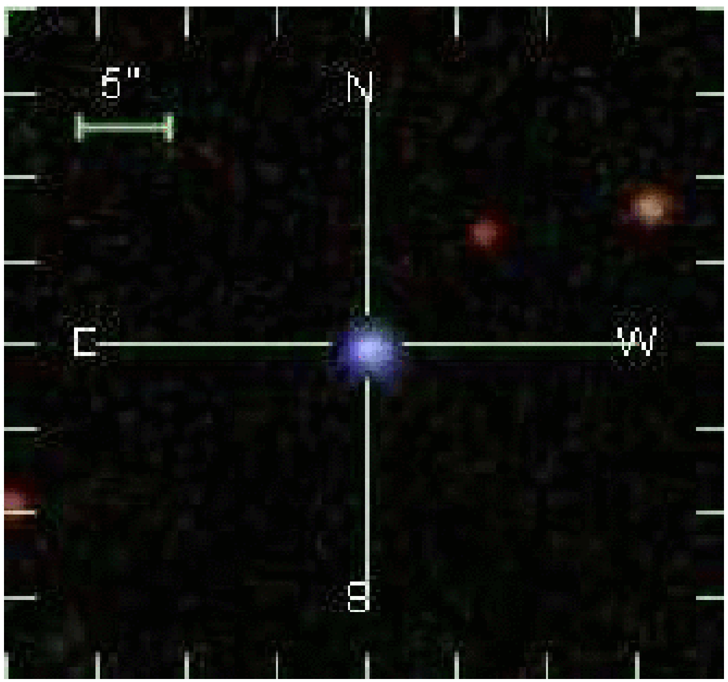
Object ID (objID): 587730773876604962

Right ascension (ra): 331.48760872

Declination (dec): 12.21102915

Image Data
(What does it look like?)

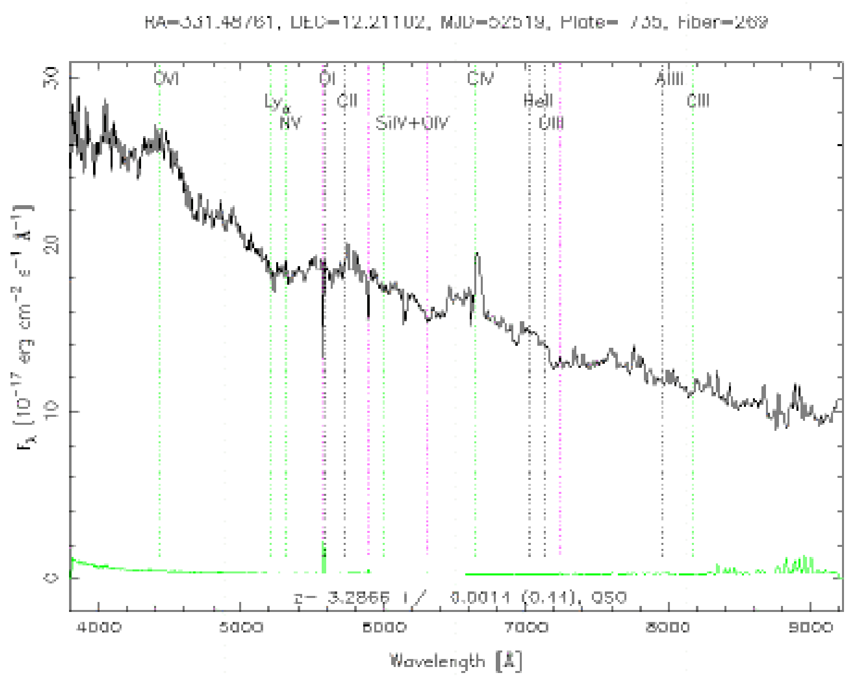
iew image (click to go to Navigate tool)



Object Type (type): STAR

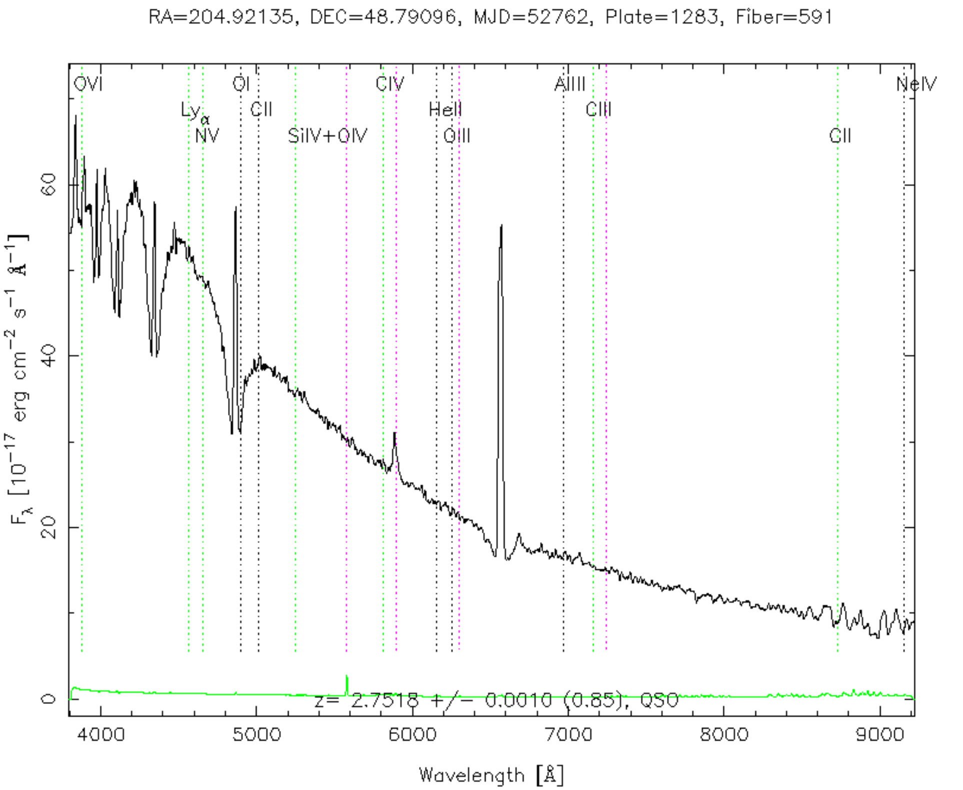
Spectrum Data
(What does its spectrum look like?)

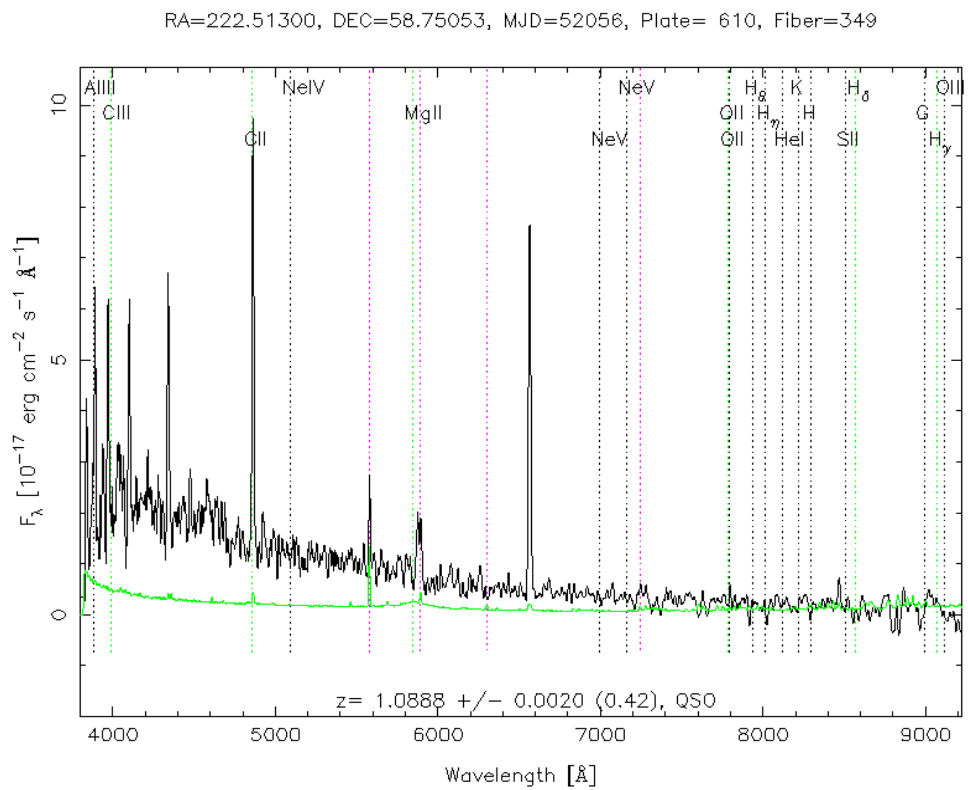
Preview spectrum (click for a larger version)

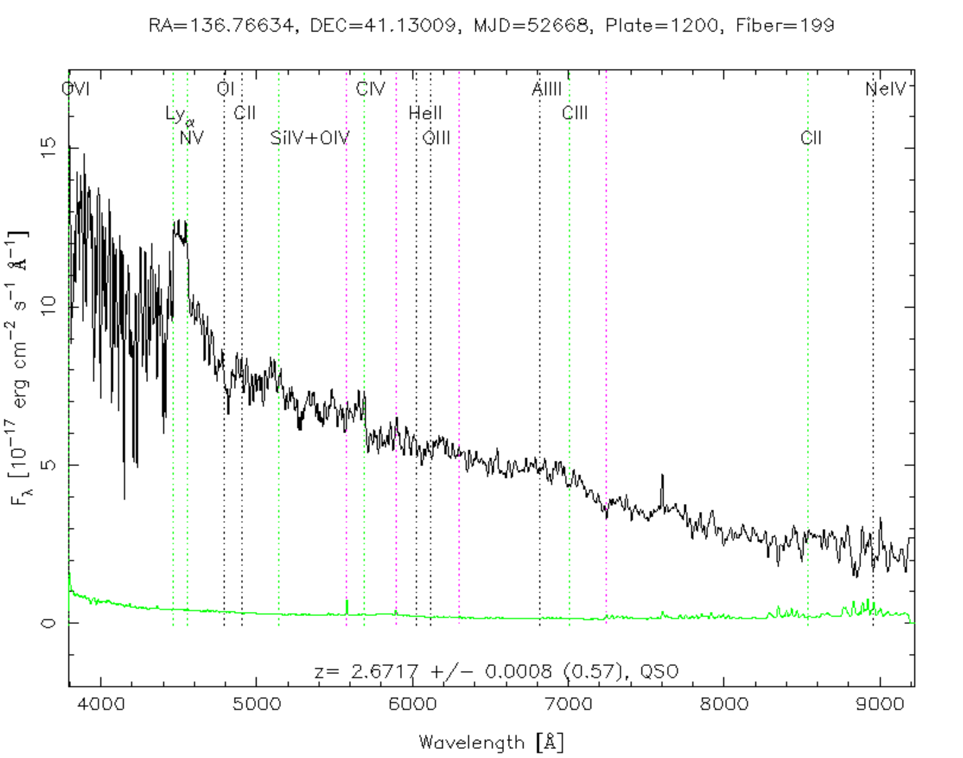


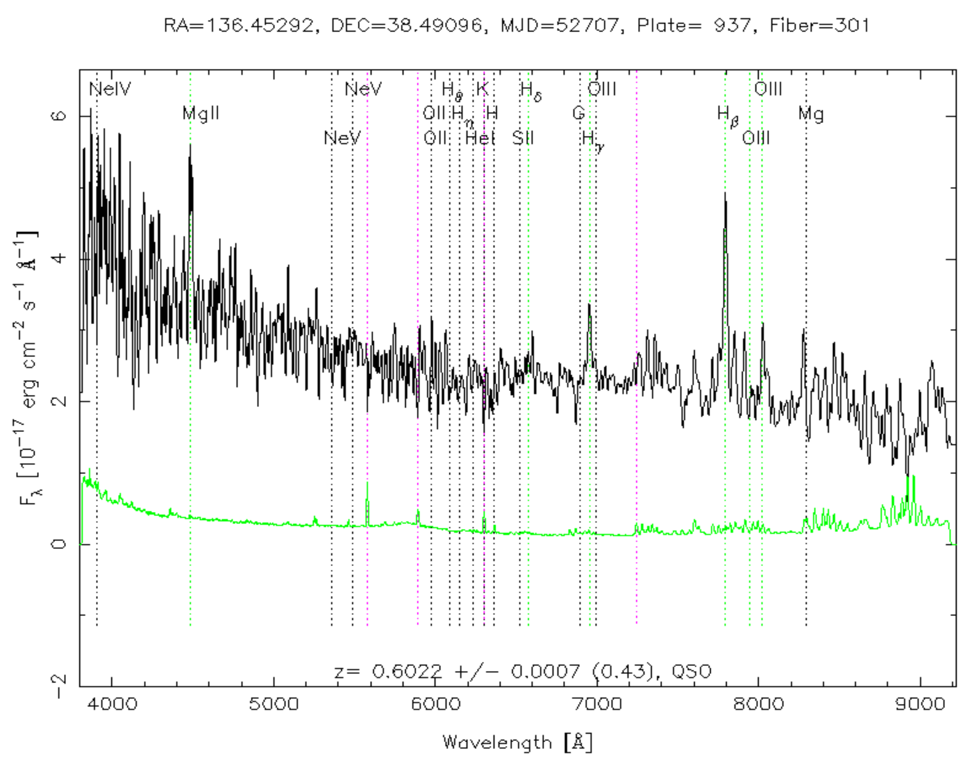
Spectral classification (specClass): QSO

Incorrect Spectroscopic Classification


RA=204.92135, DEC=48.79096, MJD=52762, Plate=1283, Fiber=591
OVI Lyα NV OI CII SiIV+OIV CIV HeII OIII AlIII CIII CII NeIV
z= 2.7518 +/- 0.0010 (0.85), QSO
Wavelength [Å]
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]


RA=222.51300, DEC=58.75053, MJD=52056, Plate= 610, Fiber=349
AlIII CIII CII NeIV MgII NeV NeV OII OII Hθ Hη HeI K H Hδ SII G Hγ OIII
z= 1.0888 +/- 0.0020 (0.42), QSO
Wavelength [Å]
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]


RA=136.76634, DEC=41.13009, MJD=52668, Plate=1200, Fiber=199
OVI Lyα NV OI CII SiIV+OIV CIV HeII OIII AlIII CIII CII NeIV
z= 2.6717 +/- 0.0008 (0.57), QSO
Wavelength [Å]
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]


RA=136.45292, DEC=38.49096, MJD=52707, Plate= 937, Fiber=301
NeIV MgII NeV NeV OII OII Hθ Hη K H HeI Hδ SII G Hγ OIII Hβ OIII OIII Mg
z= 0.6022 +/- 0.0007 (0.43), QSO
Wavelength [Å]
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]

Spectral Classification Statistics

Total Number of Failed Objects with Spectra :	1647
Objects with wrong classification labels :	283

How to reduce failures?

Define training samples with accurate labels

Increase the parameter space with multiwavelength observations

Build a hybrid spec-photo network

Usefulness of the method

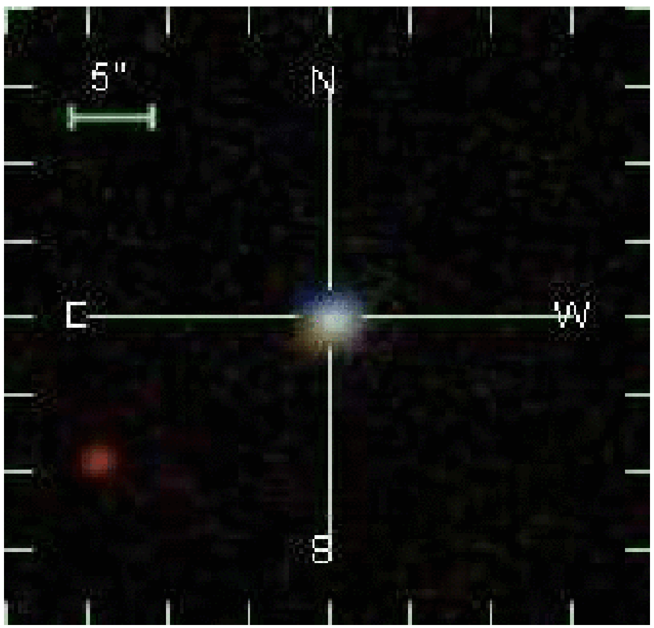
Object ID (objID): 587726033846075436

Right ascension (ra): 150.11433595

Declination (dec): 2.99756903

Image Data
(What does it look like?)

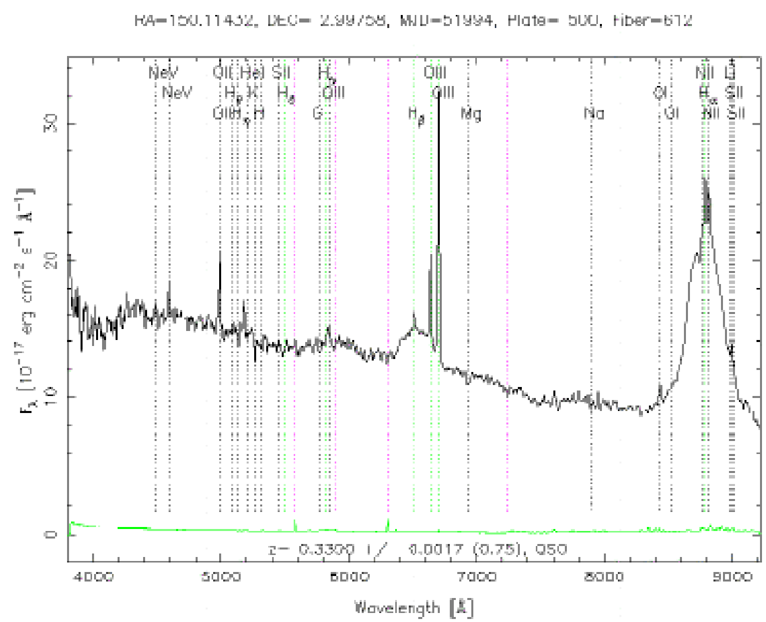
iew image (click to go to Navigate tool)



Object Type (type): STAR

Spectrum Data
(What does its spectrum look like?)

Preview spectrum (click for a larger version)



Spectral classification (specClass): QSO

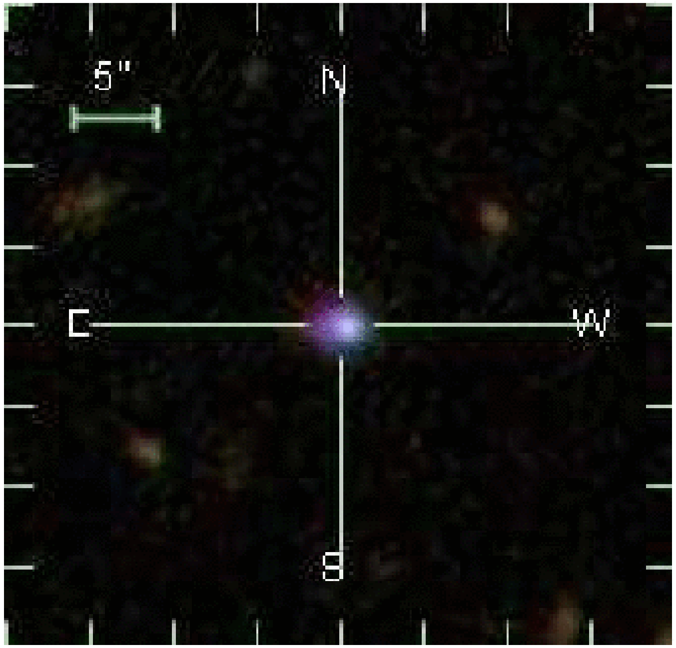
Outliers detection – objects with peculiar colors are automatically assigned low confidence values by DBNN.

Usefulness of the method

Object ID (objID):	Right ascension (ra):	Declination (dec):
588007005806854207	252.17182864	41.28061611

Image Data
(What does it look like?)

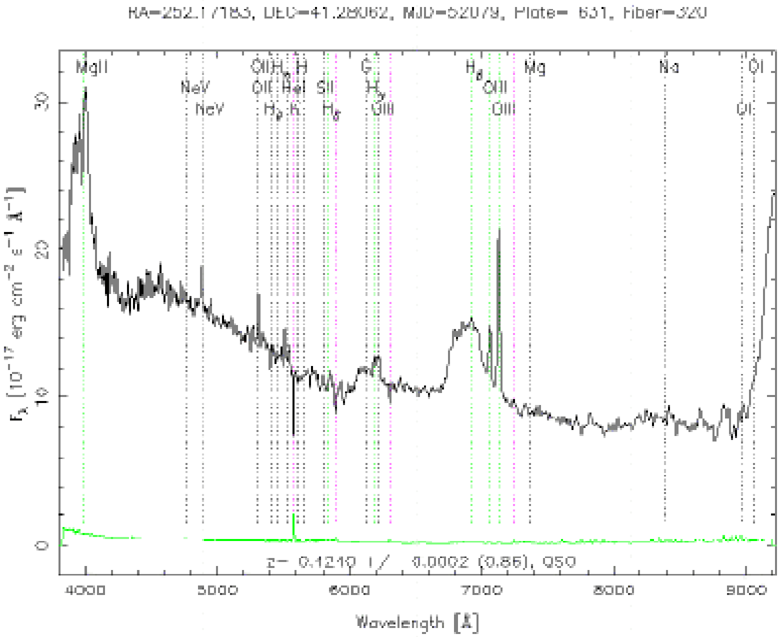
iew image (click to go to Navigate tool)



Object Type (type): STAR

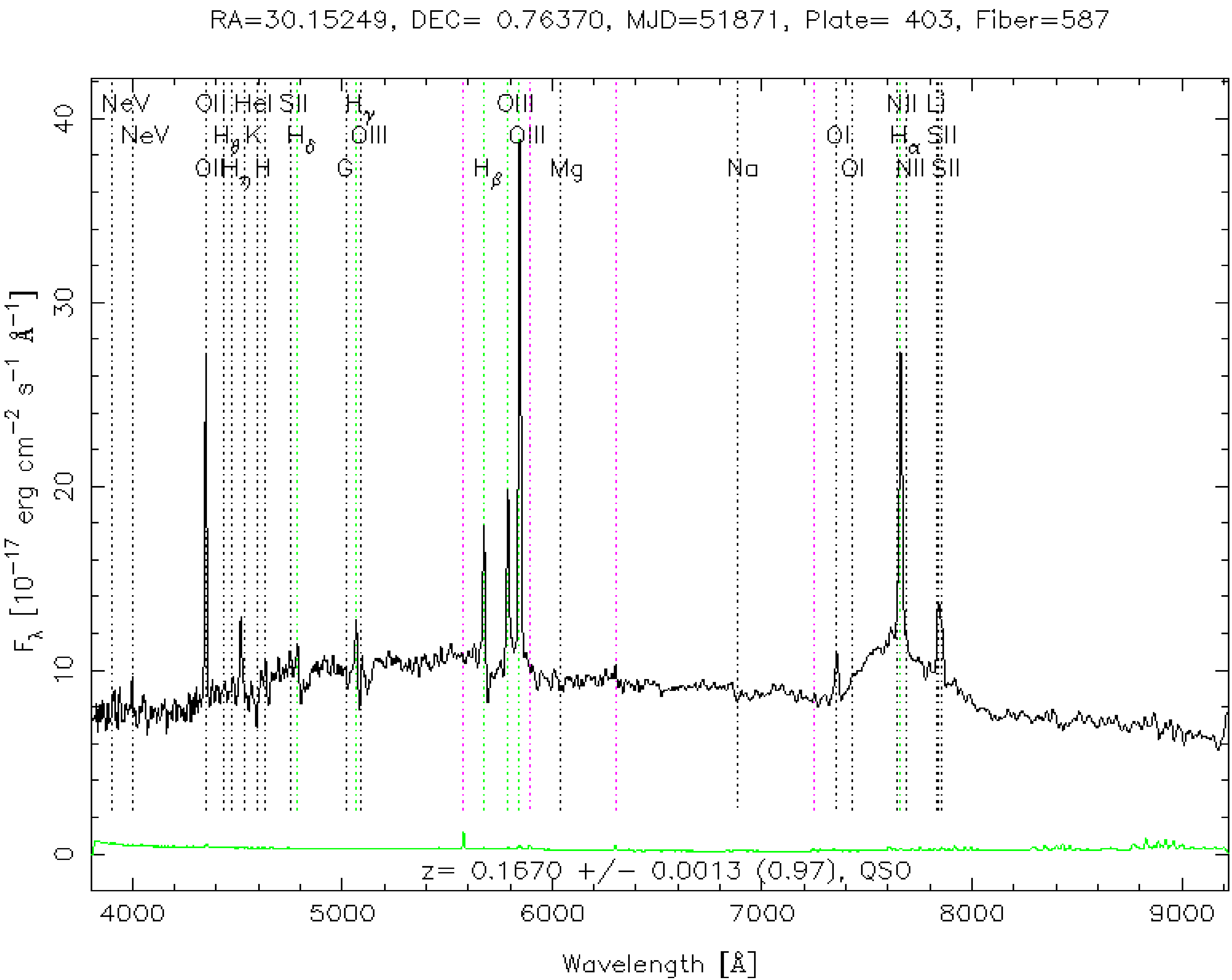
Spectrum Data
(What does its spectrum look like?)

Preview spectrum (click for a larger version)

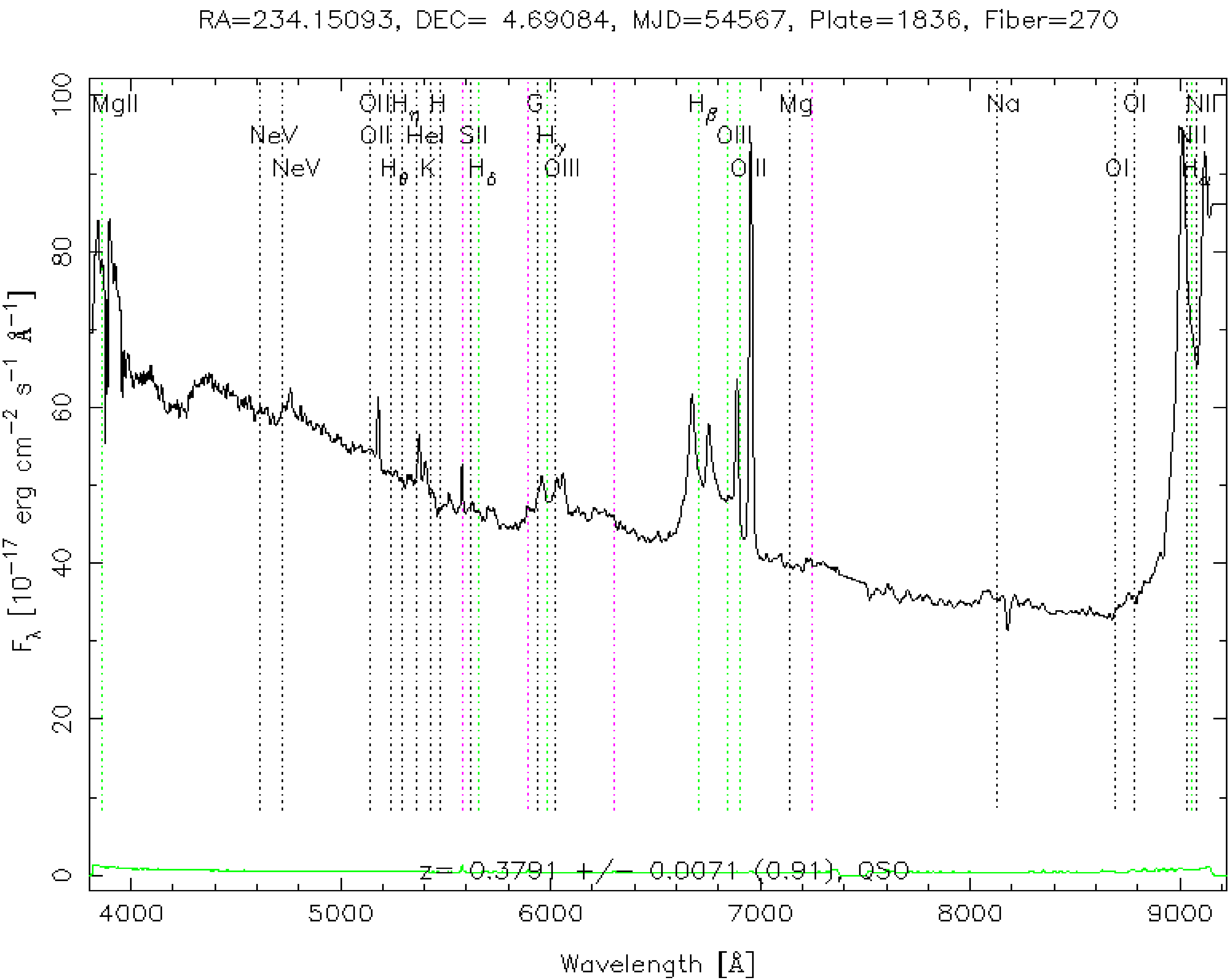


Spectral classification (specClass): QSO

Outliers detection – objects with peculiar colors are automatically assigned low confidence values by DBNN.



RA=30.15249, DEC= 0.76370, MJD=51871, Plate= 403, Fiber=587
NeV
NeV
OII
OII
HeI
SII
K
H
G
OIII
OIII
OIII
Mg
Na
OI
OI
NII
NII
Li
SII
SII
0
10
20
30
40
4000
5000
6000
7000
8000
9000
z= 0.1670 +/- 0.0013 (0.97), QSO
Wavelength [Å]
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]



RA=234.15093, DEC= 4.69084, MJD=54567, Plate=1836, Fiber=270
MgII
NeV
NeV
OII
OII
H_η
H_θ
H
HeI
K
SII
H_δ
G
H_γ
OIII
H_β
OIII
OIII
Mg
Na
OI
OI
NII
NII
H_α
z= 0.3791 +/- 0.0071 (0.91) QSO
F_λ [10^{-17} erg cm^{-2} s^{-1} Å^{-1}]
Wavelength [Å]
0
20
40
60
80
100
4000
5000
6000
7000
8000
9000

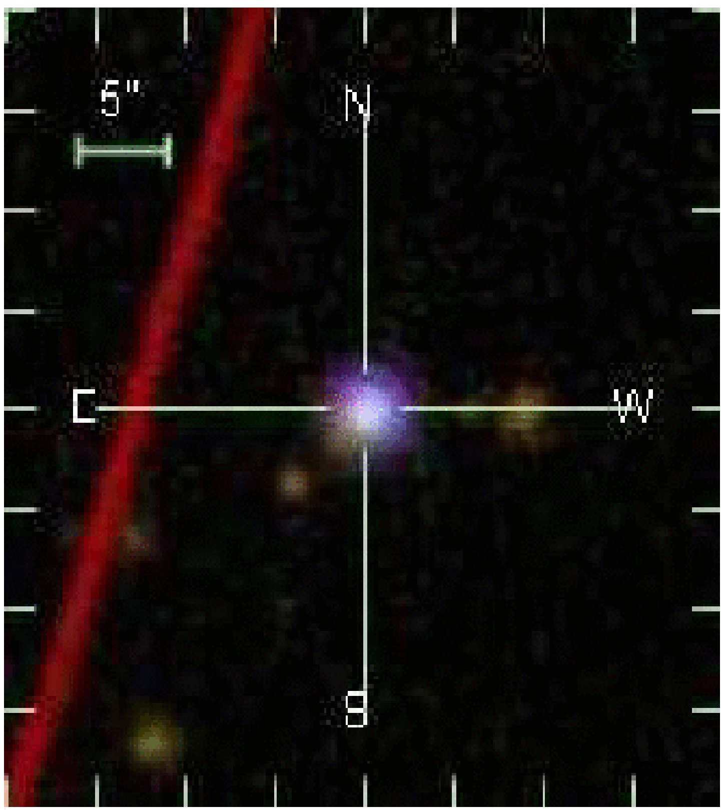
Object ID (objID): 587730022261129248

Right ascension (ra): 234.15093049

Declination (dec): 4.69083722

Image Data
(What does it look like?)

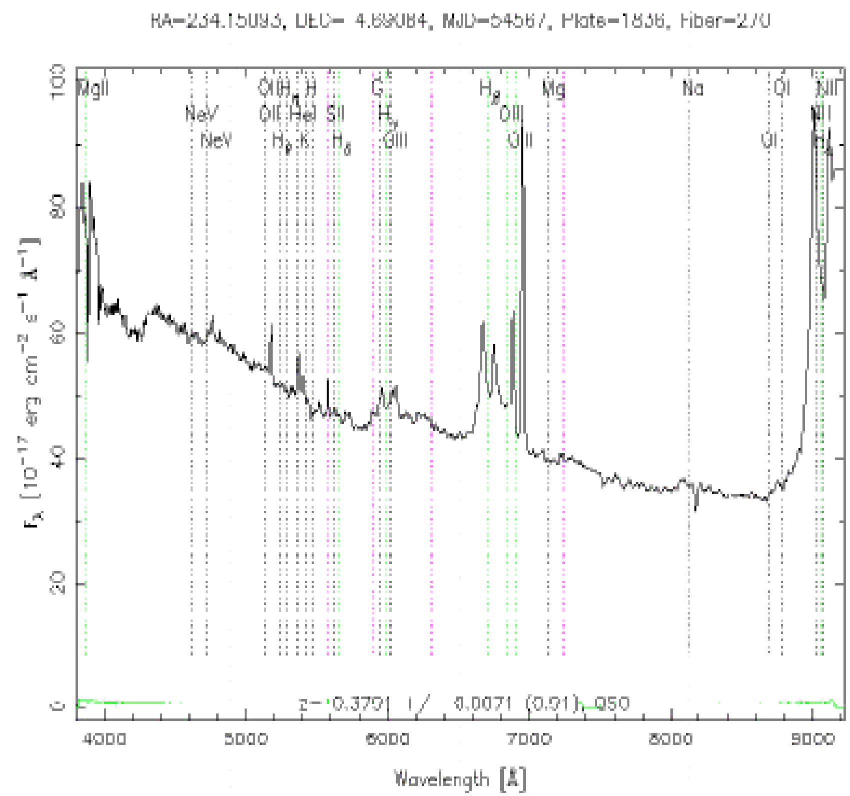
iew image (click to go to Navigate tool)



Object Type (type): STAR

Spectrum Data
(What does its spectrum look like?)

Preview spectrum (click for a larger version)



Spectral classification (specClass): QSO

Objects with Interesting Spectra

Objects with interesting spectra in the list of failed 1647 objects: 119

Usefulness of the method

Most suited for candidate selection for follow up observations

The assigned probability can be used to compromise between completeness and contamination on trained networks

Very fast – can classify a million objects in one second on a modern PC

Question Time

Thank You

nspp@iucaa.ernet.in

http://www.iucaa.ernet.in/~nspp