#### 数据描述

1.数据集介绍：

4个生产设备的传感器数据，分别存放在Fan\_BJX、JRQ、NJSB、YFJB\_JX的文件夹中。其中每个设备数据包含train.csv训练集和test.csv测试集。train.csv训练集包含该模型过去一段时间内所产生的数据信息，由于设备异常情况属于少数，故训练集默认均为正常，或包含较少异常；test.csv测试集包含该模型当前一段时间内所产生的数据信息，需要对其进行判断，找出异常段并告警。

2.数据介绍：

Fan\_BJX设备数据包含以下11个测点：

PDSFD\_1000\_02\_FDJ\_20CGB01-MW,PDSFD\_1000\_02\_YCFJ\_20HLA20CP101,PDSFD\_1000\_02\_YCFJ\_20HLA20CT301,PDSFD\_1000\_02\_YCFJ\_20HLB20CT312,PDSFD\_1000\_02\_YCFJ\_20HLB20CT315,PDSFD\_1000\_02\_YCFJ\_20HLB20CT318,PDSFD\_1000\_02\_YCFJ\_20HLB20CY101A-A,PDSFD\_1000\_02\_YCFJ\_20HLB20CY101AB-A,PDSFD\_1000\_02\_YCFJ\_20HLB20ZZ101-ZT,PDSFD\_1000\_02\_YCFJ\_20HLV21CT301,PDSFD\_1000\_02\_YCFJ\_20HLV21CT302

JRQ设备数据包含以下15个测点：

PDSFD\_1000\_02\_FDJ\_20CGB01-MW,PDSFD\_1000\_02\_GJ\_02\_DT2g,PDSFD\_1000\_02\_GJ\_02\_DT3g,PDSFD\_1000\_02\_GJ\_02\_Tdc1gs,PDSFD\_1000\_02\_GJ\_02\_Tdc1gx,PDSFD\_1000\_02\_GJ\_02\_Tdc2gs,PDSFD\_1000\_02\_GJ\_02\_Tdc2gx,PDSFD\_1000\_02\_GJ\_02\_Tdc3gs,PDSFD\_1000\_02\_GJ\_02\_Tdc3gx,PDSFD\_1000\_02\_GJ\_20LAB50CT611,PDSFD\_1000\_02\_GJ\_20LAB60CT601,PDSFD\_1000\_02\_GJ\_20LCH21ZZ101-ZT,PDSFD\_1000\_02\_GJ\_20LCH31ZZ101-ZT,PDSFD\_1000\_02\_GSB\_20LAB40CT601,PDSFD\_1000\_02\_GS\_20LAB70CT601

NJSB设备数据包含以下14个测点：

PDSFD\_1000\_01\_NJSB\_10LCA00CT301,PDSFD\_1000\_01\_NJSB\_10LCA20CP101,PDSFD\_1000\_01\_NJSB\_10LCA50CF101\_2S,PDSFD\_1000\_01\_NJSB\_10LCB10AP001-A,PDSFD\_1000\_01\_NJSB\_10LCB10AP001BP-AI7,PDSFD\_1000\_01\_NJSB\_10LCB10AP001BP-AI8,PDSFD\_1000\_01\_NJSB\_10LCB10CT351,PDSFD\_1000\_01\_NJSB\_10LCB10CT353,PDSFD\_1000\_01\_NJSB\_10LCB10CT355,PDSFD\_1000\_01\_NJSB\_10LCB10CT361,PDSFD\_1000\_01\_NJSB\_10LCB10CT363,PDSFD\_1000\_01\_NJSB\_10LCB10CT364,PDSFD\_1000\_01\_NJSB\_10LCB10CT365,PDSFD\_1000\_01\_XHSB\_10PCB20CT301

YFJB\_JX设备数据包含以下9个测点：

JT\_KFFD\_630\_01\_0100001\_1DCSOPC\_AI02154,JT\_KFFD\_630\_01\_0601212\_1DCSOPC\_AI01064,JT\_KFFD\_630\_01\_0603201\_1DCSOPC\_AI01116,JT\_KFFD\_630\_01\_0603202\_1DCSOPC\_AI01143,JT\_KFFD\_630\_01\_0603203\_1DCSOPC\_AI01144,JT\_KFFD\_630\_01\_0603205\_1DCSOPC\_AI01131,JT\_KFFD\_630\_01\_0603207\_1DCSOPC\_AI01141,JT\_KFFD\_630\_01\_0603209\_1DCSOPC\_AI01126,JT\_KFFD\_630\_01\_0603235\_1DCSOPC\_AI01129

其中各个测点（表格中的列）表示该模型的温度、气压……等，一行为一个时间节点产生的数据，并且该模型的各个测点不是相互独立的。

#### 任务描述

对测试集中的数据进行判断，找出模型哪个时间节点发生异常，即是该模型的某行任何测点异常均认为该时间节点异常需报警。

#### 评价方式

|  |  |  |
| --- | --- | --- |
|  | 实际为真的个数 | 实际为假的个数 |
| 测试为真的个数 | TP | FP |
| 测试为假的个数 | FN | TN |

此外，goudtruth文件夹包含各个模型的每个时间节点（每行）的报警情况，其中0表示正常，1表示异常。

根据groudtruth和测试结果分别计算每个模型的recall(召回率)、precision(精度)、F1 score。

同学们输出异常检测的结果（行为时间点、列为1或0，1表示异常，0表示正常），助教会根据groudtruth，计算大家的三个指标。