GPS 室内外判别
GPS indoor / outdoor classification

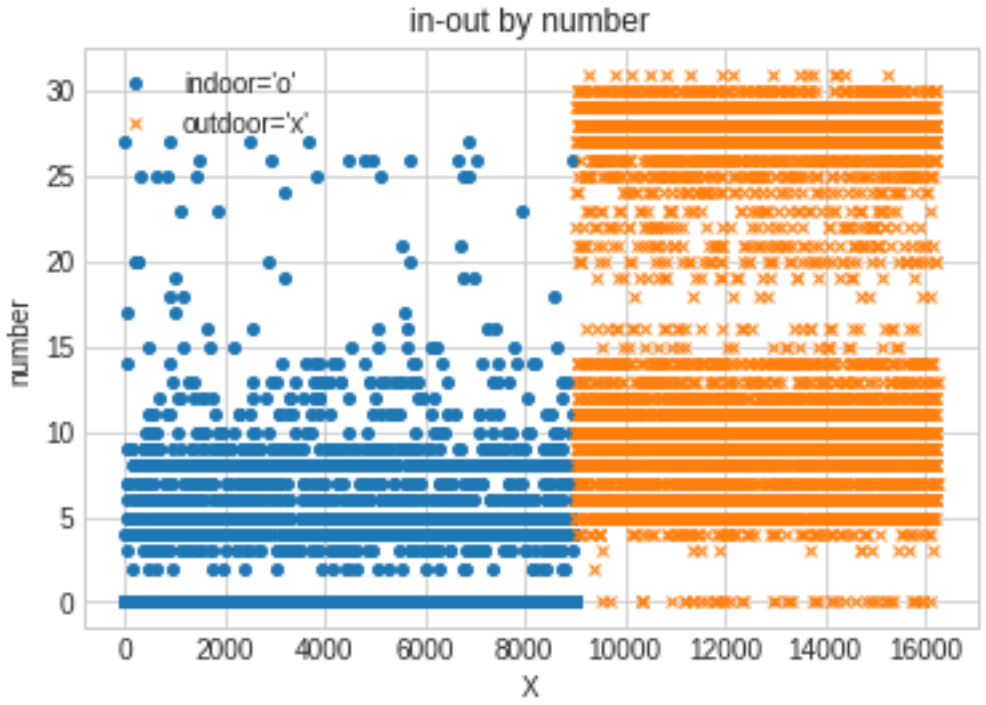
- 特征分析 feature analysis
- 模型 调试 model trials

- Notice: for this project, training data & testing data collected by different people. So there is a need to compare the features between them before deciding which model will be suitable.

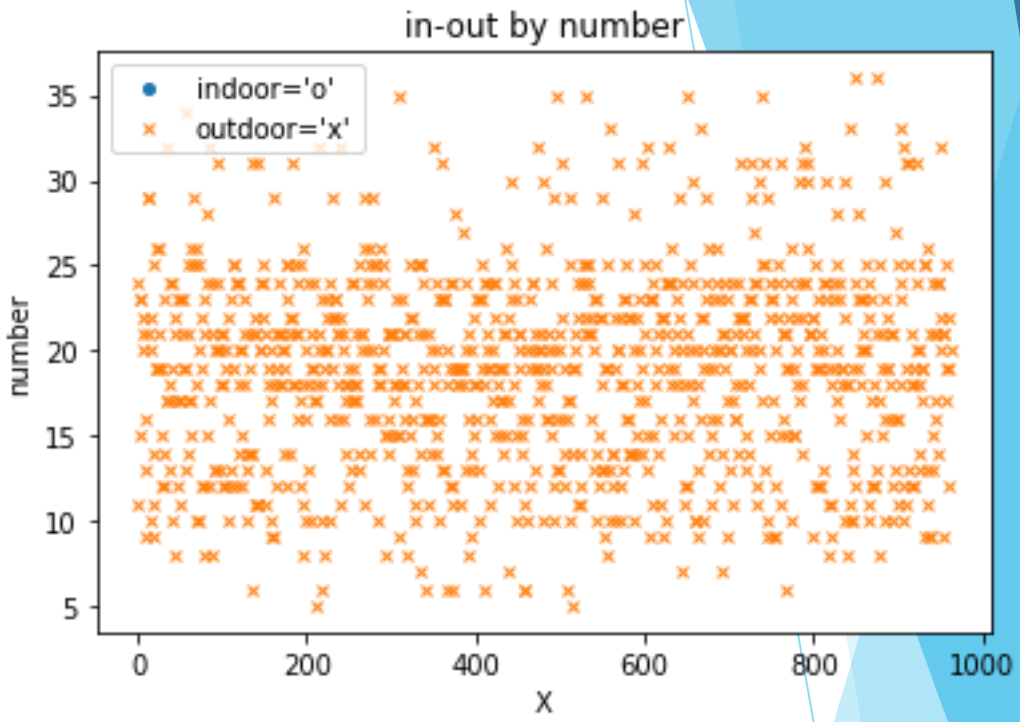
特征分析 - 采集的数据对比
feature analysis, comparisons between training data & testing data

- 12 个 特征的值域， 分布对比 compare the features between training data & testing data
- 地铁室外跑 组 the testing data when subway-train running outdoors
- 地铁站内 组 the testing data when in the subway stations (indoors)
- 其他 组 other main scenes such as streets, communities, stores, buildings, etc.

地铁室外跑 the testing data when subway-train running outdoors

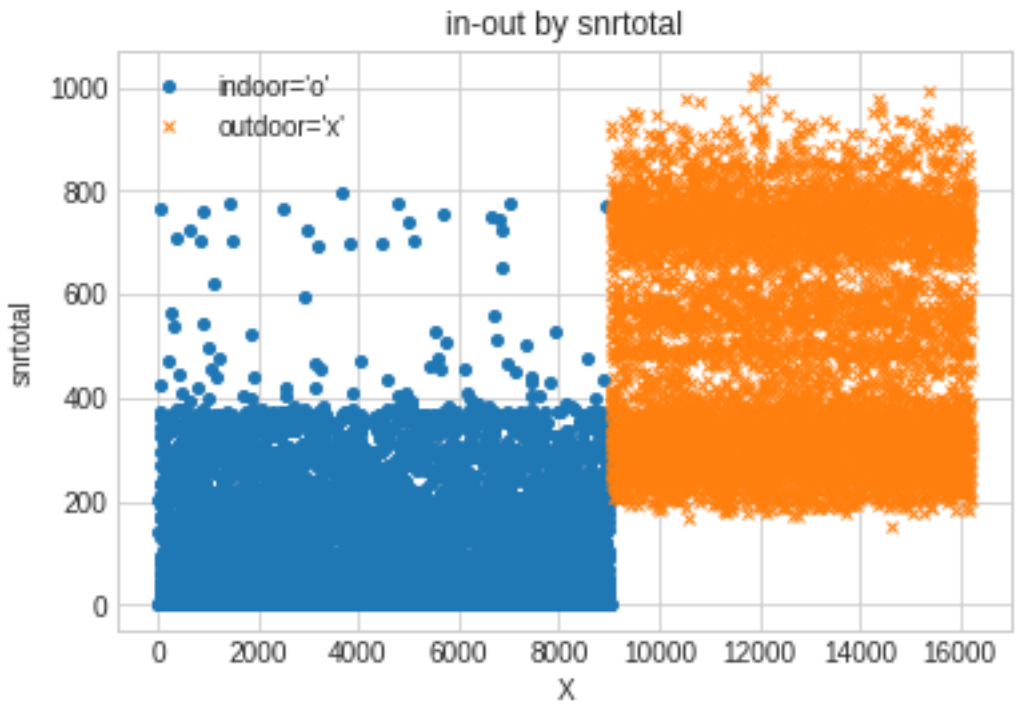


Training data

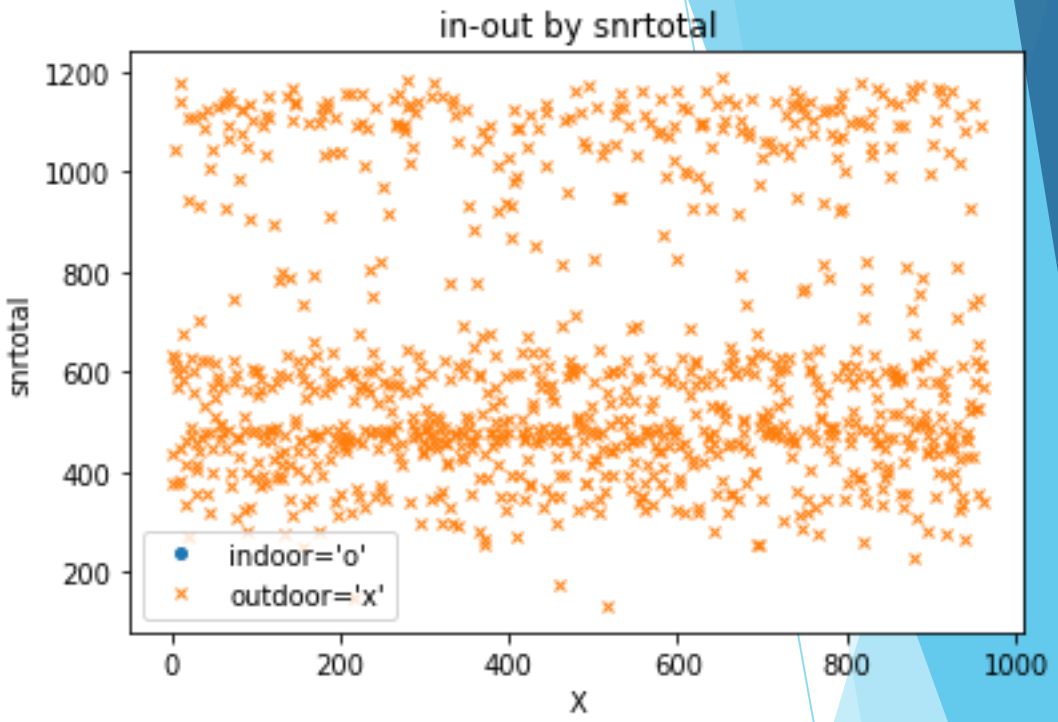


Testing data

特征 1 / feature 1
number: 手机将卫星用于定位的数量



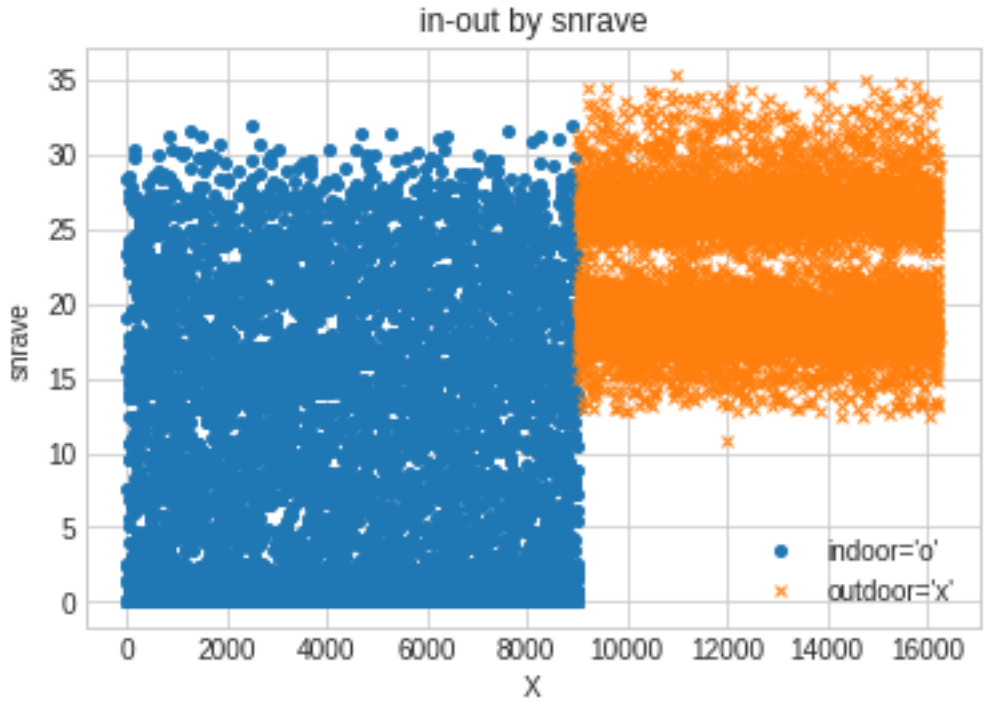
Training data



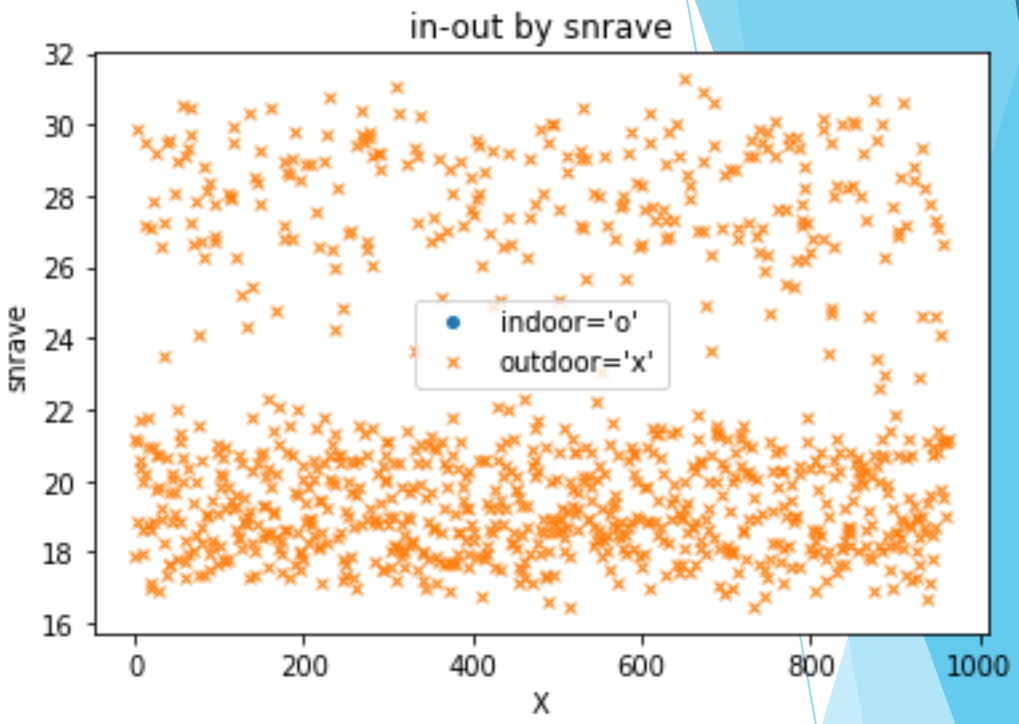
Testing data

特征 2 / feature 2

snrtotal :所有能检测到卫星的信噪比总和



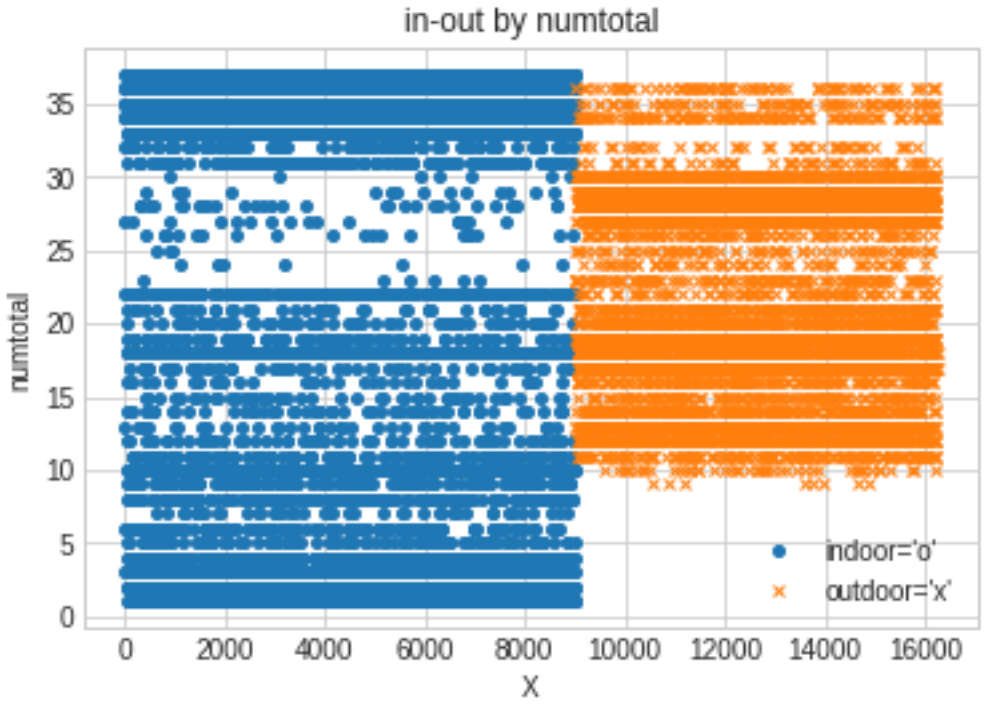
Training data



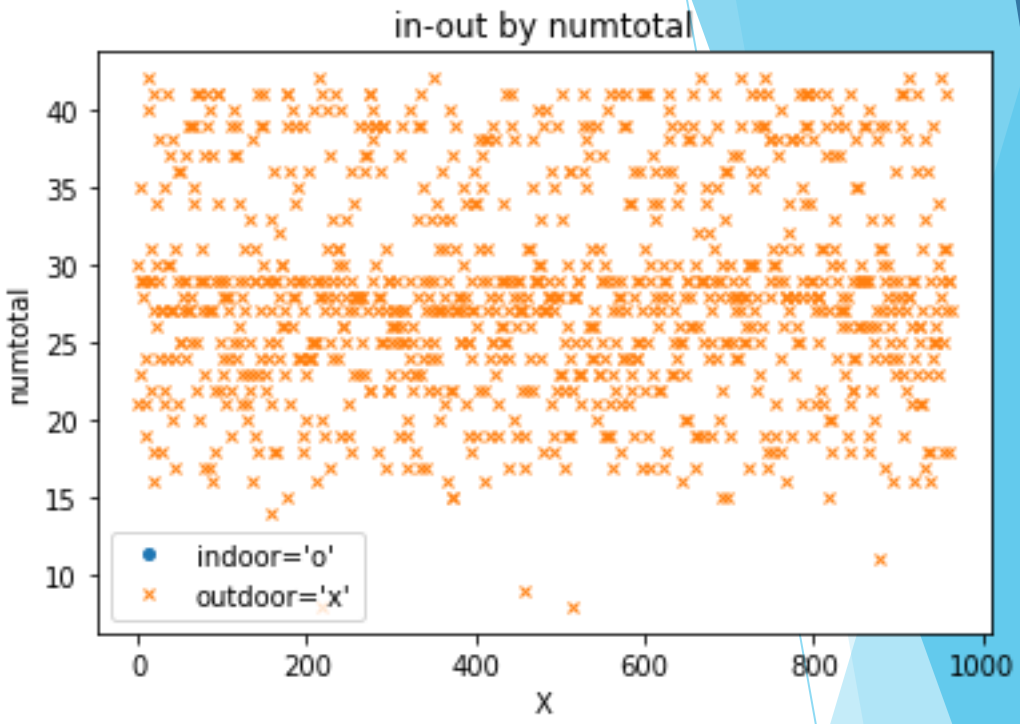
Testing data

特征 3 / feature 3

snrave :所有能检测到卫星的信噪比的平均值



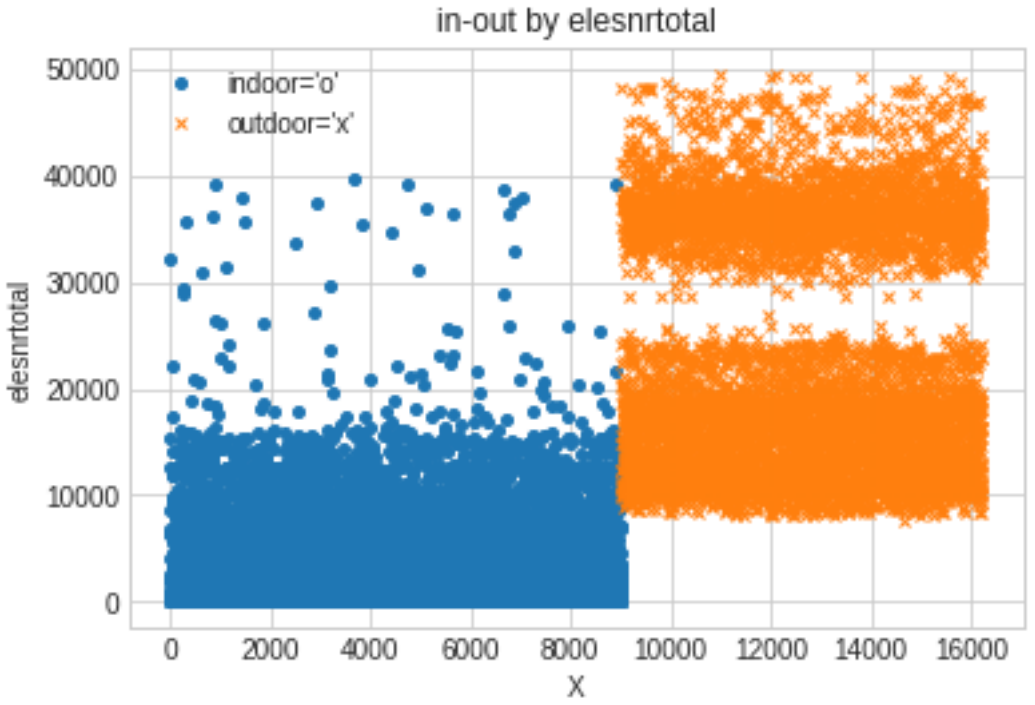
Training data



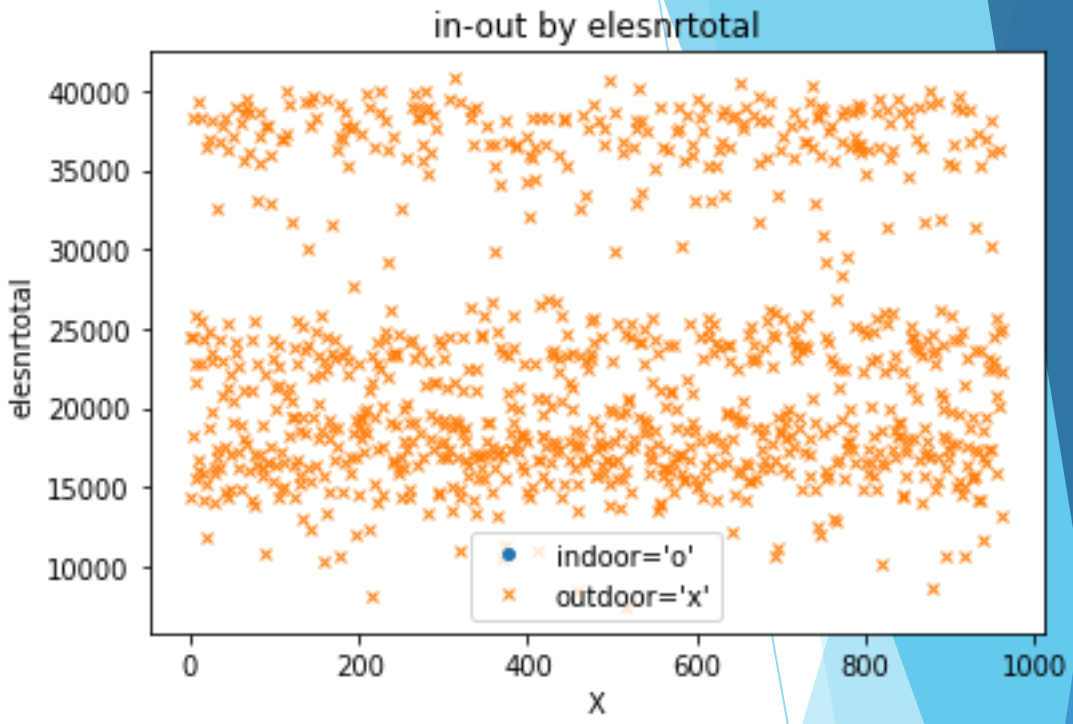
Testing data

特征 4 / feature 4

numtotal：当前手机所能检测的所有卫星数量 (包括信号强度低的卫星)



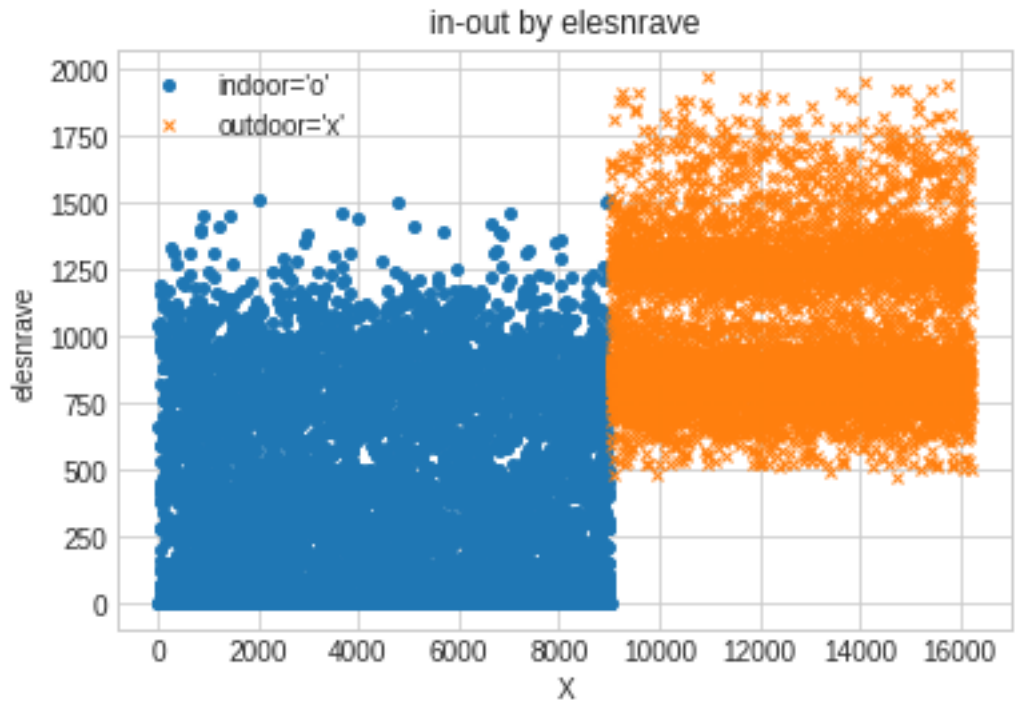
Training data



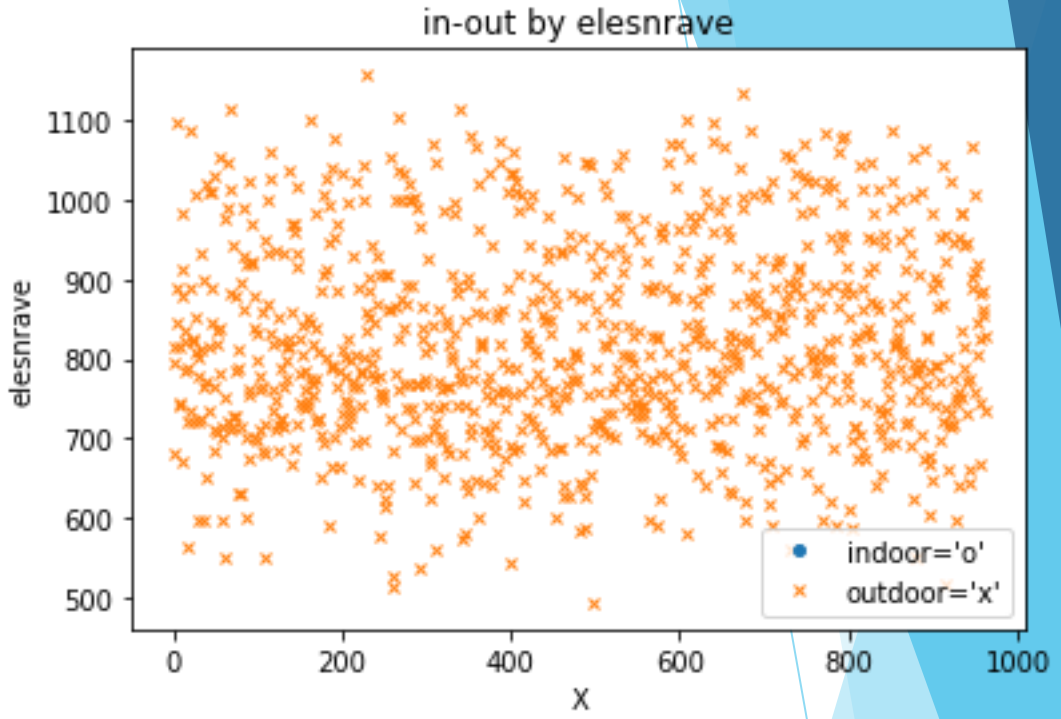
Testing data

特征 5 / feature 5

elesnrtotal :所有能检测到卫星的俯仰角和信噪比 乘积的总和



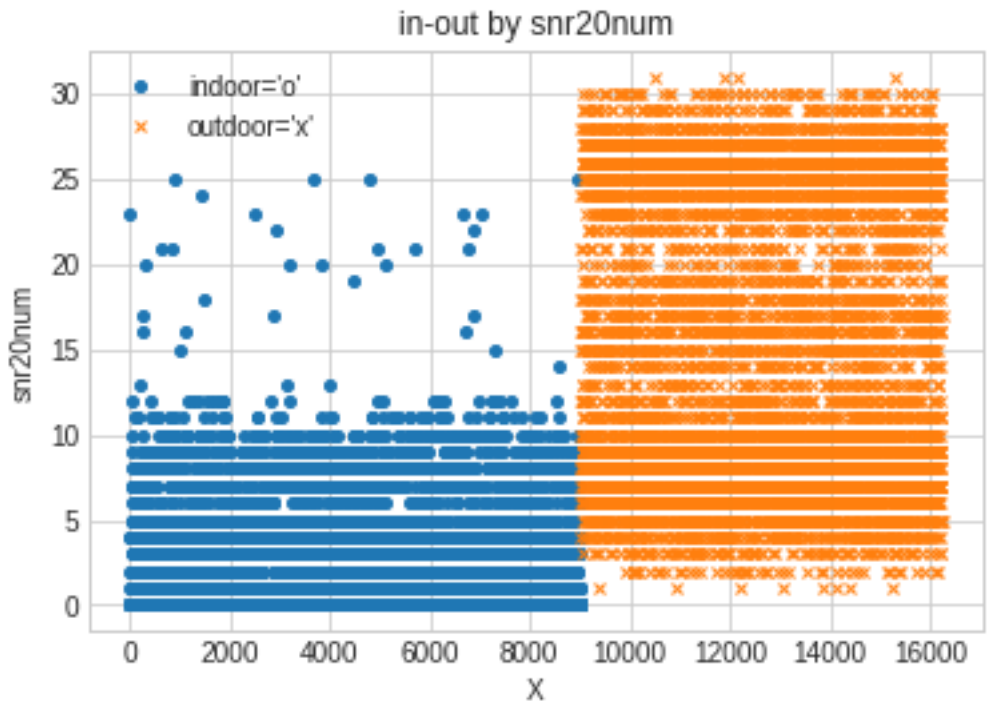
Training data



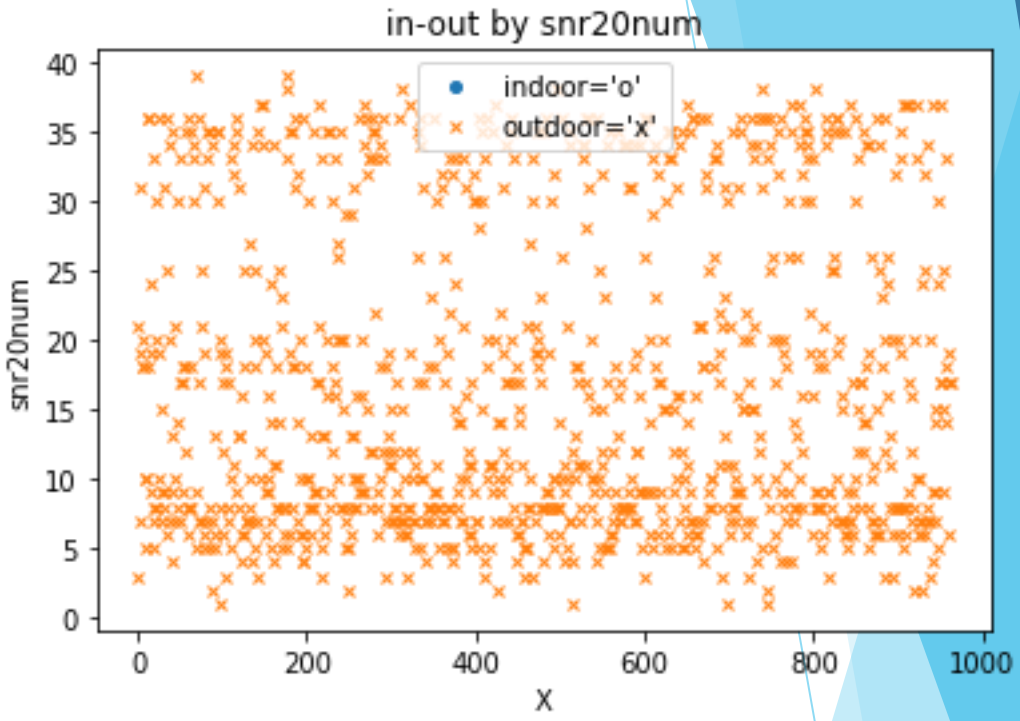
Testing data

特征 6 / feature 6

elesnrave :所有能检测到卫星的俯仰角和信噪比 乘积的平均值



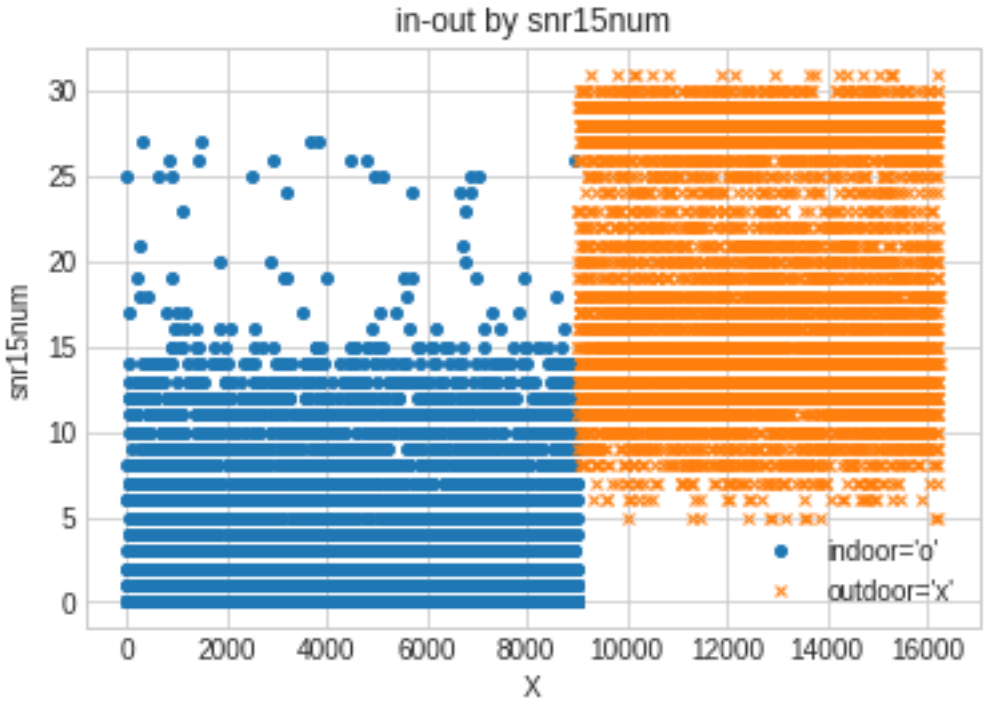
Training data



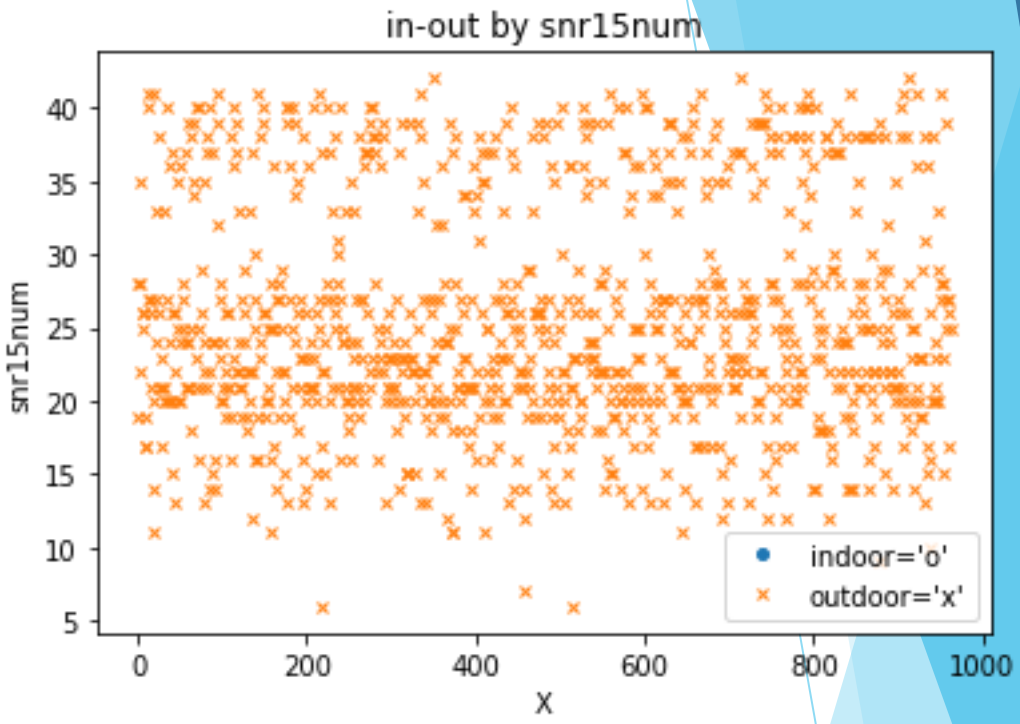
Testing data

特征 7 / feature 7

snr20num：手机检测到信噪比大于等于 20 的卫星 数量



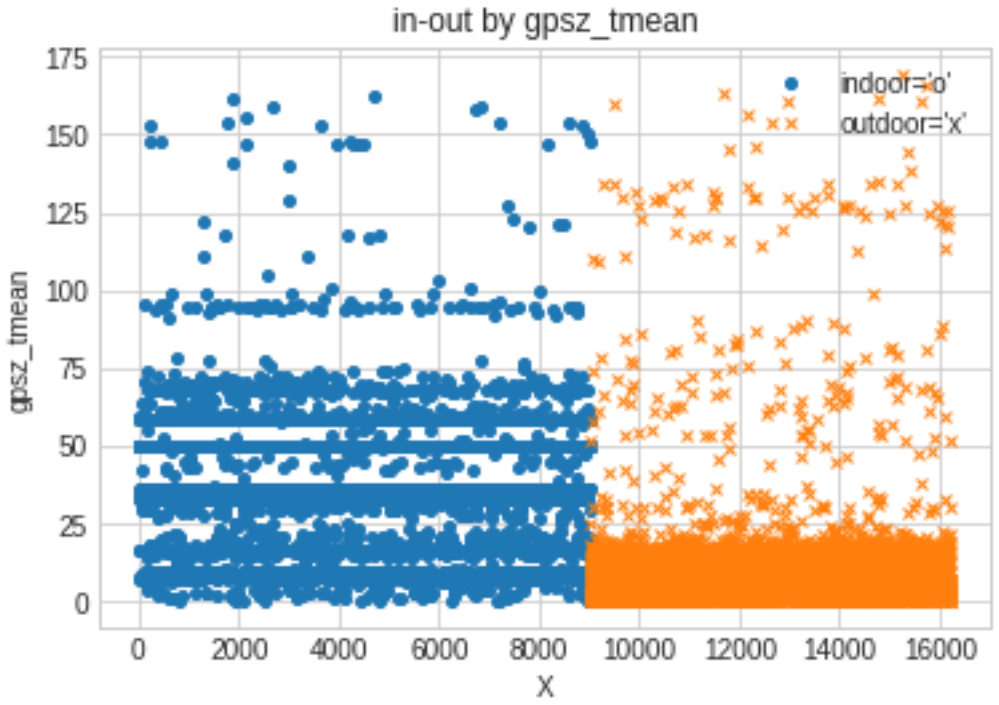
Training data



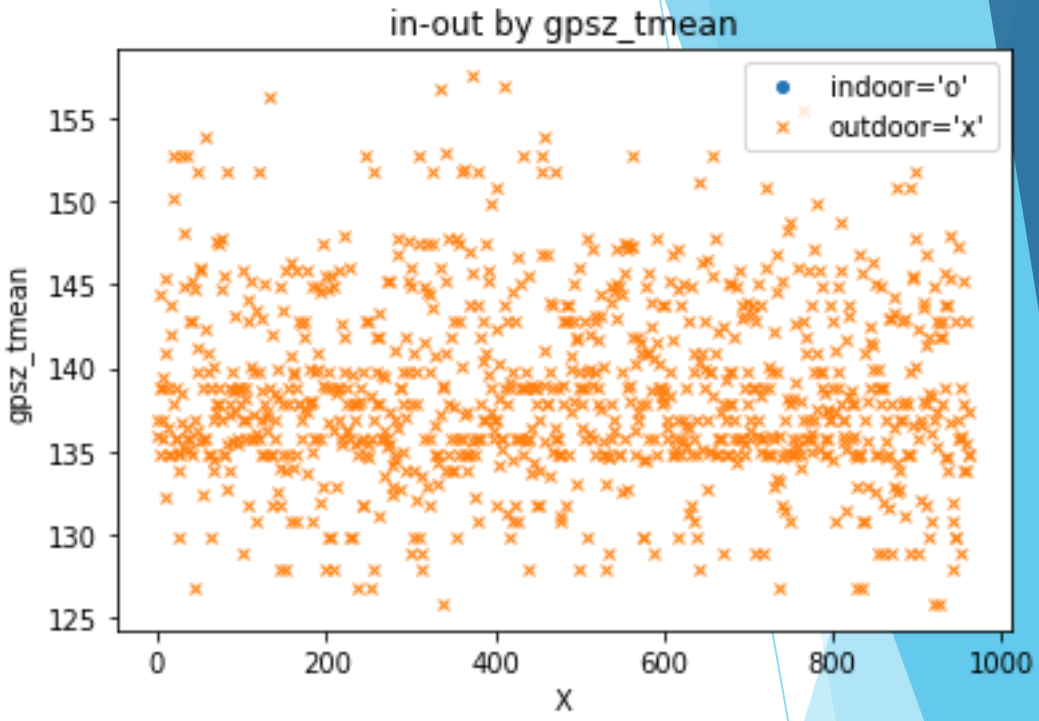
Testing data

特征 8 / feature 8

snr15num :手机检测到信噪比大于等于 15 的卫星 数量



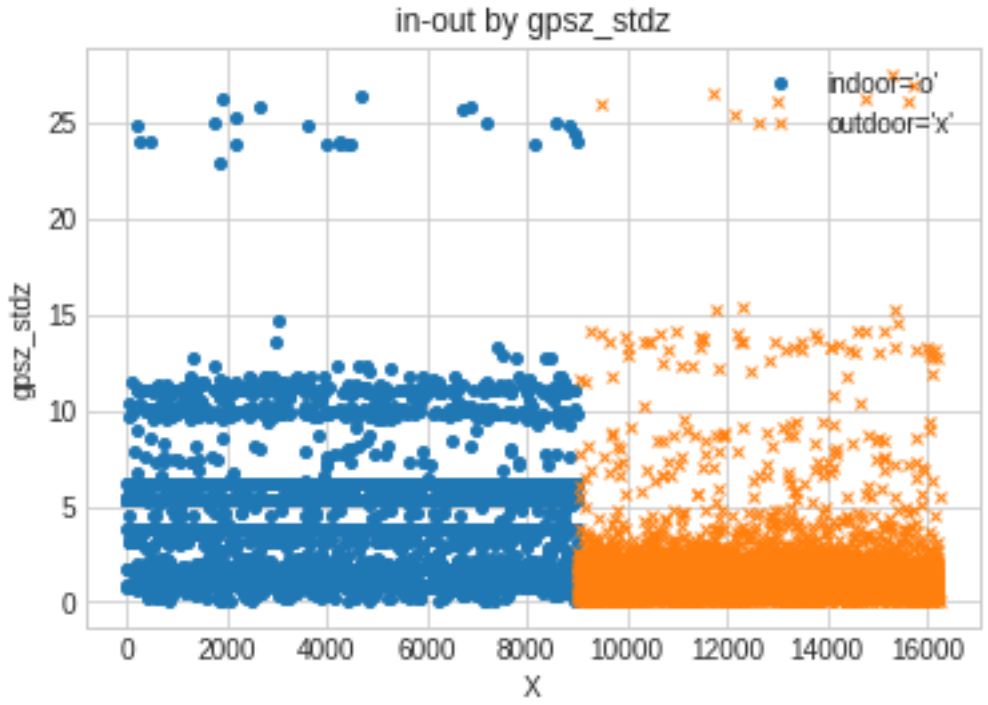
Training data



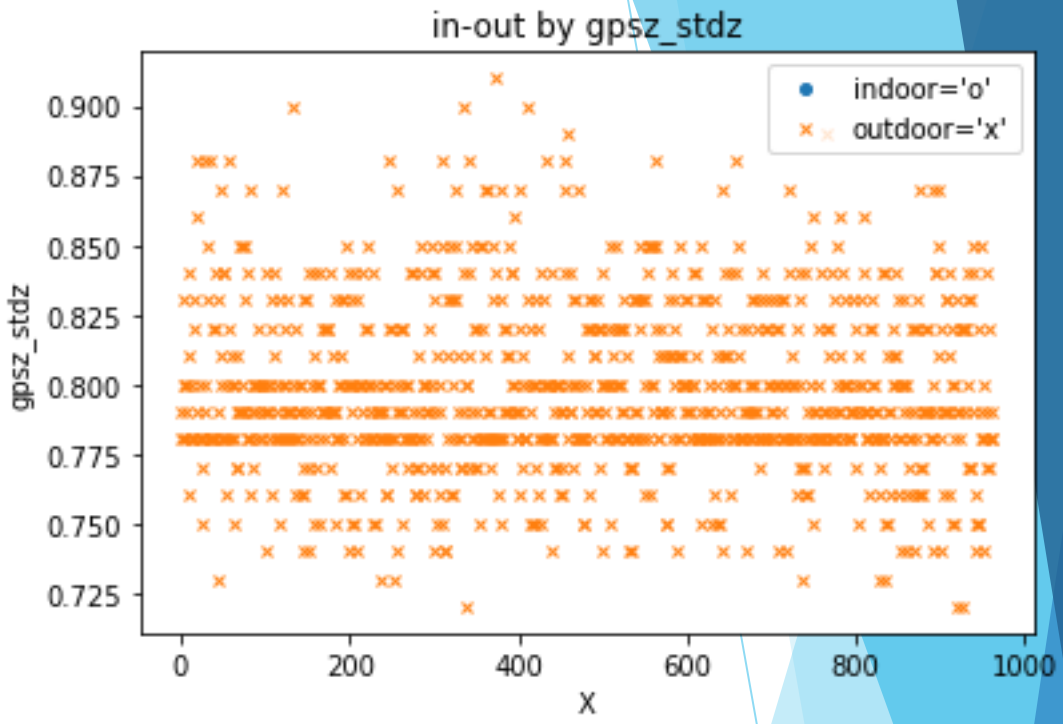
Testing data

特征 9 / feature 9

gpsz_tmean：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值



Training data

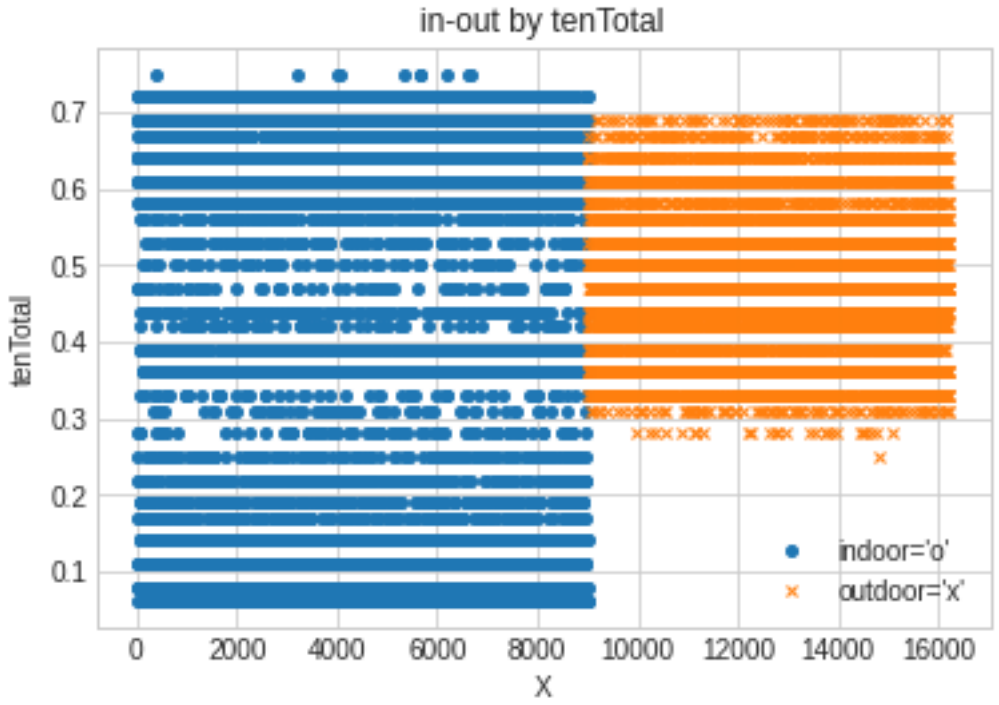


Testing data

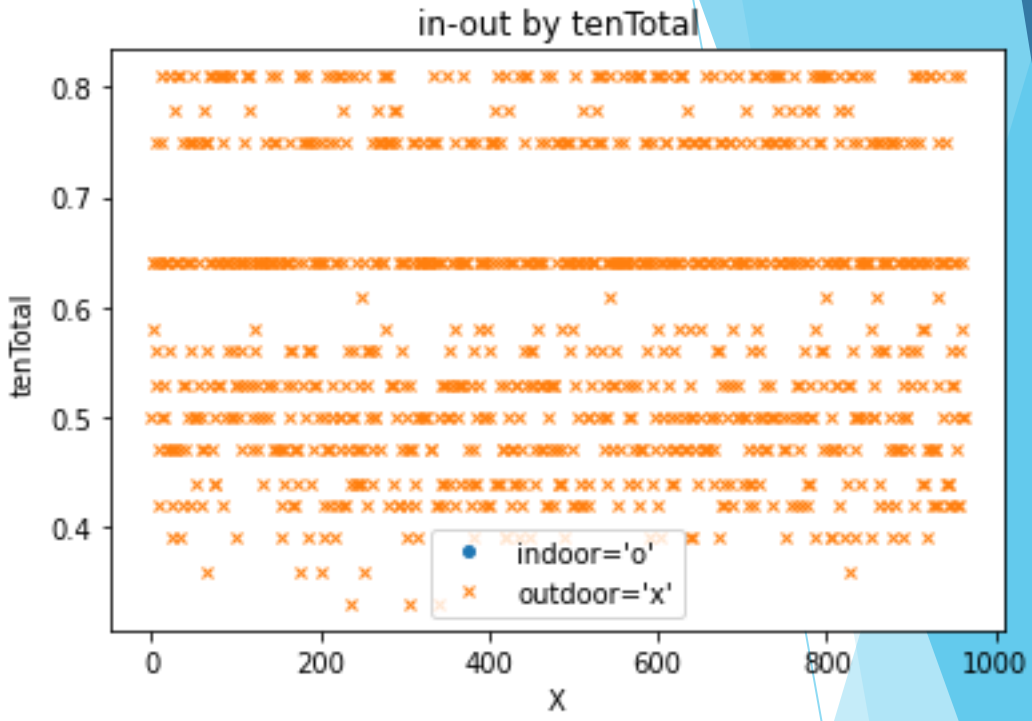
gpsz_stdz 特征处理明显不一致， 所以后面去掉这个特征

特征 10 / feature 10

gpsz_stdz：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值是否小于标准差 (0 or 1)



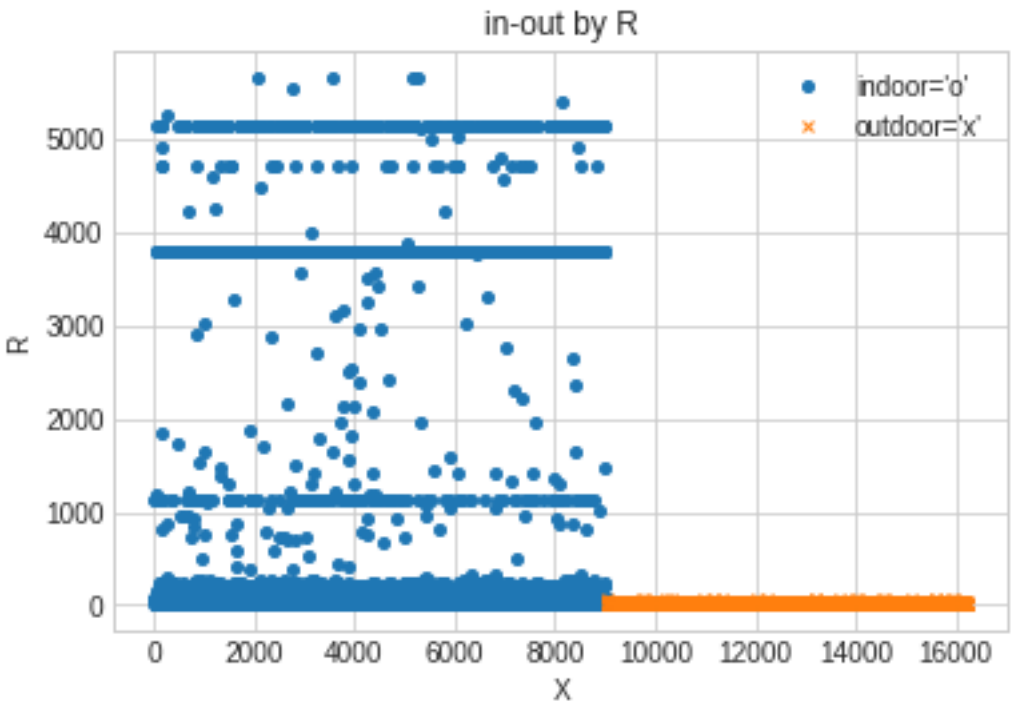
Training data



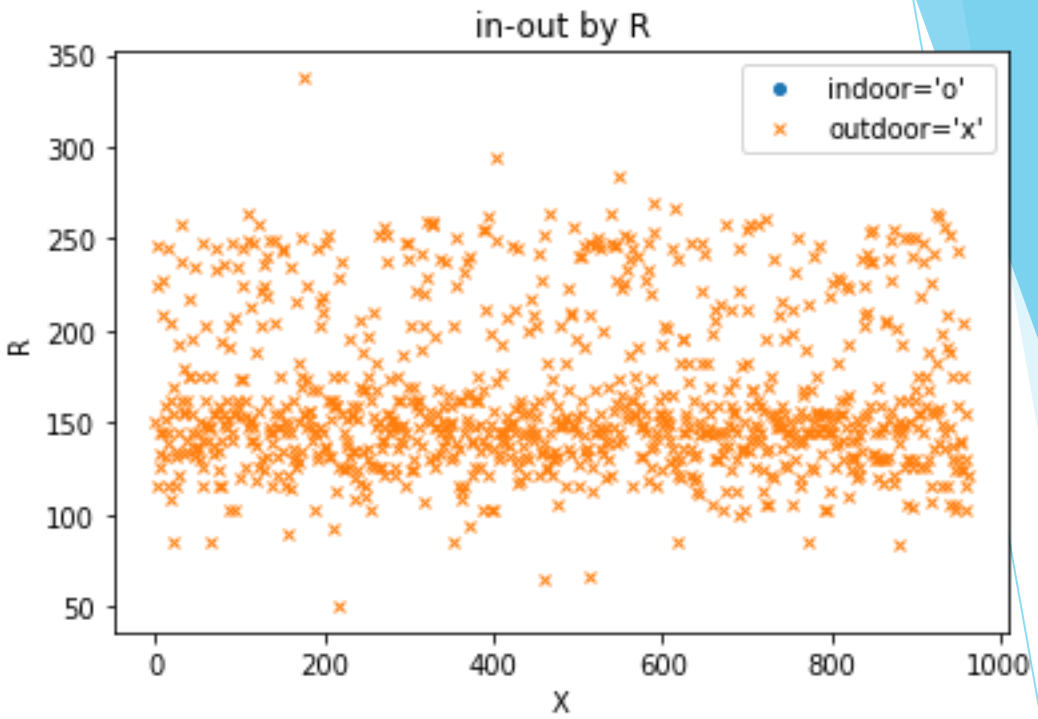
Testing data

特征 11 / feature 11

tenTotal :将 360° 分为 36 份(10°)，包含卫星的份数总和与 36 之比



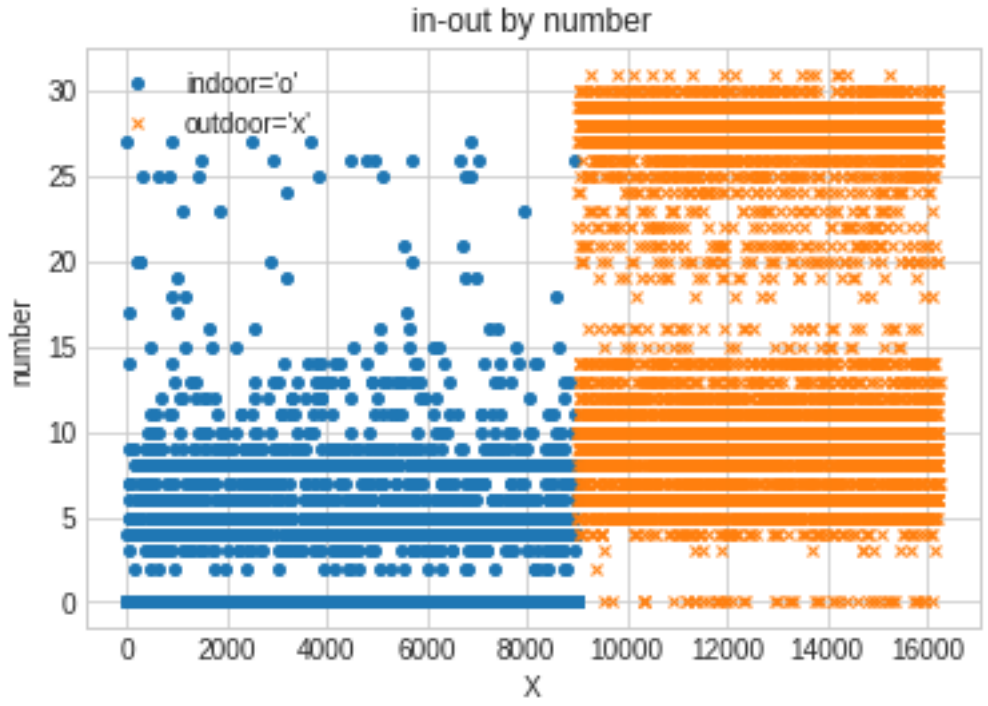
Training data



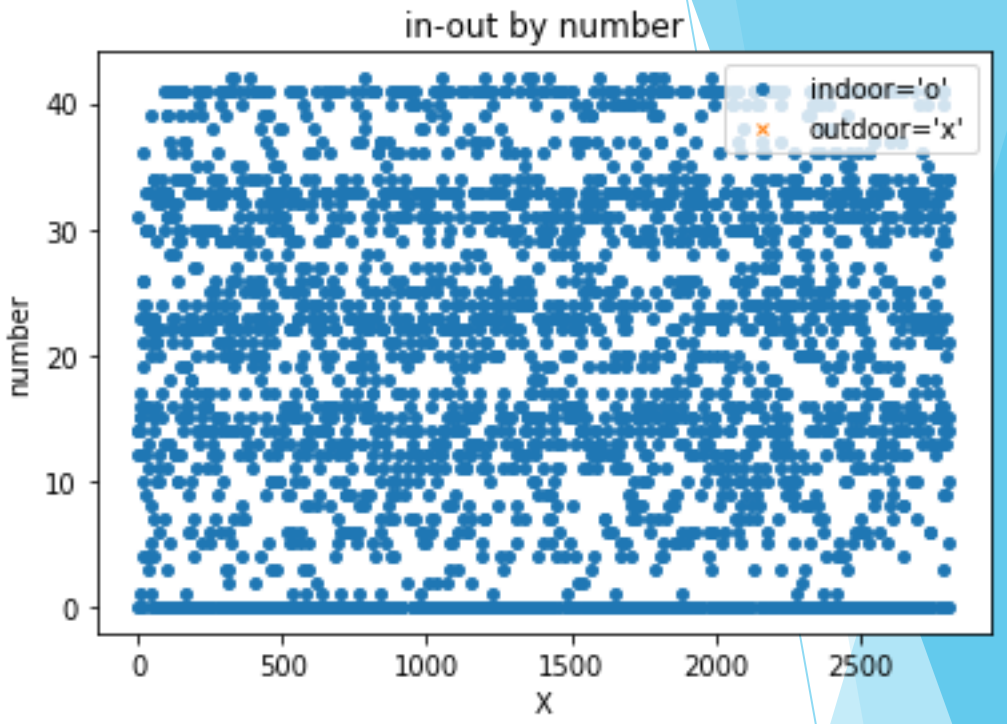
Testing data

特征 12 / feature 12
R :GPS 预测误差半径

地铁站内 the testing data when in the subway stations (indoors)

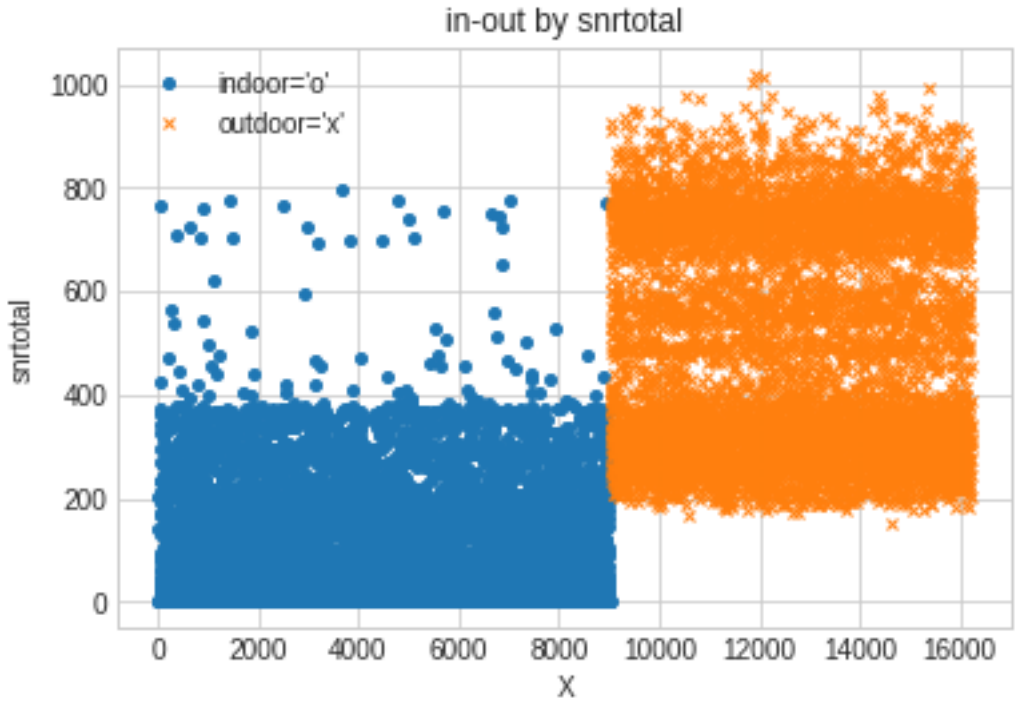


Training data

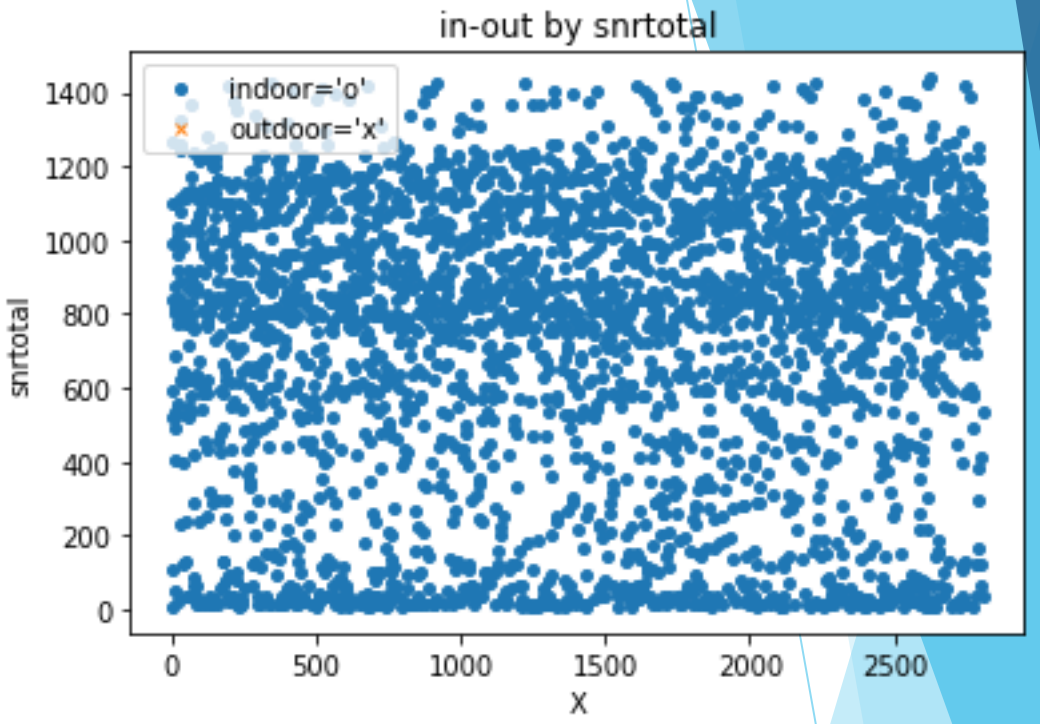


Testing data

特征 1 / feature 1
number: 手机将卫星用于定位的数量



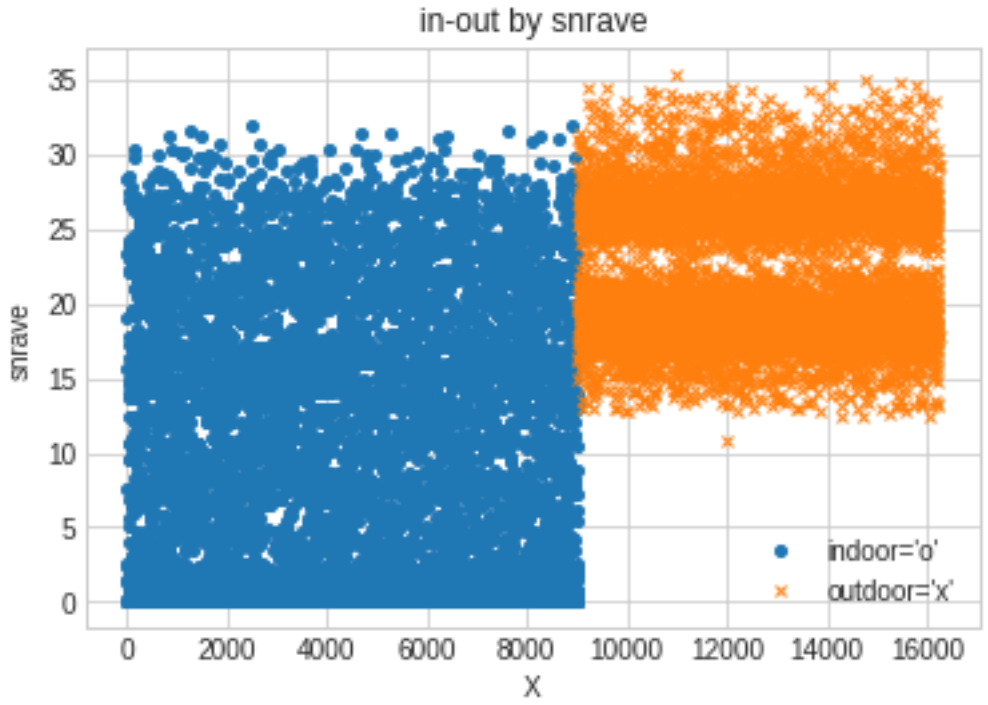
Training data



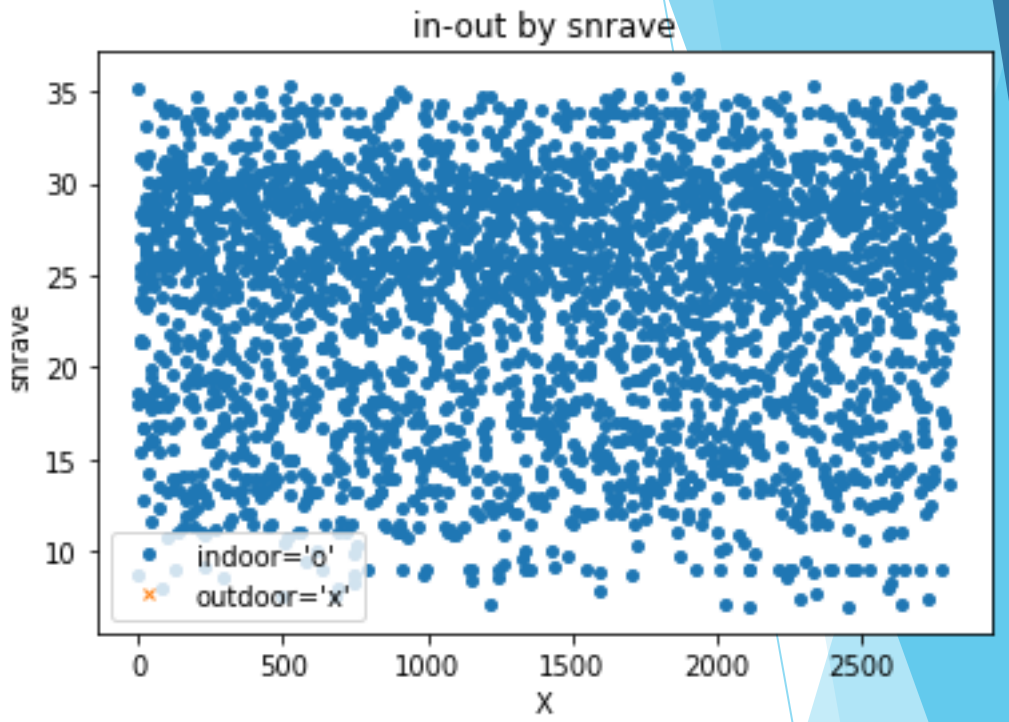
Testing data

特征 2 / feature 2

snrtotal :所有能检测到卫星的信噪比总和



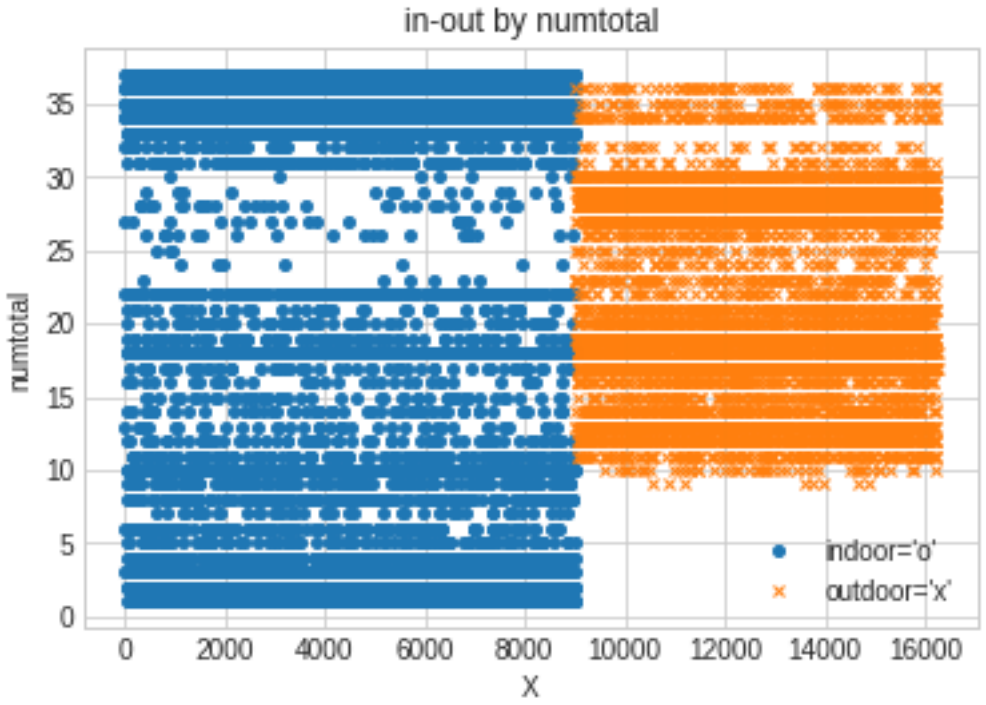
Training data



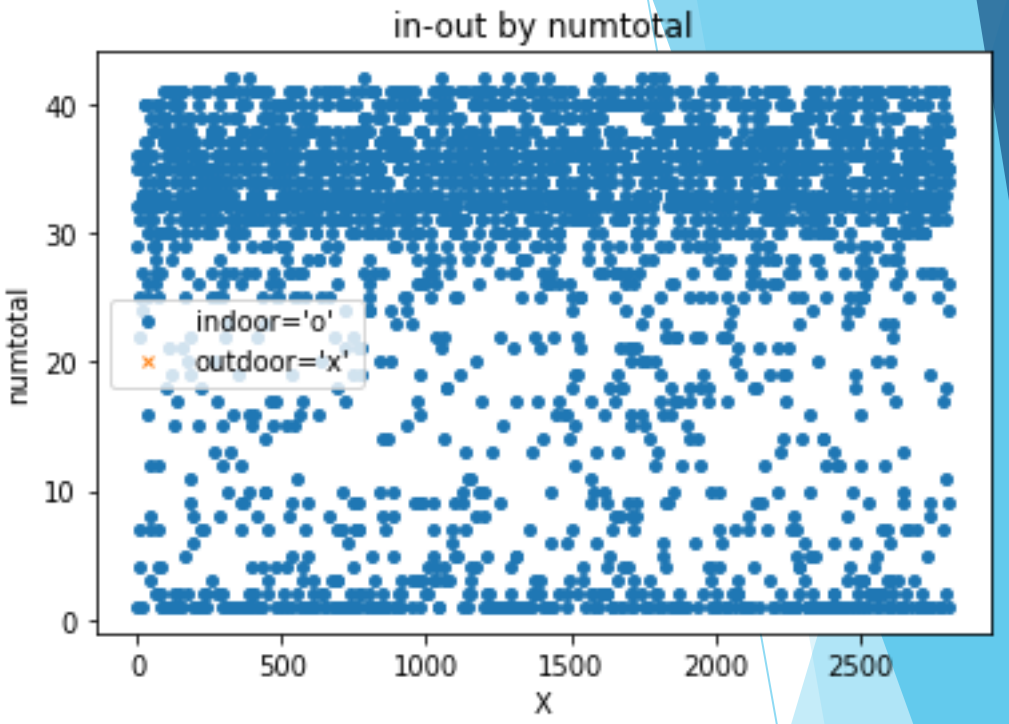
Testing data

特征 3 / feature 3

snrave :所有能检测到卫星的信噪比的平均值



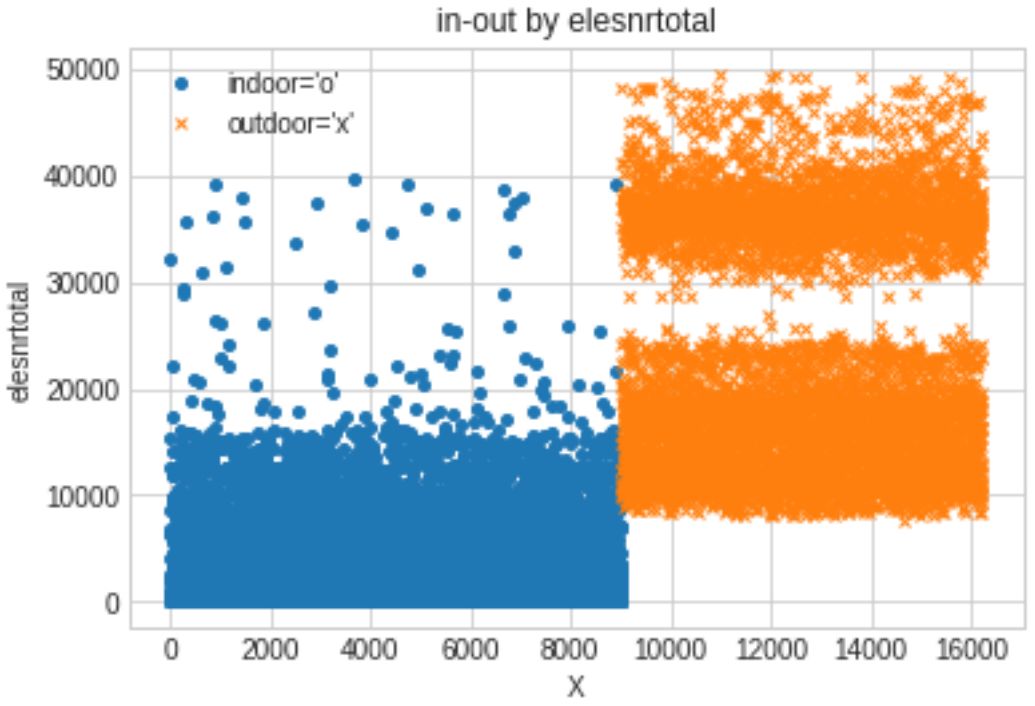
Training data



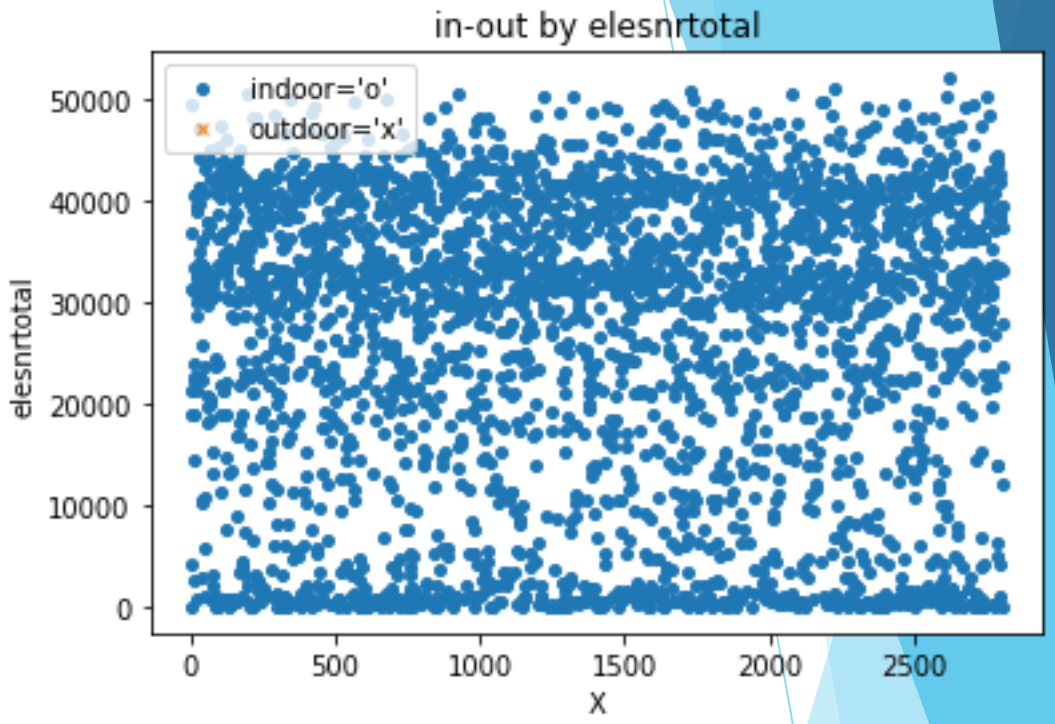
Testing data

特征 4 / feature 4

numtotal：当前手机所能检测的所有卫星数量 (包括信号强度低的卫星)



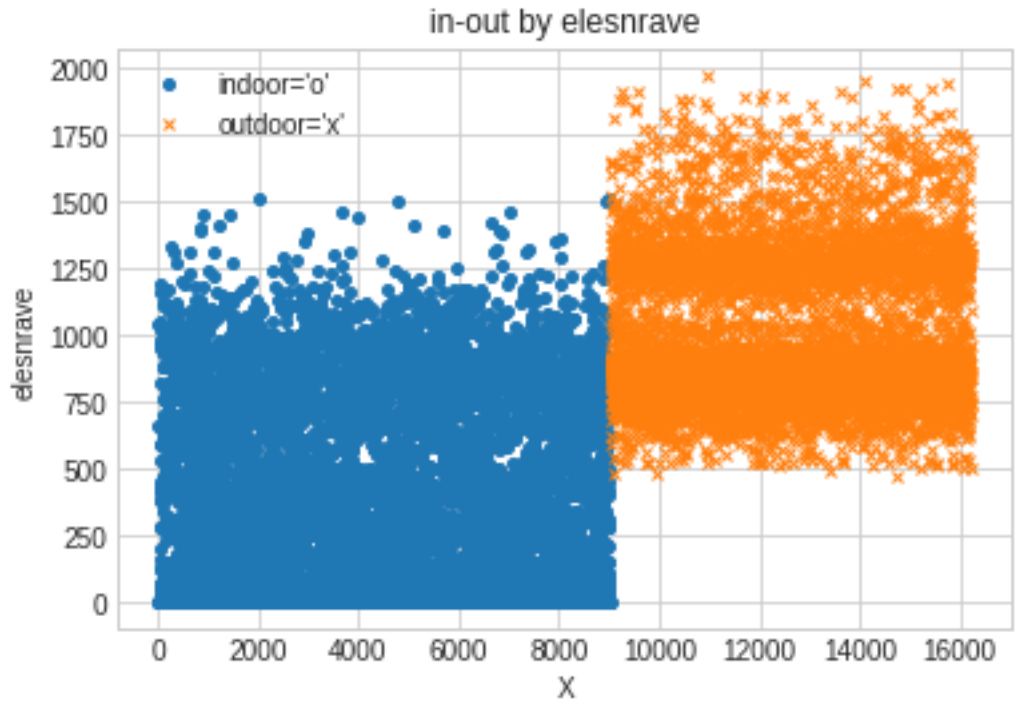
Training data



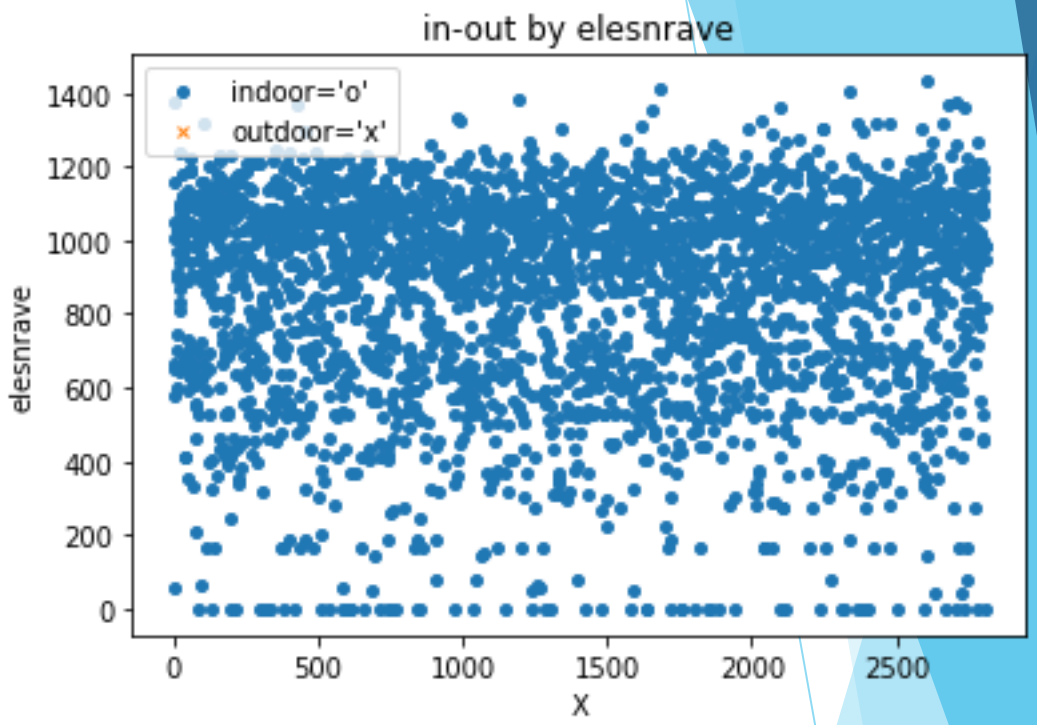
Testing data

特征 5 / feature 5

elesnrtotal :所有能检测到卫星的俯仰角和信噪比 乘积的总和



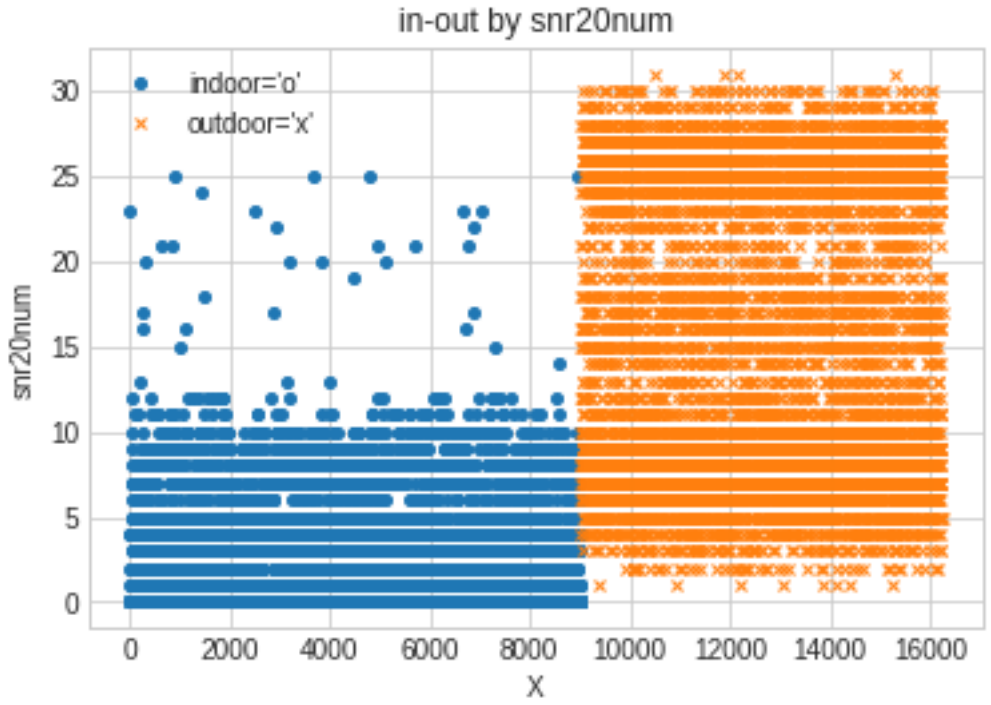
Training data



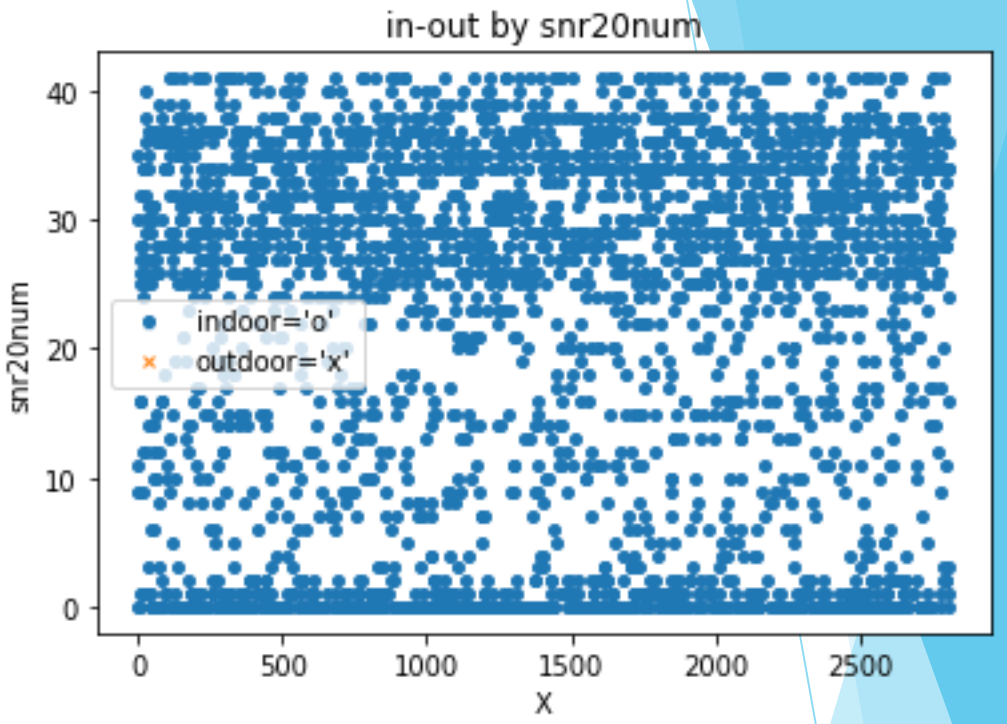
Testing data

特征 6 / feature 6

elesnrave :所有能检测到卫星的俯仰角和信噪比 乘积的平均值



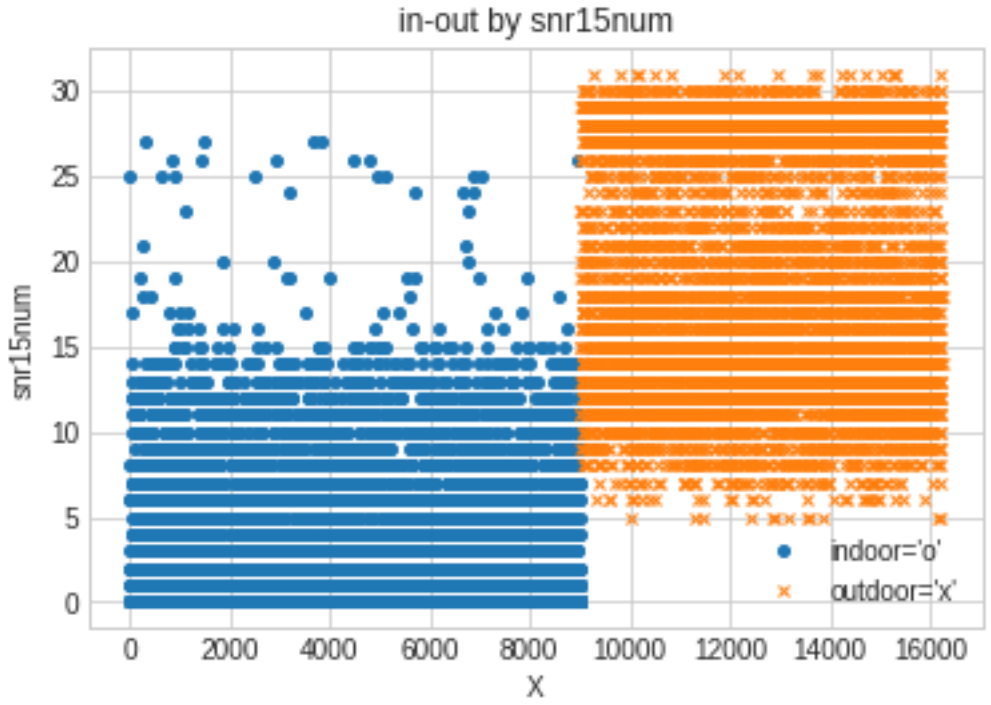
Training data



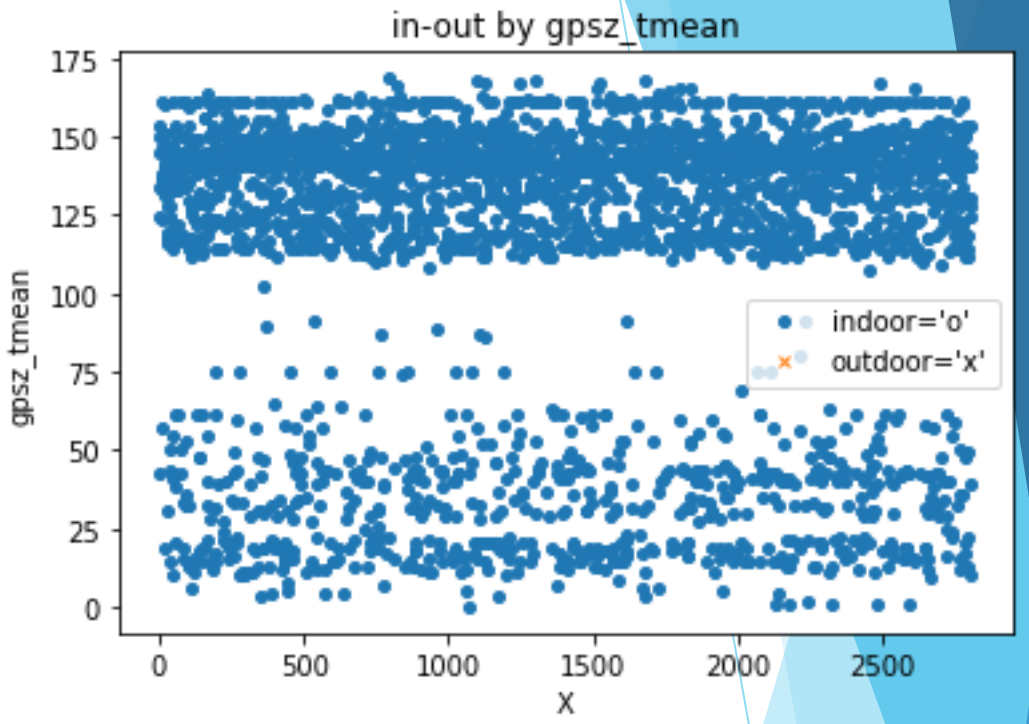
Testing data

特征 7 / feature 7

snr20num：手机检测到信噪比大于等于 20 的卫星 数量



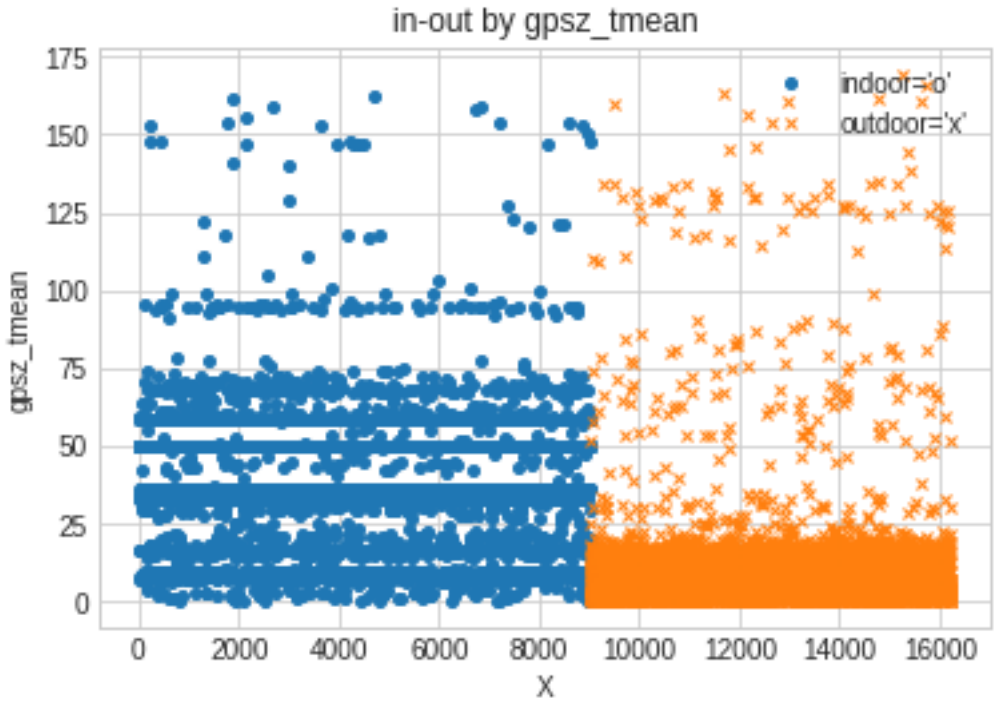
Training data



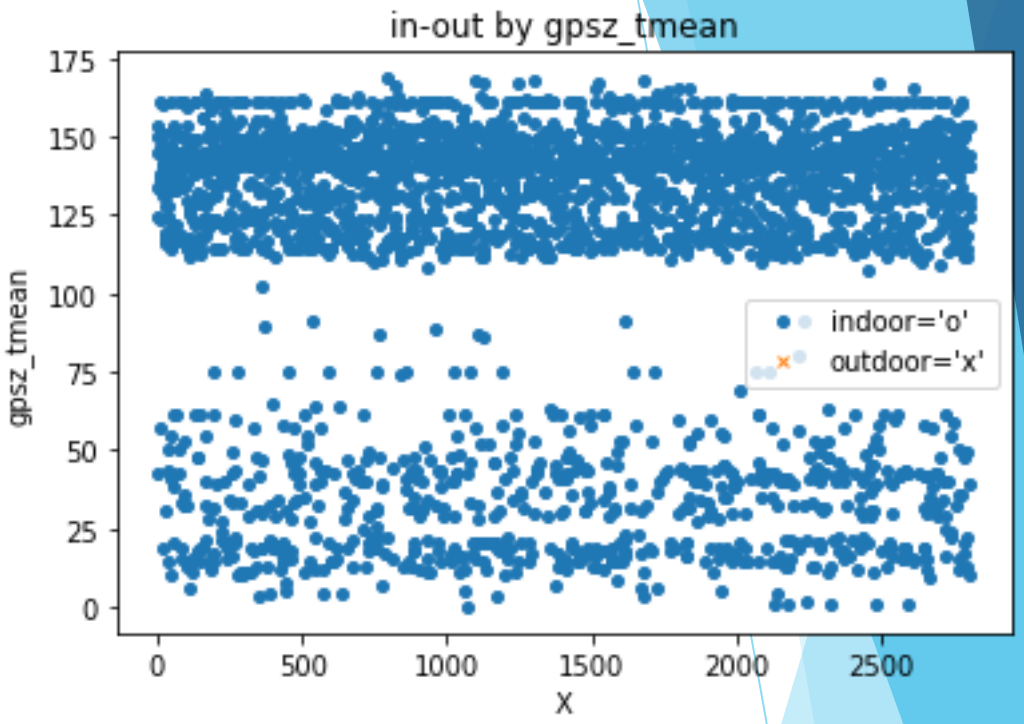
Testing data

特征 8 / feature 8

snr15num：手机检测到信噪比大于等于 15 的卫星 数量



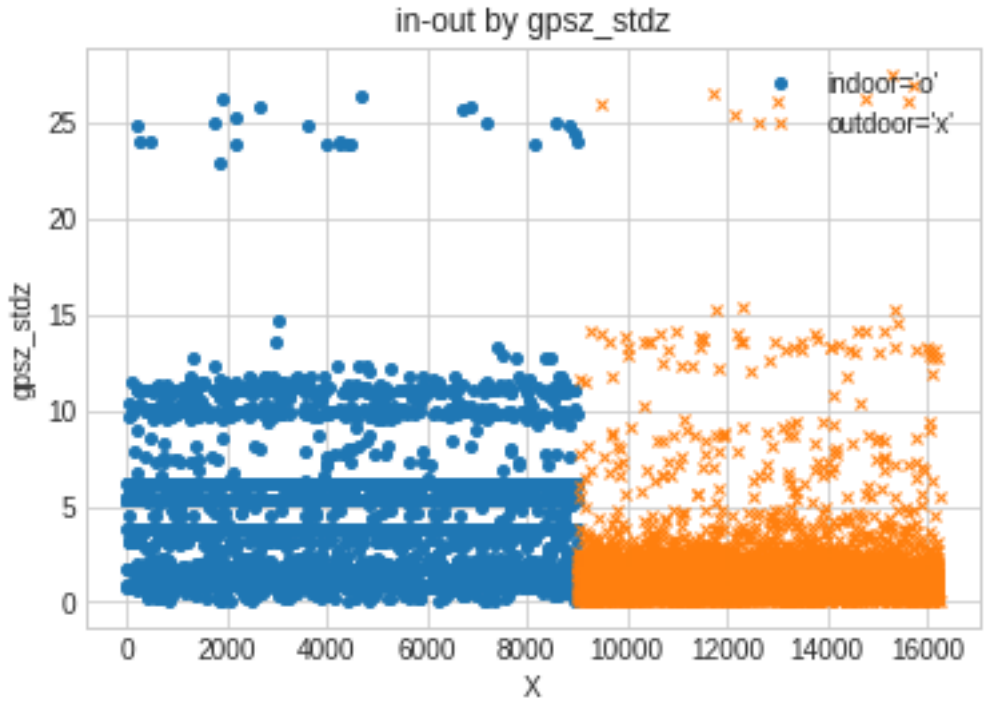
Training data



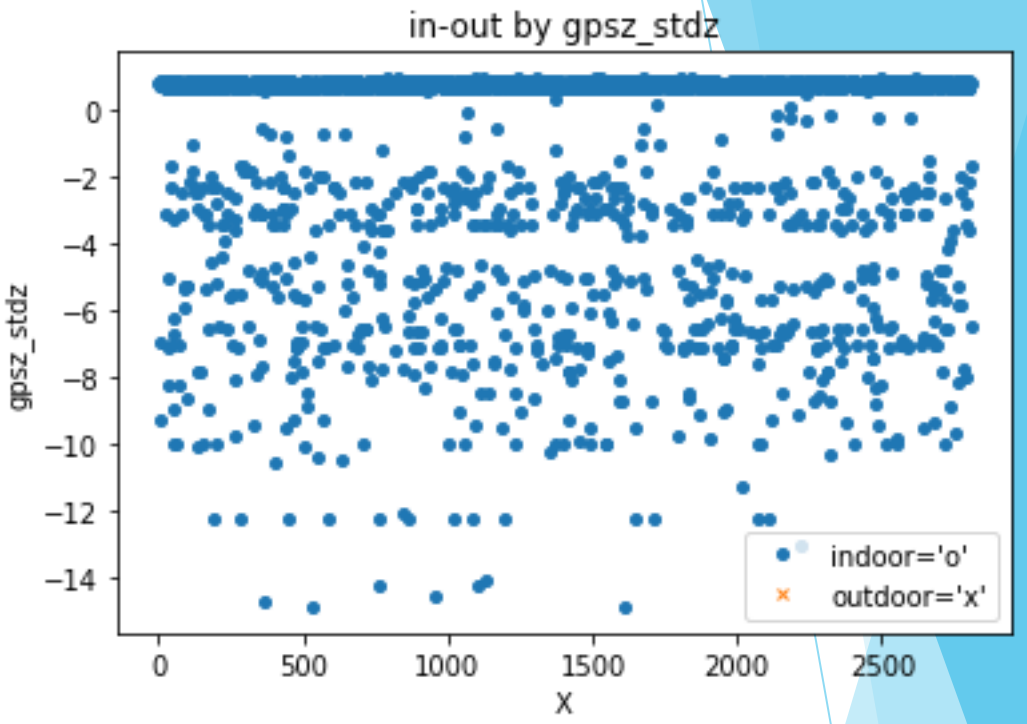
Testing data

特征 9 / feature 9

gpsz_tmean：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值



Training data

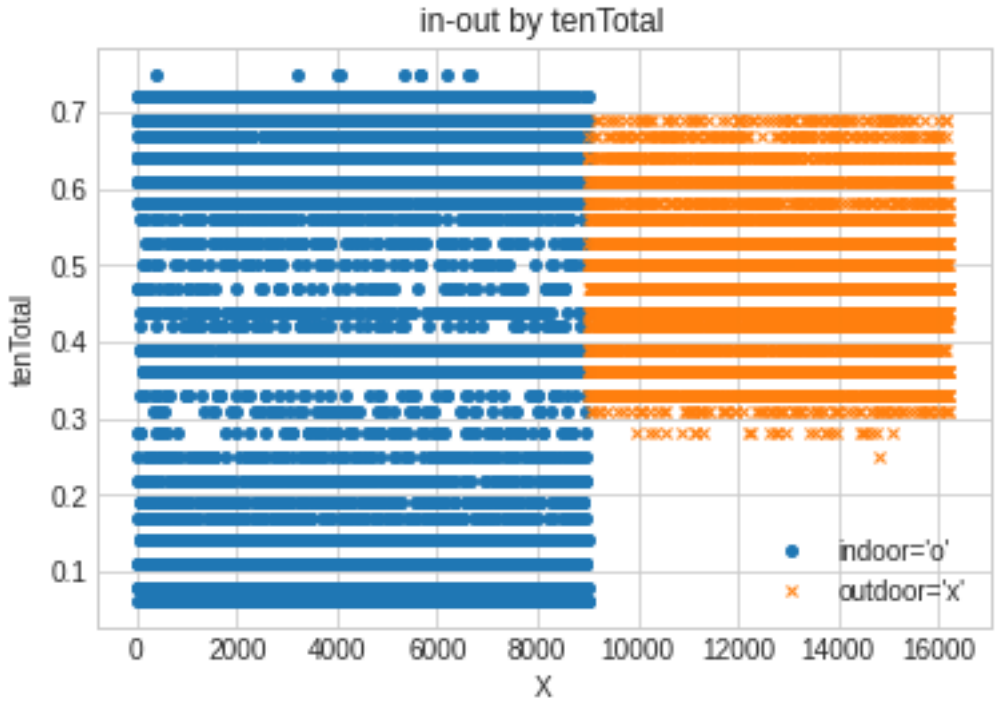


Testing data

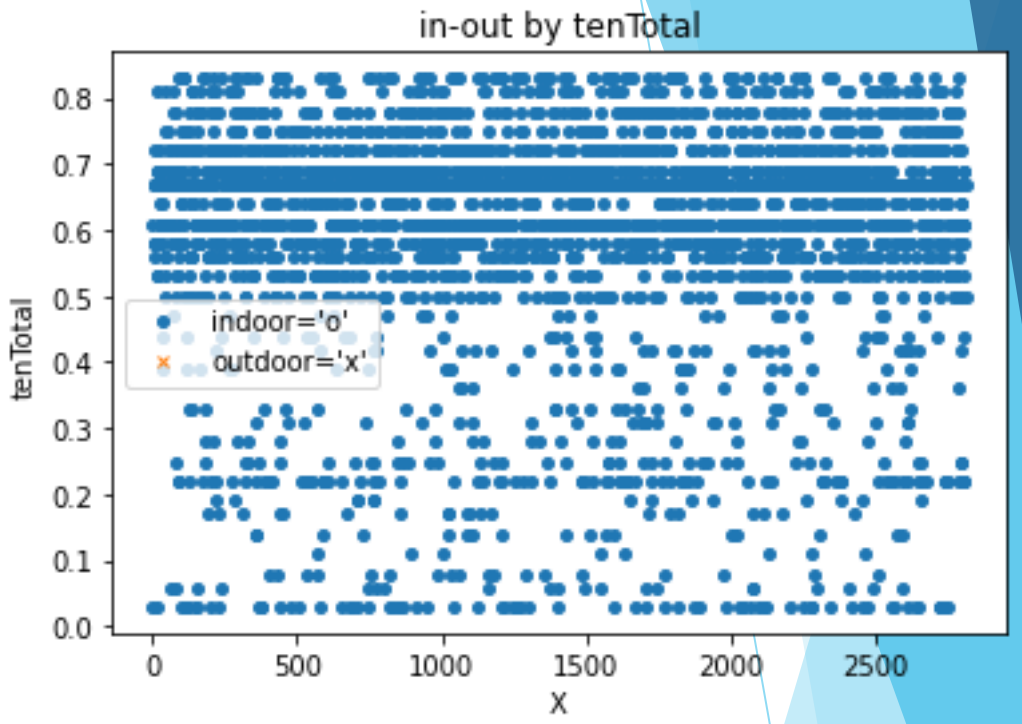
gpsz_stdz 特征处理明显不一致， 所以后面去掉这个特征

特征 10 / feature 10

gpsz_stdz：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值是否小于标准差 (0 or 1)



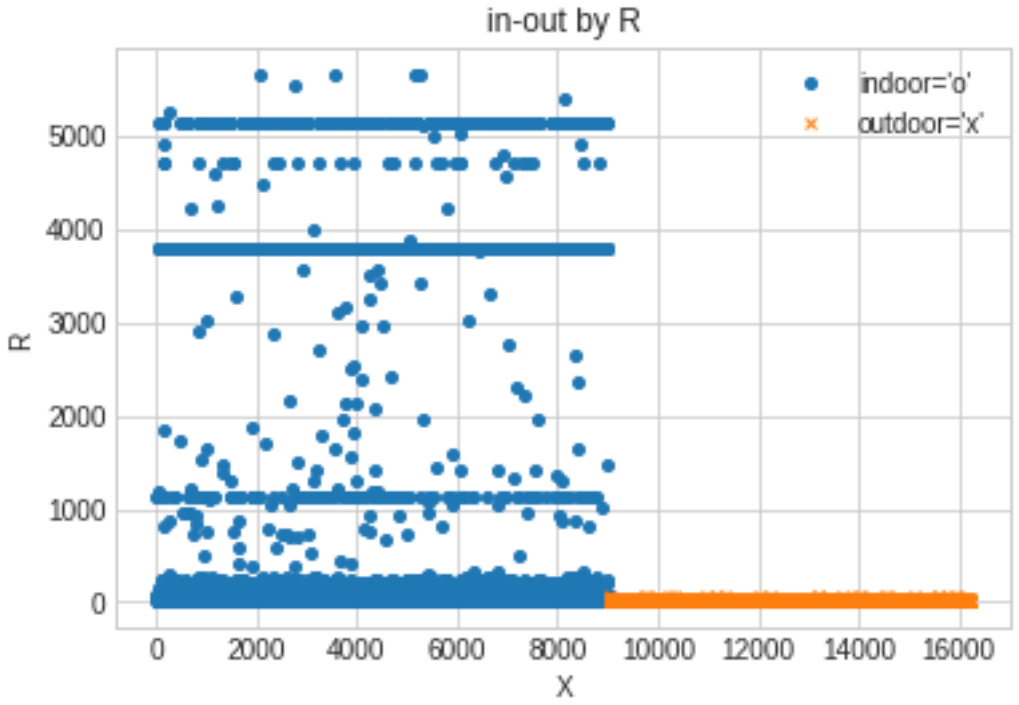
Training data



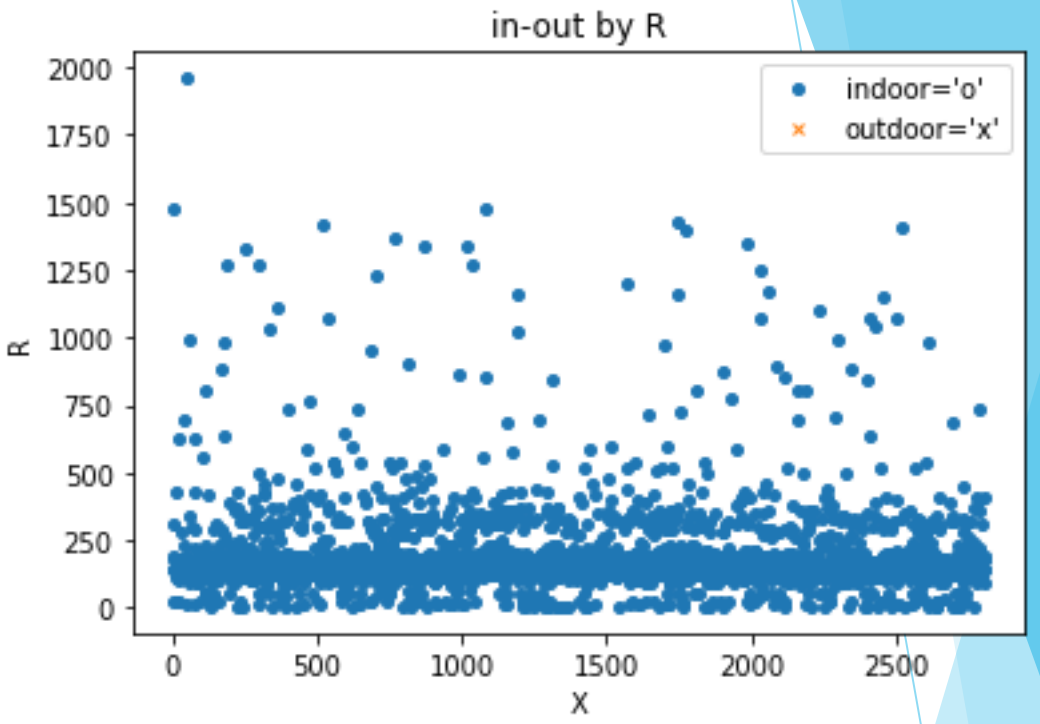
Testing data

特征 11 / feature 11

tenTotal :将 360° 分为 36 份(10°)，包含卫星的份数总和与 36 之比



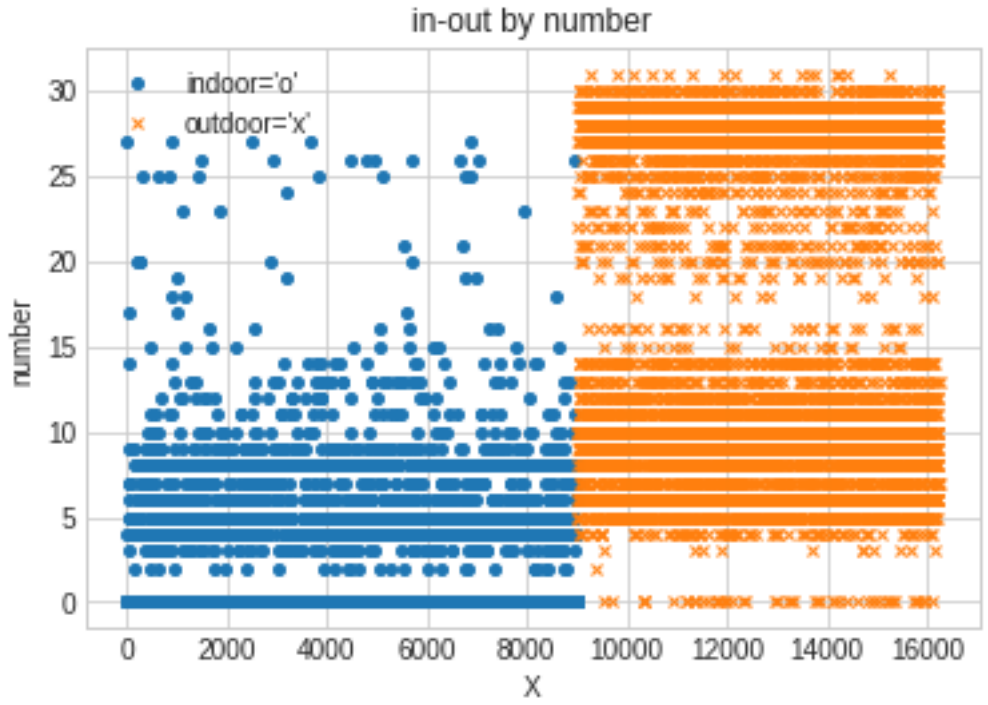
Training data



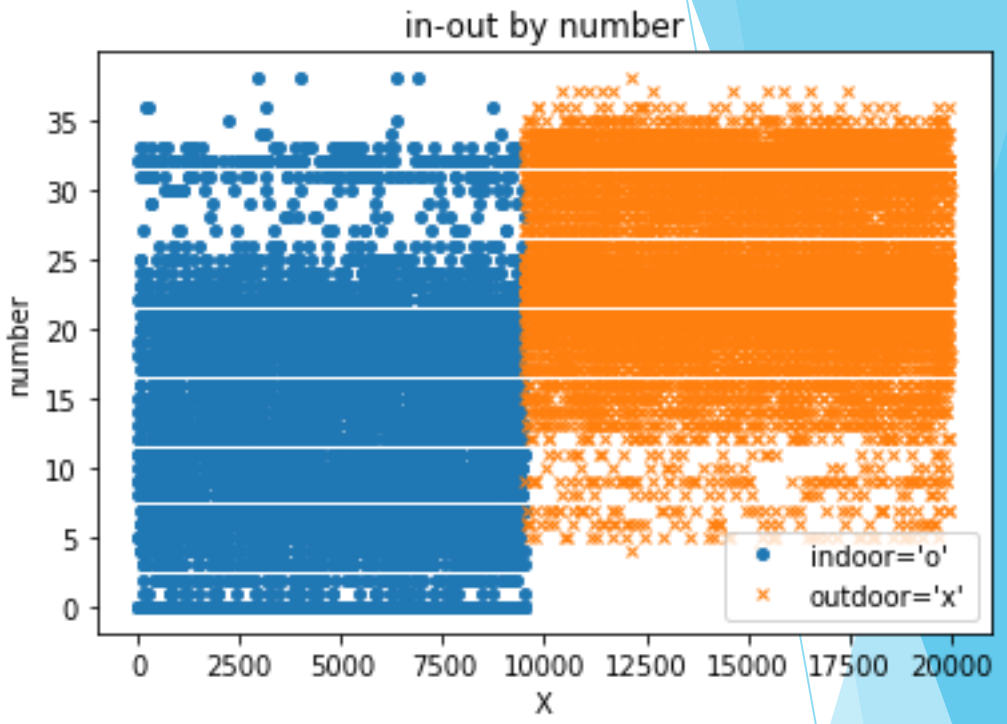
Testing data

特征 12 / feature 12
R :GPS 预测误差半径

其他 - 删掉了 地铁室外跑 & 地铁站内
other main scenes such as streets, communities, stores, buildings, etc. excluding scenes related to Subway

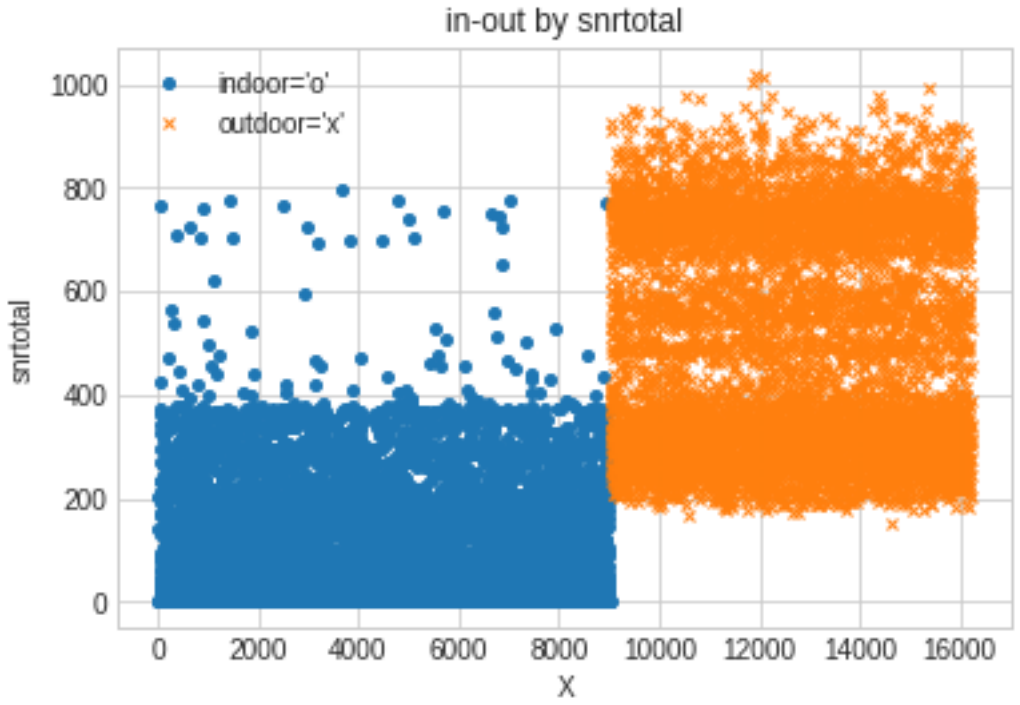


Training data

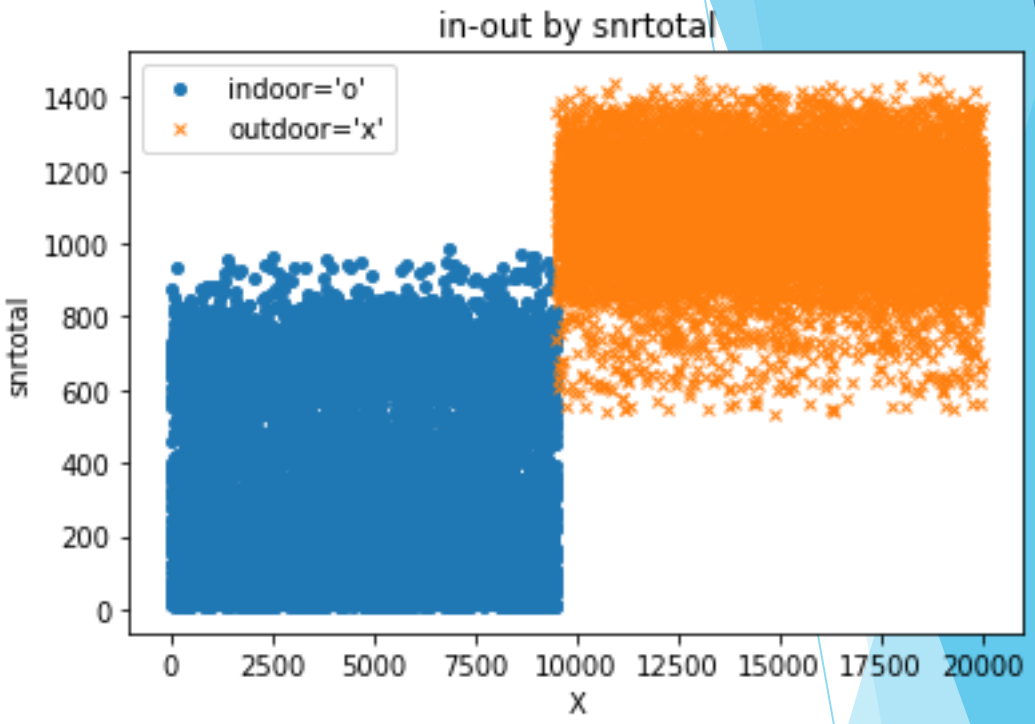


Testing data

特征 1 / feature 1
number: 手机将卫星用于定位的数量



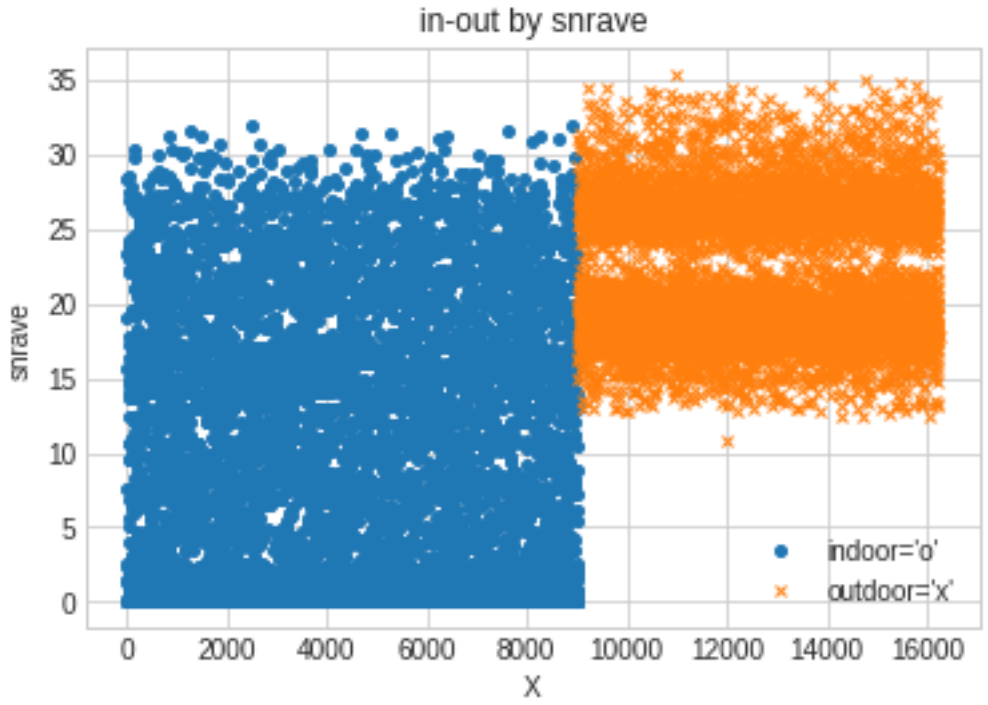
Training data



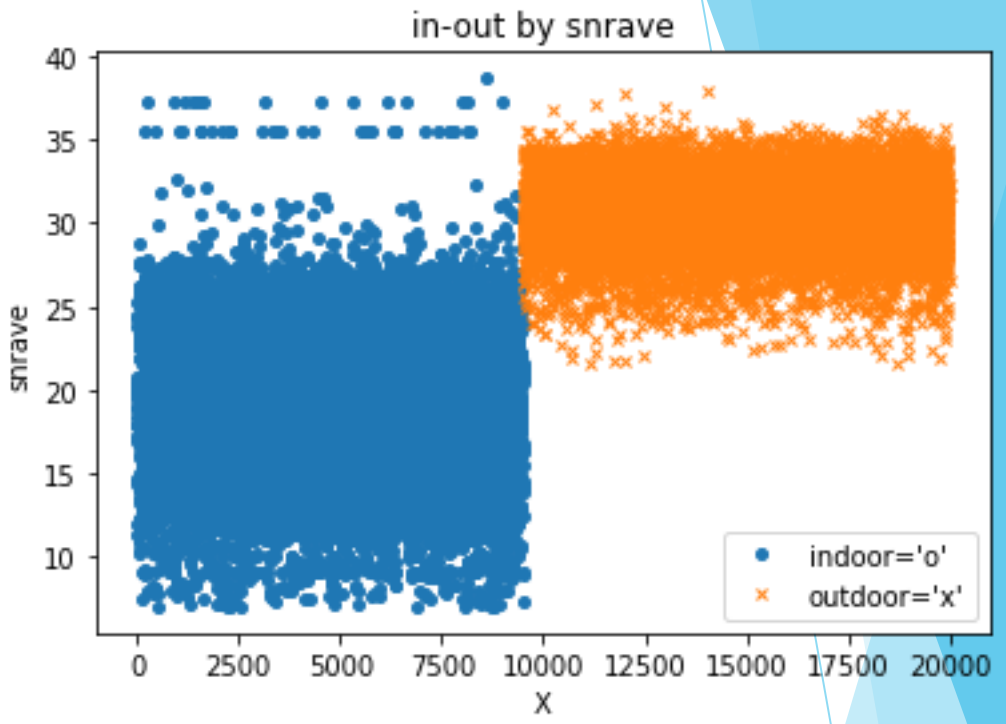
Testing data

特征 2 / feature 2

snrtotal :所有能检测到卫星的信噪比总和



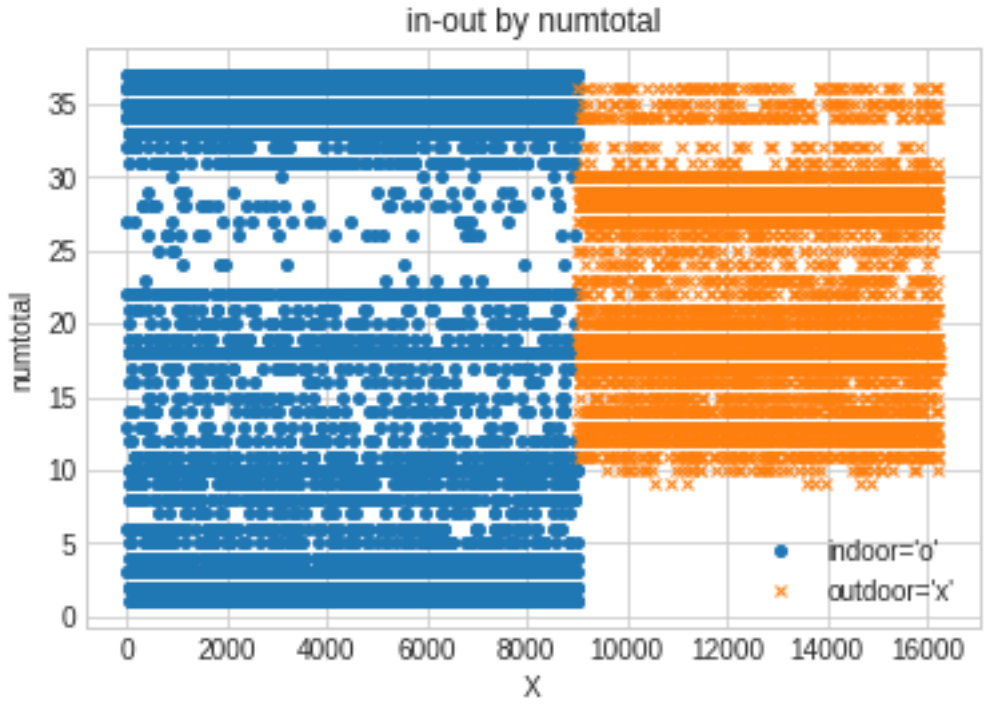
Training data



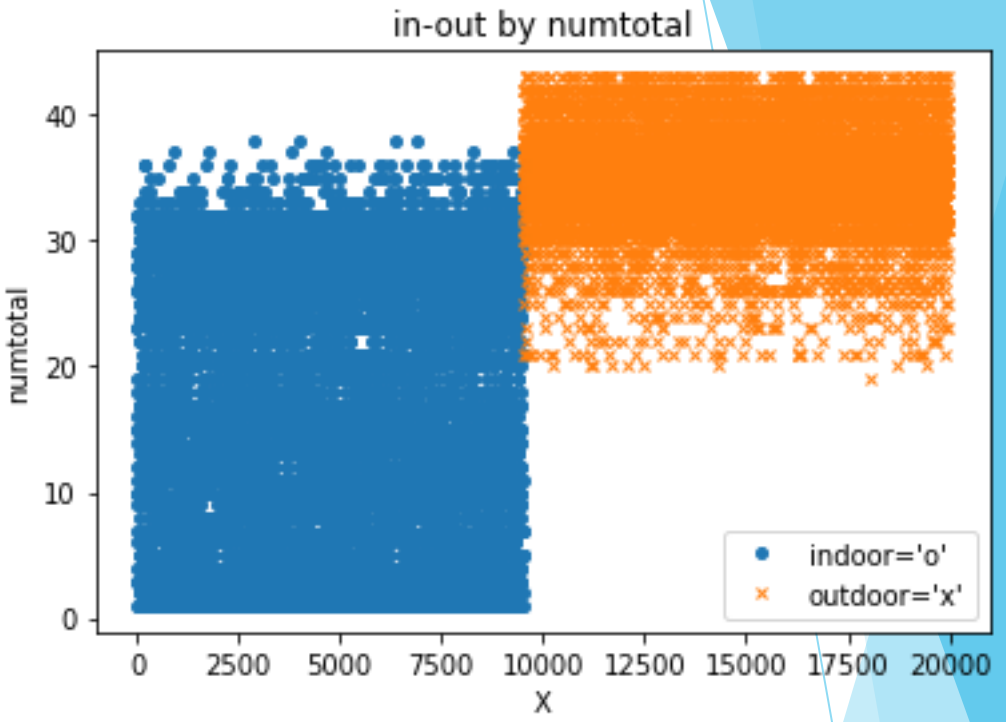
Testing data

特征 3 / feature 3

snrave :所有能检测到卫星的信噪比的平均值



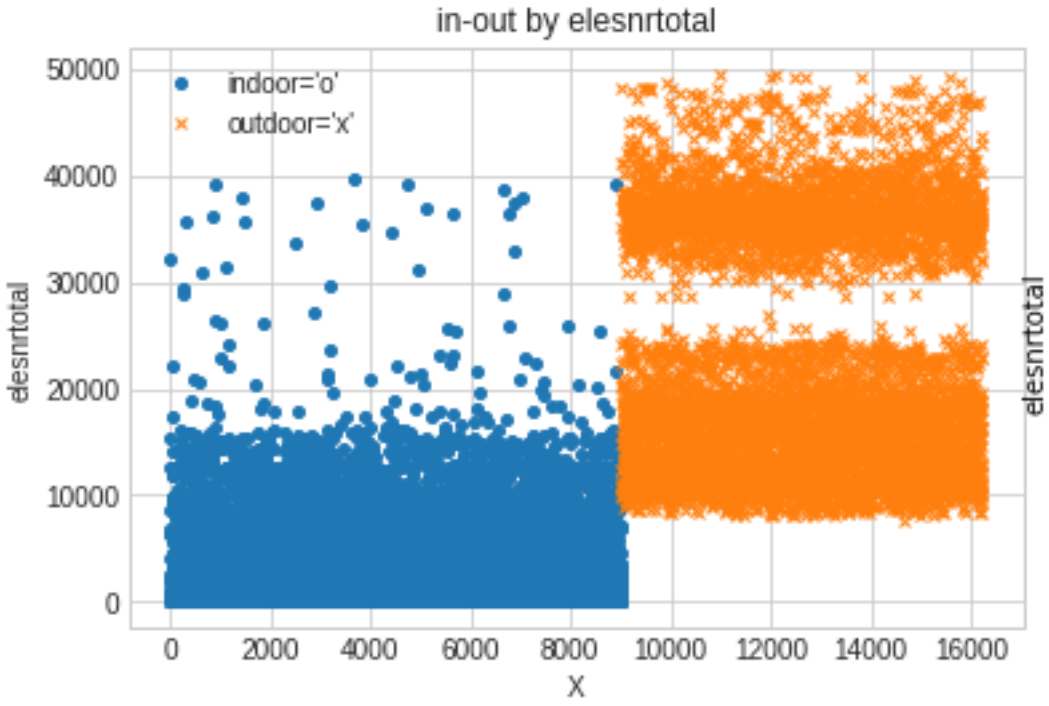
Training data



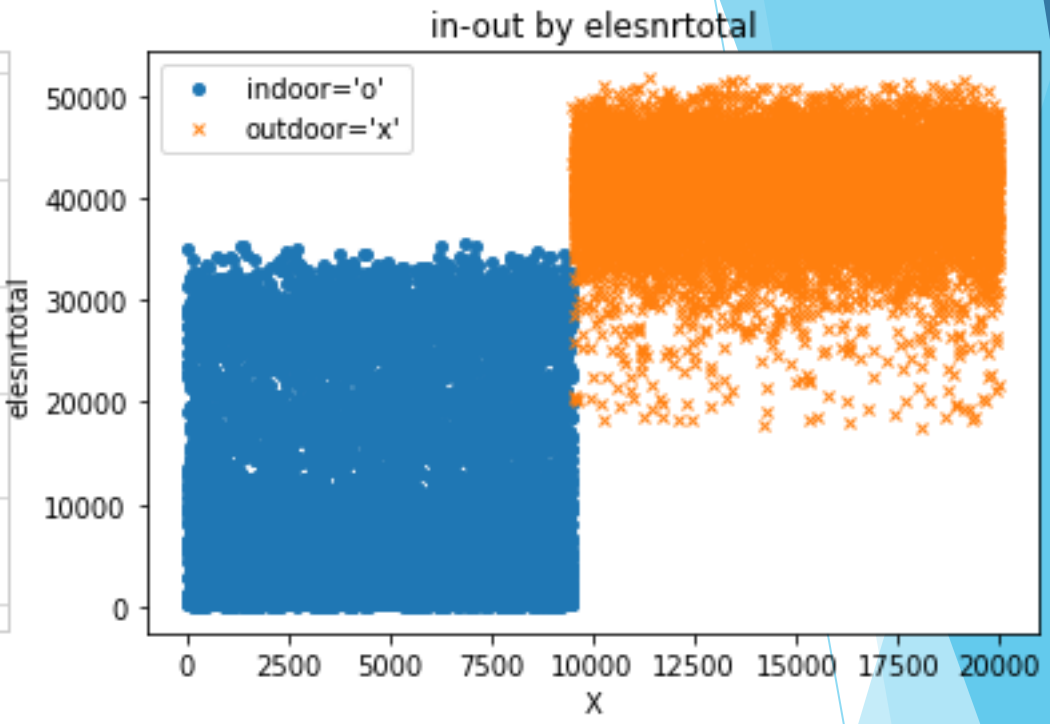
Testing data

特征 4 / feature 4

numtotal：当前手机所能检测的所有卫星数量 (包括信号强度低的卫星)



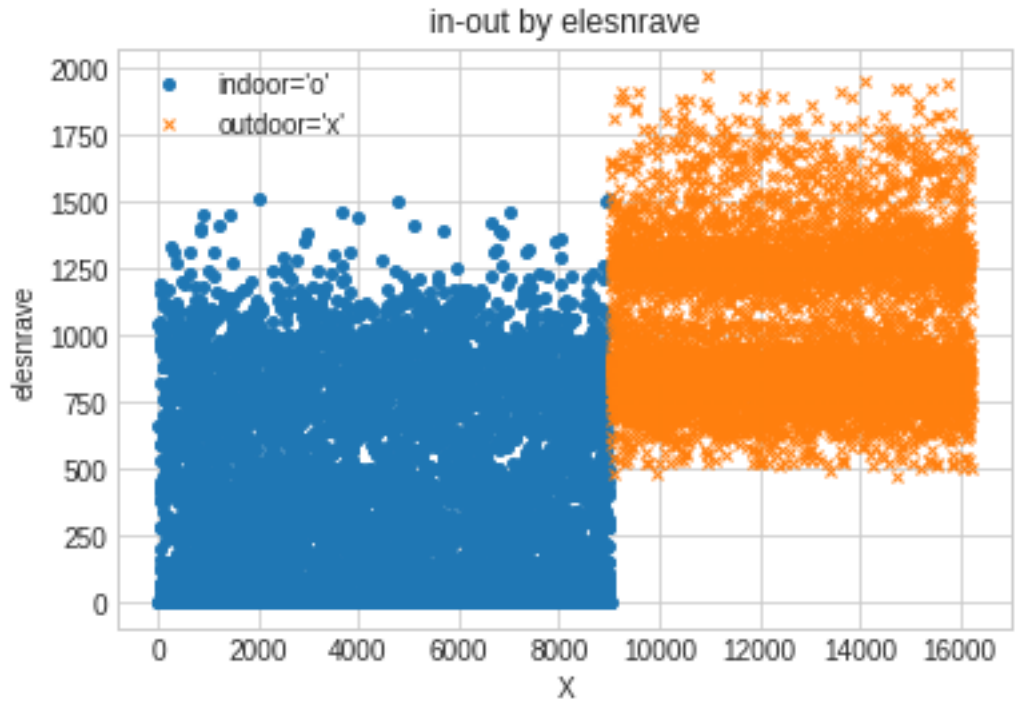
Training data



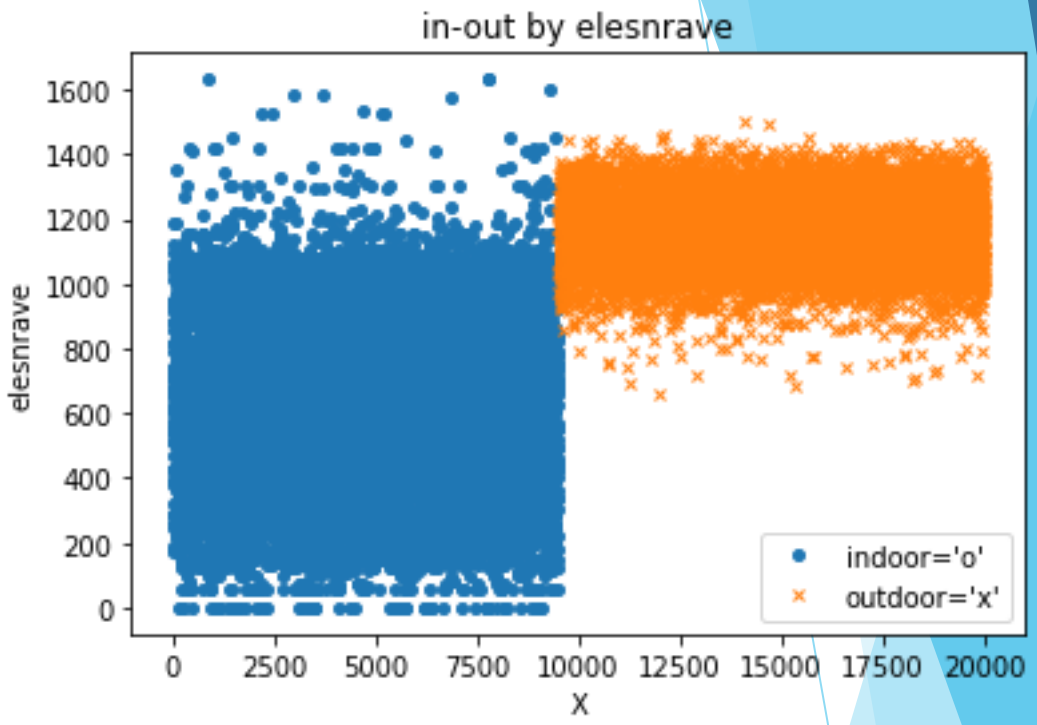
Testing data

特征 5 / feature 5

elesnrtotal :所有能检测到卫星的俯仰角和信噪比 乘积的总和



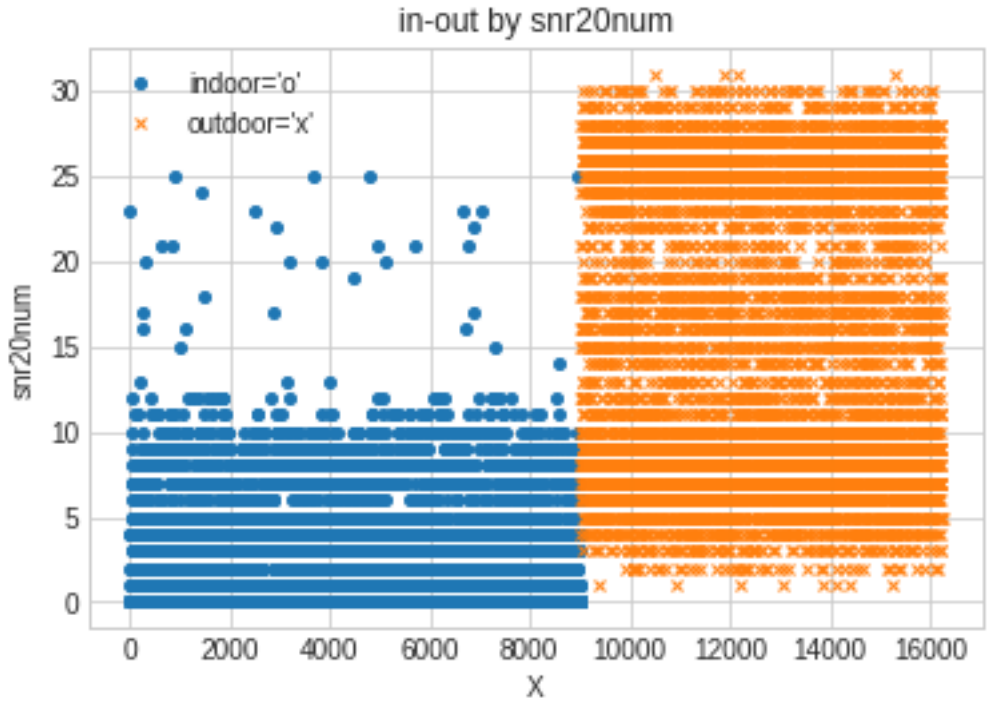
Training data



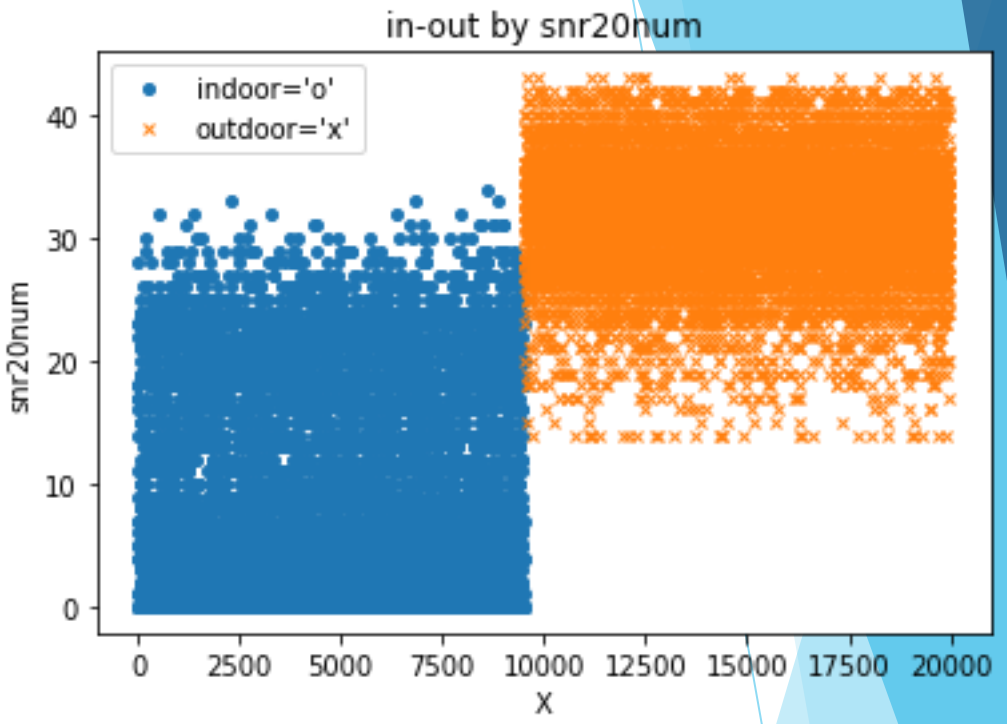
Testing data

特征 6 / feature 6

elesnrave :所有能检测到卫星的俯仰角和信噪比 乘积的平均值



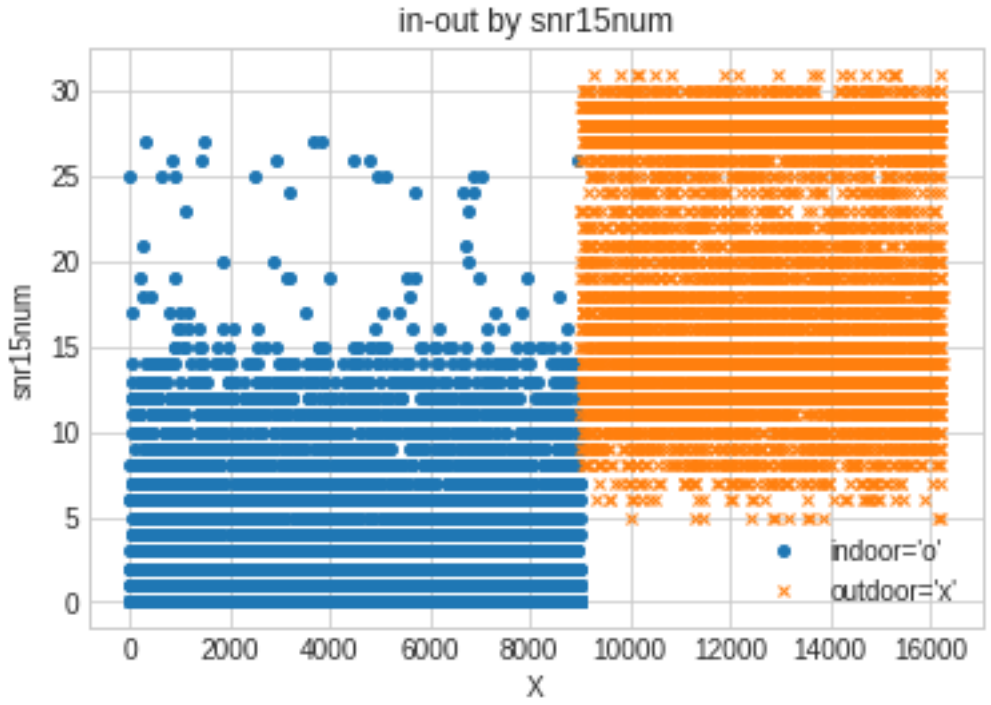
Training data



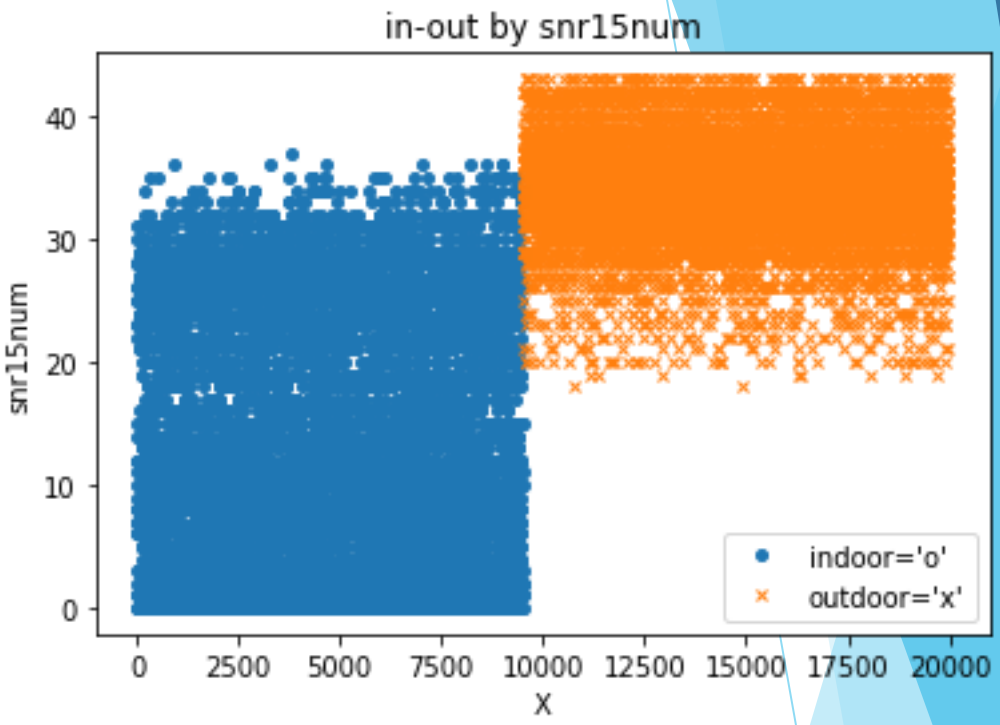
Testing data

特征 7 / feature 7

snr20num：手机检测到信噪比大于等于 20 的卫星 数量



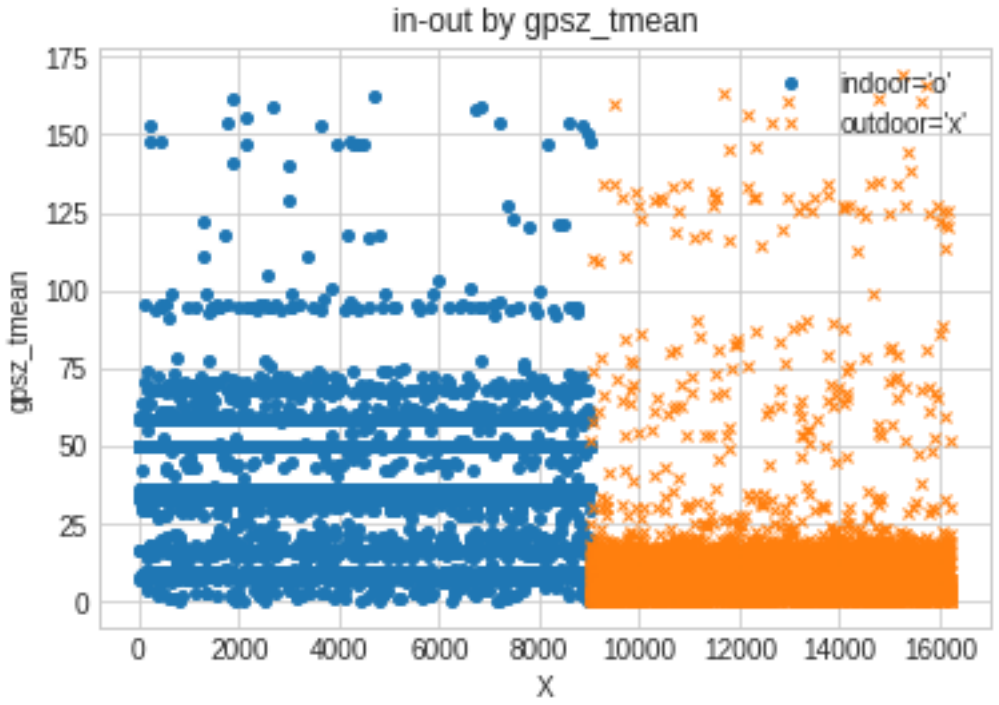
Training data



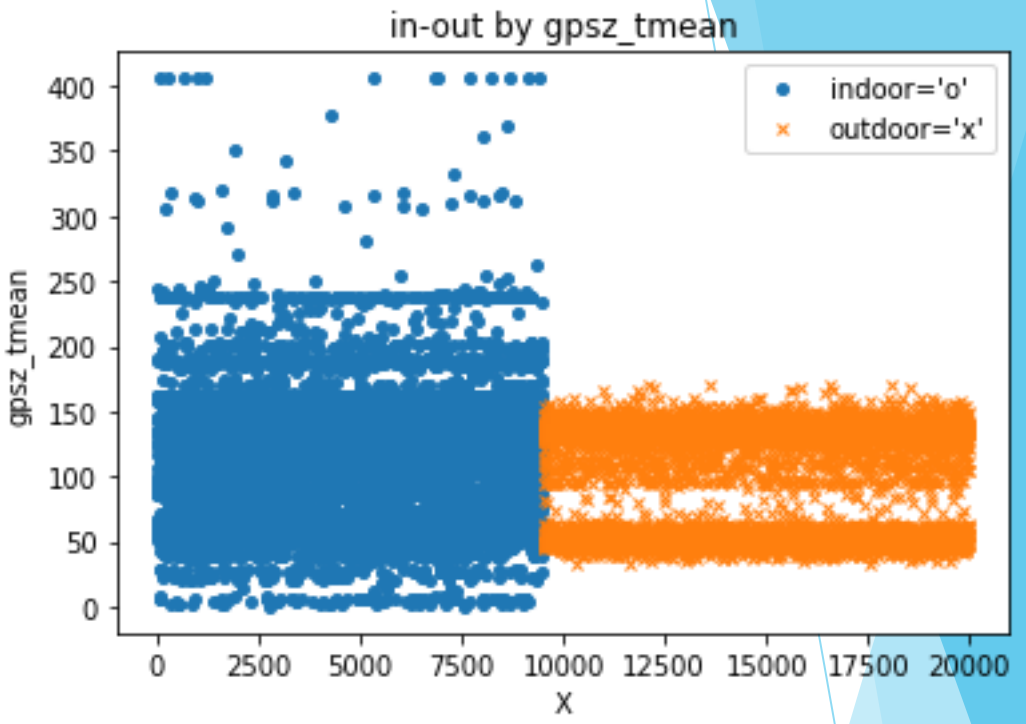
Testing data

特征 8 / feature 8

snr15num :手机检测到信噪比大于等于 15 的卫星 数量



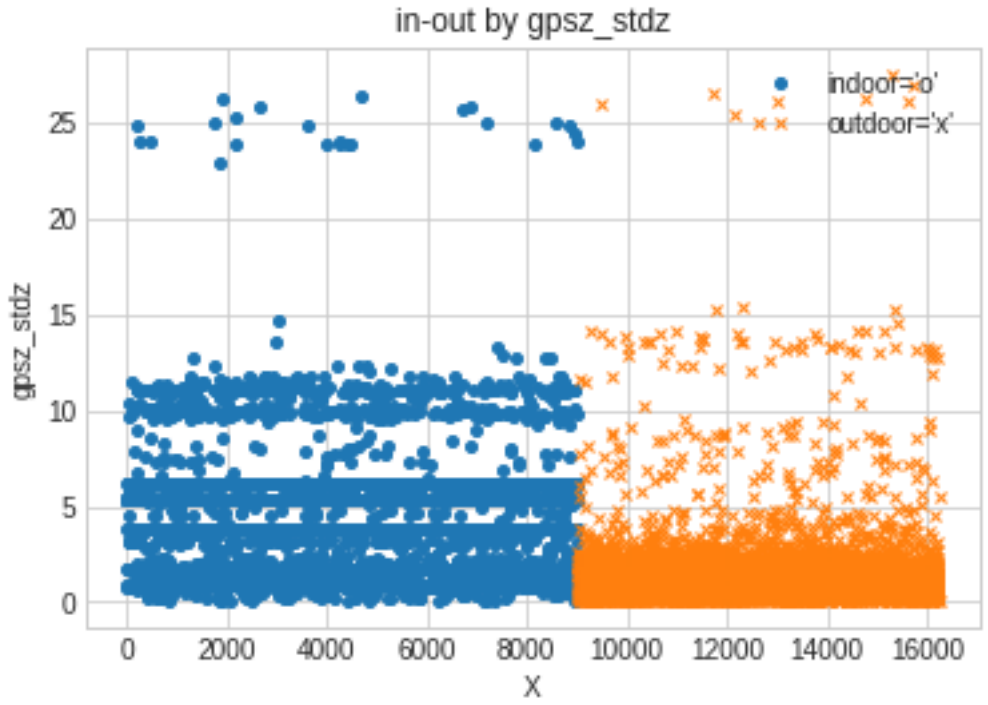
Training data



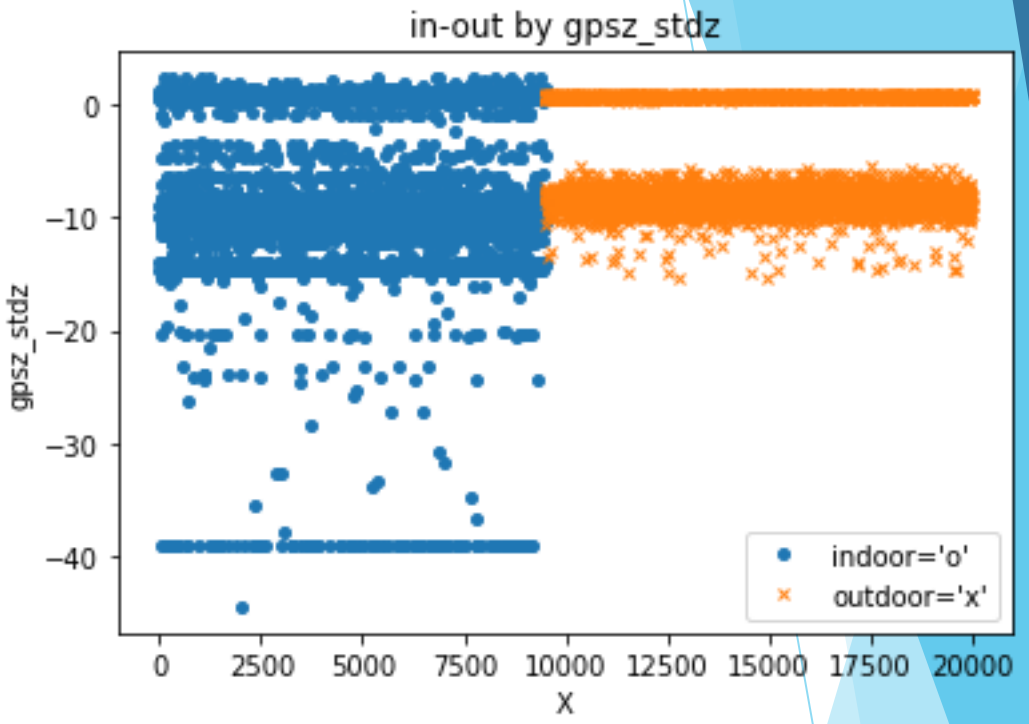
Testing data

特征 9 / feature 9

gpsz_tmean：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值



Training data

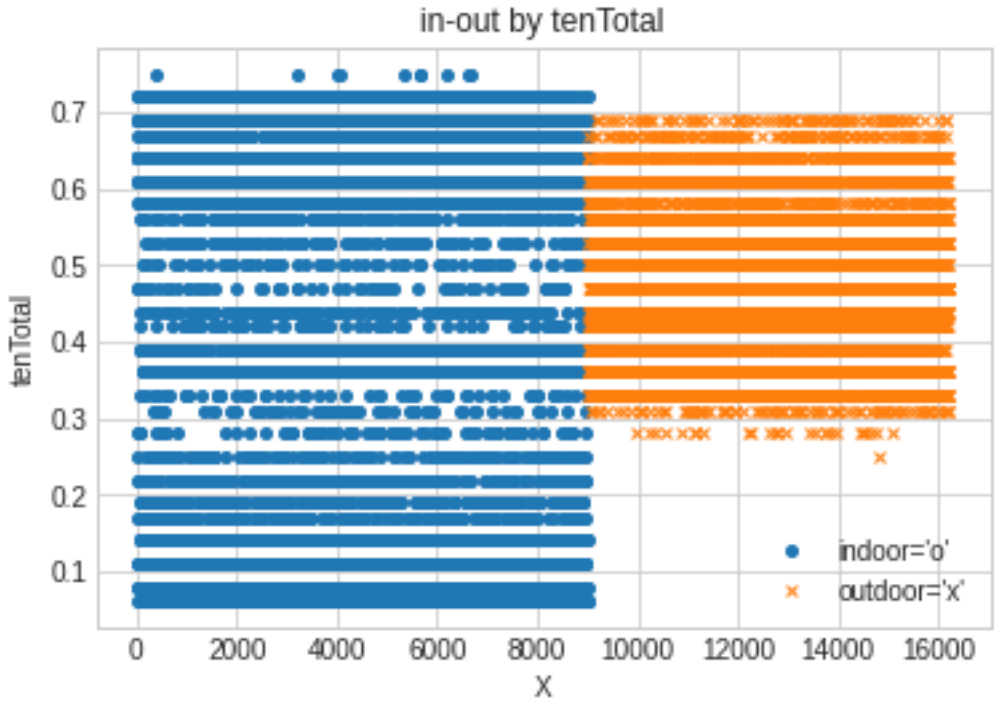


Testing data

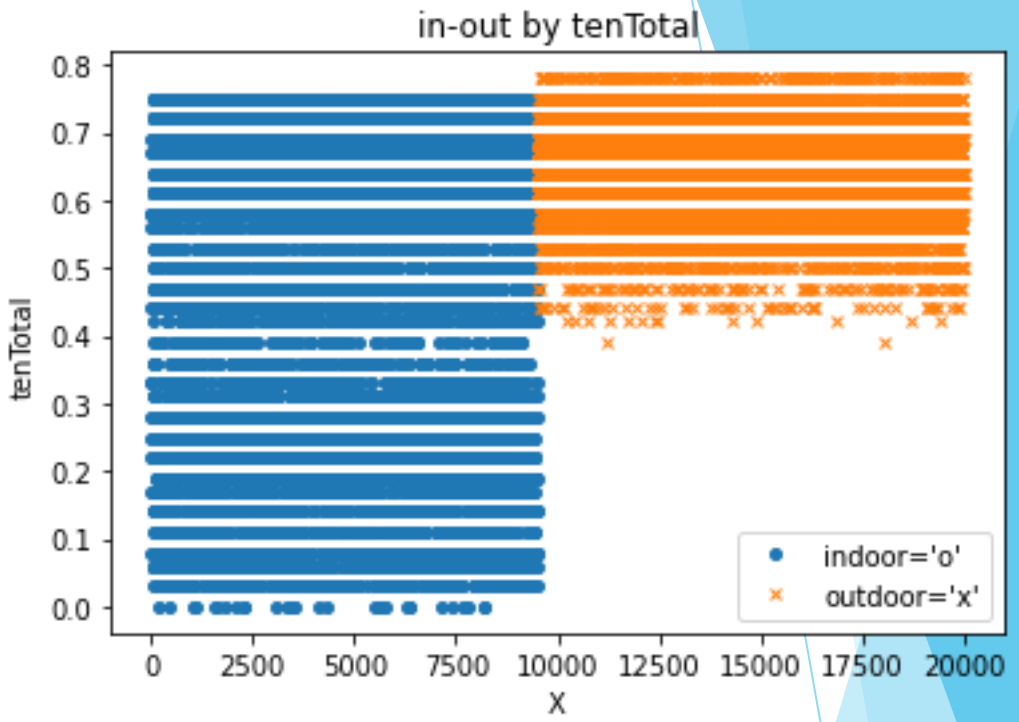
gpsz_stdz 特征处理明显不一致， 所以后面去掉这个特征

特征 10 / feature 10

gpsz_stdz：当前位置的 GPS 海拔高度与手机测试 海拔高度差的绝对值是否小于标准差 (0 or 1)



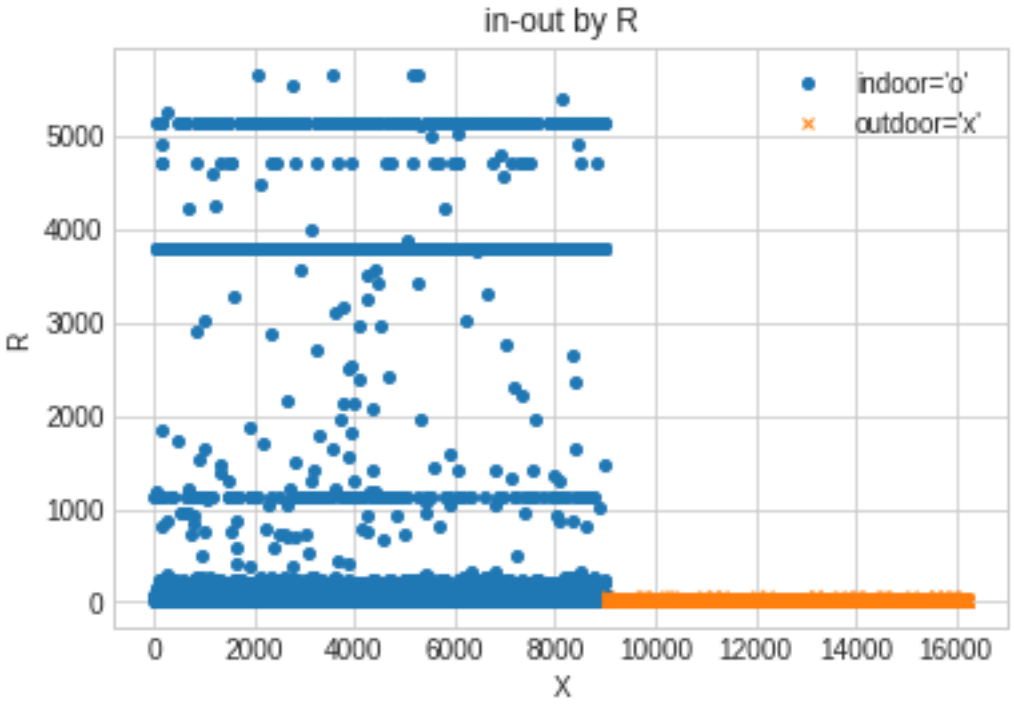
Training data



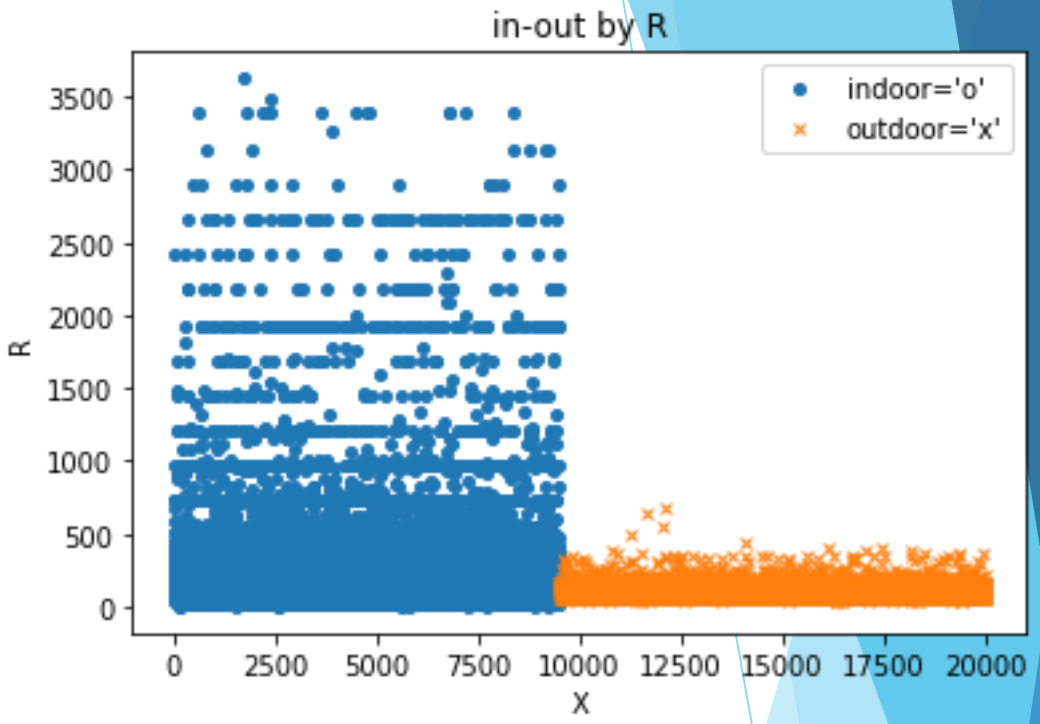
Testing data

特征 11 / feature 11

tenTotal :将 360° 分为 36 份(10°)，包含卫星的份数总和与 36 之比



Training data



Testing data

特征 12 / feature 12
R :GPS 预测误差半径

模型 调试 try different models

- 特征个数减少 reduce training features from 12 to 11, 8.
- 参数调整 parameters adjustments.

采集的训练数据集 zhuyining

采集的测试数据集 guihan

去掉 gpsz_stdz, 用 11 个特征， 4个子模型， 深度6

地铁室外跑

```
(963, 13)
(963, 11) (963,)
0.993959938366718 0.4818276220145379
              precision    recall  f1-score   support

      indoor       0.00      0.00      0.00         0
     outdoor       1.00      0.48      0.65       963

    accuracy                           0.48       963
   macro avg       0.50      0.24      0.33       963
weighted avg       1.00      0.48      0.65       963
```

地铁站内

```
(2809, 13)
(2809, 11) (2809,)
0.993959938366718 0.33855464578141686
              precision    recall  f1-score   support

      indoor       1.00      0.34      0.51      2809
     outdoor       0.00      0.00      0.00         0

    accuracy                           0.34      2809
   macro avg       0.50      0.17      0.25      2809
weighted avg       1.00      0.34      0.51      2809
```

其他
(删掉 地铁室外跑& 地铁站内)

```
(19990, 13)
(19990, 11) (19990,)
0.993959938366718 0.9173586793396699
              precision    recall  f1-score   support

      indoor       0.98      0.85      0.91      9502
     outdoor       0.88      0.98      0.93     10488

    accuracy                           0.92     19990
   macro avg       0.93      0.91      0.92     19990
weighted avg       0.92      0.92      0.92     19990
```

采集的训练数据集 zhuyining

采集的测试数据集 guihan

去掉 gpsz_stdz，用 11 个特征， 4个子模型， 深度6

用 12 个特征， 4个子模型， 深度10

地铁室外跑

	precision	recall	f1-score	support
indoor	0.00	0.00	0.00	0
outdoor	1.00	0.48	0.65	963
accuracy			0.48	963
macro avg	0.50	0.24	0.33	963
weighted avg	1.00	0.48	0.65	963

	precision	recall	f1-score	suppor
indoor	0.00	0.00	0.00	
outdoor	1.00	0.26	0.41	96
accuracy			0.26	96
macro avg	0.50	0.13	0.21	96
weighted avg	1.00	0.26	0.41	96

地铁站内

	precision	recall	f1-score	support
indoor	1.00	0.34	0.51	2809
outdoor	0.00	0.00	0.00	0
accuracy			0.34	2809
macro avg	0.50	0.17	0.25	2809
weighted avg	1.00	0.34	0.51	2809

	precision	recall	f1-score	suppor
indoor	1.00	0.38	0.55	280
outdoor	0.00	0.00	0.00	
accuracy			0.38	280
macro avg	0.50	0.19	0.28	280
weighted avg	1.00	0.38	0.55	280

其他

	precision	recall	f1-score	support
indoor	0.98	0.85	0.91	9502
outdoor	0.88	0.98	0.93	10488
accuracy			0.92	19990
macro avg	0.93	0.91	0.92	19990
weighted avg	0.92	0.92	0.92	19990

	precision	recall	f1-score	suppo
indoor	0.97	0.91	0.94	95
outdoor	0.92	0.98	0.95	104
accuracy			0.95	199
macro avg	0.95	0.94	0.94	199
weighted avg	0.95	0.95	0.94	199

Reference

- Zhu, Y., Luo, H., Wang, Q., Zhao, F., Ning, B., Ke, Q. and Zhang, C., 2019. A Fast Indoor/Outdoor Transition Detection Algorithm Based on Machine
- Learning. Sensors, 19(4), p.786.
- Huang, H., Zeng, Q., Chen, R., Meng, Q., Wang, J. and Zeng, S., 2018, March. Seamless Navigation Methodology optimized for Indoor/Outdoor Detection Based on WIFI. In 2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS) (pp. 1-7). IEEE.
- Cho, H., Song, J., Park, H. and Hwang, C., 2014, November. Deterministic indoor detection from dispersions of GPS satellites on the celestial sphere. In The 11th international symposium on location based services.
- Chen, K. and Tan, G., 2017, May. SatProbe: Low-energy and fast indoor/outdoor detection based on raw GPS processing. In IEEE INFOCOM 2017-IEEE Conference on Computer Communications (pp. 1-9). IEEE.
- Okamoto, Masayuki, and Cheng Chen. "Improving GPS-based indoor-outdoor detection with moving direction information from smartphone." In Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 257-260. ACM, 2015.

Thank you!