

## F class

F class max = 0.269

<https://machinelearningmastery.com/feature-selection-machine-learning-python/>

### F\_class 3 個 fea

```
GB accuracy
0.5803782505910166
GB CFmap
[[443 49 8 41 22 80]
 [ 6 28 20 16 15 18]
 [ 1 7 5 6 3 5]
 [ 2 10 3 6 3 4]
 [ 2 3 3 5 5 8]
 [ 1 6 2 3 3 4]]
15  Satisfication_Score  GB f1 score
19  Tenure_in_Months    0.25465866969105316
35  Contract
```

### F\_class 5 個 fea

```
GB accuracy
0.6122931442080378
GB CFmap
[[466 65 13 38 22 39]
 [ 3 33 15 8 24 20]
 [ 2 6 4 5 6 4]
 [ 2 9 1 6 6 4]
 [ 1 6 5 5 5 4]
 [ 1 7 2 1 4 4]]
15  Satisfication_Score  GB f1 score
18  Number_of_Referrals
19  Tenure_in_Months
35  Contract              0.26095757890384785
42  Total_Long_Distance_Charges
```

### F\_class 8 個 fea

```
GB accuracy
0.6252955082742316
GB CFmap
[[476 46 12 31 35 43]
 [ 5 33 11 13 22 19]
 [ 2 7 4 5 3 6]
 [ 2 9 2 3 9 3]
 [ 0 5 3 6 6 6]
 [ 0 3 0 2 7 7]]
7   Number_of_Dependents
15  Satisfication_Score
18  Number_of_Referrals
19  Tenure_in_Months
35  Contract
39  Total_Charges
42  Total_Long_Distance_Charges
43  Total_Revenue
GB f1 score
0.2691947434082061
```

### F\_class 10 個 fea

```
GB accuracy
0.6158392434988179
GB CFmap
[[477 31 15 33 31 56]
 [ 5 26 14 14 23 21]
 [ 3 6 6 5 2 5]
 [ 2 8 3 2 9 4]
 [ 1 6 3 3 6 7]
 [ 0 2 3 4 6 4]]
4   Senior_Citizen
6   Dependents
7   Number_of_Dependents
15  Satisfication_Score
18  Number_of_Referrals
19  Tenure_in_Months
35  Contract
39  Total_Charges
42  Total_Long_Distance_Charges
43  Total_Revenue
GB f1 score
0.25477398138605084
```

### F\_class 20 個 fea

2	Age	
4	Senior_Citizen	
6	Dependents	
7	Number_of_Dependents	
15	Satisfaction_Score	
17	Referred_a_Friend	
18	Number_of_Referrals	
19	Tenure_in_Months	
20	Offer	GB accuracy
21	Phone_Service	0.6264775413711584
24	Internet_Service	GB CFmap
25	Internet_Type	[[493 21 28 33 36 32]
27	Online_Security	[ 4 21 21 20 20 17]
30	Premium_Tech_Support	[ 2 7 4 6 3 5]
35	Contract	[ 2 8 5 2 6 5]
36	Paperless_Billing	[ 1 5 3 2 8 7]
38	Monthly_Charge	[ 1 2 5 4 5 2]]
39	Total_Charges	GB f1 score
42	Total_Long_Distance_Charges	0.23997181400270784
43	Total_Revenue	

## LogisticRegression

### sklearn.feature\_selection.SelectFromModel

LogisticRegression max = 0.155

[https://scikit-](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectFromModel.html)

[learn.org/stable/modules/generated/sklearn.feature\\_selection.SelectFromModel.ht](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectFromModel.html)  
ml

```
GB accuracy
0.34515366430260047
GB CFmap
[[260 68 88 79 72 76]
 [ 16 18 17 19 20 13]
 [  6  5  5  4  2  5]
 [  6  7  4  3  6  2]
 [ 10  4  4  1  3  4]
 [  0  3  3  5  5  3]]
10 City
12 Lat_Long
14 Longitude
38 Monthly_Charge
39 Total_Charges
42 Total_Long_Distance_Charges
43 Total_Revenue
GB f1 score
0.15506166978661637
```

## RFE

### sklearn.feature\_selection.RFE ¶

RFE max = 0.2848

[https://scikit-](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html#sklearn.feature_selection.RFE)

[learn.org/stable/modules/generated/sklearn.feature\\_selection.RFE.html#sklearn.fea](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html#sklearn.feature_selection.RFE)  
[ture\\_selection.RFE](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html#sklearn.feature_selection.RFE)

n = 5

```
GB accuracy
0.5874704491725768
GB CFmap
[[463 19 42 41 48 30]
 [  3 16 28 19 21 16]
 [  1  3  7  8  4  4]
 [  0  6  6  5  6  5]
 [  3  4  7  1  5  6]
 [  0  2  8  6  2  1]]
4 Senior_Citizen
7 Number_of_Dependents
15 Satisfaction_Score
32 Streaming_Movies
35 Contract
GB f1 score
0.22661760268980635
```

N = 10

```
GB accuracy
0.6382978723404256
GB CFmap
[[487 20 23 35 24 54]
 [  5 28 16 20 15 19]
 [  2  7  3  8  3  4]
 [  1  3  1  9  3 11]
 [  3  5  1  7  7  3]
 [  1  2  6  4  0  6]]
4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
9 State
15 Satisfaction_Score
27 Online_Security
30 Premium_Tech_Support
32 Streaming_Movies
35 Contract
36 Paperless_Billing
GB f1 score
0.2848737826695433
```

N = 12

```

4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
9 State
15 Satisfication_Score
16 Quarter
23 Multiple_Lines
27 Online_Security
30 Premium_Tech_Support
32 Streaming_Movies
35 Contract
36 Paperless_Billing

GB accuracy
0.6158392434988179
GB CFmap
[[475 19 25 38 32 54]
 [ 5 30 21 14 17 16]
 [ 1 3 1 9 6 7]
 [ 1 5 3 5 5 9]
 [ 2 4 4 5 5 6]
 [ 1 4 4 3 2 5]]
GB f1 score
0.2524406407377126

```

N = 15

```

4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
9 State
15 Satisfication_Score
16 Quarter
21 Phone_Service
23 Multiple_Lines
27 Online_Security
30 Premium_Tech_Support
31 Streaming_TV
32 Streaming_Movies
34 Unlimited_Data
35 Contract
36 Paperless_Billing

GB accuracy
0.6170212765957447
GB CFmap
[[472 28 25 31 44 43]
 [ 8 29 18 16 19 13]
 [ 2 3 3 10 4 5]
 [ 1 4 5 9 4 5]
 [ 1 4 5 3 6 7]
 [ 1 4 5 5 1 3]]
GB f1 score
0.2652504867652091

```

N = 20

```

4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
9 State
15 Satisfication_Score
16 Quarter
17 Referred_a_Friend
18 Number_of_Referrals
21 Phone_Service
23 Multiple_Lines
25 Internet_Type
27 Online_Security
29 Device_Protection_Plan
30 Premium_Tech_Support
31 Streaming_TV
32 Streaming_Movies
33 Streaming_Music
34 Unlimited_Data
35 Contract
36 Paperless_Billing

GB accuracy
0.6323877068557919
GB CFmap
[[487 26 25 33 28 44]
 [ 4 24 25 18 13 19]
 [ 1 6 6 6 1 7]
 [ 1 10 3 5 3 6]
 [ 0 4 8 3 8 3]
 [ 0 4 3 6 1 5]]
GB f1 score
0.27506493778870206

```

嵌入式 (embedded) Max = 0.269

SelectFromModel 選取特征

<https://www.itread01.com/content/1527535241.html>

1	Gender	
2	Age	
3	Under_30	
4	Senior_Citizen	
5	Married	
6	Dependents	
7	Number_of_Dependents	
10	City	
12	Lat_Long	
13	Latitude	
14	Longitude	
15	Satisfaction_Score	
16	Quarter	
17	Referred_a_Friend	
18	Number_of_Referrals	
19	Tenure_in_Months	
20	Offer	
21	Phone_Service	
22	Avg_Monthly_Long_Distance_Charges	
23	Multiple_Lines	
24	Internet_Service	
25	Internet_Type	
26	Avg_Monthly_GB_Download	
27	Online_Security	
28	Online_Backup	
30	Premium_Tech_Support	
32	Streaming_Movies	GB accuracy
34	Unlimited_Data	0.6252955082742316
35	Contract	GB CFmap
36	Paperless_Billing	[[477 34 28 44 39 21]
37	Payment_Method	[ 2 31 16 16 19 19]
38	Monthly_Charge	[ 0 5 6 11 3 2]
39	Total_Charges	[ 1 9 4 6 5 3]
40	Total_Refunds	[ 1 5 4 5 6 5]
41	Total_Extra_Data_Charges	[ 0 4 2 4 6 3]]
42	Total_Long_Distance_Charges	GB f1 score
43	Total_Revenue	0.2692123198894921

---

## KNeighbors

`sklearn.feature_selection.SequentialFeatureSelector`

KNeighbors **max = 0.30** -> **kaggle:0.27**

– K = 10

[https://scikit-](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html#sklearn.feature_selection.SequentialFeatureSelector)

[learn.org/stable/modules/generated/sklearn.feature\\_selection.SequentialFeatureSelector.html#sklearn.feature\\_selection.SequentialFeatureSelector](https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html#sklearn.feature_selection.SequentialFeatureSelector)

```
0 Count GB accuracy
3 Under_30 0.6028368794326241
7 Number_of_Dependents GB CFmap
9 State [[461 34 24 47 22 55]
15 Satisfaction_Score [ 4 28 25 16 16 14]
16 Quarter [ 0 7 6 6 4 4]
21 Phone_Service [ 2 7 0 6 4 9]
25 Internet_Type [ 3 4 5 4 6 4]
35 Contract [ 0 6 3 7 0 3]]
40 Total_Refunds GB f1 score
0.26211437182969793
```

– K = 15

```
0 Count
4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
15 Satisfaction_Score GB accuracy
16 Quarter 0.6288416075650118
17 Referred_a_Friend GB CFmap
18 Number_of_Referrals [[479 17 32 42 27 46]
27 Online_Security [ 3 23 23 16 19 19]
30 Premium_Tech_Support [ 2 4 5 8 6 2]
31 Streaming_TV [ 1 4 3 9 4 7]
34 Unlimited_Data [ 1 1 5 2 12 5]
35 Contract [ 0 3 3 7 2 4]]
36 Paperless_Billing GB f1 score
40 Total_Refunds 0.2894165329684166
```

K = 18 //pred(4)

```
0 Count
4 Senior_Citizen
6 Dependents
7 Number_of_Dependents
8 Country
9 State
15 Satisfaction_Score
16 Quarter
17 Referred_a_Friend GB accuracy
24 Internet_Service 0.6382978723404256
27 Online_Security GB CFmap
28 Online_Backup [[483 28 20 37 32 43]
29 Device_Protection_Plan [ 5 26 24 11 12 25]
30 Premium_Tech_Support [ 3 8 5 5 2 4]
32 Streaming_Movies [ 1 3 3 12 6 3]
35 Contract [ 3 4 7 1 8 3]
36 Paperless_Billing [ 1 2 4 4 2 6]]
40 Total_Refunds GB f1 score
0.3002158434124354
```

Name	Submitted	Wait time	Execution time	Score
pred (4).csv	just now	1 seconds	0 seconds	0.27477

Complete

K = 20

0	Count	
3	Under_30	
7	Number_of_Dependents	
8	Country	
9	State	
15	Satisfaction_Score	
16	Quarter	
17	Referred_a_Friend	
18	Number_of_Referrals	
23	Multiple_Lines	GB accuracy
27	Online_Security	0.6323877068557919
28	Online_Backup	GB CFmap
29	Device_Protection_Plan	[[494 34 17 32 29 37]
30	Premium_Tech_Support	[ 6 19 25 14 23 16]
31	Streaming_TV	[ 1 5 4 6 4 7]
32	Streaming_Movies	[ 1 8 2 8 4 5]
33	Streaming_Music	[ 1 5 6 1 5 8]
35	Contract	[ 1 2 4 5 2 5]]
36	Paperless_Billing	GB f1 score
40	Total_Refunds	0.25870611261495713

## Tree-based

---

### 1.13.4.2. Tree-based feature selection

[https://scikit-learn.org/stable/modules/feature\\_selection.html#univariate-feature-selection](https://scikit-learn.org/stable/modules/feature_selection.html#univariate-feature-selection)

`n_estimator = 50`

```
2  Age
10 City
11 Zip_Code GB accuracy
12 Lat_Long 0.5957446808510638
13 Latitude GB CFmap
14 Longitude [[455 34 35 44 45 30]
15 Satisfication_Score [ 7 23 16 18 20 19]
19 Tenure_in_Months [ 1 3 8 9 2 4]
25 Internet_Type [ 2 8 4 4 6 4]
35 Contract [ 1 2 6 3 10 4]
38 Monthly_Charge [ 1 4 3 3 4 4]]
43 Total_Revenue GB f1 score
0.264491121363186
```

`n_estimator = 40`

```
2  Age
6  Dependents
10 City
11 Zip_Code GB accuracy
12 Lat_Long 0.607565011820331
13 Latitude GB CFmap
14 Longitude [[463 28 39 51 36 26]
15 Satisfication_Score [ 7 27 16 17 18 18]
19 Tenure_in_Months [ 2 5 8 7 2 3]
35 Contract [ 3 7 5 5 5 3]
38 Monthly_Charge [ 0 3 6 6 7 4]
39 Total_Charges [ 0 4 4 5 2 4]]
43 Total_Revenue GB f1 score
0.2699155845526327
```

`n_estimator = 30`

```
2  Age
10 City
11 Zip_Code GB accuracy
12 Lat_Long 0.607565011820331
13 Latitude GB CFmap
14 Longitude [[465 30 35 37 43 33]
15 Satisfication_Score [ 6 26 16 18 19 18]
18 Number_of_Referrals [ 0 3 8 9 3 4]
19 Tenure_in_Months [ 2 9 4 3 5 5]
25 Internet_Type [ 0 3 6 4 8 5]
35 Contract [ 1 6 1 3 4 4]]
38 Monthly_Charge GB f1 score
43 Total_Revenue 0.2640508997238404
```

`n_estimator = 20`



```

2 Age
4 Senior_Citizen
10 City
11 Zip_Code
12 Lat_Long
13 Latitude
14 Longitude
15 Satisfication_Score
19 Tenure_in_Months
35 Contract
38 Monthly_Charge
39 Total_Charges
42 Total_Long_Distance_Charges
43 Total_Revenue

GB accuracy
0.5756501182033097
GB CFmap
[[440 30 47 54 41 31]
 [ 8 23 15 21 20 16]
 [ 1 6 6 10 3 1]
 [ 3 11 5 4 4 1]
 [ 2 3 3 4 9 5]
 [ 0 4 3 3 4 5]]
GB f1 score
0.2553160739826658

```

**n\_estimator = 10**

```

2 Age
4 Senior_Citizen
10 City
11 Zip_Code
12 Lat_Long
13 Latitude
14 Longitude
15 Satisfication_Score
18 Number_of_Referrals
19 Tenure_in_Months
35 Contract
38 Monthly_Charge
43 Total_Revenue

GB accuracy
0.6182033096926713
GB CFmap
[[472 28 36 35 39 33]
 [ 6 25 14 17 21 20]
 [ 2 4 7 10 2 2]
 [ 2 11 4 3 4 4]
 [ 0 3 5 4 11 3]
 [ 0 6 1 4 3 5]]
GB f1 score
0.2748428732124128

```

