Discrete Source Classifier in CU8-Apsis

Kester Smith, MPIA Gaia team

Task of DSC is classification of all Gaia sources.

Supervised approach

Available data;

- BPRP photometry
- Astrometry: π , proper motions
- Variability
- RVS spectrum
- Possible morphology
- Position, magnitude

Task of DSC is classification of all Gaia sources.

Supervised approach

Available data;

- BPRP photometry
- Astrometry: π , proper motions
- Variability
- RVS spectrum
- Possible morphology
- Position, magnitude

Modular design with subclassifiers

- Modular design with subclassifiers
- 'Hierarchical' photometric classifier based on Support Vector Machines (SVM)

- Modular design with subclassifiers
- 'Hierarchical' photometric classifier based on Support Vector Machines (SVM)
- Astrometric subclassifier now based on Gaussian mixture model

- Modular design with subclassifiers
- 'Hierarchical' photometric classifier based on Support Vector Machines (SVM)
- Astrometric subclassifier now based on Gaussian mixture model
- Position G Mag subclassifier based on Kernel Density Estimator

$$P(C|D_1...D_N) = \frac{\prod_{n=1}^{n=N} P(C|D_n)}{P(C)^{N-1}}$$

[Bailer Jones & Smith, GAIA-C8-TN-MPIA-CBJ-053]







Grid	star	wd	bin	qso	gal	???
APec	94.44	_	_	5.55	0.00	0.00
Fastrot	98.26	_	_	0.69	0.00	1.04
Phoenix N	95.73	_	_	0.02	0.54	3.69
Phoenix R	98.67	_	_	0.15	0.85	0.33
SDSS Stars	99.58	_	_	0.20	0.21	0.00
UCD Cond	70.52	_	_	0.45	0.76	28.27
UCD Dust	98.30	_	_	0.00	0.00	1.70
WR	76.74	_	_	9.30	0.00	13.95
SDSS QSOs	0.20	_	_	95.85	2.33	1.60
SDSS galaxies	0.15	—	-	0.50	99.01	0.33
MARCS	82.05	2.80	5.75	4.05	1.90	3.45
0	78.80	11.00	0.40	5.80	1.20	2.80
В	86.20	5.80	1.00	4.40	0.80	1.80
А	89.10	3.20	0.90	3.60	1.00	2.20
Be	82.75	1.72	0.00	10.92	4.59	0.00
C stars	89.48	0.23	1.40	7.24	1.63	0.00
Binaries	29.27	1.60	57.40	4.67	2.32	4.72
WDA	18.20	58.42	0.52	16.10	1.95	4.80
WDB	15.67	58.20	0.55	19.37	1.95	4.25

Grid	star	wd	bin	qso	gal	???
APec	94.44	_	_	5.55	0.00	0.00
Fastrot	98.26	_	_	0.69	0.00	1.04
Phoenix N	95.73	_	_	0.02	0.54	3.69
Phoenix R	98.67	_	_	0.15	0.85	0.33
SDSS Stars	99.58	_	_	0.20	0.21	0.00
UCD Cond	70.52	_	_	0.45	0.76	28.27
UCD Dust	98.30	_	_	0.00	0.00	1.70
WR	76.74	_	_	9.30	0.00	13.95
SDSS QSOs	0.20	_	_	95.85	2.33	1.60
SDSS galaxies	0.15	_	_	0.50	99.01	0.33
MARCS	82.05	2.80	5.75	4.05	1.90	3.45
0	78.80	11.00	0.40	5.80	1.20	2.80
В	86.20	5.80	1.00	4.40	0.80	1.80
А	89.10	3.20	0.90	3.60	1.00	2.20
Be	82.75	1.72	0.00	10.92	4.59	0.00
C stars	89.48	0.23	1.40	7.24	1.63	0.00
Binaries	29.27	1.60	57.40	4.67	2.32	4.72
WDA	18.20	58.42	0.52	16.10	1.95	4.80
WDB	15.67	58.20	0.55	19.37	1.95	4.25

Grid	star	wd	bin	qso	gal	???
APec	94.44	_	_	5.55	0.00	0.00
Fastrot	98.26	_	_	0.69	0.00	1.04
Phoenix N	95.73	_	_	0.02	0.54	3.69
Phoenix R	98.67	_	_	0.15	0.85	0.33
SDSS Stars	99.58	—	—	0.20	0.21	0.00
UCD Cond	70.52	_	_	0.45	0.76	28.27
UCD Dust	98.30	_	_	0.00	0.00	1.70
WR	76.74	_	_	9.30	0.00	13.95
SDSS QSOs	0.20	—	—	95.85	2.33	1.60
SDSS galaxies	0.15	—	_	0.50	99.01	0.33
MARCS	82.05	2.80	5.75	4.05	1.90	3.45
0	78.80	11.00	0.40	5.80	1.20	2.80
В	86.20	5.80	1.00	4.40	0.80	1.80
А	89.10	3.20	0.90	3.60	1.00	2.20
Be	82.75	1.72	0.00	10.92	4.59	0.00
C stars	89.48	0.23	1.40	7.24	1.63	0.00
Binaries	29.27	1.60	57.40	4.67	2.32	4.72
WDA	18.20	58.42	0.52	16.10	1.95	4.80
WDB	15.67	58.20	0.55	19.37	1.95	4.25

Grid	star	wd	bin	qso	gal	???
APec	94.44	_	_	5.55	0.00	0.00
Fastrot	98.26	_	—	0.69	0.00	1.04
Phoenix N	95.73	_	—	0.02	0.54	3.69
Phoenix R	98.67	_	—	0.15	0.85	0.33
SDSS Stars	99.58	_	_	0.20	0.21	0.00
UCD Cond	70.52	_	_	0.45	0.76	28.27
UCD Dust	98.30	_	_	0.00	0.00	1.70
WR	76.74	_	_	9.30	0.00	13.95
SDSS QSOs	0.20	_	_	95.85	2.33	1.60
SDSS galaxies	0.15	—	_	0.50	99.01	0.33
MARCS	82.05	2.80	5.75	4.05	1.90	3.45
0	78.80	11.00	0.40	5.80	1.20	2.80
В	86.20	5.80	1.00	4.40	0.80	1.80
А	89.10	3.20	0.90	3.60	1.00	2.20
Be	82.75	1.72	0.00	10.92	4.59	0.00
C stars	89.48	0.23	1.40	7.24	1.63	0.00
Binaries	29.27	1.60	57.40	4.67	2.32	4.72
WDA	18.20	58.42	0.52	16.10	1.95	4.80
WDB	15.67	58.20	0.55	19.37	1.95	4.25

Hierarchical classification



Hierarchical classification



Hierarchical classification



























































Position



Class fractions



log(Fc)

Class fractions



log(Fc)

Class fractions



Summary

- Several different types of information to combine
- Large number of objects, vast majority of same class
- Small subgroups within classes. Challenge for class definition
- Must match noise characteristics in training model grid