**大规模智能电网用户的精细化电能分解与用电行为分析**

**文献综述**

1. **前言**

家庭用户消耗的用电量占总用电的三分之一以上。智能电网用户的用电数据隐含着丰富的信息，可以用来研究用户行为模式，提高能源利用率。然而，目前存在的电能分解方法需要借以在家庭中安装嵌入式传感器从而会需要消耗昂贵的硬件资源。针对上述问题，本论文通过分析用电数据中总用电量和用电器用电量之间的关系，并建立非嵌入式的模型对用户的总用电量进行分解，预测用户各用电器的用电量。同时根据用户的用电数据分析用户的家庭属性，包括家庭面积、房间数和家庭人数等。

1. **研究现状**

**2.1基于嵌入式的用电分解**

现在的用电分解主要是基于嵌入式的用电分解，这种方式需要人工安装传感器，需要在每个家庭用电器接电处安装传感器，每隔相同时间收集每个传感器的数据。虽然基于嵌入式的用电分解操作简单，但此种方法存在严重的弊端，此方法不仅需要消耗昂贵的人力物力，更为重要的是分解的精确度不高，是因为在收集传感器数据时会存在误差。

**2.2基于非嵌入式的用电分解**

如今也存在基于非嵌入式的用电分解，也就是不用在每个用电器上安装传感器，通过收集一个家庭的总消耗用电量，通过分析用电数据中总用电量和用电器用电量之间的关系，并建立非嵌入式的模型对用户的总用电量进行分解，预测用户各用电器的用电量。

现有的基于非嵌入式用电分解主要是矩阵分解，矩阵分解主要是将矩阵分解成两个具有特殊含义的矩阵，但此种方法在一定意义上存在着问题，因为此种方法把数据解释为二维的，但事实上，这种解释性不强。除此之外，现有的基于非嵌入式的用电分解存在不高效和精确度低的问题，而造成这种问题的原因主要是上述所说的二维矩阵可解释性不强和对矩阵分解的优化函数不合适。

**3.1矩阵分解**

·**矩阵分解(matrix factorization)**

矩阵分解是将矩阵拆解为数个矩阵的乘积。在矩阵运算中，把矩阵分解成形式比较简单或具有某种特性的一些矩阵的乘积，在矩阵理论的研究和应用中，具有重要的意义。一方面，矩阵分解能够明显反映出原矩阵的某些数值特征，如矩阵的秩、行列式、特征值及奇异值，另一方面，分解的方法和过程往往提供了某些有效的数值计算方法和理论分析根据。

·**非负矩阵分解(Non-negative Matrix Factorisation)**

非负矩阵分解[1-6]是由Lee和Seung于1999年在自然杂志上提出的一种矩阵分解方法，它使分解后的所有分量均为非负值。

**3.2回归模型**

回归模型重要的基础或者方法就是回归分析，回归分析是研究一个变量（被解释变量）关于另一个（些）变量（解释变量）的具体依赖关系的计算方法和理论，是建模和分析数据的重要工具。在这里，我们使用曲线/线来拟合这些数据点，在这种方式下，从曲线或线到数据点的距离差异最小。下面是回归分析的几种常用方法 ：

·**Linear Regression线性回归**[7-8]

它是最为人熟知的建模技术之一。线性回归通常是人们在学习预测模型时首选的技术之一。在这种技术中，因变量是连续的，自变量可以是连续的也可以是离散的，回归线的性质是线性的。线性回归使用最佳的拟合直线（也就是回归线）在因变量（Y）和一个或多个自变量（X）之间建立一种关系。

·**Ridge Regression岭回归**[9-10]

岭回归分析是一种用于存在多重共线性（自变量高度相关）数据的技术。在多重共线性情况下，尽管最小二乘法（OLS）对每个变量很公平，但它们的差异很大，使得观测值偏移并远离真实值。岭回归通过给回归估计上增加一个偏差度，来降低标准误差。

·**Lasso Regression套索回归**[11]

它类似于岭回归，Lasso （Least Absolute Shrinkage and Selection Operator）也会惩罚回归系数的绝对值大小。此外，它能够减少变化程度并提高线性回归模型的精度。

**3.3深度学习**

·[**深度学习**](https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0)

[深度学习](https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0)的概念源于[人工神经网络](https://baike.baidu.com/item/%E4%BA%BA%E5%B7%A5%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C)的研究。含多隐层的[多层感知器](https://baike.baidu.com/item/%E5%A4%9A%E5%B1%82%E6%84%9F%E7%9F%A5%E5%99%A8)就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，以发现数据的分布式特征表示[12-14]。

深度学习的概念由Hinton等人于2006年提出。基于深度置信网络(DBN)提出非监督贪心逐层训练算法，为解决深层结构相关的优化难题带来希望，随后提出多层自动编码器深层结构。此外Lecun等人提出的卷积神经网络是第一个真正多层结构学习算法，它利用空间相对关系减少参数数目以提高训练性能。

[深度学习](https://baike.baidu.com/item/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0)是机器学习中一种基于对数据进行[表征](https://baike.baidu.com/item/%E8%A1%A8%E5%BE%81)学习的方法。观测值（例如一幅图像）可以使用多种方式来表示，如每个[像素](https://baike.baidu.com/item/%E5%83%8F%E7%B4%A0)强度值的向量，或者更抽象地表示成一系列边、特定形状的区域等。而使用某些特定的表示方法更容易从实例中学习任务（例如，[人脸识别](https://baike.baidu.com/item/%E4%BA%BA%E8%84%B8%E8%AF%86%E5%88%AB)或面部表情识别）。深度学习的好处是用非监督式或半监督式的特征学习和分层特征提取高效算法来替代手工获取[特征](https://baike.baidu.com/item/%E7%89%B9%E5%BE%81)。

·**神经网络（Neural Network）**[15-16]

神经网络是一种运算模型，由大量的节点（或称神经元）之间相互联接构成。每个节点代表一种特定的输出函数，称为激励函数（activation function）。每两个节点间的连接都代表一个对于通过该连接信号的加权值，称之为权重，这相当于人工神经网络的记忆。网络的输出则依网络的连接方式，权重值和激励函数的不同而不同。而网络自身通常都是对自然界某种算法或者函数的逼近，也可能是对一种逻辑策略的表达。

人工神经网络是由大量处理单元互联组成的非线性、自适应信息处理系统。它是在现代神经科学研究成果的基础上提出的，试图通过模拟大脑神经网络处理、记忆信息的方式进行信息处理。

**4.研究内容**

**4.1分析Dataport公开数据集，对数据进行预处理**

利用python语言、numpy和panda数据处理库，对Dataport公开数据集的数据进行预处理，筛选和处理成本文需要的特定数据格式。

**4.2采用tensorflow框架实现矩阵分解进行电能分解**

利用tensorflow库搭建矩阵分解的模型框架，通过输入经过预处理的数据，最终通过优化损失函数达到分解矩阵也就是补全矩阵中缺失数据的目的。

**4.3采用tensorflow框架实现回归模型、神经网络进行属性特征预测**

利用tensorflow库搭建回归模型、神经网络的模型框架，通过输入经过预处理的数据，最终通过优化损失函数达到预测家庭属性特征的目的。

**5.总结**

本文旨在节省用电分解的硬件花销和提高用电分解的高效性和精确度。基于国内外用电分解的现状，综合各种在用电分解方面具有优势的算法和工具，通过分析用电数据中总用电量和用电器用电量之间的关系，并建立非嵌入式的模型对用户的总用电量进行分解，预测用户各用电器的用电量。同时根据用户的用电数据分析用户的家庭属性，包括家庭面积、房间数和家庭人数等。

**6．参考文献**

[1]汪鹏.非负矩阵分解:数学的奇妙力量[J].计算机教育,2004(10):38-40.

[2]栗茂林,梁霖,陈元明,徐光华,何康康.基于聚类优化的非负矩阵分解方法及其应用[J].中国机械工程,2018,29(06):720-725.

[3]何梦娇,杨燕,王淑营.一种基于非负矩阵分解的聚类集成算法[J].计算机科学,2017,44(09):58-61.

[4]孙静,蔡希彪,孙福明.基于特征融合的多约束非负矩阵分解算法[J].计算机应用,2017,37(10):2834-2840.

[5]何海洋,王勇,蔡国永.基于用户类别偏好相似度和联合矩阵分解的推荐算法[J].数据采集与处理,2018,33(01):179-185.

[6]王建芳,刘冉东,刘永利.一种带偏置的非负矩阵分解推荐算法[J].小型微型计算机系统,2018,39(01):69-73.

[7].Combined model based on optimized multi-variable grey model and multiple linear regression[J].Journal of Systems Engineering and Electronics,2011,22(04):615-620.

[8].Statistical Inference for Partially Linear Regression Models with Measurement Errors[J].Chinese Annals of Mathematics,2008(02):207-222.

[9]Xiuli WANG,Mingqiu WANG.Combination of Nonconvex Penalties and Ridge Regression for High-Dimensional Linear Models[J].Journal of Mathematical Research with Applications,2014,34(06):743-753.

[10]Huan Luo. A Prediction Method Based on Improved Ridge Regression[A]. The Institute of Electrical and Electronics Engineers、IEEE Beijing Section.Proceedings of 2017 IEEE 8th International Conference on Software Engineering and Service Science[C].The Institute of Electrical and Electronics Engineers、IEEE Beijing Section:,2017:4.

[11]张勇进. 带弹性网的稀疏主成分分析[D].华中科技大学,2006.

[12]熊红凯,高星,李劭辉,徐宇辉,王涌壮,余豪阳,刘昕,张云飞.可解释化、结构化、多模态化的深度神经网络[J].模式识别与人工智能,2018,31(01):1-11.

[13]王月岭. 基于支持向量机与概率输出网络的深度学习[D].西安理工大学,2017.

[14]刘晓曈. 基于深度学习的分类预测方法研究及应用[D].东南大学,2017.

[15].Real estate appraisal system based on GIS and BP neural network[J].Transactions of Nonferrous Metals Society of China,2011,21(S3):626-630.

[16]焦李成,杨淑媛,刘芳,王士刚,冯志玺.神经网络七十年:回顾与展望[J].计算机学报,2016,39(08):1697-1716.