昇科论文复现大赛报告

这一文档用以说明本工程所使用的方法，以及实验过程中的一些发现。希望可以为贵公司的相关工作人员提供参考，为进一步提升容量预测精度提供支持。

# 最终的复现精度

首先给出结论，本组所能做到的精度为：

表1 本文方法精度

|  |  |  |  |
| --- | --- | --- | --- |
|  | 训练集MAPE | 验证集MAPE | 测试集MAPE |
| 论文中结果 | 9.8% | 10.1% | 8.6% |
| CNN法 | 3.6% | 9.3% | 9.2% |

另外本组成员还探索了一种测试集精度有望达到**8.7%**的方法，但该方法仅能基于某些假设条件实现，故严格上不能作为本组成果。这种方法将在第3节详细介绍，仅供参考。

# CNN法

基于卷积神经网络的方法是目前所能做到精度最高的方法，CNN的输入为原文中提到的ΔQ100-10曲线，输出即为电池寿命（CYC）。CNN的具体结构为：

表2 CNN结构

|  |  |  |  |
| --- | --- | --- | --- |
|  | 输入通道/神经元数 | 输出通道/神经元数 | 其他信息 |
| 一维卷积层 | 1 | 6 | 卷积核大小为10 |
| 一维卷积层 | 6 | 6 | 卷积核大小为10 |
| 全连接层 | 504 | 100 | 激活函数为Relu |
| 全连接层 | 100 | 1 |  |

初始学习率设置为0.001，并随训练衰减。Batch size为8。代码详见withCNN.py。

按照比赛官方划定的训练集、测试集、验证集，训练后的CNN的预测结果与真值的比较如图1所示。

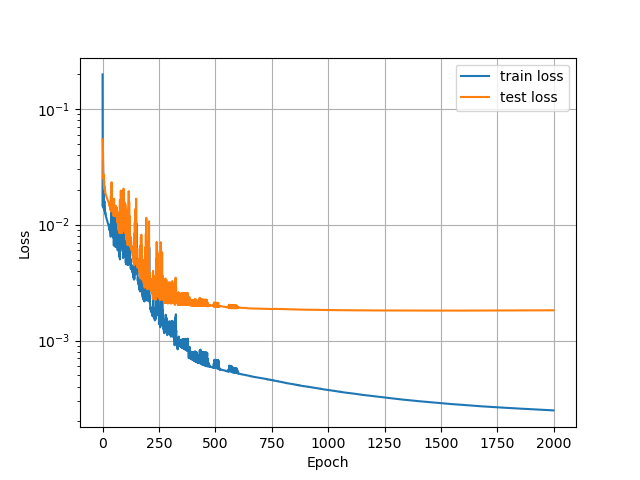
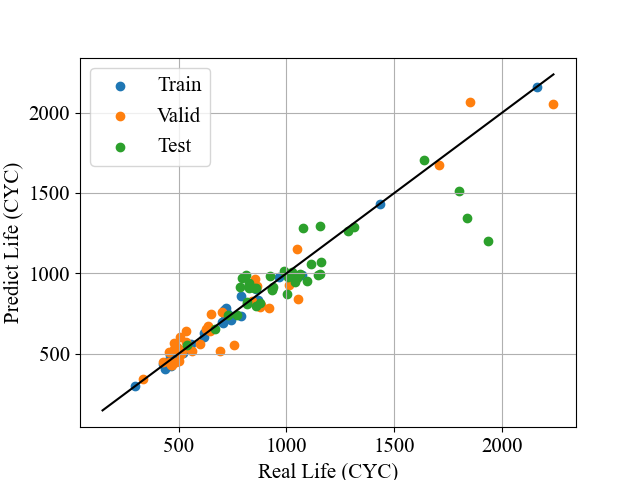


图1 （a）CNN训练结果与（b）训练过程中的误差变化

此外，为进一步提高精度，本组成员还尝试了其他的CNN方法，包括以ΔQx-10（x=20, 30, 40, …, 100）二维矩阵为输入的方法以及CNN+LSTM的方法。然而这些方法都仅能达到10%附近的测试集精度，故这里不再赘述，可参见withCNN\_2d.py。

# 长短寿命模型分离的弹性网络法

## 3.1 现象1：训练集与测试集的不同

在CNN法基础上，测试集误差之所以难以进一步下降，究其原因是训练集电池与测试集电池存在某些不同。这里观察到两个现象：

1. 训练集、验证集的电池寿命总体偏低，而测试集的电池寿命整体偏高。例如，训练集中，寿命超过1000 CYC的电池仅占9%，而测试集中占48%。这便导致基于训练集很难学习到长寿命电池的某些特性。
2. 无论是本文的图1，还是原论文中的图3，以及使用原论文数据进行研究的参考文献[1]中的图7，都能观察到测试集寿命最长的3个电池的估计精度是测试集中最差的。且三篇文章中的模型均低估了这3个电池的寿命。这也印证了现象（1）中提到的观点。

因此我们认为，在不改变训练集、测试集划分方式的前提下，很难通过设计机器学习方法或选取合适的特征，使得测试集误差有明显的下降。

## 3.2 现象2：长寿命电池与短寿命电池的不同

此外我们通过特征相关性分析的方法发现一种现象：若衡量某早期特征对寿命的相关性，会发现对于长寿命电池和短寿命电池，这种相关性很可能是不同的，甚至是相反的。例如，原文中提到的“第2~100 CYC的容量衰减曲线的线性斜率”特征，对于长寿命电池呈现出很小的负相关性（相关系数为-0.086），而对短寿命电池呈现出较大的正相关性（相关系数为0.49）。更具体的分析请见表3。（源码在Elanet.py）

表3 电池寿命与原文中各特征的相关系数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 特征 | 总数据集 | 长寿命电池  (大于1200) | 短寿命电池(小于1200) |
| 1 | ΔQ100-10最小值 | -0.885 | -0.666 | -0.829 |
| 2 | ΔQ100-10方差 | -0.909 | -0.662 | -0.869 |
| 3 | ΔQ100-10峰度 | 0.167 | 0.256 | 0.118 |
| 4 | ΔQ100-10偏度 | 0.05 | 0.317 | 0.195 |
| 5 | CAP2~100线性斜率 | 0.4 | -0.086 | 0.49 |
| 6 | intercept\_lin\_fit\_2\_100 | -0.12 | 0.508 | -0.18 |
| 7 | 第2周放电容量 | -0.061 | 0.479 | -0.09 |
| 8 | diff\_discharge\_capacity\_max\_2 | 0.036 | 0.143 | 0.005 |
| 9 | 第2~6周平均放电时间 | 0.476 | 0.637 | 0.244 |
| 10 | 第2~100周最小内阻 | -0.329 | -0.31 | -0.15 |
| 11 | diff\_IR\_100\_2 | 0.238 | 0.375 | 0.24 |
| 12 | ΔQ5-4最小值 | -0.194 | -0.14 | -0.26 |
| 13 | ΔQ5-4方差 | -0.558 | 0.448 | -0.68 |

注：

橙色：对长短寿命电池均有较高相关性，且相关性相同

绿色：仅对短寿命电池有较高相关性

蓝色：仅对长寿命电池有较高相关性

黄色：对长短寿命电池均有较高相关性，但相关性相反

这也说明了，短寿命电池与长寿命电池之间存在某种特征相关性上的不同，应当使用不同的模型分别表征。若将长寿命和短寿命电池一并归于同一模型，则有可能导致与物理机理完全相悖的训练结果，例如表3中，若将所有电池混在一起，则“第2周放电容量”将与电池寿命呈负相关特性（相关系数为-0.061），这显然是有违常识的。

## 3.3 长短寿命双模型的训练与测试

根据3.2节介绍的内容，应当分别训练两个模型，一个用于长寿命电池，一个用于短寿命电池。然而在实际使用中，如何判断某电池是长寿命还是短寿命呢？如何判断该用哪个模型来估计这个电池的寿命呢？这一问题将在3.4节讨论。我们暂时假设认为我们已知了所有电池为长寿命or短寿命。

如3.1所说，想要训练一个适用于长寿命电池的模型是困难的，这是因为训练集中寿命大于1000 CYC的仅有4个样本。为此我们使用形式简单的Elastic Net来训练长寿命模型。根据表3，选取特征1、特征2、特征9、特征13作为模型的输入特征。详见（double\_model\_expect.py）

而对于短寿命电池模型，其样本量相对充足，我们使用CNN进行训练（详见short\_life\_model\_CNN.py）。其输入仍为ΔQ100-10曲线。

长短寿命电池模型、及其总体误差的统计如表4所示。其中总体误差通过长短寿命模型的误差按长短寿命样本个数加权得到，这保证了其与单模型方法误差的可比性：



表4 双模型误差

|  |  |  |  |
| --- | --- | --- | --- |
|  | 训练集MAPE | 验证集MAPE | 测试集MAPE |
| 长寿命样本（Elastic Net） | 3.77% | 6.72% | 10.06% |
| 短寿命样本  （CNN） | 3.97% | 10.67% | 7.46% |
| 综合误差 | 3.95% | 10.09% | 8.69% |

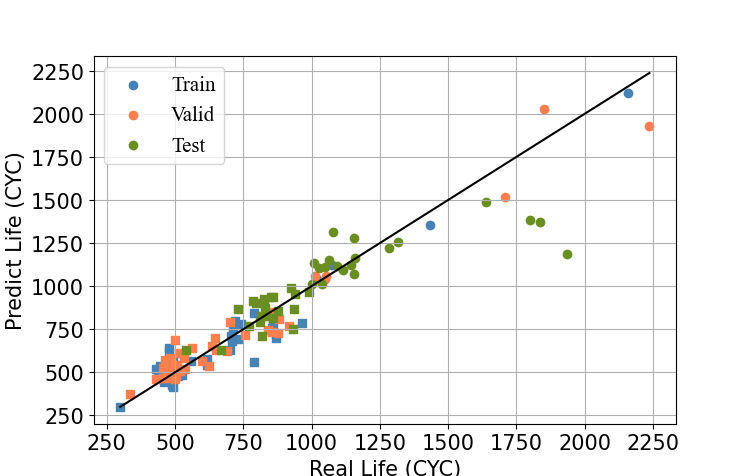


图2 长短寿命双模型下的预测值与真值对比（方形为短寿命模型，圆圈为长寿命模型）

从表4可见，分双模型的方法有助于测试集误差的下降。

## 3.4 电池寿命长短的初步分类

本节回答3.3节提到的，如何确定该用哪个模型来预测某电池的寿命的问题。也即电池的寿命长短的初分类方法。只有这个初分类完全正确，才可能达到表4所示的精度。

本组尝试了如下方法：使用CNN初步预测电池的寿命，并根据CNN的预测结果判断属于长寿命电池还是短寿命电池。而若CNN的预测结果落在长短寿命的分界线附近（本文取1000 CYC），则使用长短寿命模型同时预测，并加权得到最终的预测结果。这一整体流程如图3所示。

在图3所示的框架下，测试集的MAPE仅能达到9.28%（参见double\_model\_real.py）。即初步寿命估计模型的设计依然不够准确，导致无法实际达到表4的高精度。但这仍比仅使用单个弹性网络的方法具有更高的精度（虽然原文中给出的最优测试集精度为9.1%，但我们组用单个弹性网络所能复现出的最高精度仅有12%左右，详见Elanet.py），即使将短寿命模型的CNN也替换为简单的弹性网络也是如此（10.26%的MAPE）。

后续可考虑改进初步分类模型的设计，例如通过集成学习的方法获取电池为长寿命或短寿命的概率分布。

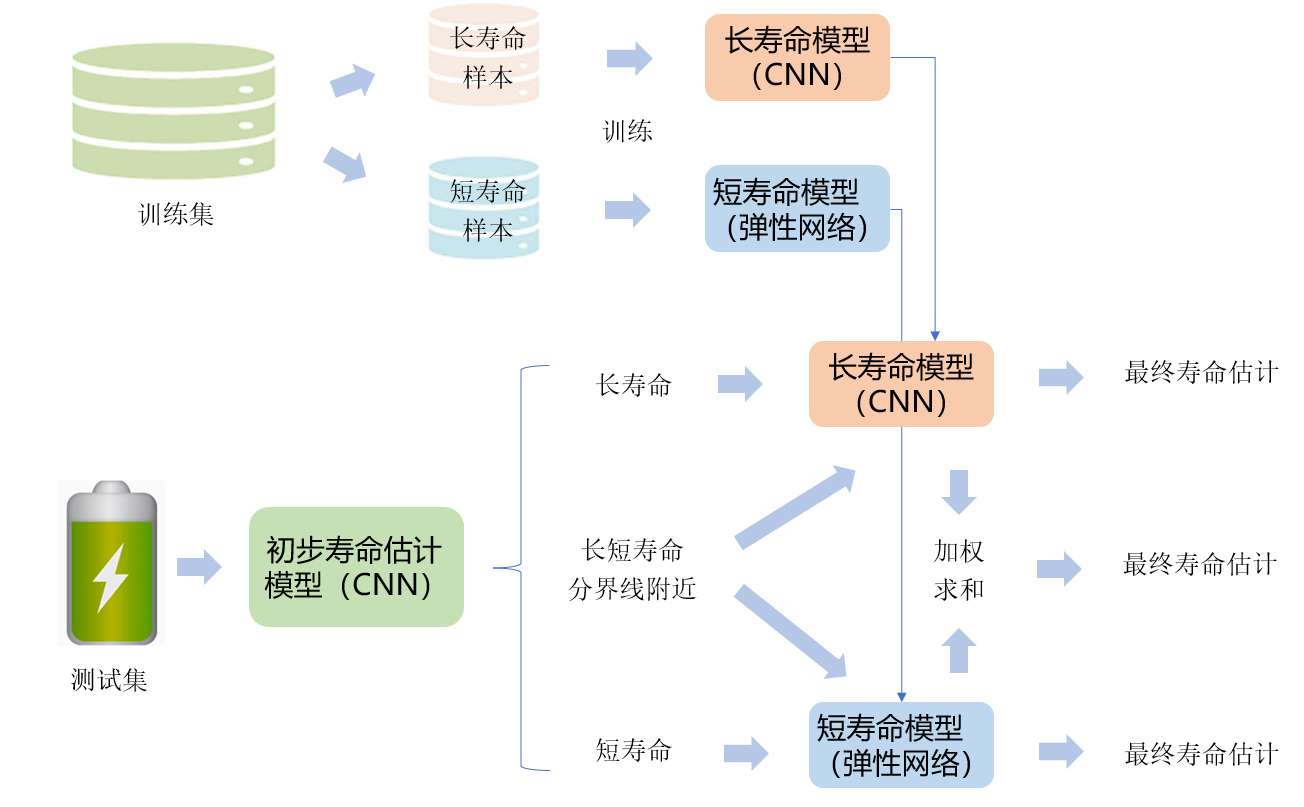


图3 双模型法的整体框图

# 参考文献

1. Saxena S, Ward L, Kubal J, et al. A convolutional neural network model for battery capacity fade curve prediction using early life data[J]. Journal of Power Sources, 2022, 542: 231736.