20170510说明文档

该文件夹内的函数主要功能为对数据格式进行整理，并提取特征值，生成特征值矩阵。

1. **combinedata.m文件**

该段程序的主要功能为对2009年挑战项目中所给的训练集测试集数据进行整理，整理前，数据按照变量名进行排序，以训练集为例，存储的是HR、ABPSys、ABPDias、ABPMean等以生理参数名称命名的mat文件，每个mat文件的维数为660×60，660表示共11个小时的数据，60表示训练集中共有60个样本。运行combinedata.m程序后，将数据整理成按照样本编号进行存储，上述训练集数据将存储为60个mat文件，即每个样本都对应一个mat文件，每个mat文件的维数为660\*7，660表示11个小时长度的数据，7表示共使用7个生理参数。

1. **latest.m文件**

该程序实现对挑战项目中训练集、测试集、以及MIMIC下载并筛选出的样本进行特征值的提取，样本总数为358个，程序实现过程如下：

1. **首先构造特征值矩阵。**样本数为358个，对于每个样本共有心率、动脉收缩压、动脉舒张压、动脉平均压、脉搏、呼吸、血氧7个生理参数，每个生理参数提取均值、中位数、标准差、偏度、百分位数、峰度、四分位数、平均绝对偏差、极差、方差、协方差共11个统计参量。因此最终生成的特征值矩阵为358×78的矩阵，第78列为标签。此处标签用1位表示，标签为‘1’表示发生急性低血压，标签为‘2’表示未发生急性低血压。
2. **数据处理，提取特征值。**按照训练集、测试集、筛选出的AHE、筛选出的非AHE的顺序，依次根据拉依达即3σ准则去除异常数据（xigma.m）、去除正常生理参数范围外的数据（mmMissingValues.m）、提取特征值（tezhengzhi.m）。
3. **Create\_eigen.m**

该文件和上面的latest.m文件功能一致，都是从11小时的数据段中生成特征值矩阵，只不过没有考虑挑战项目中的训练集和测试集，因为没法提取相应的临床参数

1. **resample.m及reSample60.m文件**

resample.m程序是30个点求一个平均值进行降采样，主要是用来对动脉压数据半小时采样一次的情况进行实验。

reSample60.m程序是60个点求一个平均值进行降采样，主要是用来对挑战项目中给定的测试集数据（1s钟采样一次）进行降采样，也变成1min采样一次。

4．数据说明

程序中用到的AHE和nonAHE两个数据文件夹在最外层的Data文件夹中