使用数据驱动建模方法预测电池寿命

Vikram C Patil，北卡罗来纳州立大学，美国罗利。

摘要

如今，电池无处不在，应用范围从智能手机、手表和笔记本电脑，延伸到电动汽车、无人机和电动飞机。锂离子电池因其能量密度高、可充电且生命周期成本低，被广泛应用于这些领域。了解锂离子电池的寿命，对于在许多领域有效使用电池至关重要。本研究探讨了数据驱动建模方法，利用各种可测量的电池参数预测锂离子电池的寿命。研究使用了 NASA 电动飞机实验的电池数据集，其中包括 17 个预测变量，以及作为响应变量、用于表示电池寿命的剩余飞行时间。该数据集包含超过 4,000,000 行数据。然而，原始数据集提供的有关电池随时间使用情况的直接有效信息有限，因此通过特征工程生成了信息量更丰富的变量。此外，研究采用主成分分析（PCA）进行降维，通过选取较少数量的主成分作为模型开发的预测变量，降低计算成本和模型复杂度。研究利用工程化特征，探索了随机森林和神经网络模型在电池寿命预测中的应用。研究评估了多种神经网络配置，包括具有不同节点数量的单隐藏层和双隐藏层架构。研究采用测试数据集上的均方误差（MSE）作为模型比较的性能指标。结果表明，数据驱动建模方法能够有效预测电池寿命；根据均方误差指标，神经网络模型的表现优于其他模型。此外，神经网络在处理高维电池数据时表现出良好的稳健性。总的来说，利用数据驱动的机器学习技术，可以有效地对电池的复杂行为建模，并以合理的准确度预测电池寿命。未来的相关研究有望进一步提高预测准确率、计算效率和模型可靠性。

1. 引言

电池，尤其是可充电锂离子电池，广泛应用于各种领域，例如智能手机、笔记本电脑、无人机、电动滑板车、电动汽车、公共汽车和飞机等（Pistoia, G. 2013年）。电池还通过减少化石燃料的使用、增加可再生能源资源的利用，在遏制全球变暖和气候变化方面发挥重要作用（Climate Central, 2019年）。与其他电池类型相比，锂离子电池具有能量密度高、寿命长等特点，因此在许多应用中更受青睐（Yoshio, M., Brodd, R.J. and Kozawa, A. 2009年）。对于锂离子电池的许多应用而言，预测电池寿命对于可靠使用和评估财务可行性至关重要。例如，汽车制造商或电动汽车车主希望了解汽车所用电池的使用寿命。如果电池寿命明显短于汽车的使用寿命，就需要规划更换电池。相反，如果电池的使用寿命超过智能手机本身的寿命，那么这种长寿命可能并无必要，还会给用户带来过重成本。因此，针对各种电池应用预测锂离子电池寿命，对于合理使用电池具有重要价值。此外，了解寿命预测中的不确定性，对于确定这些电池的保修期限也很有帮助。通过量化寿命预测中的不确定性或概率，锂离子电池制造商可以为电池提供故障概率较低的保修方案，从而降低因电池故障导致保修索赔的财务风险。

通常有多种方法可用于对电池寿命进行建模。图1（Hosen 等，2021年，第102060页）展示了三种主要方法及其特点。第一种是电化学模型，也称为基于物理的建模，因为它以物理定律为基础，用于描述电池在使用过程中老化的各种机制。这种基于物理的方法只需有限的实验数据，就能构建出准确度很高的模型；不过，该方法较为复杂，计算成本高，而且不太适用于现实条件。第二种方法是半经验建模方法，它使用较为简单的数学方程，以简化形式描述电池内部的物理过程。半经验模型通常需要高质量且

数量充足、涵盖多种应用条件的实验数据。这些数据用于调整模型中的参数。与电化学模型相比，半经验模型计算上更简便，但其模型准确度可能明显较低。随着数据科学领域的发展，第三种受到积极探索的主要建模方法是数据驱动建模。在数据驱动建模方法中，采用先进的数据科学技术来对电池老化和寿命进行建模。数据驱动建模只需有限的电池物理机制知识，却能构建高度复杂的模型，适用于多种应用场景并具有很高的准确度。事实证明，数据驱动模型易于开发，但通常需要大量涵盖各种运行条件的数据。

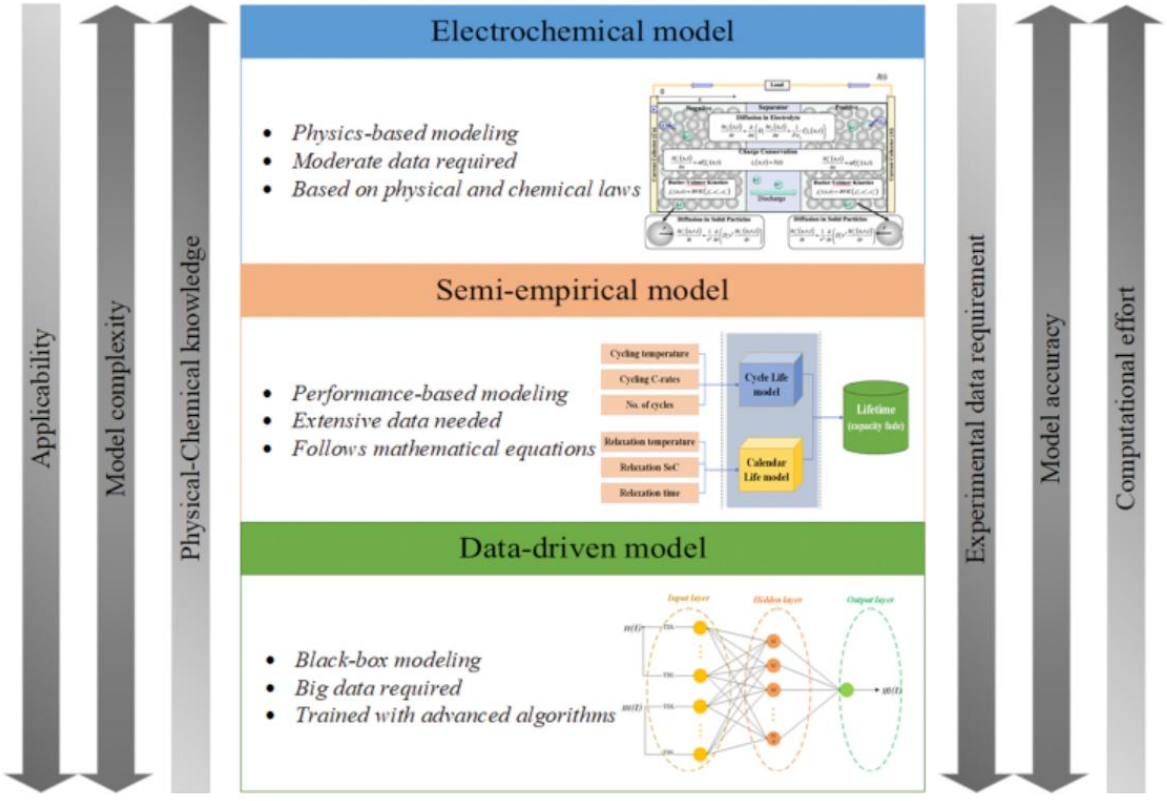


图1。 电池寿命建模方法一览。（来源：Hosen 等，2021年，第102060页）

随着技术进步以及人们对可再生能源的认识不断提高，电动汽车日益普及。然而，电动商用飞机的梦想仍远未实现。使用电力推进系统为商用飞机提供动力面临诸多挑战，例如缺乏实用的冷却系统、发动机重量过大等。但主要的限制因素是电池，尤其是如何估算电池的剩余运行时间，进而估算飞机的剩余飞行时间。必须预测剩余飞行时间，以便在电池荷电状态（SOC）距离低于安全阈值 30% 还有两分钟时，提醒飞行员启动着陆程序。美国国家航空航天局（NASA）正在开展相关研究以解决这一问题。本项目将利用电动飞机的电池数据，通过数据驱动机器学习模型预测电池寿命。

1.1. 目标

利用电池现场运行数据，开发并确定适用于预测锂离子电池寿命的数据驱动机器学习模型。

1.2. 研究目标

i) 分析并识别电池寿命数据中适用于以合理精度预测寿命的特征。ii) 开发并测试各种机器学习模型，以预测电池寿命。iii) 根据不同的模型评估指标比较并评估各种模型的有效性。iv) 评估各种机器学习模型在预测电池寿命方面的有效性和局限性。v) 识别并评估适用于数据驱动电池寿命建模方法的潜在不确定性估计技术。

1.3. 研究问题

i) 数据驱动建模是否是以合理的准确度预测电池寿命的有效方法？

ii) 是否存在能够在训练数据集所涵盖条件之外有效预测电池寿命的机器学习模型?

iii) 数据驱　　动模型能否为电池寿命预测提供不确定性度量?

2. 文献综述

过去，由于数据可用性有限、数据生成成本较高以及数据驱动模型难以解释，数据驱动建模方法很少用于电池建模。许多早期关于电池寿命建模的研究基于半经验建模方法，使用在覆盖不同运行条件的电池上生成的实验数据。Zhang, Liang, and Zhang（2017，第012134页）概述了电池寿命预测的不同方法。

机器学习技术最近已被探索用于锂离子电池的寿命预测应用。Severson等人（22019，第383-391页）采用数据驱动建模方法，利用显著容量衰减之前的数据来预测电池循环寿命。在Severson等人的工作中，早期循环的放电电压曲线被用于按循环寿命预测和分类电池单体。通过对原始数据进行线性和非线性变换所创建的特征被用于训练弹性网络。该模型在124组商用锂离子电池数据上预测电池循环寿命时显示出有效性，测试误差为9.1%。

Zhu、Zhao 和 Sha（2019，第 e98 页）的另一项研究使用了 138 个磷酸铁锂/石墨电池单体的记录数据。他们构建了一个分类模型，将电池单体预测为“低寿命”和“高寿命”两个不同类别。研究拟合并比较了多种机器学习方法，包括决策树、K近邻、神经网络、AdaBoost、高斯过程回归、随机森林、支持向量机和朴素贝叶斯，以获得更高的分类准确率。图 2 比较了该研究中各种机器学习方法在同一数据集上的预测准确率。许多方法都表现出较高的准确率，其中决策树模型的测试准确率指标最佳。尽管该方法能够以合理的准确率将电池单体分为低寿命或高寿命类别，但无法预测以年为单位的电池寿命。

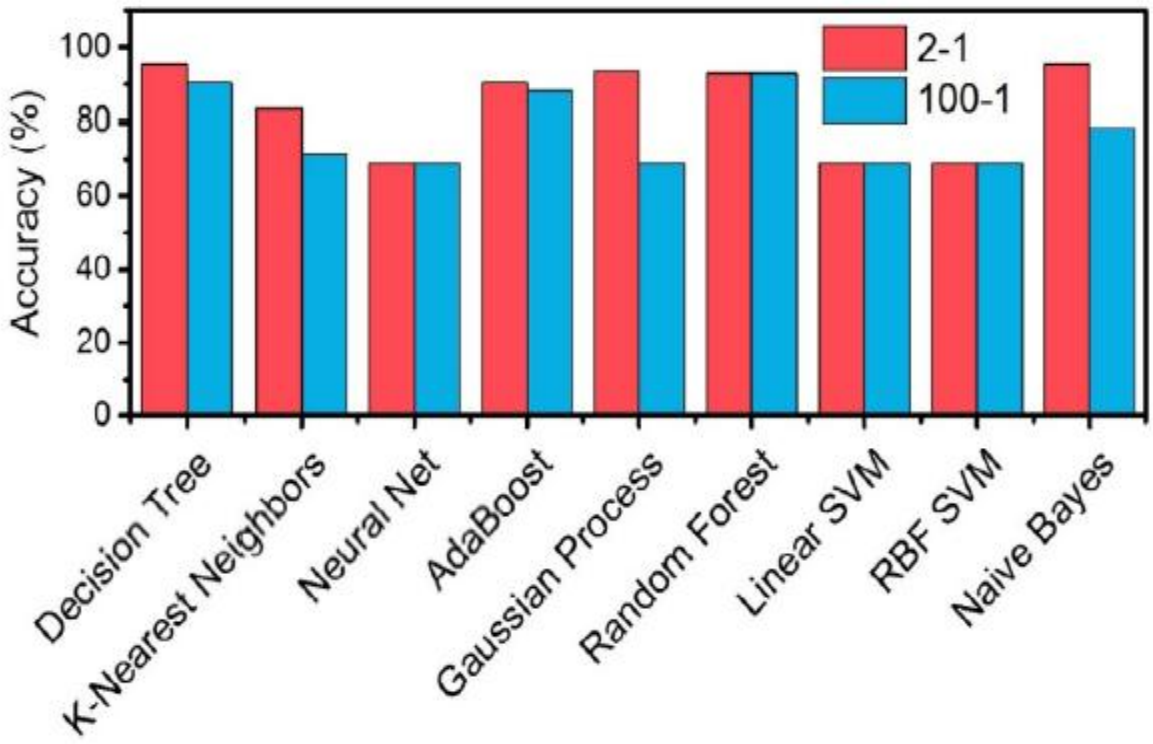


图 2。不同机器学习算法在同一测试数据集上的预测准确率比较。图例“2-1”和“100-1”表示两组特征组合。（来源：Zhu, Zhao and Sha 2019年，第 e98 页）。

在最近的一项研究中，Hosen 等人（2021，第 102060 页）比较了多种电池寿命建模方法。研究将两种基于机器学习的模型——分别基于高斯过程回归（GPR）和人工神经网络（ANN）——与半经验模型进行了比较。研究构建了采用指数核的高斯过程回归模型，并使用了带外部输入的非线性自回归动态网络（NARX）人工神经网络模型。模型评估结果表明，在均方根误差（RMSE）、平均绝对误差（MAE）和计算时间指标方面，NARX 人工神经网络模型表现最佳。值得注意的是，这两种数据驱动的机器学习建模方法的准确率均高于半经验模型。

在另一项近期研究中，Fie 等人（2021，第 120205 页）提出了一个全面的机器学习（ML）框架，包括特征提取、特征选择，以及基于早期循环数据的机器学习电池寿命预测。研究针对仅从前 100 次循环中提取的 42 个特征，考察了六种基于机器学习的模型，包括弹性网络、高斯过程回归（GPR）、支持向量机（SVM）、梯度提升（GB）回归树、随机森林和神经网络。结果显示，弹性网络、GPR 和 SVM 模型的测试结果优于 GB 回归树、随机森林和神经网络。此外，基于包装器的特征选择已被证明能够提升所有模型的预测性能。尽管许多早期研究已展示数据驱动模型在电池寿命预测方面的有效性，但许多此类方法尚未证明其能够可靠地预测模型开发所用数据范围之外的运行条件。对于发展迅速的电池行业而言，亟需利用有限数据预测电池训练数据运行条件范围之外的寿命。

此外，锂离子电池数据驱动寿命建模这一主题也受到许多致力于推动电池技术发展的研究人员关注。美国国家可再生能源实验室正致力于开发一种基于物理的机器学习方法，以提高寿命预测准确率（美国能源部，2022）。此外，美国阿贡国家实验室（ANL）正在开发一种全面的人工智能驱动方法，旨在利用有限的数据预测电池寿命和退化趋势（阿贡国家实验室，2022）。随着电池用途日益广泛，许多应用都要求电池拥有更长的寿命，因此，准确预测电池寿命对电池行业和学术界都具有重要意义。

3. 数据集

目前公开可用的电池数据集有限。一篇文章（BatteryBits，2020）评述并比较了来自不同来源的14个公开电池数据集。斯坦福大学的循环寿命预测数据集非常适合用于测试数据驱动方法。图3（BatteryBits，2020）展示了该数据集中135个在不同运行条件下接受测试的锂离子电池单体的容量老化数据。锂离子电池的寿命表示为

达到电池容量初始容量某一百分比（通常为70%或80%）所需的时间或循环次数。电池在使用或静置过程中，电池容量会降低。图中的黑色虚线表示达到初始容量80%的寿命条件，实心圆点表示各曲线对应的寿命终止时间。可以看出，该数据集中的寿命从几百次循环到几千次循环不等。这些差异可能是电池在使用过程中经历的运行条件不同所致。该数据集仅提供了公开数据集中电芯数量最多（共135个）的电池的循环寿命信息，并包含丰富的特征信息。为了获得包含丰富信息的真实世界数据集，重要的是获取应用中电池的现场数据。

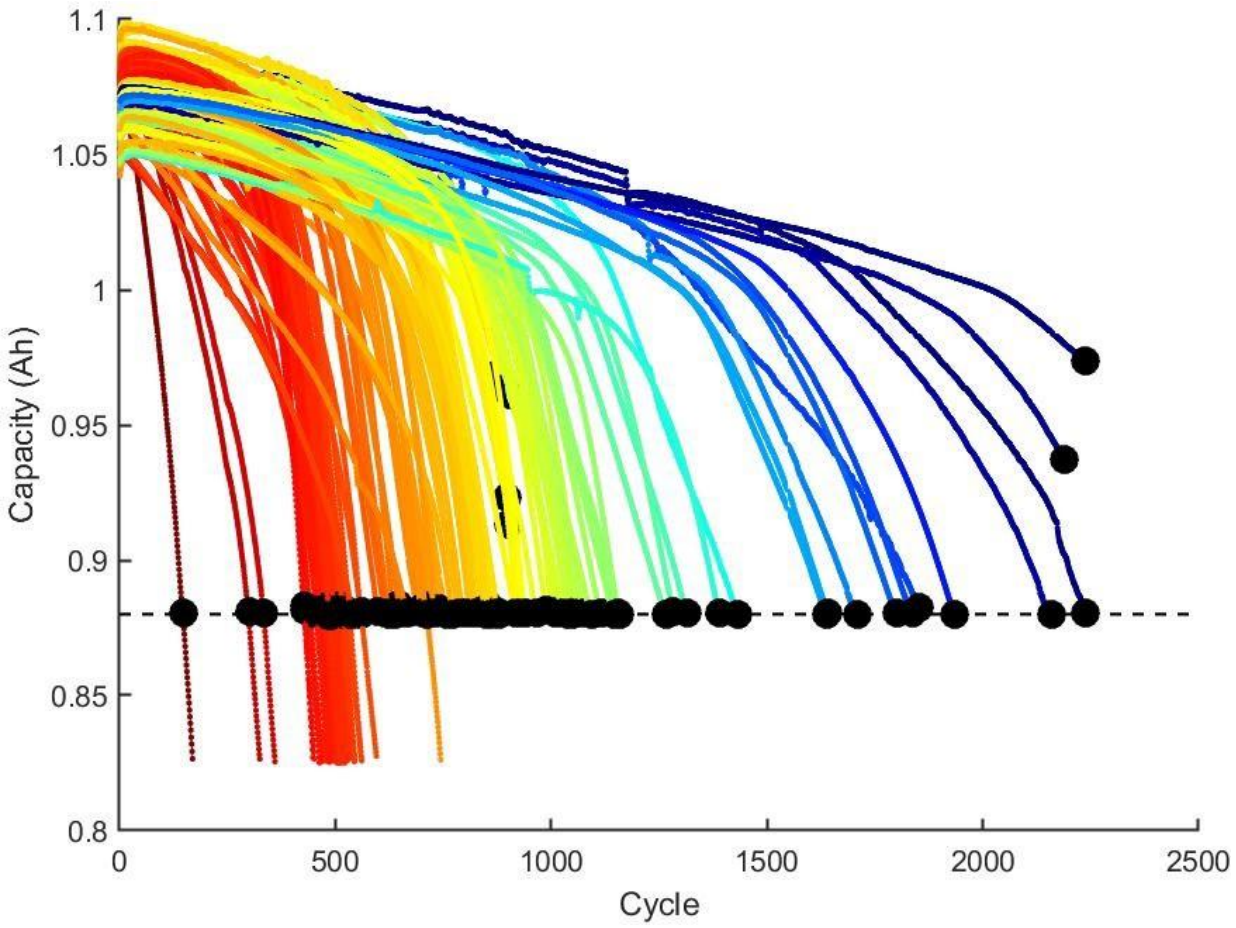


图3。在不同运行条件下测试的135个锂离子电池单体的容量老化。（来源：BatteryBits，2020）

我们尝试从各种来源获取数据，并将其转换为可用格式。然而，许多电池数据以编码格式提供，难以使用。此外，只有少数来源提供了数据变量的详细说明。获取包含电池寿命等详细信息的数据非常困难。一个易于获取但需要大量数据处理才能转换为正确格式的独特数据集来自NASA的预测卓越中心实验室。

电动飞机通常会记录大量详细的电池数据。从电动飞机记录的电池数据不仅信息丰富，而且高度属于真实世界数据。该应用要求更高的安全标准，因此数据质量通常较高。因此，本项目考虑使用来自电动飞机的数据。具体而言，数据来自NASA对Electric Edge 540飞机开展的地面测试实验（NASA故障预测数据仓库）。这些是特技飞行器，实验中使用的是其缩比版本。这些实验可获得大量数据。在我们的分析中，我们使用了来自9次实验的数据，约4,000,000行。

3.1. 变量

该数据集包含 18 个变量，用于描述其中包含的三类主要测量值：i) 飞机的剩余飞行时间估计——该变量即电池寿命，因此是本分析的响应变量。其单位为秒。ii) 飞机状态——三个变量，即“每分钟转数（RPM）”、“前电机受控传感器”和“后电机受控传感器测量值”，用于测量飞机在特定时刻的状态。iii) 电池 SOC 估计——14 个变量用于描述飞机所用不同电池的荷电状态（SOC）。电池的荷电状态由特定时刻的电压、电流和温度测量值来描述。

这些变量通常作为电池管理系统记录的一部分进行测量和存储。这些变量既包括直接测量的变量，也包括间接计算的变量。响应变量“剩余飞行时间”和飞机状态是直接测量的。电池SOC估计是根据电池电压、电流和温度的直接测量计算得到的变量。直接变量通常不存在显著的测量

误差。间接变量很多时候既存在因用于计算该间接变量的直接测量变量而产生的测量误差，也存在因估计而产生的误差。

全部18个变量的名称列表如下。

响应变量：

- 剩余飞行时间估计，即电池寿命（以秒为单位）

描述飞行器状态的变量：

- 每分钟转数（RPM）
- 前向电机控制器传感器（FMC）
- AMC – 后向电机控制器传感器

描述电池荷电状态的变量：

- 左下前方 LLF20V – 电池电压
- 左上后方 ULA20V – 电池电压
- LRF40V – 右下前方 – 串联组合电压
- URA40V – 右上后方 – 串联组合电压
- LRF20V – 右下前方 – 电池电压
- URA20V – 右上后方 – 电池电压
- LLF20C – 左下前方 – 电池电流
- ULA20C – 左上后方 – 电池电流
- LRF40C – 右下后方 – 串联组合电流
- URA40C – 右上后方 – 串联组合电流
- LLF20T – 左下前方 – 温度
- ULA20T – 左上后方 – 温度
- LRF40T – 右下前方 – 串联组合温度
- URA40T – 右上后方 – 串联组合温度

4. 方法论

将可用数据集划分为训练、测试和验证数据集，用于训练、测试和验证模型。使用训练数据集训练各种预测机器学习模型，以预测电池的循环寿命。首先测试一些简单模型，例如回归、决策树和单层小型神经网络，以了解简单模型在预测电池寿命方面的有效性和准确度。结果发现，简单模型效果不佳，其预测结果并不优于随机预测。此外，还尝试了高斯过程回归、随机森林、深度学习等复杂模型，以检验复杂机器学习模型能否提高模型准确度。然而，初步结果显示，复杂机器学习模型在预测电池寿命方面同样效果不佳。

经过多次建模和预测尝试均未成功后，我们发现有必要进行特征选择。我们对各个输入变量之间的相关性进行了分析，发现许多变量存在冗余且高度相关。此外，可提供有助于提高预测效果的有用信息的特征数量有限。因此，我们深入开展了特征选择，以确定具有影响力的变量。部分飞行器状态变量需要经过预处理，以纳入电池使用的整体历史，而不只是瞬时测量值。因此，我们推导出包含累积状态变量的特征，以评估电池使用历史。此外，我们还探索了使用主成分分析（PCA）减少无影响特征的可能性。该数据集规模庞大，包含超过 4,000,000 行和 18 个变量。在基于电池数据集开展数据驱动机器学习建模时，特征选择被证明是至关重要的过程。此外，我们还测试了生成和使用合成数据集的可能性及其在电池寿命预测中的有效性。然而，我们发现没有必要使用合成数据集，因为原始数据集规模已经足够大，能够较好地对电池寿命进行建模。

经过特征选择后，随机森林和深度学习模型开始展现出更高的预测准确率。随后进一步调整这些模型，以优化各个

模型的准确率。最后，基于均方误差、绝对误差和计算时间等性能比较指标，对两种模型在新的测试数据集上的预测误差进行比较，以确定最适用于电池寿命预测的建模框架。

以下是本项目采用该方法论预计取得的成果列表。

i) 影响电池寿命的重要特征。这些信息有助于识别有用的数据集，并通过减少不必要的信息来缩短计算时间。此外，这将指导通过设计包含重要输入特征的实验来生成新数据，从而在未来构建更优模型。

ii) 用于预测寿命的各种数据驱动模型，以及在未见数据上采用共同指标进行的比较。

iii) 将使用表现最佳的模型来预测电池寿命。对于指定电池工况的新输入特征，机器学习模型将提供寿命估计值。

4.1. 伦理考量

所使用的数据集可在 NASA 网站上的 NASA 故障预测数据仓库中公开获取。数据集来源已在参考文献中注明。此外，该数据集不包含任何可能违反伦理规范的信息。为避免数据集中的偏置，我们从大量电池文件中随机抽取文件，并检查变量是否存在不一致。未发现数据集中的变量存在不一致或偏置。此外，在特征选择过程中，我们发现不同电芯上识别出的特征相同，且数据集中未观察到偏置。我们还针对状态变量进行了特征工程，以确保所得变量不存在任何额外偏置。项目执行期间，我们充分考虑了伦理考量，并恰当地呈现了所有可用信息。此外，项目实施过程中使用了开放评估计算库。计算工作使用个人计算平台完成，未对任何实体或个人造成损害。我们根据项目数据中可用的信息开展了细致分析，并如实传达项目结果。

4.2. 挑战

电池寿命受多种因素以非线性且复杂的方式影响，因此难以建模。公开可用的电池寿命数据集非常有限，因为测试电池既成本高昂又耗时。此外，受竞争、版权和许可限制，许多电池制造商不会公开提供或发布寿命数据。因此，公开可用的数据既不丰富，也不够多样。此外，可用数据还需要包含许多重要特征。电池由高度复杂的机制支配，要在数据驱动模型中捕捉其中的大部分机制，可能需要包含丰富特征细节的数据集。另外，数据集中存在少量错误值和缺失值。处理这些值颇具挑战，因为删除相关数据会进一步减少数据量。此外，包含重复测量的数据对于量化预测中的不确定性（包括样本间变异）很有价值。然而，由于没有重复测量数据，因此无法进行不确定性评估。

所使用的数据集是真实电池数据集，其中包含许多可能不会直接影响电池寿命的原始测量值。在未进行任何特征工程、直接基于原始数据进行模型拟合时，拟合效果和预测结果都不理想，这一点得到了验证。难点在于进行特征工程，以识别出能够以适当形式提供信息、适用于模型拟合并预测电池寿命的特征。在深入了解输入变量和响应变量的属性后，确定了一种合适的特征工程方法，用于创建新的衍生输入变量，从而纳入电池的运行历史。这一点至关重要，因为电池寿命不仅取决于瞬时变量，还取决于各种变量的历史变化。

5. 结果与讨论

5.1. 数据预处理

从数据源获取的数据集仅以每项电池测试对应一个独立文件的形式提供。此外，变量并未存储在有组织的表格或数据框格式中。加载数据时发现，数据以字典格式载入。这是因为原始格式的数据采用mat文件格式。除此之外，与电池荷电状态估计值相关的变量还存储在单独的结构中。因此，有必要进行数据预处理，将这些非结构化数据转换为结构化表格或Python数据框格式。在数据预处理中，将电池荷电状态变量追加到其他变量中，以便在表格格式中包含全部数据。每个可用的数据文件都针对每项电池实验执行了这一操作。随后，使用Matlab脚本将表格格式的数据转换为CSV文件。每项实验对应的CSV文件用于在Python环境中访问数据并进行进一步分析。加载并合并每项实验的数据文件后，数据规模显著增大，最终超过4,000,000行和18列。

5.2. 特征工程

使用原始形式的数据时，模型拟合效果较差。因此，我们发现，原始形式的因变量无法为构建良好模型提供有价值的信息。需要进行特征工程，将数据变量转换为能够包含有关电池寿命状态的有价值信息的形式。原始数据格式包含有关电池使用的瞬时信息，但电池寿命取决于使用历史。因此，有必要将有关电池使用的瞬时信息转换为能够反映电池使用历史的变量。

由于该数据集中的响应变量依赖于预测参数的函数，函数型数据分析（FDA）似乎是一个自然的选择。FDA是一组专门为分析曲线数据而开发的统计技术（Frøslie，2013年）。然而，Yogev（2014）提到，使用简单的汇总度量（例如曲线下面积（AUC））从函数中获取信息是一种常见做法。因此，在本次分析中，使用自变量的估计曲线下面积来预测剩余电池寿命。

使用曲线下面积来确定剩余电池寿命的理念是，所有预测变量都必须是功率的函数，并随时间造成电池消耗。由于功率-时间图下的面积提供了能量消耗的估计，预测变量-时间曲线下的面积将给出电池中随时间消耗的锂离子比例。将这一方法论应用于所有预测变量，可提供有关电池中随时间消耗的锂离子数量的信息。

因此，我们对这些参数进行了特征工程，计算其曲线下面积（即累积和），并将其作为后续预测建模的变量。图4展示了原始自变量每分钟转数（RPM）随时间变化的曲线。可以看到，RPM 值变化显著，范围从 0 到约 5000。RPM 越高，运行期间的锂离子损失越多；单独使用这一瞬时变量的取值，并不能有效指示寿命。将 RPM 变量转换为曲线下面积汇总度量，以纳入该变量的历史信息后，它便能有效用于寿命预测，因为曲线下面积汇总度量能够很好地表示锂离子的消耗量。因此，对 RPM 随时间进行累积求和，就通过特征工程创建了一个新变量。图5展示了与图4中相同的瞬时 RPM 数据随时间累积求和的结果。可以看到，RPM 的累积和表示瞬时 RPM 曲线下面积，因此能够提供有助于电池寿命预测的汇总度量。

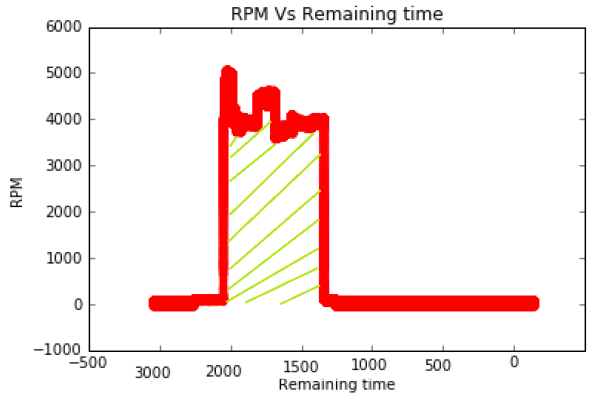


图4： RPM响应随时间的变化

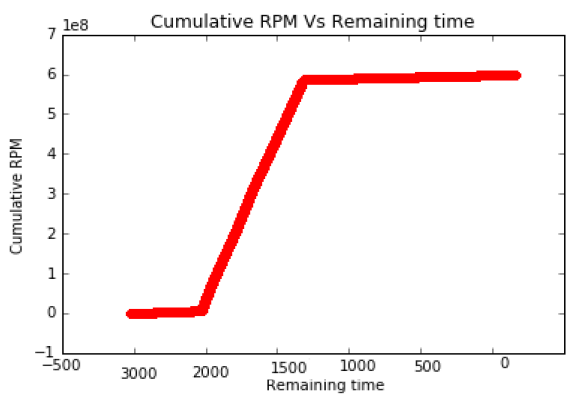


图5： RPM随时间的累积和

5.3. 相关性分析

由于许多预测变量都是电压、电流和温度的函数，因此这些参数之间存在较强的相关性。图3展示了不同预测变量之间的相关性图。深蓝色表示强正相关，深红棕色表示强负相关。数据中有 17 个预测变量，可以看到其中许多变量彼此高度相关。预测变量之间的相关性主要为正。这是预期的，因为在电池系统中，一个变量增加时，其他变量往往也会随之增加。例如，RPM 较高是电池输出电流较大的结果。还可以注意到，所有预测变量之间均呈正相关。前七个预测变量之间以及其余十个预测变量之间的相关性都很强。因此，我们识别出预测变量中的两个主要簇，可用于降维。

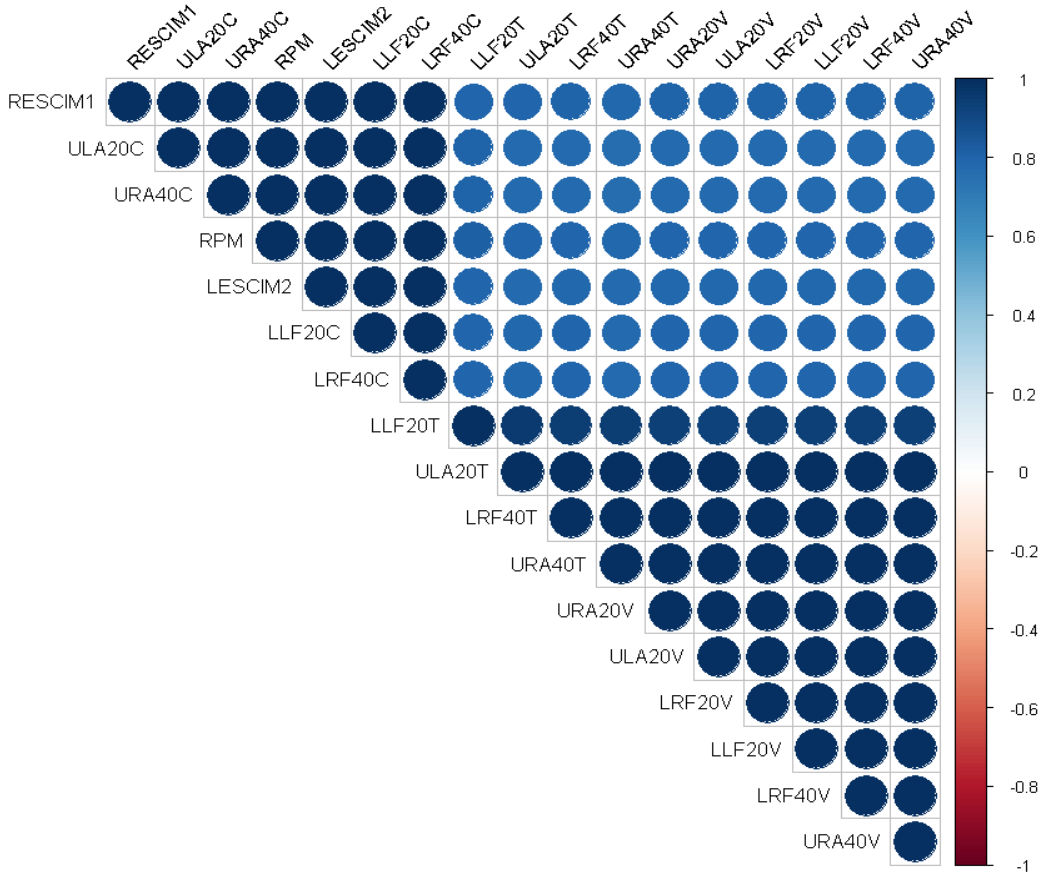


图 6： 协变量之间的相关性图。

5.4. 主成分分析（PCA）

当预测变量高度相关时，主成分分析（PCA）是一种用于降维的有效技术。通过PCA，可以用数量显著更少的变量保留原有多个变量中的相同信息。降维是必要的，因为预测变量越多，机器学习模型就需要估计越多的模型参数，这会增加模型训练所需的时间。降低预测变量的维度还能简化模型，减少模型参数的数量。

为了降低数据集的维度，我们进行了 PCA，以确定较少参数能够捕获多少方差。图 4 展示了 PCA 结果的肘部图。可以看出，前两个主成分捕获了 99% 的方差。增加其他 PCA 变量只能略微提高方差解释率。数据中的噪声也会影响方差，因此，只有在增加 PCA 变量能够显著提高方差解释率时，才强烈建议这样做。显然，仅用 2 个 PCA 变量即可解释数据集 99% 的变异。因此，利用 PCA 可将预测变量的维度降至仅 2 个转换后的变量，同时不会损失数据集中的太多信息。由于所用数据集包含 4,000,000 多个数据点，通过 PCA 将预测变量从 17 个降至 2 个，可显著降低计算成本。

如果考虑 5 个主成分，则数据集中的 99.999% 变异都得到了保留。因此，在不丢失原始预测变量信息的情况下，可以将 17 个变量缩减为仅 5 个主成分分析变量。额外的 3 个主成分分析变量仅能提供预测变量中约 1% 变异的信息；不过，对于增加少量预测变量不会带来显著计算成本的预测模型，可以考虑使用 5 个主成分分析变量。

主成分分析是一种强大的降维技术，有助于显著缩短模型训练过程中的计算时间。然而，由于主成分分析变量是原始预测变量的线性组合，因此需要额外的预处理，将原始预测变量转换为主成分分析变量。此外，为了进行模型测试，在测试数据集中也必须保持原始预测变量到主成分分析变量的转换方式一致。

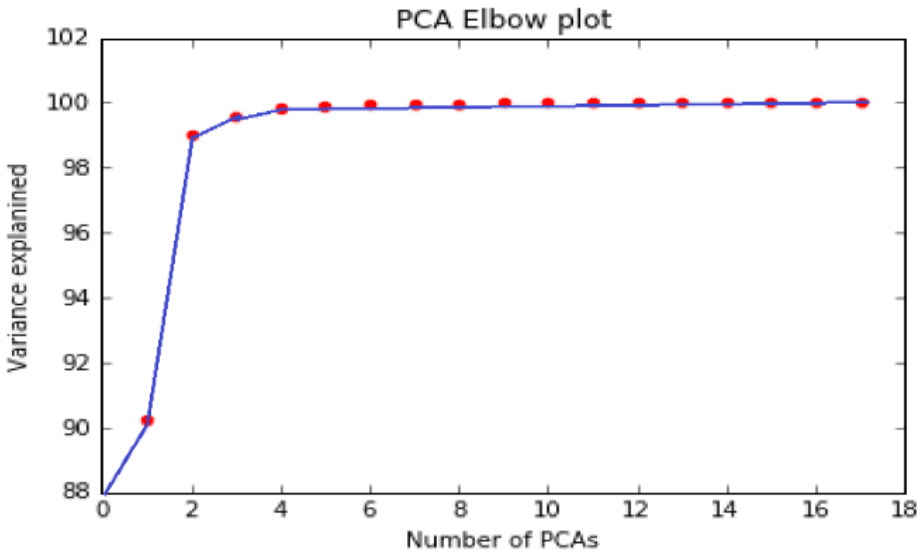


图 7： 主成分数量所解释的累积方差。

5.5。 预测建模

完成特征工程和主成分分析后，数据集便已准备好用于应用预测算法。图 8 展示了响应变量的分布。该分布较为独特：初始范围内的正值呈均匀分布，而较高值则呈指数分布。显然，基于已知分布的广义线性模型可能不适合对这一响应变量建模。此外，根据对电池和飞行技术的科学认识，预计协变量之间存在非线性关系，并包含许多显著的高阶交互作用。要成功预测电池寿命，需要能够捕捉具有高阶交互作用的复杂非线性关系的预测模型。因此，在数据驱动的预测建模中，我们研究了基于随机森林和神经网络的两种预测模型。这两种机器学习模型适用范围广且稳健，并且都以善于捕捉复杂非线性关系而著称。此外，真实世界电池数据可能含有大量噪声，而这些模型借助监督学习能够很好地处理噪声。

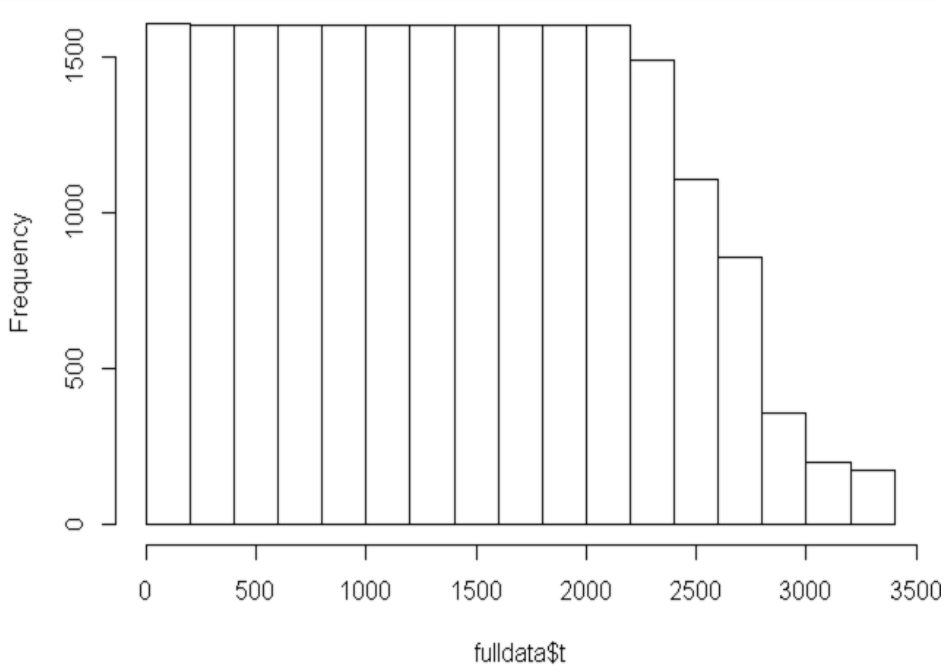


图 8：剩余飞行时间响应的直方图。

5.5.1. 随机森林

随机森林也称为随机决策森林，是一种集成监督学习模型，利用多棵决策树进行模型拟合。剩余使用寿命是一个连续变量，因此采用随机森林回归模型，根据预测变量进行模型拟合和训练。为减少计算时间，模型调优时仅将 2 个 PCA 变量作为预测变量，并将剩余使用寿命作为响应变量。模型训练时，使用了基于 2 个主成分的 50 棵决策树。此外，来自 8 次实验的数据作为训练数据集，来自 1 次实验的数据作为测试数据集。该模型在训练集上的表现良好，均方误差为 27431；然而，在测试数据上，模型预测值与实际值存在一定偏差。图 9 展示了测试数据上预测值与实际响应变量的拟合情况。显然，模型对测试数据的预测值与实际响应变量之间存在显著偏差。

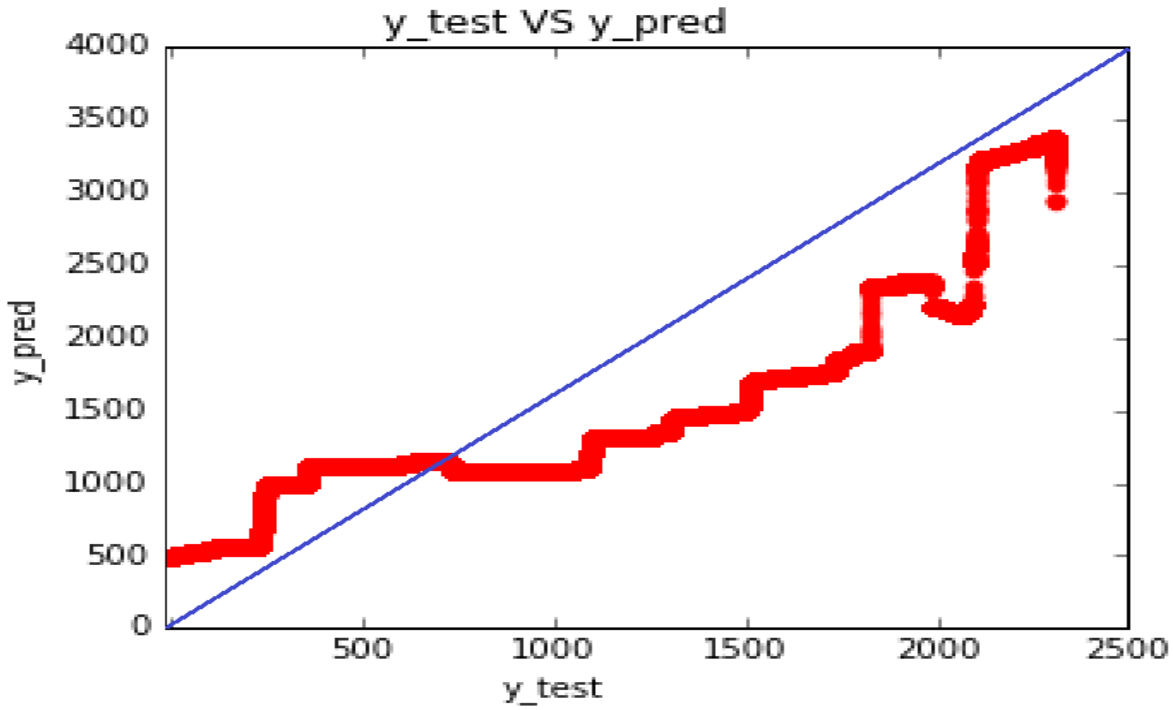
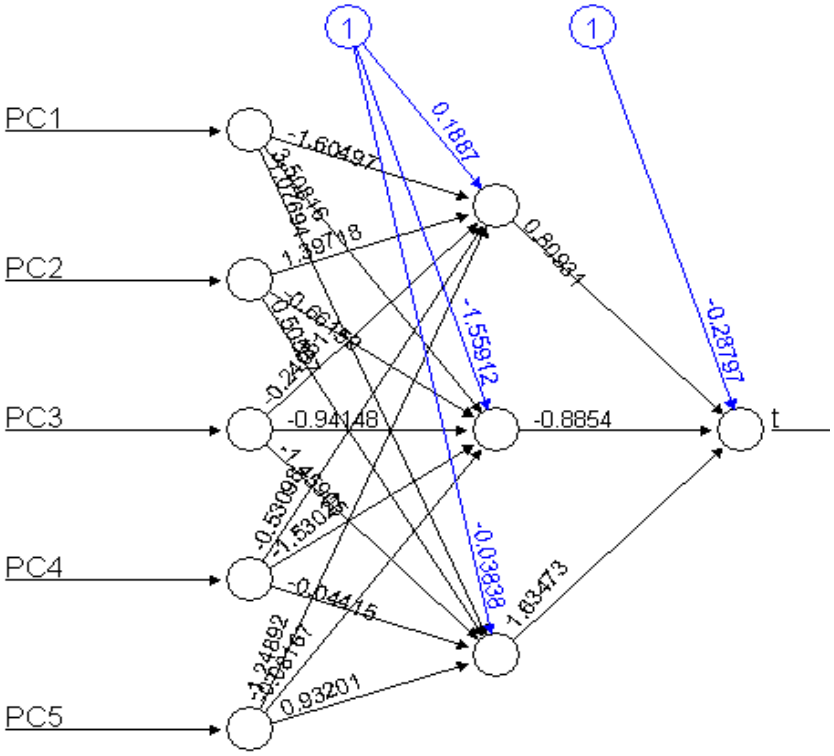


图9：使用随机森林模型预测的剩余飞行时间与实际值对比。

5.5.2. 神经网络

神经网络模型是功能强大的预测模型，能够很好地从大型数据集中捕捉复杂关系。因此，我们考虑使用神经网络构建预测模型。由于神经网络模型能够通过分配适当的权重自行进行变量选择，因此将能够解释 99.999% 方差的前五个主成分用作该模型的输入预测变量。通过调整隐藏层和节点的数量，确定拟合效果更好的模型。

共考虑了四种神经网络模型配置。首先构建了一个包含 3 个节点的单隐藏层模型。选择包含 3 个节点的单隐藏层作为起始网络，是因为这是一种简单的网络配置。该神经网络模型训练后的权重与偏置如图10-(a)所示，使用该模型对新的电池飞行测试数据进行预测的拟合优度如图10-(b)所示。预测值与实际值略有偏差。可以观察到，该模型大多数时候都低估了结果，仅在末尾出现了轻微的高估。该模型预测的均方误差为 35763。该模型似乎不够复杂，无法捕捉预测变量之间的关系。



Fig图10-(a)：具有3个节点的单隐藏层神经网络中的权重与偏置

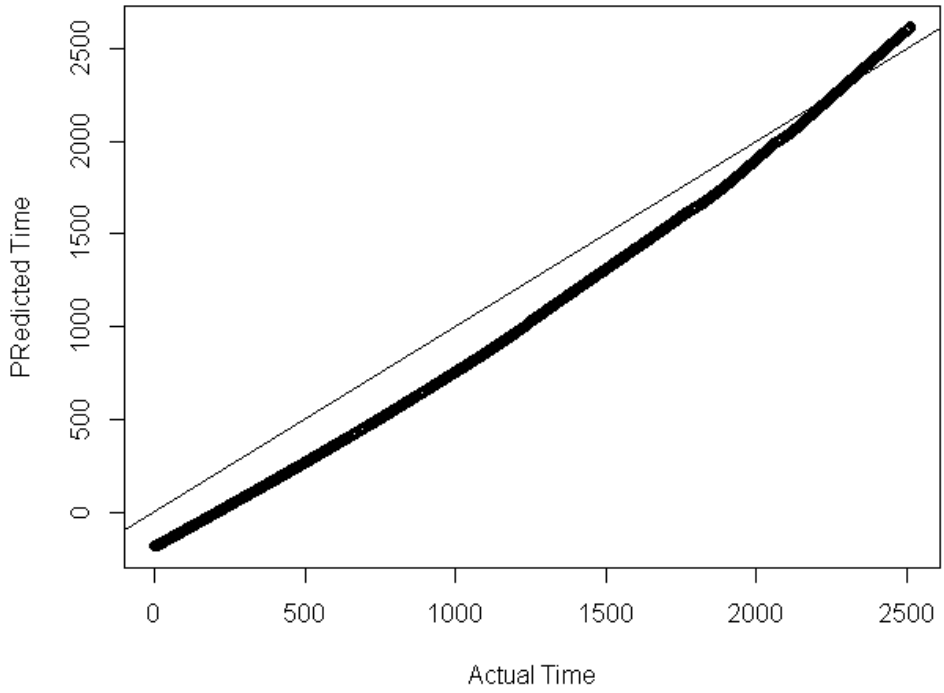
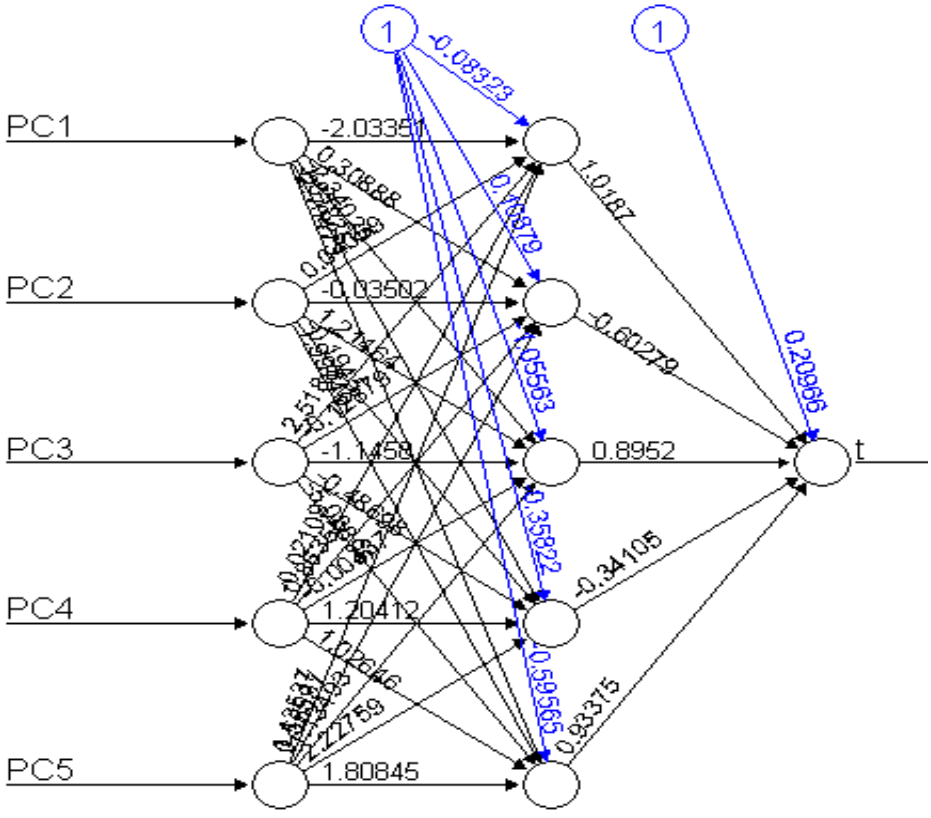


图10-(b)：具有3个节点的单隐藏层神经网络的拟合优度

此外，我们还考虑了一个具有单隐藏层、包含5个节点的第二个模型。我们希望模型中增加的节点能够更好地捕捉预测变量之间的复杂关系和交互作用。增加节点会为模型引入更多权重与偏置参数，因此模型会更复杂，训练所需的计算时间也会更长。图11-(a)展示了训练后的神经网络权重与偏置，图11-(b)展示了该神经网络模型的拟合优度预测图。可以观察到，与之前包含3个节点的模型相比，包含5个节点的模型预测准确率显著提高。预测时间始终与实际时间非常接近。图11-(b)清楚地显示，预测误差显著降低，预测时间值与实际值高度吻合。预测的均方误差显著降低至2717。显然，在神经网络中增加2个节点有助于将均方误差降低92%。增加2个节点后模型复杂度有所提高，但均方误差的大幅降低充分证明了这种复杂度提升是合理的。



Fi图 11-(a)： 5 个节点的单隐藏层神经网络中的权重与偏置

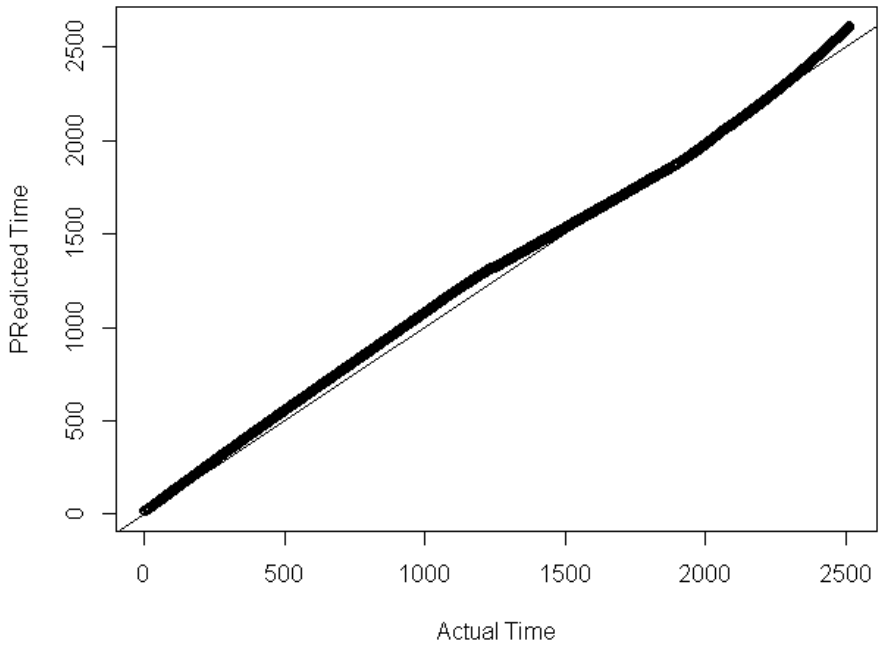


图11-(b)：具有5个节点的单隐藏层神经网络的拟合优度

此外，我们还探索了复杂的深度学习神经网络配置。我们训练了具有两个隐藏层的神经网络配置，其中第一层包含5个节点，第二层包含1个节点。图12-(a)展示了模型配置以及训练后的权重和偏置值。图12-(b)还展示了具有两个隐藏层的模型的拟合优度，其中第一层包含5个节点，第二层（单层）包含1个节点。可以观察到，与包含5个节点的单层网络相比，该模型的预测准确率有所下降。测试数据上的均方误差为11896。显然，与包含5个节点的单层模型的均方误差相比，该模型的均方误差高出338%。

尽管双隐藏层模型的拟合效果较差，在测试数据上的均方误差较高，但为了确认并了解在第二层增加节点是否有助于改善模型拟合效果，我们进一步探索了另一个包含更多节点的双隐藏层模型：第一层有5个节点，第二层有3个节点。图13-(a)显示了训练后网络配置中的权重与偏置值，图13-(b)显示了该网络配置在测试数据上的拟合优度。第一层有5个节点、第二层有3个节点的模型预测性能也较差，均方误差为15709。结果表明，与单层5节点模型相比，双层模型的预测效果明显更差。尽管复杂神经网络的训练误差有所降低，但观察发现，该模型在测试数据上的预测误差

最低的是具有单层5个节点的模型。因此，采用神经网络建模方法时，单层5节点模型是最佳模型。

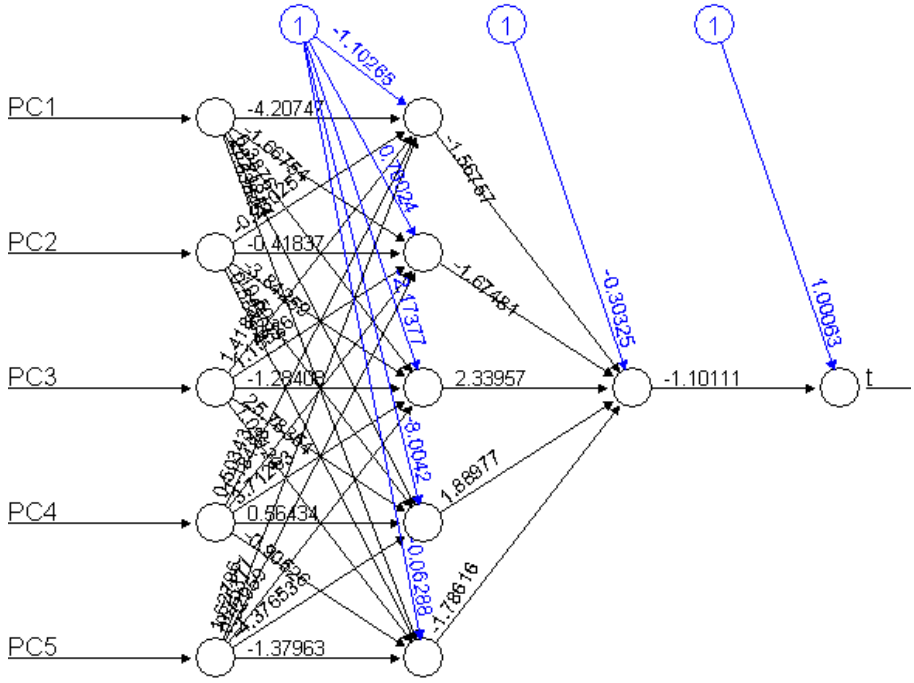


图12-(a)：具有两层的神经网络的权重与偏置，第一层有5个节点，第二层有1个节点。

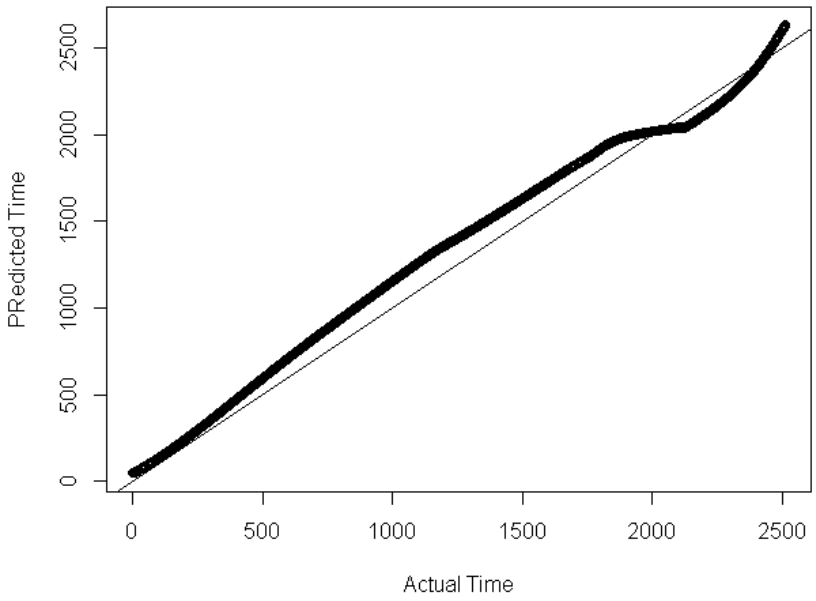


图12-(b)：具有两层的神经网络的拟合优度，第一层有5个节点，第二层有3个节点。

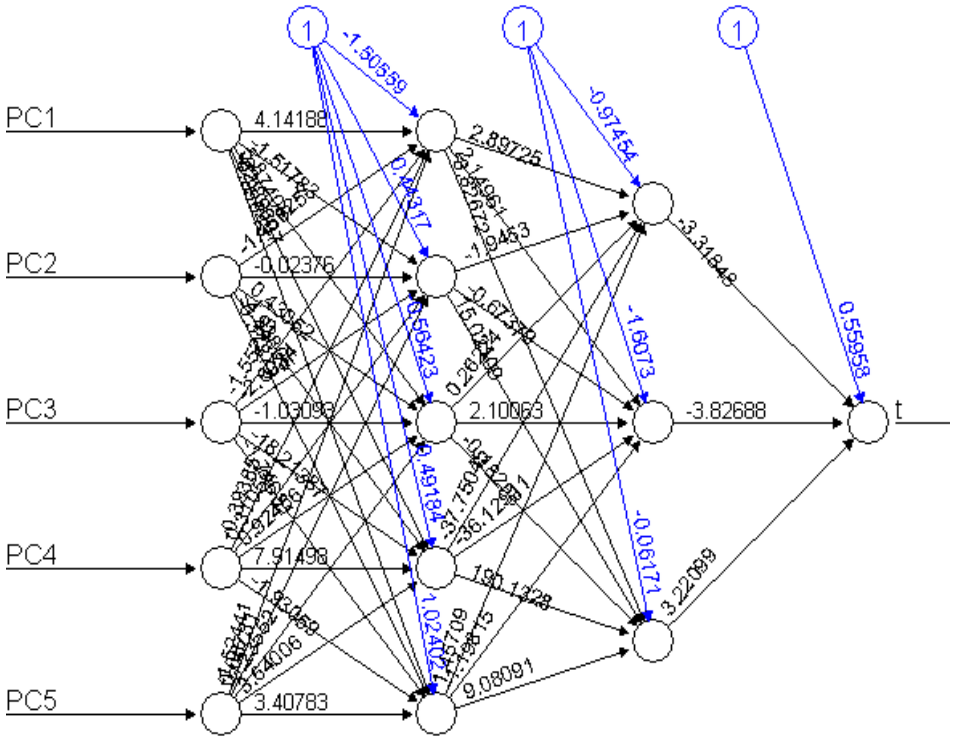


图13-(a)：2层神经网络的权重与偏置，第一层有5个节点，第二层有3个节点。

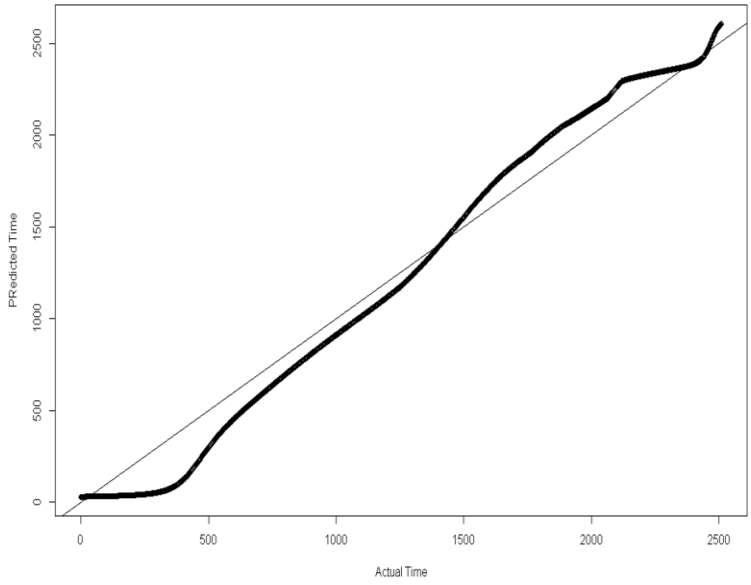


图13-(b)：2层神经网络的拟合优度，第一层有5个节点，第二层有3个节点。

5.5.3. 模型比较

本文从预测准确率、计算时间以及模型处理高维数据的稳健性等方面，对随机森林和神经网络的预测性能进行了比较。模型比较时，通过测试数据上的均方误差（MSE）衡量预测准确率。将训练模型所需的大致时间作为计算时间的衡量指标。模型的稳健性则根据模型无需降维即可处理高维数据的能力进行评估。表1列出了本研究中考察的随机森林模型及各种神经网络模型的性能指标。最佳神经网络模型的预测均方误差为2717，而随机森林模型的预测均方误差为27431。显然，神经网络建模方法的预测误差显著较小。然而，即使使用缩减后的数据集，训练神经网络所需的计算时间仍显著更长（约为3倍）。此外，神经网络能够处理特征选择，而随机森林在构建模型之前需要进行特征选择。因此，随机森林使用了排名前两位的主成分，而神经网络模型则使用了排名前五位的主成分。与随机森林模型相比，神经网络模型在处理高维数据方面更为稳健。仅根据测试数据上的预测均方误差来看，具有一个隐藏层且包含5个节点的神经网络模型是预测电池寿命的最佳模型。

表1：随机森林和各种神经网络（NN）模型在电池数据集上的比较。

模型	**测试均方误差**	**计算时间**	**鲁棒性**
随机森林	27431	~ 10 分钟	Low
3节点单隐藏层神经网络	35763	~ 30 分钟	High
神经网络-单层5节点	2717	~ 35 分钟	High
神经网络-双层（5,1）节点	11896	~ 50 分钟	High
神经网络-双层（5,3）节点	15709	~ 55 分钟	High

6. 结论

本文探讨了数据驱动建模方法，利用电动飞机的电池数据，根据飞行器状态和电池在使用过程中的SOC，预测锂离子电池的寿命。研究发现，构建预测模型需要进行特征工程。需要能够提供电池运行期间使用情况信息的特征。因此，本文根据曲线下面积指标，从原始的瞬时电池使用情况信息中生成新特征，以反映电池过去运行期间的使用情况。此外，许多特征之间存在高度相关性，因此本文采用主成分分析探索降维方法。

在完成特征工程和主成分分析之后，基于随机森林和神经网络模型开发了两种预测模型。对于随机森林模型，使用训练数据构建了基于50棵决策树的集成模型。测试数据上的预测均方误差被用作模型比较的性能指标。此外，还考察了拟合优度图，以便直观地查看模型预测与实际寿命数据的拟合情况。在探索神经网络模型时，本研究考虑了四种神经网络配置。这四种模型配置包括：具有3个节点的单层、具有5个节点的单层、第一层5个节点且第二层1个节点的双层，以及第一层5个节点且第二层3个节点的双层。四个神经网络模型均以5个主成分作为输入、并使用全部训练数据进行训练，以评估模型的权重与偏置参数。观察发现，具有5个节点的单层神经网络模型在测试数据的均方误差性能指标上表现最佳。显然，节点过少会导致欠拟合，而神经网络中节点过多则会导致过拟合。训练数据上的均方误差随节点数增加而持续下降；然而，在达到最优节点配置之后，测试数据上的均方误差开始上升。这清楚地表明，最优节点配置应根据测试数据上的均方误差来确定。

总体而言，表现最佳的神经网络模型的均方误差为 2717，而随机森林模型在测试数据上的预测均方误差为 27431。显然，从均方误差来看，神经网络模型表现更佳，对电池寿命的预测也更准确。此外，训练神经网络模型所需的计算时间是随机森林模型的 3 倍

。不过，最佳神经网络模型将均方误差降低了 90% 以上，因此较高的计算时间是值得的。此外，神经网络模型的鲁棒性更强，能够处理高维数据，而无需进行大量降维。模型比较清楚地展现了神经网络模型在电池寿命预测方面的准确性和鲁棒性均优于随机森林模型。因此，可以采用数据驱动建模方法，利用真实世界电池数据以合理的准确度预测电池寿命。

7. 未来工作

本研究主要考察了两种适用性最广的机器学习模型，除随机森林和神经网络模型外，还可以进一步探索更多模型。此外，还可以采用数据驱动建模方法，探索不同应用场景下的电池寿命预测。对于电池寿命预测的数据驱动建模方法而言，特征工程和预测变量之间的高度相关性是重要的考量因素。为了准确预测寿命，准确测量电池相关数据至关重要。因此，可以进一步开展分析，探索在测量电池数据时降低数据维度的方法，从而将信息量更丰富的变量纳入电池管理系统的测量内容中。

未来可进一步研究和开发其他模型，例如函数回归、结合基于物理的模型与机器学习模型的混合模型，以更准确、稳健地估算电池寿命。此外，还可以探索如何量化预测的不确定性度量，从而找出预测不确定性较低的数据驱动模型。对于关键应用而言，这一点尤为重要，因为此类应用要求电池寿命预测具有较低的不确定性。总体而言，本研究仅探讨了电池寿命预测的若干可能建模方法中的一小部分。仍有许多方法有待探索，以实现更准确、计算效率更高且更可靠的电池寿命预测。

参考文献

阿贡国家实验室（2022），《基于机器学习的电池寿命与性能预测》，网址：https://www.anl.gov/egs/machine-learning-prognosis-for-battery-life-and-performance（访问日期：2022年6月10日）

BatteryBits（2020），《锂离子电池测试开放数据集比较》，Medium，12月12日。网址：https://medium.com/batterybits/comparison-of-open-datasets-for-lithium-ion-battery-testing-fd0de091ca2（访问日期：2022年6月11日）

Climate Central（2019年），《Climate Central 解决方案简报：电池储能》，ClimateCentral，11月13日。网址：https://www.climatecentral.org/news/climate-central-solutions-brief-battery-energy-storage（访问日期：2022年6月11日）

Fei, Z., Yang, F., Tsui, K.L., Li, L. 和 Zhang, Z., 2021年。基于机器学习框架的电池寿命早期预测。《Energy》，225，p.120205。

Frøslie, K. F., Røislien, J., Qvigstad, E., Godang, K., Bollerslev, J., Voldner, N., & Veierød, M. B.（2013年）。葡萄糖曲线中的形状信息：函数型数据分析与传统汇总指标的比较。《BMC医学研究方法学》，13(1)，6。

Hosen, M.S.、Jaguemont, J.、Van Mierlo, J.和Berecibar, M.，2021年。电池寿命预测与不同建模方法的性能评估。Iscience，24(2)，p.102060。

NASA故障预测数据仓库，B. Saha和K. Goebel（2007）。“电池数据集”，NASA故障预测数据仓库，NASA艾姆斯研究中心，莫菲特机场，加利福尼亚州

Pistoia, G. 2013年，《锂离子电池：进展与应用》。Newnes。

Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D. 和 Bazant, M.Z.，2019年。《容量衰减前电池循环寿命的数据驱动预测》。Nature Energy，4(5)，第383-391页。

美国能源部（2021），《NREL 利用基于物理的机器学习推进电池研究，加速电池性能、寿命与安全性的表征》，Cleantechnica，2021年2月19日。可在以下网址查阅：https://cleantechnica.com/2021/02/19/nrel-advances-in-battery-research-with-physics-based-machine-learning-accelerates-characterization-of-cell-performance-lifetime-safety/（查阅日期：2022年6月11日）。

Yogev Y, Ben-Haroush A, Chen R, Rosenn B, Hod M, Langer O：肥胖和正常体重的非糖尿病孕妇的昼夜血糖曲线。美国妇产科杂志。2004年，191：949-953。10.1016/j.ajog.2004.06.059。

Yoshio, M., Brodd, R.J. 和 Kozawa, A. 2009年，《锂离子电池》（第1卷）。纽约：施普林格。

Zhang, P., Liang, J. 和 Zhang, F.，2017年5月。电池寿命预测的不同方法概述。载于《IOP会议系列：材料科学与工程》（第199卷，第1期，第012134页）。IOP出版社。

Zhu, S., Zhao, N. 和 Sha, J.，2019年。利用机器学习和早期循环数据预测电池寿命。《能源储存》，1(6)，第e98页。