Write RF info run (using no spaces!): **crcTME\_InitialCheck**

Input files:

S-phase\_subset.txt

G0\_subset.txt

G1\_subset.txt

apoptosis\_subset.txt

M-phase\_subset.txt

G2\_subset.txt

Downsampling taking place. Check output folder for more info

3 | G1\_subset.txt 563146

2 | G0\_subset.txt 307004

4 | apoptosis\_subset.txt 147238

1 | S-phase\_subset.txt 134501

6 | G2\_subset.txt 15281

5 | M-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

Working with 5306 cells per file\_origin

5 | M-phase\_subset.txt 5306

3 | G1\_subset.txt 5306

2 | G0\_subset.txt 5306

6 | G2\_subset.txt 5306

4 | apoptosis\_subset.txt 5306

1 | S-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

(30244, 5) (30244,)

(1592, 5) (1592,)

Feature ranking:

1. feature 2 (0.216383) -> 142Nd\_cCaspase 3

2. feature 3 (0.215195) -> 150Nd\_pRB

3. feature 1 (0.195091) -> 127I\_IdU

4. feature 4 (0.190953) -> 176Yb\_Cyclin B1

5. feature 0 (0.182379) -> 89Y\_pHH3

Score agains test data 1.0

**Predictions on original non-downsampled data: 0.8324589017138859**

precision recall f1-score support

0 1.00 1.00 1.00 263

1 1.00 1.00 1.00 268

2 1.00 1.00 1.00 243

3 1.00 1.00 1.00 270

4 1.00 1.00 1.00 259

5 1.00 1.00 1.00 289

accuracy 1.00 1592

macro avg 1.00 1.00 1.00 1592

weighted avg 1.00 1.00 1.00 1592

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
|  |  |  |  |  |
| **0** | **0.42** | **1.00** | **0.59** | **2421** |
| **1** | **0.94** | **0.83** | **0.88** | **17394** |
| **2** | **0.66** | **0.58** | **0.62** | **1680** |
| **3** | **0.99** | **1.00** | **1.00** | **113** |
| **4** | **1.00** | **0.83** | **0.91** | **6960** |
| **5** | **0.92** | **1.00** | **0.96** | **22** |
|  |  |  |  |  |

accuracy 0.83 28590

macro avg 0.82 0.87 0.83 28590

weighted avg 0.89 0.83 0.85 28590

Write RF info run (using no spaces!): **crcTME\_epi**

Input files:

S-phase\_subset.txt

G0\_subset.txt

G1\_subset.txt

apoptosis\_subset.txt

M-phase\_subset.txt

G2\_subset.txt

Downsampling taking place. Check output folder for more info

3 | G1\_subset.txt 563146

2 | G0\_subset.txt 307004

4 | apoptosis\_subset.txt 147238

1 | S-phase\_subset.txt 134501

6 | G2\_subset.txt 15281

5 | M-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

Working with 5306 cells per file\_origin

4 | apoptosis\_subset.txt 5306

3 | G1\_subset.txt 5306

5 | M-phase\_subset.txt 5306

1 | S-phase\_subset.txt 5306

6 | G2\_subset.txt 5306

2 | G0\_subset.txt 5306

Name: file\_origin, dtype: int64

(30244, 5) (30244,)

(1592, 5) (1592,)

Feature ranking:

1. feature 2 (0.217623) -> 142Nd\_cCaspase 3

2. feature 3 (0.214948) -> 150Nd\_pRB

3. feature 1 (0.194857) -> 127I\_IdU

4. feature 4 (0.190856) -> 176Yb\_Cyclin B1

5. feature 0 (0.181715) -> 89Y\_pHH3

Score agains test data 1.0

**Predictions on original non-downsampled data: 0.9056866622007218**

precision recall f1-score support

0 1.00 1.00 1.00 242

1 1.00 1.00 1.00 272

2 1.00 1.00 1.00 275

3 1.00 1.00 1.00 271

4 1.00 1.00 1.00 271

5 1.00 1.00 1.00 261

accuracy 1.00 1592

macro avg 1.00 1.00 1.00 1592

weighted avg 1.00 1.00 1.00 1592

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
|  |  |  |  |  |
| **0** | **0.43** | **1.00** | **0.60** | **10216** |
| **1** | **0.89** | **0.85** | **0.87** | **73435** |
| **2** | **0.68** | **0.90** | **0.77** | **44456** |
| **3** | **1.00** | **1.00** | **1.00** | **184881** |
| **4** | **1.00** | **0.77** | **0.87** | **108497** |
| **5** | **0.97** | **0.99** | **0.98** | **9928** |

accuracy 0.91 431413

macro avg 0.83 0.92 0.85 431413

weighted avg 0.93 0.91 0.91 431413

Write RF info run (using no spaces!): **crcTME\_macro**

Input files:

S-phase\_subset.txt

G0\_subset.txt

G1\_subset.txt

apoptosis\_subset.txt

M-phase\_subset.txt

G2\_subset.txt

Downsampling taking place. Check output folder for more info

3 | G1\_subset.txt 563146

2 | G0\_subset.txt 307004

4 | apoptosis\_subset.txt 147238

1 | S-phase\_subset.txt 134501

6 | G2\_subset.txt 15281

5 | M-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

Working with 5306 cells per file\_origin

5 | M-phase\_subset.txt 5306

1 | S-phase\_subset.txt 5306

6 | G2\_subset.txt 5306

2 | G0\_subset.txt 5306

3 | G1\_subset.txt 5306

4 | apoptosis\_subset.txt 5306

Name: file\_origin, dtype: int64

(30244, 5) (30244,)

(1592, 5) (1592,)

Feature ranking:

1. feature 3 (0.216792) -> 150Nd\_pRB

2. feature 2 (0.216503) -> 142Nd\_cCaspase 3

3. feature 1 (0.194688) -> 127I\_IdU

4. feature 4 (0.191273) -> 176Yb\_Cyclin B1

5. feature 0 (0.180744) -> 89Y\_pHH3

Score agains test data 1.0

**Predictions on original non-downsampled data: 0.7488388333850077**

precision recall f1-score support

0 1.00 1.00 1.00 288

1 1.00 1.00 1.00 238

2 1.00 1.00 1.00 256

3 1.00 1.00 1.00 268

4 1.00 1.00 1.00 264

5 1.00 1.00 1.00 278

accuracy 1.00 1592

macro avg 1.00 1.00 1.00 1592

weighted avg 1.00 1.00 1.00 1592

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
|  |  |  |  |  |
| **0** | **0.20** | **1.00** | **0.34** | **5289** |
| **1** | **0.80** | **0.76** | **0.78** | **80710** |
| **2** | **0.69** | **0.94** | **0.79** | **108498** |
| **3** | **0.99** | **1.00** | **0.99** | **2448** |
| **4** | **1.00** | **0.57** | **0.73** | **134768** |
| **5** | **0.92** | **0.97** | **0.95** | **496** |

accuracy 0.75 332209

macro avg 0.77 0.87 0.76 332209

weighted avg 0.84 0.75 0.76 332209

Write RF info run (using no spaces!): **crcTME\_fibro**

Input files:

S-phase\_subset.txt

G0\_subset.txt

G1\_subset.txt

apoptosis\_subset.txt

M-phase\_subset.txt

G2\_subset.txt

Downsampling taking place. Check output folder for more info

3 | G1\_subset.txt 563146

2 | G0\_subset.txt 307004

4 | apoptosis\_subset.txt 147238

1 | S-phase\_subset.txt 134501

6 | G2\_subset.txt 15281

5 | M-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

Working with 5306 cells per file\_origin

1 | S-phase\_subset.txt 5306

2 | G0\_subset.txt 5306

4 | apoptosis\_subset.txt 5306

5 | M-phase\_subset.txt 5306

6 | G2\_subset.txt 5306

3 | G1\_subset.txt 5306

Name: file\_origin, dtype: int64

(30244, 5) (30244,)

(1592, 5) (1592,)

Feature ranking:

1. feature 2 (0.215210) -> 142Nd\_cCaspase 3

2. feature 3 (0.213684) -> 150Nd\_pRB

3. feature 1 (0.195764) -> 127I\_IdU

4. feature 4 (0.191285) -> 176Yb\_Cyclin B1

5. feature 0 (0.184058) -> 89Y\_pHH3

Score agains test data 1.0

**Predictions on original non-downsampled data: 0.8702930544625112**

precision recall f1-score support

0 1.00 1.00 1.00 282

1 1.00 1.00 1.00 263

2 1.00 1.00 1.00 271

3 1.00 1.00 1.00 264

4 1.00 1.00 1.00 261

5 1.00 1.00 1.00 251

accuracy 1.00 1592

macro avg 1.00 1.00 1.00 1592

weighted avg 1.00 1.00 1.00 1592

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
|  |  |  |  |  |
| **0** | **0.38** | **1.00** | **0.55** | **1522** |
| **1** | **0.85** | **0.79** | **0.82** | **10838** |
| **2** | **0.86** | **0.97** | **0.91** | **35840** |
| **3** | **1.00** | **1.00** | **1.00** | **4931** |
| **4** | **1.00** | **0.70** | **0.83** | **21372** |
| **5** | **0.97** | **0.97** | **0.97** | **466** |

accuracy 0.87 74969

macro avg 0.84 0.91 0.85 74969

weighted avg 0.90 0.87 0.87 74969

Write RF info run (using no spaces!): **crcTME\_ALLcells**

Input files:

S-phase\_subset.txt

G0\_subset.txt

G1\_subset.txt

apoptosis\_subset.txt

M-phase\_subset.txt

G2\_subset.txt

Downsampling taking place. Check output folder for more info

3 | G1\_subset.txt 563146

2 | G0\_subset.txt 307004

4 | apoptosis\_subset.txt 147238

1 | S-phase\_subset.txt 134501

6 | G2\_subset.txt 15281

5 | M-phase\_subset.txt 5306

Name: file\_origin, dtype: int64

Working with 5306 cells per file\_origin

5 | M-phase\_subset.txt 5306

1 | S-phase\_subset.txt 5306

2 | G0\_subset.txt 5306

4 | apoptosis\_subset.txt 5306

6 | G2\_subset.txt 5306

3 | G1\_subset.txt 5306

Name: file\_origin, dtype: int64

(30244, 5) (30244,)

(1592, 5) (1592,)

Feature ranking:

1. feature 2 (0.216165) -> 142Nd\_cCaspase 3

2. feature 3 (0.214662) -> 150Nd\_pRB

3. feature 1 (0.194152) -> 127I\_IdU

4. feature 4 (0.191166) -> 176Yb\_Cyclin B1

5. feature 0 (0.183856) -> 89Y\_pHH3

Score agains test data 1.0

**Predictions on original non-downsampled data: 0.8410679341896109**

precision recall f1-score support

0 1.00 1.00 1.00 266

1 1.00 1.00 1.00 266

2 1.00 1.00 1.00 260

3 1.00 1.00 1.00 288

4 1.00 1.00 1.00 248

5 1.00 1.00 1.00 264

accuracy 1.00 1592

macro avg 1.00 1.00 1.00 1592

weighted avg 1.00 1.00 1.00 1592

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
|  |  |  |  |  |
| **0** | **0.32** | **1.00** | **0.48** | **17027** |
| **1** | **0.84** | **0.80** | **0.82** | **165054** |
| **2** | **0.71** | **0.94** | **0.81** | **189006** |
| **3** | **1.00** | **1.00** | **1.00** | **192445** |
| **4** | **1.00** | **0.67** | **0.80** | **264116** |
| **5** | **0.98** | **0.99** | **0.98** | **10943** |

accuracy 0.84 838591

macro avg 0.81 0.90 0.82 838591

weighted avg 0.89 0.84 0.85 838591