# 1   竞赛背景

轴承是在机械设备中具有广泛应用的关键部件之一。由于过载，疲劳，磨损，腐蚀等原因，轴承在机器操作过程中容易损坏。事实上，超过50％的旋转机器故障与轴承故障有关。实际上，滚动轴承故障可能导致设备剧烈摇晃，设备停机，停止生产，甚至造成人员伤亡。一般来说，早期的轴承弱故障是复杂的，难以检测。因此，轴承状态的监测和分析非常重要，它可以发现轴承的早期弱故障，防止故障造成损失。

最近，轴承的故障检测和诊断一直备受关注。在所有类型的轴承故障诊断方法中，振动信号分析是最主要和有用的工具之一。

在这次比赛中，我们提供一个真实的轴承振动信号数据集，选手需要使用机器学习技术判断轴承的工作状态。

# 2   竞赛任务

轴承有3种故障：外圈故障，内圈故障，滚珠故障，外加正常的工作状态。如表1所示，结合轴承的3种直径（直径1,直径2,直径3），轴承的工作状态有10类：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 外圈故障 | 内圈故障 | 滚珠故障 | 正常 |
| 直径1 | 1 | 2 | 3 | 0 |
| 直径2 | 4 | 5 | 6 |
| 直径3 | 7 | 8 | 9 |

                                                                    表1轴承的故障类别

参赛选手需要设计模型根据轴承运行中的振动信号对轴承的工作状态进行分类。

# 3   数据说明

可供下载使用的有2个文件：

1.train.csv，训练集数据，1到6000为按时间序列连续采样的振动信号数值，每行数据是一个样本，共792条数据，第一列id字段为样本编号，最后一列label字段为标签数据，即轴承的工作状态，用数字0到9表示。

2.test\_data.csv，测试集数据，共528条数据，除无label字段外，其他字段同训练集。

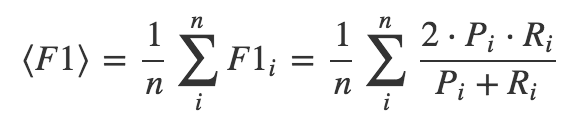
总的来说，每行数据除去id和label后是轴承一段时间的振动信号数据，选手需要用这些振动信号去判定轴承的工作状态label。

注意：同一列的数据不一定是同一个时间点的采样数据，即不要把每一列当作一个特征。

# 4   评分算法

binary-classification  
评分标准

采用各个品类F1指标的算术平均值，它是Precision 和 Recall 的调和平均数。



其中，Pi是表示第i个种类对应的Precision， Ri是表示第i个种类对应Recall。

# 5   结果提交

#### 提交说明：

1) 以csv格式提交，编码为UTF-8，第一行为表头；

2) 内含2列，一列为id，另1列为label ；

3) id对应测试集中样本的id，label为参赛者的模型预测结果

**格式如下：**

id,label

1,1

2,2

3,3

4,4

5,5

……