**部件/产品资源模型的预测模型使用说明**

1. 模型用途

给定一个部件/产品模型，获知制造该部件/产品所需的时间、设备、材料信息等。输入一个部件/产品模型对应的特征向量，输出预测的该部件/产品所需的时间、设备、材料信息等（资源模型）。

1. 数据说明

输入数据有两个文件，一个是训练文件，共有m行n列，每一行代表一个模型，模型的前n列为模型特征向量在每个维度的值，接下来分为三个部分，第一个部分是时间，占一列，第二个部分和第三个部分分别是所需设备和原材料，每一列代表一种设备或材料，0代表需要该设备或材料，1代表不需要该设备或材料；另一个是需要进行预测部件的特征向量存储文件，导入后可转化为测试集。

模型训练的输出文件为两个模型文件，分别以.pkl和.h5结尾。预测的输出文件是一个txt文件，其中的每一部分对应于测试文件每一行的预测结果。

1. 算法原理

部件/产品资源模型的预测模型的核心思想是利用线性回归和卷积神经网络分别对部件/产品所需的时间和设备以及原材料进行预测。本模型采用的神经网络还有两个卷积层，用relu作为激活函数，卷积之后进行一个dropout，再通过两个全连接层得到最终的输出。输出的结果维度等于设备种类与材料种类之和

1. 数据预处理

用Pandas将训练文件导入，将时间对应的列作为线性回归模型的标签，将设备以及原材料对应的列作为卷积神经网络的标签。

1. 数学模型

线性模型的形式如下



我们假设输入的特征为，对应的样本值为，用模型估计出的值为，估计值与真实值之间的误差表示为



成为损失函数，损失函数的自变量为，需要找到最小时的取值。

1. 使用方法

训练部分具体输入命令：

**python train\_model.py --train\_dir训练数据存放地址 –model\_name 训练模型保存名称 --epoch 训练迭代次数 --feature\_length 特征向量维度 --device\_num 设备种类数 --material\_num 材料种类数**

例如：

**python train\_model.py --train\_dir ./sample\_train.txt –model\_name sample--feature\_length 20 --device\_num 10 --material\_num 10**

预测部分具体输入命令：

**python predict.py --test\_dir测试数据存放地址 --result\_dir 结果保存地址 --time\_model\_dir 线性回归模型存放地址 --model\_dir 神经网络模型存放地址 --feature\_length 特征向量维度 --device\_num 设备种类数 --material\_num 材料种类数**

例如：

**python predict.py --test\_dir ./sample\_test.txt --result\_dir ./sample\_result.txt --time\_model\_dir ./sample\_time.pkl --model\_dir ./sample\_model.h5 --feature\_length 20 --device\_num 10 --material\_num 10**

预测结果会存放在**结果保存地址**中