深度学习技术在私有企业用户电力定价中的应用

**中文摘要**

随着我国的经济步入高速发展阶段，电力需求增速换挡，近些年新能源的快速发展，是我国摆脱了供不应求的局面。然而，在某些负荷需求高峰期，仍然避免不了负荷短缺的情况，如果通过增加负荷容量以满足高峰负荷需求，显然是不具备经济性的。我国开始的新一轮电力改革的核心是电价，通过改变以往完全由国家定价，各种社会性的加价的模式，进而以根据需求制定电价，通过电价引导用户在用电高峰期少用电，低谷期多用点，提高用电效率，优化用电方式，来缓解用电压力，降低供电成本。首先，梳理了需求响应动态定价策略研究的国内外现状。介绍了电力需求响应的基本概念及分类；介绍了电力负荷预测的常用方法并对比。对电力需求响应环境下的定价方法进行比较，并说明了基于强化学习的需求响应动态定价方法所具有的优点。 其次，针对前馈型神经网络不能处理序列间关联信息与传统循环神经网络无法记忆久远关键信息的缺陷，提出了基于长短期记忆循环神经网络的负荷预测模型。最后，提出了一种适用于私有企业用户的动态定价方法并进行了算例验证。根据零售电力市场中主场主体间的关系，构建了兼顾用户用电成本和能源收益综合最优为目标函数、并计及用户满意度要求、零售价和批发价之间价差限制等多种约束在内的居民需求响应动态定价模型。在模型求解过程中，将居民的需求响应的动态定价问题转化为有限的马尔科夫决策过程，从而以较小的计算代价获取较优解。结合 Q-learning 算法实现动态定价模型的求解。通过实际仿真算例，对选取的私有企业用户的零售电价进行定价，并且进行仿真，验证了所构建的模型及动态定价方法的可靠性。

**关键词：动态定价；负荷预测；强化学习；LSTM**

1. **引言**
   1. **选题背景及研究意义**
      1. **选题背景**

随着全球变暖，全国的电力负荷出现逐年快速增加，多地出现用电紧张。工业大省四川情况尤为严峻，2022年7月以来，用电负荷接连6次创历史新高。为保障居民用电，四川省内的所有工业电力用户，生产全停。此外，多省也面临电力供应紧张的局面。与个人用户相比，企业用户用电巨大，尤其是工业企业。根据2022年发布的中国电力消费情况，全国全社会用电量86372亿千瓦时，全国工业用电量56000亿千瓦时，占全社会用电量的比重为64.8%，可见合理调控工业用电对电网的稳定运行影响更大。

通过随着电力体制改革的大力推行，改变以往完全由国家定价，各种社会性的加价的模式，进而以根据需求制定电价，通过电价引导用户在用电高峰期少用电，低谷期多用点，提高用电效率，优化用电方式，来缓解用电压力，降低供电成本是本轮电力体制改革的核心。截至2020年8月底，国家电网经营范围内注册的售电公司数量达到3700家。然而由于业务开展存在困难、抗风险能力差等因素，近60％处于观望状态，未跻身市场真正参与购售电交易。在电力市场中售电公司主要承担将变化的电力批发成本平抑后传导至用户侧，提供基本电力销售以及相应的附加增值服务。打破售电侧的垄断局面，如何开展基于需求侧管理的增值业务，采用合理有效的定价和营销手段整合用户需求资源、挖掘用户需求响应潜力，已成为其在购售电业务之外，寻找新的盈利增长点的关键问题。

2023年7月19日，国务院发布了《中共中央国务院关于促进民营经济发展壮大的意见》，指出民营经济是推进中国式现代化的生力军，是高质量发展的重要基础，是推动我国全面建成社会主义现代化强国、实现第二个百年奋斗目标的重要力量。电网的正常稳定运行，持续稳定的供电，是保障民营企业正常发展的最基本问题。民营经济的主要成分是私营企业、个体工商户和农民专业合作社，其中，私营企业和个体工商户在民营经济中又占据了绝大部分。

所以在用电负荷快速增加，高峰期用电紧张的情况下，保障电网正常运行，持续供电；保障民营经济的主力军私有民营企业的正常电力运行，减少用电成本；打破电价垄断，发展新电改是本文研究的核心。

* + 1. **研究意义**

随着经济的快速发展、企业机械化动力占比逐年升高，随之对电的消耗也在逐年递增，保障电的正常供应对我国经济的发展至关重要；全球变暖的影响，天气恶劣，空调等制冷设备成为了夏日必备，保障居民用电的正常供应是基本民生。根据2022年发布的中国电力消费情况，全国全社会用电量86372亿千瓦时，全国工业用电量56000亿千瓦时，占全社会用电量的比重为64.8%，对工业用电的宏观调控，是保障电网稳定运行的重要手段。工业的发展往往存在周期性，在快速发展期，订单较多，用电飙升，在衰败期，订单较少，用电较少。所以其用电趋势也存在一定的周期性，可以根据历史用电量预测未来短期用电量。我国是分时电价，对于企业的订单，可以根据交付时效来进行规划排产，对于订单交期不紧张的订单，可以排在电价低的时间段进行生产，进而减少企业的用电成本，扩大利润。

所以通过对私有工业企业历史用电量的分析，依赖科学工具预测未来短期的用电量，再依赖需求侧对不同电价的响应，制定合理的电价，进而通过电价干预高峰时期的用电量，达到削峰填谷，提高电网的稳定性，保障电网系统的有序运行的目的。针对不同的用户定制不同的电价定价策略，有助于减少企业用电成本，有助于改善电荷供应商的服务质量，实现用户和售电商的共赢。

* 1. **国内外研究现状**
     1. **短期负荷预测研究现状**

电力负荷预测是根据某一地区的历史能耗数据，预测该地区未来一段时间的用电情况。准确的预测可以为电力建设和电网运行提供有效、可靠的指导。目前，在电力工业的发展中，电力负荷预测的研究越来越发挥重要作用。准确地对电力负荷进行预测，既可以保证电力供应的稳定，降低用电成本，也在提高供电质量发挥重要作用。电力系统短期负荷预测可以分为超短期、短期、中期和长期预测，其中短期预测是负荷预测的重要组成部分，其准确性直接影响各主体调度和决策的正确性和经济性[1]

目前短期负荷预测有现代预测和经典预测两大类方法。经典预测方法由传统数学模型完成，包括回归分析法、指数平滑法、时间序列法等。而随着电力系统日益复杂化，基于智能算法的现代预测方法逐渐兴起，主要包括专家系统、支持向量机（SVM）以及人工神经网络（Artificial Neural Network，即 ANN）等。人工神经网络自提出以来，一直都是人工智能方向的研究热点。主要包括常见的BP 神经网络，及在进一步发展中出现的 RBF 神经网络、Elman 神经网络、CNN网络，RNN 网络及 RNN 的变种 LSTM、GRU 神经网络等。

随着需求响应的日益普及，准确的短期负荷预测(STLF)在智能电网运行中发挥着至关重要的作用。Kong Xiangyu[6]等以传统算法的结果为基准，采用两阶段方法提高STLF的精度。通过对初始预测偏差进行模态分解和重组，构建目标DR偏差序列。使用动态模态分解(DMD)获得DR引起的偏差序列，并通过构造矩阵来简化这一过程。将传统算法的预测结果与得到的偏差序列相结合，提高了最终的预测精度。通过对灾备试验区数据的多个实例研究以及与现有模型的比较，表明基于现有算法的需求响应偏差校正有效提高了电力负荷预测精度，具有良好的泛化性。针对传统负荷预测策略的不足，综合考虑各种气象因素在负荷预测中的应用，[Di](https://ieeexplore.ieee.org/author/37088846493) Shuai[10]等提出了一种基于最小二乘支持向量机(LSSVM)模型的电力系统短期负荷预测方法。首先，通过相关分析为选择建模变量提供理论依据，并利用相似原理对负荷数据进行处理，优化样本质量;其次，建立天气LSSVM负荷预测模型，并利用某地区历史负荷数据对模型进行训练和测试。最后，通过仿真验证了预测模型的准确性和有效性。Hinde[4]等提出了一种新的、成熟的联合概率聚类和回归模型的理论框架，它不同于现有的独立处理这两个过程的模型。通过在训练过程中同时使用输入数据和预测目标来增强聚类过程。因此，该模型能够获得比其他方法更好的聚类，从而获得更多信息丰富的数据概要，同时保持或提高预测性能。

针对传统负荷预测方法在处理大规模非线性时间序列数据时的局限性，Du Jie[2]等提出了一种基于注意力的双向长短期记忆(Attention - bilstm, Attention- based Bidirectional Long - short- Memory, Attention- bilstm)网络来进行准确的短期负荷预测。该模型基于对时间序列数据建模具有较高鲁棒性的Bi LSTM递归神经网络和能够突出输入数据中在负荷预测中起关键作用的关键特征的注意机制。在一定区域的真实数据验证实验表明，该模型在预测精度和算法鲁棒性方面优于其他模型。Tang De Dong[3]等提出了一种基于两长短期记忆层神经网络的电力负荷预测方法。基于EUNITE提供的实际电力负荷数据，构建了一种基于LSTM的电力负荷预测方法。建立了单点预测模型和多点预测模型，对未来一小时和半天的电力进行预测。实验结果表明，LSTM网络单点预测模型的平均绝对百分比误差为1.806，多点预测模型的平均绝对百分比误差为2.496。为了提高短期负荷预测的准确性，解决短期负荷数据的时间序列、非线性和多维性问题，Guo, Xifeng[5]等提出了一种基于强化学习深度Q网络(DQN)和长短期记忆网络(LSTM)的多步短期负荷预测方法。首先，采用时间滑动窗口法对归一化后的历史数据进行采样，构造特征映射作为输入。然后，以LSTM为主体构建预测模型，捕获长、短时间序列特征。DQN与LSTM网络结合形成一个跳跃层来捕获超长时间序列特征，使DQN-LSTM能够捕获不同时间尺度的特征。最后，引入全连接层，融合各层提取的多尺度时间特征，并采用自回归模型(AR)作为线性分量，得到最终的预测结果。实验结果表明，所提预测方法的MAPE和RMSE均低于其他算法，且能有效捕捉电力负荷数据的多尺度时间序列特征，提高了预测精度。Liu Min[7]等提出了一种将深度交叉与时间卷积网络(DCN)和时间卷积网络(TCN)相结合的混合网络，称为深度交叉与时间卷积网络(DTCN)。该模型能够自动提取数据集中包含的有效信息，并自动进行特征提取和特征融合。通过“澳大利亚电力负荷数据”验证了模型的有效性。通过对多时段不同试验结果的比较，表明该模型能够准确预测电力负荷预测结果。具有结构清晰、泛化程度高、数据容忍度强等特点。Li Fan[8]等分析了现有采用KNN算法的电力负荷预测方法的不足，并利用相关的能源软件，通过Portelt技术从数据访问、总体结构和功能划分等方面对电力负荷预测系统进行了设计。然后，对数据异常预警、数据分析等功能领域的应用效果进行了观察。本文设计的基于knn的电力负荷预测方法能够准确地分析和预测短时间内的电力负荷，具有较高的预测精度。Cui Can[9]等采用LSTM预测模型，得到更准确的电力负荷预测结果。根据电力负荷的时间序列规律，建立了用于负荷预测的LSTM预测模型，并通过实验验证了该方法的有效性。

从上述国内外的研究发现，通过神经网络进行短期电力预测是当前热点，而LSTM改善了RNN中存在的长期以来问题，且表现通常比时间递归神经网络更好。LSTM通过各种门函数，能够有效减缓长序列问题中可能出现的梯度消失或爆炸问题，而且具有更好的准确性和模型表达能力。

* + 1. **需求响应动态定价研究现状**

目前，主流的需求响应定价方法主要有基于优化理论、博弈论、强化学习的需求响应定价方法等。首先对国内外动态定价学习的相关研究进行分析。目前，动态定价问题运用于经济学、计算机科学、运筹学等多个领域。其中机器学习方法在动态定价问题中的应用，与本文研究内容较为相关。在机器学习领域，目前研究一般在神经网络、演化算法、强化学习等方面。其中，强化学习理论在动态定价中的应用是当前研究热点

Wen[11]等利用博弈论模型探讨了能源服务提供商(ESP)与电力消费者之间的相互作用，证明了纳什均衡的存在性。建立基于动态价格的需求响应(DR)模型，并以某个商业和住宅集群的实际数据验证了所提出的DR模型。Hassan[12]等对基于价格的动态需求响应(DR)进行优化建模，该动态需求响应包括弹性负荷和非弹性负荷以及可再生能源的有效利用即微电网(MG)中的光伏(pv)和风力涡轮机(WTs)。利用粒子群算法求解了柔性和非柔性负载的利润最大化优化问题。Hassan[13]等提出了一种差分私有需求响应增强动态定价(DRDP)策略，该策略结合了差分私有和基于使用的动态计费的优点。该策略有效地保护了用户隐私，并通过激励参与的智能家居来增强动态定价。Tang[14]等通过Stackelberg博弈来研究电网和耗电建筑物基于纳什均衡的交互策略，结果表明电网可以优化电价实现净利润最大化，减少需求波动;建筑物可以通过优化小时电力需求，减少使用电费。此外，并提出了增强的鲁棒交互作用来处理不确定性对交互作用的影响。Zhong[15]等采用Weber-Fechner定律和聚类算法构建定量响应特征模型，利用DQN构建了负荷聚合器的动态补贴价格生成框架，考虑了需求响应参与者的用户行为量化和用户之间的差异，使参与者通过动态补贴价格获得最大的收益。Zeng[16]等为解决天然气消耗与社会发展之间的供需矛盾日益突出的问题，基于价格弹性理论和用户满意度，着眼于天然气用户和天然气供应商参与DR的目的，构建了动态定价视角下的天然气DR模型。采用K-means算法对天然气负荷峰、平、谷进行分类，采用多目标粒子群优化(MOPSO)方法求解天然气负荷恢复模型，提出了一种有效的天然气负荷恢复策略。减缓了日负荷波动，提高了用户满意度，实现了用户和供气商利益最大化，有效平衡了系统需求。Hassan[17]等设计了一种基于线性回归的动态电价方案，使可变负荷和不变负荷用户利润最大化。通过粒子群优化技术求解需求响应优化问题，并对拟议的动态电价方案进行了评估。仿真结果验证了在两种情况下，与固定电价方案相比，所提出的动态电价方案能够维持可变负荷和不变负荷用户的利润空间和舒适度。

从上述研究可以发现，机器学习领域的强化学习理论在很多领域都进行了相关应用，在解决动态定价的问题方面也有较好的研究基础，其中 Q-learning 算法作为强化学习算法之一，也被国内外相关研究者使用。但目前应用于电力需求响应动态定价，尤其是私有企业用户响应动态定价方面的研究较少。因此本文将私有企业用户的需求响应的动态定价问题转化为有限的马尔科夫决策过程，结合 Q-learning 算法实现动态定价模型的求解。

1. **研究基础**

电力定价问题是一个交叉混合问题，在实际研究前需要对相关概念以及关键技术进行细致探讨。本章分析了私有企业用户侧负荷的相关特点，并结合其运行特点选择合适的数学模型对其进行精准刻画，并得出用户侧整体负荷模型的特性，梳理了几种常用的负荷预测方法的特点，最后，本章梳理了当前常见的几种需求响应动态定价方法，从而为后续的基于强化学习的需求响应动态定价方法打下一定的基础。

**2.1 电力需求响应类型分析**

需求响应（Demand Response，简称DR）即电力需求响应的简称，是指当电力批发市场价格升高或系统可靠性受威胁时，电力用户接收到供电方发出的诱导性减少负荷的直接补偿通知或者电力价格上升信号后，改变其固有的习惯用电模式，达到减少或者推移某时段的用电负荷而响应电力供应，从而保障电网稳定，并抑制电价上升的短期行为。

需求侧响应策略主要分为两种:基于价格和基于激励，如图2.1所示。其中基于价格的需求侧响应策略分为分时电价、尖峰电价和实时电价。分时电价是国内较为常见的一种电价策略，能有效反映电网不同时段供电成本差别的电价机制，其措施主要是在高峰数段适当提高电价，在低谷时期适当降低电价，降低负荷峰谷差，改善用户用电，达到削峰填谷的作用。

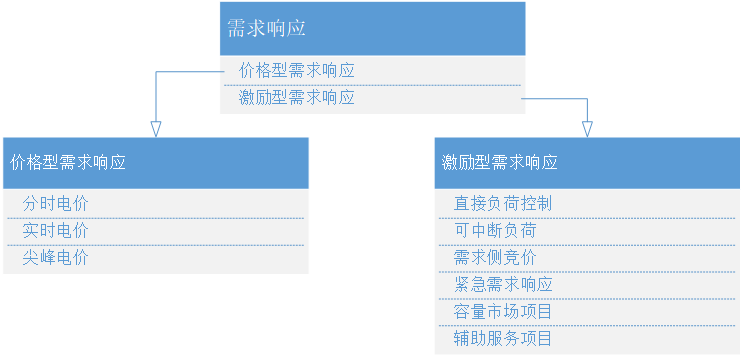


图2.1 需求响应分类

基于激励的DR是指DR实施机构根据电力系统供需状况制定相应政策，用户在系统需要或电力紧张时减少电力需求，以此获得直接补偿或其他时段的优惠电价，包括直接负荷控制（Direct Load Control，DLC）、可中断负荷（Interruptible Load，IL）、需求侧竞价（Demand Side Bidding，DSB）、紧急需求响应（Emergency Demand Response，EDR）、容量市场项目和辅助服务项目等。参与用户获得的激励一般有两种方式：一是独立于现有电价政策的直接补偿；二是在现有电价基础上给予折扣优惠。在需求响应计划实施前，通常DR实施机构要与参与用户提前签订合同，在合同中约定需求响应的内容（削减用电负荷大小及核算标准、响应持续时间、合同期内的最大响应次数等），提前通知时间、补偿或电价折扣标准、以及违约的惩罚措施等。

**2.1.1价格型需求响应**

基于价格的需求响应(price-based demand response, PBDR)是指随着电力价格的变化，电力用户自主的调节用电行为，主要分为三种形式：分时电价、实时电价和尖峰电价，主要调节对象是可调节负荷部分。当电价升高时，用户会减少可控负荷的使用量，将其调节到电价相对较低的时段使用，依次达到缩减电费开支。

1. 分时电价（Time-of-Use Pricing，TOU）机制

分时电价是基于电网系统的运行情况，按照尖、峰、平、谷四个等级，对一天24小时进行划分，根据所处的时刻不同收取相应的电价。在用电高峰期，系统需要高负荷运行，收取的电费相对较高，在用电低谷，系统低负荷运行，维护成本低，收取电费相对较低，电费价格按照尖、峰、平、谷四个等级依次降低。通过价格激励的方式鼓励用户在高峰期少用电，低谷期多用电，达到避峰削谷的目的。

1. 实时电价（Real-Time Pricing，RTP）机制

实时电价是在电力系统能正常运行和市场条件满足的情况下，结合电力长期成本和短期成本为定价依据，在给定的极短时段（如30min、15min、5min）内向用户提供电能的边际成本。实时电价能反映短期的生产成本及用电量信息，能够指导用户优化用电。和分时电价类似，在电力系统负荷较高时，电价相对较高，电力系统负荷较低时，电价相对降低，不同之处在于实时电价的电价变化频率更高。

1. 尖峰电价（Critical Peak Pricing，CPP）机制

尖峰电价是在前两种电价形式的基础上，增加了一个应急电价，在系统出现紧急情况下触发。尖峰电价的发生是不定时的，且出现频率很低。CPP主要有4种模式：固定时段CPP(CPP-F)、变动时段CPP(CPP-V)、变动峰荷定价和尖峰补贴电价

**2.1.2激励型需求响应**

基于激励的需求响应(incentive-based demand response, IBDR)是指实部门通过制定相关政策，调控用户在电力供应进展或电价较高时，积极响应并减少负荷需求，包括多种形式：需求侧竞价、直接负荷控制、可中断负荷、紧急需求响应、容量以及辅助服务计划。参加基于计划的激励型需求响应项目的用户通常要和实施部门签署有关协议，告知用户激励和处罚措施。

（1）直接负荷控制

直接负荷控制项目（Directed Load Control，DLC）是指电力系统运营机构在配电网负荷高峰时期，在不提前告知用户或只是短时间（15min以内）暂时通知用户，就采用遥控的方式调整或者关闭用户用电设备，并且给予用户一定的补偿机制，例如电费折扣或现金补偿等方式。这种方式适用于无重要设施的小型用户，而大型企业用户往往由于存在启动一次需要付出较大成本的大型机器设备而不适用。由于直接负荷控制的不确定性，往往会事先和用户签订相关条款，条款表明无需用户同意即可对其设备进行断电。DLC的目的通常是在价格激励无效的情况下，主动干预来削减电网负荷，保持电网正常运行。

（2）可中断负荷

可中断负荷通过签订经济合同，在[电网](https://baike.baidu.com/item/%E7%94%B5%E7%BD%91/1250763?fromModule=lemma_inlink)[高峰时段](https://baike.baidu.com/item/%E9%AB%98%E5%B3%B0%E6%97%B6%E6%AE%B5/18816595?fromModule=lemma_inlink)或紧急状况下，电力用户可以中断的负荷部分，一般针对大型工商业用户。一方面可以在负荷高峰期电网系统高负荷运行，另一方面电力用户可以通过中断相对不重要的负荷获取响应的经济补偿。可中断负荷是削峰的主要手段，对需求侧管理的实施有着重要的作用。随着用户侧分布式负荷调控技术的日趋成熟，可削减/可转移负荷目前已经逐步向居民用户开放，允许短时间范围内对一定容量的可控负荷进行调控，且需要满足一定的舒适度要求。

（3）需求侧投标项目

需求侧投标是一种较为新颖的 IBDR 项目，其目的在于激励大型工商业用户参与投标，投标策略为其在某价格下愿意削减的负荷量以及在削减一定量的电力负荷期望获取的补贴电价。在投标前，各参与电力用户需要准确评估自身的负荷削减量、持续时间以及响应速度，并基于这些基本特性进行投标，投标信息被市场运行商（Market Operator，MO）接受后，市场运行商依据参与者的投标电价以及投标电量进行排序，根据规则，中标者将按照最高投标电价获取补偿。

1. 紧急需求响应

紧急需求响应是电力用户在电力系统发生紧急情况时，通过减少负荷以避免系统运行安全风险及产生尖峰电力价格的紧急调整用电事件。容量市场／辅助服务是指用户提供増减负荷量作为系统的备用资源，以此来缩减对发电机组的调节需求的一种备用形式。在实施部门发出紧急需求响应指令后，大型用户自主参与负荷削减，实施部门将奖励用户很高的经济补偿。由于用户是自主参加，所以削减高峰期负荷电量有很大随机性，但是在丰厚的经济补偿刺激下，参加此计划的效果还是比较明显。

**2.1.3需求响应类型分析小结**

由上可知两种需求响应方式在维护电网正常运行、调控电力用户用电行为方面不仅相辅相成，还存在内部联系。价格型需求响应主要是通过制度不同的电价，引导电力用户根据自身情况自行规范用电行为，主角是用户，具有自主性。激励性需求响应主要是售电商或负荷系统以宏观调控负荷，电力用户处于被动，其往往是在特殊情况或PBDR已无法维持负荷正常管理得情况下进行的。

**2.2 私有企业用户负荷特性分析**

本文研究私有企业用户的电力定价，其中主要研究私有工业企业。根据2022年发布的中国电力消费情况，全国全社会用电量86372亿千瓦时，全国工业用电量56000亿千瓦时，占全社会用电量的比重为64.8%，工业用电是社会总用电量的重要组成部分，保障工业企业用电是重中之重。

**2.2.1私有工业企业负荷分类**

私有工业企业用电设备主要分为三种：主要生产负荷、辅助生产负荷和非生产负荷

主要生产负荷通常是值大型生产机器进行生产活动所消耗的负荷，比如车床、空压机、熔炼炉等，该类负荷中断后生产设备将会停止甚至造成损失，再次启动需要清理机器中的生产杂物、耗费大量人力物力。

辅助生产负荷是指不直接进行生产，对生产进行辅助的设备所需负荷，比如传送带、恒温器等，为正常的生产活动提供必要的环境。

非生产负荷指办公区的中央空调、热水器、照明等

**2.2.2私有工业企业负荷特性**

工业用电具有用电负荷大，规律性高的特点，在需求响应方面有很大的潜力。

**2.3 电力负荷预测方法分析**

电力用户在用电过程中，受天气变化、昼夜更替和用电习惯的影响，负荷往往会出现周期性、规律性的变化。短期负荷预测就是根据负荷以往的曲线特性和相关因素的影响，来预测未来短时间内的负荷用量，所以选择合适的方法和手段来刻画用户用电画像对预测结果的准确性尤为重要。按照预测方法来划分，短期负荷预测往往分为一下两种。

1. 经典方法预测

经典方法预测由传统的数学模型来完成的，在早期神经网络出现之前，由于电力系统相对简单，可以选择合适的数学统计方法进行建模来进行预测。主要的方法包括回归分析法、指数平滑法等。回归分析法的步骤一般分为三步，首先分析历史数据，确定自变量和因变量之间的线性和非线性关系，然后基于变量的数量，建立相应的一元或多元回归方程，最后对负荷进行预测。指数平滑法是一种加权移动平均法，适用于规律性较强的时间序列数据，通过观察历史负荷确定不同的权值，按照权值求得对应移动值，最终求得预测值。只根据最近一期的真实值和误差值即可求得预测值，有较高的局限性。

1. 智能预测

智能预测是主要有支持向量机、专家系统和神经网络等。专家系统通过模拟人类专家的思考和判断，融合人工智能技术，通过学习的相关领域的专家知识和经验对未来短时间内的负荷进行预测。其核心在于将人类能读懂的知识提炼出来，转化成机器能识别的符号，然后再根据知识库中大量专家知识进行推理，以此来获得预测的负荷。而支持向量机则是寻找一个平面，将数据进行分割，经常用来作为分类器。神经网络是由大量的神经元组成的一个网状结构，神经元之间相互连接，通过模拟人脑的对信息的选择性“遗忘”和“记忆”，在存储信息的同时对信息进行自主学习，以此达到预测的目的。常见的神经网络包括BP神经网络，循环神经网络（RNN），长短期记忆神经网络（LSTM），门控神经网络（GRU）。RNN包含输入层、隐藏层和输出层，用于独特的信号反馈结构，当前层的输出可以和上一层的输入相关联，具有记忆和传输的能力，从而可以进行时间序列的预测，但是同一层的神经元不会相连，在传输中会出现误差。而LSTM引入了灵活的记忆和遗忘模式，可以充分解决RNN在传播中产生的梯度爆炸和梯度消失问题。GRU将LSTM的遗忘门和更新们整合成一个新的更新门，训练的效率更快，训练的误差也很小，但随着数据集的增加，性能就不如LSTM。本文的研究对象负荷，是一种数据量大，且随着时间的推移会越来越大，且特征复杂的数据集，为追求更好的预测准确性，将会采用LSTM算法来预测未来短期的负荷。

**2.4 需求响应定价方法分析**

**2.4.1基于优化理论的需求响应定价方法**

基于优化理论的需求响应定价即是综合考虑用户的参与需求响应的成本和收益，求解用户最佳需求响应策略，建立的需求响应定价模型。用户参加DR的经济成本是指因参与DR而削减的那部分负荷的费用和响应补贴，随着需要负荷的削减，对应的舒适度或者满意度（舒适成本）也会下降，是和负荷的平方成正比的。用户的舒适成本不是一成不变的，也不仅仅和负荷相关，同时会受电费补贴的影响，电费补贴高，用户的不满会下降，电费补贴低，用户的不满会升高。用户参加DR的收益主要是削减部分的负荷所减少的花销和获得的经济补贴。目前基于优化理论的需求响应方法有直接对定价方法进行优化的，也会通过优化成本函数，减少用户成本等方式。

文献[20]基于数据中心工作负载的服务质量(QoS)要求的现实成本函数对数据中心进行建模，并提出了成本函数参数估计的逆优化方法，以实现精确和有效的定价。文献[21]设计了一个DR价格信号，根据消费者对峰值需求的贡献来区分消费者。提出了日前市场的两层定价机制，其中价格由固定的基本价格和随时间变化的惩罚价格组成。消费者需求出价分为两个部分:基本部分和惩罚部分。基本价格税率适用于基本成分，处罚价格税率适用于处罚成分。在电力消费者消费电费以后，对具有惩罚性成分的消费者进行补偿。文献[22]提出一种基于分布式计算的需求响应定价方法, 配电网运营商采用迭代分布式算法计算使用户净能耗成本最小，并沟通激励价格，从而鼓励客户积极参与需求响应计划。

**2.4.2基于博弈论的需求响应定价方法**

基于博弈论的需求响应定价方法是通过双方或多方之间的竞争关系，改变响应策略，以此达到最小化自身成本。有售电商和用户之间的博弈，也有用户与用户之间的博弈，这个博弈关系是一种非合作博弈的竞争关系。

用户和售电商之间的博弈关系可以描述为以下步骤：

1. 售电商作为领导，发布参与需求响应的电价，目标是使自身利益最大化。
2. 用户收到电价后，结合自身情况，决定响应的负荷量，目标是保障舒适度的情况下最小化用电成本。

售电商先给出电价，拥有决策先行性，是主动方。而用户作为电价的承受者，是被动方。是一种标准的主从博弈问题，该问题的优化最常见的求解方法即为将下层模型的优化问题用其 Karush-Kuhn-Tucker（KKT）最优性条件替代，进而将整个问题转化为混合整数线性规划问题，并通过成熟商业软件或者利用启发式算法进行求解。

而用户与用户之间的博弈则要更复杂一些，参与需求响应的用户还要考虑其他用户的决策，每个用户都以自身最大收益为目的进行决策，循环迭代，用户不断更新自己的决策，若干轮博弈后达到均衡。

目前基于博弈论的需求响应定价方法有很多，文献[18] 提出一种多维度保密竞价博弈，是一种公平且保护隐私的激励定价方法。根据用电用户的用电负荷特点，可以将其划分为不同的群组。然后采用积分的公平机制来保证所有电力用户之间的公平。同时考虑电力公司的收益和电力用户的风险，使电力用户实现均衡削峰。文献[19] 认为需求响应定价是一个双层优化问题，在优惠券激励的需求响应方案下，制定了负荷服务实体与消费者之间的Stackelberg博弈来实现电力定价。文献[23] 利用博弈论模型探讨了能源服务提供商与电力消费者之间的相互作用，证明了纳什均衡的存在性。能源服务提供商对电力定价进行优化，以最大限度地提高其效用，减少负荷波动，同时最大限度地减少电费和电力消费者的不满。

**2.4.3基于强化学习的需求响应定价方法**

随着参与需求响应业务的用户不断增多，业务场景也变得复杂起来，原始的方法已不适用于新的业务，基于强化学习的需求响应定价方法被提了出来。强化学习的主要思想是通过智能体与环境的不断响应，从环境中不断学习，获取到最大奖励的一种策略。在智能体学习的过程中，环境通常需要被规范化为马尔科夫决策过程（Markov Decision Process，MDP）。MDP 是一种贯序决策的数学模型，通常包括状态空间、动作空间、转移函数以及奖励函数组成，可理解为当前状态下智能体选择的动作不仅对于当前奖励值有影响，还会影响下一个状态以及奖励。强化学习在DR定价的应用，使得DR实现了自动化，且准确性和响应速度得到了很大的提升，且受到了国内外大量学者的热捧。

文献[24]梳理了深度强化学习的发展历程和现状，总结了深度强化学习技术，并对深度强化学习在需求响应方面应用的可行性进行了分析，提出了深度强化学习在需求响应业务的研究框架。文献[25]将强化学习算法应用于求解分散式储能最优充放电策略，售电商在峰谷时段参与需求响应，取得周期最大收益。文献[26] 提出了基于强化学习的 DR 动态定价策略，将需求响应的动态定价问题描述为有限的马尔科夫决策过程，并采用 Q 学习算法实现对马尔科夫决策过程的求解。

1. **短期负荷预测**

电力定价是一个复合型问题，在定价之前，首先要分析用户的历史负荷数量，预测未来短期的负荷需求，进而可以准确的把握负荷的需求曲线。然后再根据价格激励的方式，引导用户规范用电行为，削峰填谷，在降低用户用电成本的同时，保持电网稳定运行。所以，进行负荷的短期预测是非常有必要的。

**3.1 数据预处理**

由于得到的数据可能包含了大量的缺失值，可能包含大量的噪音，也可能因为人工录入错误导致有异常点存在，若不进行预处理可能会导致模型训练非常耗时，模型收敛效果差等一些问题。数据预处理的主要有：数据清理、数据集成、数据规约和数据变换。

数据清洗：主要思想是通过填补缺失值、光滑噪声数据，平滑或删除离群点，并解决数据的不一致性来清理数据。本文主要采用的数据清洗方法有：删除缺失值达到80%的变量；针对缺失值较低，且重要性较低的变量采用固定值填充；删除离群点；

数据变换：对数据进行规范化，离散化，稀疏化处理，达到适用于挖掘的目的。本文采用最大-最小规范化处理，将数据映射到[0,1]区间。

**3.2循环神经网络的基本原理**

传统人工神经网络中，如图3.1所示，其输入层与隐含层、输出层与隐含层之间全连接，且各层神经元之间独立，导致其忽视了输入数据的时序相关性，进而导致这种网络结构对长时间序列问题处理能力很差，如在训练Apple这个单词的标签时，可以是水果也可以是公司，如果不结合上下文，很难翻译正确。为解决该问题，循环神经网络（Recur-rent Neural Network，RNN）应运而生。

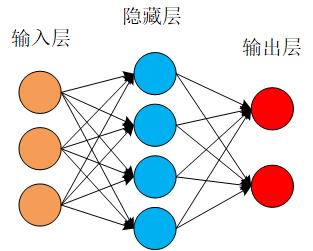


图3.1 神经网络结构图

**3.2.1 循环神经网络的基本原理**

循环神经网络（Recurrent Neural Network, RNN）是一类以[序列](https://baike.baidu.com/item/%E5%BA%8F%E5%88%97/1302588?fromModule=lemma_inlink" \t "_blank)数据为输入，在序列的演进方向进行[递归](https://baike.baidu.com/item/%E9%80%92%E5%BD%92/1740695?fromModule=lemma_inlink)且所有节点（循环单元）按链式连接的[递归神经网络](https://baike.baidu.com/item/%E9%80%92%E5%BD%92%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C/16020230?fromModule=lemma_inlink" \t "_blank)。不同于传统神经网络，RNN内部存在循环结构，可以将隐藏层的神经元连接起来，从而实现了前后信息的传递，展开的结构如图3.2所示。

由图3.2所示可知，循环神经网络主要分为三个部分：输入层、输出层和隐藏层，是t时刻的输入向量，是t时刻的隐藏层的状态，U是输入层和隐藏层的权值，V是隐藏层和输出层的权值，W是t-1时刻与t时刻的隐藏层间的权值，激活函数记为g，则t时刻的输出值的公式如3-1所示：

(3-1)

另外从图3.2中看出，t时刻的隐藏层的状态的公式如3-2所示，其中f是隐藏层的激活函数：

(3-2)

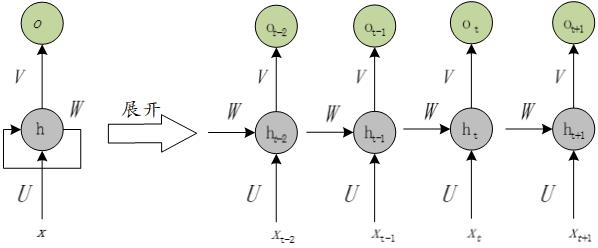


图3.2 循环神经网络

**3.2.2 长短期记忆神经网络的基本原理**

然而RNN存在一个致命的缺陷：无法做到长期依赖，当两个输入之间的步长太大时，会不可避免的出现梯度消失或梯度爆炸，长短期记忆神经网络（LSTM）应运而生。LSTM内部存在独特记忆和遗忘模式，当运用在数据规模较大的序列数据时，可以充分保留序列数据的特征，内部结构如图3.3所示。

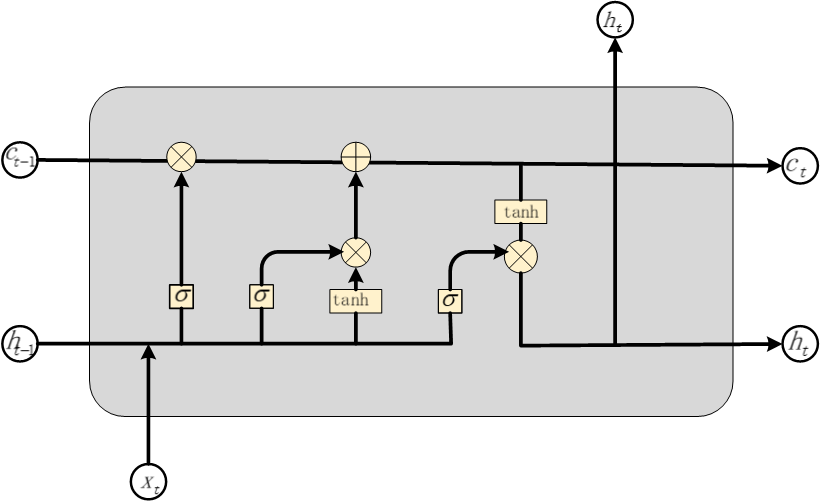


图3.3 LSTM 单元结构

LSTM由三个门来控制细胞状态，这三个门分别称为忘记门、输入门和输出门。输入门控制当前计算的新状态以及以多大程度更新到记忆单元中；遗忘门控制前一步记忆单元中的信息以多大程度被遗忘掉；输出门控制当前的输出有多大程度取决于当前的记忆单元。

遗忘门主要决定将单元中的那些信息丢弃，控制t时刻流入到该门的信息哪一部分被遗忘，然后输出一个[0,1]之前的值到上一时刻的细胞状态，其中0表示完全遗忘，1表示完全保留，对应公式如3.3所示。

(3-3)

其中是t时刻输入，是t-1时刻隐藏门的状态，是连接权重，是偏置向量，是激活函数，常用tanh或sigmoid，分解图见。

输入门是决定在单元状态中保留哪些信息，其中表示决定哪些信息将被更新的sigmoid函数，tanh表示当前时刻输出的细胞状态（公式3-4），对应公式如3.5所示。

(3-4)

(3-5)

t时刻的细胞状态，用数学公式表示为：

(3-6)

输出门是决定输出什么内容，它包含两部分，首先，运行一个sigmoid层，它决定要输出单元状态的哪部分。然后再将单元状态设置为tanh（值介于-1到1之间），最后两者相乘得出输出值。对应公式如3.7所示

(3-7)

最终输出表示为：

(3-8)

遗忘门、输入门和输出门的分解图分别对应图3.4、图3.5和图3.6。在一个训练好的网络中，当输入序列没有重要信息时，LSTM遗忘门的值接近为1，输入门接近0，此时过去的记忆会被保存，从而实现了长期记忆；当输入的序列中出现了重要信息时，LSTM会将其存入记忆中，此时输入门的值会接近于1；当输入序列出现重要信息，且该信息意味着之前的记忆不再重要的时候，输入门接近1，遗忘门接近0，这样旧的记忆被遗忘，新的重要信息被记忆。经过这样的设计，整个网络更容易学习到序列之间的长期依赖。另外在LSTM中，状态是通过累加的方式来计算的。不像RNN中的累乘的形式，这样的话，它的导数也不是乘积的形式，这样就不会发生梯度消失的情况了。

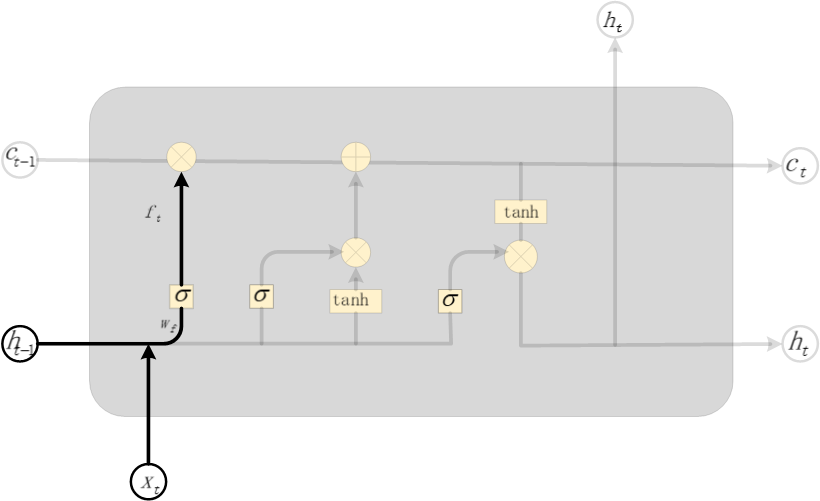


图3.4 遗忘门

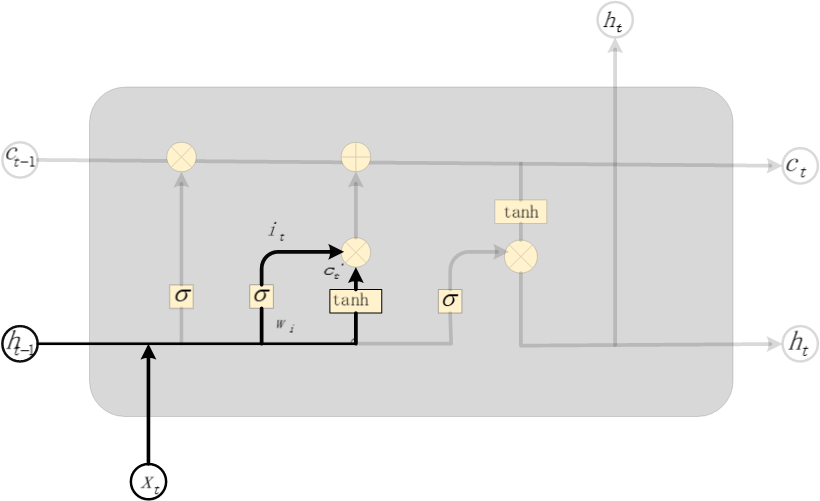


图3.5 输入门

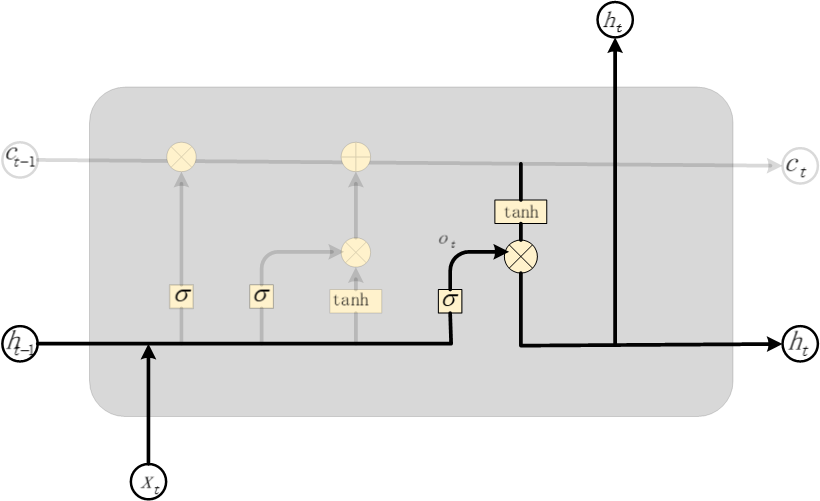


图3.6 输出门

**3.3 长短期记忆神经网络短期负荷预测**

短期负荷影响因素繁多，比如家庭用电主要和天气、温度有关，随人们的生活特征、饮食习惯变化。而本文的研究对象是私有工业企业用电，用电设备主要是大中型生产机器，其用电量随订单需求变化，具有一定的周期性、连续性和波动性，受天气温度的影响较小。排除影响较小的输入变量，主要影响因素包括历史负荷、月份、日、小时和峰平谷类型。

**3.3.1 历史负荷数据预处理**

本文选取要预测时刻前连续的历史负荷，针对一些异常情况（如系统宕机、自然灾害、电网检修等）导致的异常数据，如果不进行处理会导致收集的数据不完整不真实，进而导致负荷曲线出现非正常波动，对预测的结果产生很大影响。为保证预测的准确性，本文对异常数据进行如下处理。

针对缺省值取正常曲线或同期中位数填充；针对异常值（波动超过）可以相邻数据作为基准值，比较前后差值，变化超过我们设定的误差灵敏值太多，则认为异常，可取前后均值填充。另外常见的神经网络的激活函数工作区间很窄，如sigmoid为（0,1）,tanh为（-1,1），而负荷数据波动较大，并不能完全落到（-1,1）的区间内，会导致激活函数在超过其工作区间时发生过饱和，大大降低了网络训练效率，因此需要进行归一化处理，本文采用的是min-max归一化，其公式如下：

(3-9)

公式中是样本归一值，、、分别是样本值、样本最小值和样本最大值。样本归一值预测后，需要恢复实际值，方便判断模型的准确性和性能，反归一化公式如：

(3-10)

**3.3.2 短期负荷预测流程**

短期负荷预测流程一般分为数据预处理和模型搭建两部分，结合3.3.1得出流程如下：

步骤1：收集历史负荷数据，对数据中缺省值采用同期中位数填充，异常值前后基准值判断，并用前后均值修改。

步骤2：数据预处理，对负荷数据进行归一化处理

步骤3：取历史数据的70%作为训练数据，搭建LSTM神经网络进行循环训练

步骤4：取剩余数据的30%作为验证数据，利用训练好的模型进行预测

步骤5：将预测的数据进行反序列化，得到最终预测的实际值

步骤6：预测值与实际值进行对比，计算预测误差

其中步骤6中涉及的预测误差评价指标可按照以下两个误差公式来计算：

1. 均方根误差（Root Mean Square Error，RMSE）

(3-11)

1. 平均绝对百分比误差（Mean Absolute Percentage Error，MAPE）

MAPE = \*100% (3-12)

n是测试点的数量，是第个测试点的真实值，是第个测试点的预测值。

**3.4 预测结果及其分析**

本文选用国内广东省某一私有制造型企业2017年5月1日 1:00到2017年8月20日24：00之间的负荷数据，共计112天，间隔1小时，样本容量为2688，如图3.7所示。

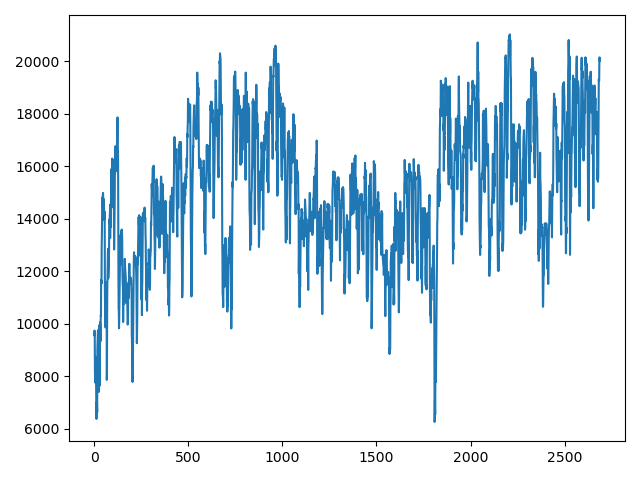
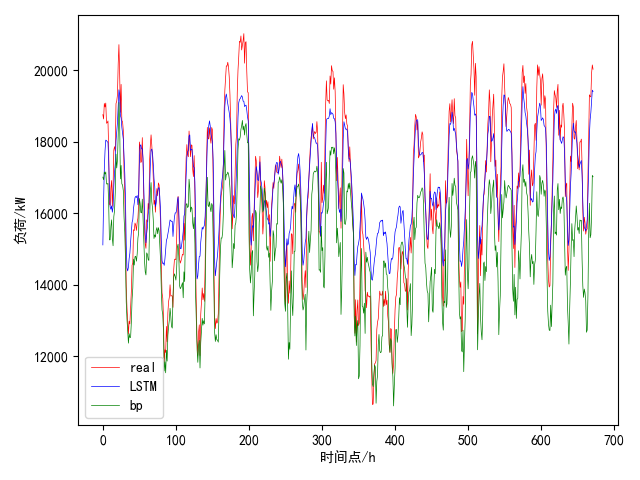
****

图3.7 负荷数据图

特征值有月份、日期、小时、峰平谷和历史负荷，一共有5个，因此LSTM模型输入层神经元是5，本文根据过去24小时的负荷来预测未来1个小时的负荷，神经网络输出层神经元个数为1。隐藏层节点数为32，激活函数是ReLU，优化器是Adam，学习率是0.001，损失函数是MSE。根据上述参数构建LSTM神经网络，安装3.3.2的预测流程，分别采用BP神经网络和LSTM神经网络进行预测，并用结果进行对比，预测结果如图3.8所示：

****图3.8 测试集预测与真实值对比图

以某天的预测结果进行分析，如图3.9所示。对比BP神经网络和LSTM神经网络可以看出，都可以大致反映真实负荷的趋势，具有较好的拟合结果。根据误差层面进行分析，对比平均绝对百分比误差，LSTM当天的误差是2.01%，而BP神经网络的误差是7.20%，可见LSTM神经网络更优。对比均方根误差，LSTM当天的误差是381，而BP是1266，LSTM神经网络更优。综合来说LSTM是最优的。

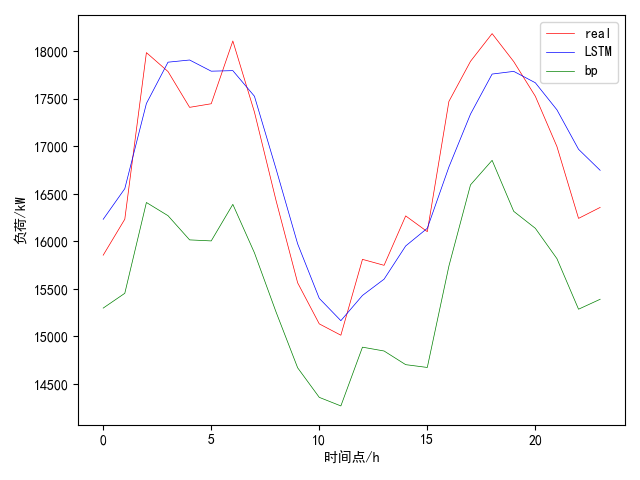
****

图3.9 某天的真实负荷与预测负荷曲线

1. **动态定价策略**

在本章中，我们提出了一种面向电力市场中用户侧需求响应的新型动态定价方法。该方法独特之处在于，它同时考虑了售电公司（EPC）和私有企业用户（PEUs）之间的成本因素，将负荷动态定价的复杂过程抽象成了一个离散有限的马尔科夫决策过程。接着，我们运用Q-learning算法，使系统能够实现最优决策。这一方法的核心优势在于，通过强化学习的框架，售电公司能够在实时学习的过程中灵活地制定零售电价，提高了对负荷不确定性和批发电价波动性的适应能力。Q-learning算法的应用使得系统能够逐步学习并优化决策，以更加有效地适应电力市场的动态变化。最终，我们通过在广东某私有企业用户的具体案例中进行验证，证明了所提出模型及算法的实用性和有效性。这种方法不仅在理论上具有创新性，而且在实际应用中得到了成功的应用，为电力市场中用户侧需求响应提供了一种可行而实用的动态定价解决方案。

**4.1 企业动态定价模型**

本文提出了一种创新性的电力市场交易模型，其中售电公司（EPC）以先从电力批发市场以批发价获取负荷的方式开始，然后根据市场波动情况灵活地制定零售电价，参与电力零售市场。在更大范围上，EPC不仅具备电力批发资格，能够从上级电网以批发电价购买电力，同时也是售电商，向下级居民用户出售电量。实际上，EPC在这个过程中充当了电能交易的中间媒介，通过不断调整零售电价，引导私有企业用户（PEUs）进行需求响应，并通过电量和电价的反馈不断调整零售电价，最终实现双方的共赢。从售电公司的角度看，通过实时调整零售电价，EPC可以有效引导私有企业用户在不同时间段内做出灵活的需求响应。这个过程是一个动态的调整过程，通过及时的电量和电价信息反馈，使零售电价更贴近市场实际情况，提高了售电公司在电力市场中的竞争力。对于私有工业用户而言，通过响应售电公司发布的零售电价，他们可以根据自身电力需求和能源供应成本，在每一时刻进行自适应的用电量和用电行为调整。这使得私有企业用户在满足负荷需求的同时，最大化地降低用电成本，实现经济效益的最优化。总体而言，这种分级电力市场模型实现了电力市场各方之间的协同作用。通过动态调整零售电价，售电公司引导私有企业用户更加灵活和经济地响应需求，最终实现了整个电力交易过程中的双赢局面。分级电力市场的具体效果如图4-1所示。

在电力需求响应的过程中，两个主体售电商和用户具有不同的决策，售电商作为需求响应先行的一方，了解电价市场波动及用户负荷等，且作为中间商往往以长期收益最大化为目的；而用户则由于只能被动接受电价，只能追求短期收益。

**4.1.1 售电商收益**

以单个用户的为例，t时刻，售电商从电力批发市场，以批发价购买电量，再结合用户参与需求响应的程度，以合理的零售价销售给用户，其收益函数可如公式4-1所示：

(4-1)

T表示时间，，将一天24小时的实际划分为 T 个时段，本文设 T=24,表示每小时作为一个响应时段。表示为t时刻的零售电价，表示为t时刻的批发电价，表示当前用户t时刻的实际总负荷。针对售电商而言，

一般零售价是大于批发价，且受市场规则和用户需求响应的限制。

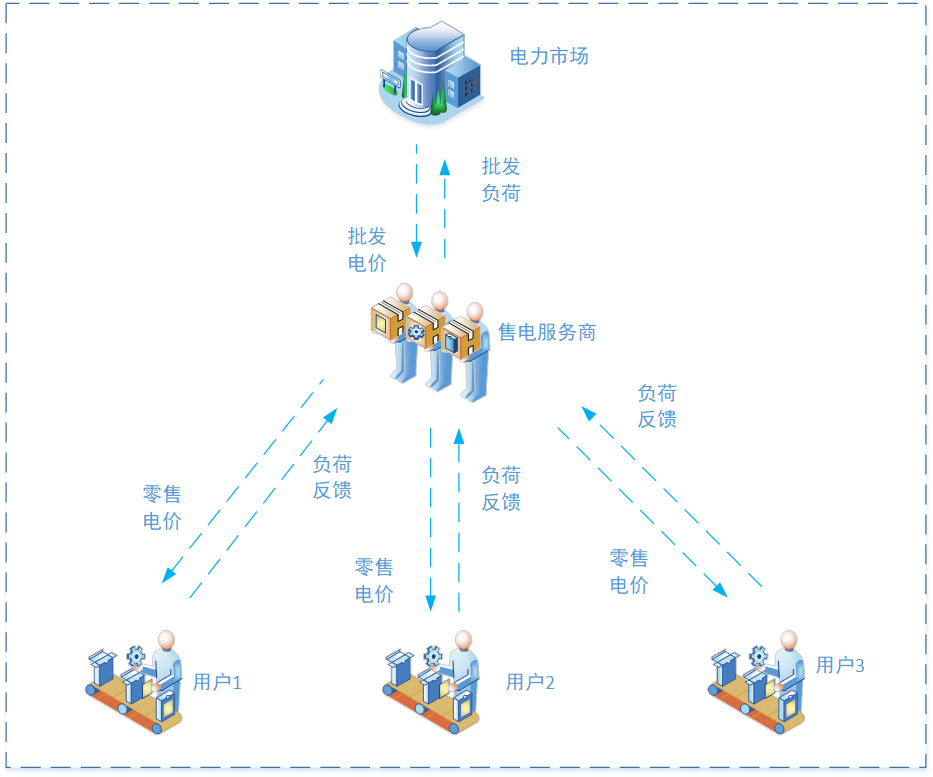


图4.1 电力市场分级图

**4.1.2 用户成本**

私有工业企业用户的负荷可分为可参与需求响应负荷和不可参与需求响应负荷，首先不可参与需求响应负荷指的是必须要消耗的负荷，如照明、办公等，其可表示为表达式4.2所示：

= k\* (4-2)

式中，为t时刻不可参与需求响应的负荷，为t时刻的负荷需求量，为t时刻实际不可参与需求响应的负荷，k为不可参与需求响应的负荷占总负荷的系统，每个用户都不相同，可通过用户的需求响应成本与响应量之间的历史数据关系辨识得出。

可参与需求响应的负荷，其负荷量会随着电价的升高而降低，电价的降低而升高，其可表示为表达式4-3所示：

(4-3)

(4-4)

(4-5)

(4-6)

式（4-3）中为可参与需求响应负荷需求量。式(4-4)中为t时刻可参与需求响应的实际负荷，表示t时刻的需求响应弹性系数，需求响应的大小随着零售电价的大小按照一定的比例系数进行弹性调整。弹性变量指的是一种经济变量对另一种变量变化速度的响应程度的度量，在智能电网中，能源需求的大小将随着价格的变化而变化，且弹性系数通常为负数，表明需求和电价成反比关系。此外，在不同时段，需求响应弹性系数也不尽相同，通常电力需求高峰时期的负荷要比电力需求低谷时期的负荷更加具有弹性，且长期弹性往往大于短期弹性。 结合文献[72]可知通过对美国三个全国性数据集进行的试验，得出弹性响应范围在[−0.8, −0.2] 区间范围内变化[68]。

可参与需求响应的负荷会随着电价的升高而减少，同时会对用户造成损失成本，其可以用二次函数表示，如式（4-7）所示。

(4-7)

(4-8)

(4-9)

式（4-6）中为用户因可参与需求响应负荷减少而产生的损失成本，和是两个损失成本的系数。式（4-8）中是可参与需求响应负荷减少的上限，可参与需求响应负荷的下限。

综上，要使用户的用电成本最低，可表现为式（4-10）。

(4-10)

综合式（4-1）和（4-10），售电商收益最高和用户用电成本最低的综合表达式如式（4-11）所示

(4-11)

式中是权值因子，，其值由售电商决策。

**4.2 基于强化学习算法的动态定价模型**

电力定价问题涉及到私有企业用户在时间序列上做出的电力购买决策，这是一个动态的过程，与前后时间状态紧密相连。为了解决这个问题，可以采用强化学习算法，这种方法具有高效的计算能力，可以在不断与环境交互中获得接近最优解。通过智能体学习零售电价与用户响应之间的关系，能够更好地适应市场电价和负荷需求的复杂多变性，并具备一定的鲁棒性。在这个问题中，售电商充当智能体的角色，私有企业用户则是环境，零售电价则是智能体在决策过程中的可选动作。私有企业用户的信息，例如负荷需求大小、各类参数系数等，构成了状态空间，而售电商的利润和私有企业用户的成本则成为奖励信号。因此，我们可以将这一动态定价问题建模为有限的马尔科夫决策过程。基于这个模型，可以利用Q-learning算法高效地求解问题。Q-learning是一种值函数驱动的强化学习算法，通过不断更新Q值来学习在特定状态下采取某个动作的最佳策略。这种方法使得售电商能够根据环境的变化调整零售电价，以最大化其利润。这个方法的优势在于，通过与环境的交互学习，智能体能够适应市场的不确定性和用户需求的变化，具备较好的计算效率和鲁棒性。当然，在实际应用中，可能需要考虑一些具体问题，如状态空间的维度、学习率的选择以及对抗性用户行为的因素。总体来说，采用强化学习算法，特别是Q-learning，对解决电力定价问题是一种合理而有效的方法，同时需要在实际应用中灵活调整以应对具体情境。

**4.2.1动态定价问题的MDP转化**

在强化学习中，我们定义了执行强化学习算法的智能体，以及智能体感知的环境。学习的过程实际上是智能体与环境之间不断互动的过程。一个有限的马尔科夫决策过程由四个主要部分组成，包括状态空间、动作空间、转移概率以及奖励函数。

其中，售电商扮演了智能体的角色，而批发市场、零售市场与用户构成了智能体与环境的交互环境。智能体首先在某个时刻的状态下观测环境，获取当前各个用户的可参与需求响应的负荷信息以及批发市场电价信息。在获得这些信息后，智能体计算在不同动作下的奖励值，并根据奖励值的大小选择相应的动作。在给定当前时刻和所选择动作的条件下，智能体进一步计算状态转移的概率，并根据这些概率更新至下一个时刻的状态。这个过程在每个时刻不断重复，直至完成一个完整的回合。以下将从状态空间、动作空间、转移概率以及奖励函数四个方面详细建立其动态定价问题的马尔科夫决策过程。

1. 状态空间

在这个MDP过程中，状态包括实时定价周期T、批发市场实时电价、居民侧负荷需求量和、电能实际消耗量和。其中，可参与需求响应负荷的实际消耗量是一个未知变量。一天的零售电价定价时段被分为T个决策周期，每个决策周期对应一个时段[t,t+1)，批发市场电价按价格大小被离散化为0-个电价等级，而t时刻在电价等级为k下的批发电价记为。未知变量按照负荷削减等级离散化为0-个状态等级，表示用户在t时刻处于第l级负荷需求等级下的负荷需求量。因此，EPC的动态定价问题在第t个决策周期的状态空间可以表示为：

（4-12）

1. 动作空间

售电商的动作空间主要是指定不同的零售电价，对其进行离散化处理，假设零售电价的等级一共级，上下线分别为、，于是离散化的动作值如下所示：

（4-13）

（4-14）

1. 状态转移概率

针对一个有限的MDP过程，状态转移概率一般为：

（4-15）

而结合EPC的动态定价问题，其状态转移概率为：

p() （4-16）

由于可参与需求响应的电量是未知的，所以在时刻 t 的转移概率p()也是未知的。鉴于本文采用基于Q学习算法来解决上述MDP过程，这是一种无模型的算法，它不需要提前了解状态转移概率的具体数值。相反，通过智能体的不断探索与学习，它能够从与环境的交互中获得经验，并隐式地学习转移概率。这使得智能体能够逐步确定不同状态下的概率转移大小，从而实现对问题的求解。这种无模型的方法使得算法更加灵活，能够适应未知的、动态变化的环境，从而更好地解决动态定价问题。

1. 奖励函数

奖励函数的设置会影响到算法的收敛速度和收敛程度，且直接关系到售电商制定零售电价的最优定价策略，当智能体处于状态，选择动作后，奖励函数可写成如下形式：

（4-17）

式中，是权重因子，表面售电商的利润和用户的成本之间的相对重要性。

**4.2.2基于 Q-learning 算法的动态定价问题求解过程**

Q-learning 算法是一种基于值的无模型强化学习算法，可用于解决模型不可知前提下的 MDP 贯序决策问题，该算法不需要依赖数学模型，智能体通过不断与环境的交互过程中获取反馈信息，基于值函数迭代不断学习，以至达到最优策略。定义表格化的 Q 函数值为在状态ts 时采取动作ta 的动作价值函数的期望：

（4-18）

智能体通过不断探索更新动作价值函数 ，逐步逼近最优动作价值，获取最优动作价值函数的迭代机制可用贝尔曼方程表示：

（4-19）

式中，为状态下的可行动作，为学习率，代表 Q 值迭代的速度。第 i 次迭代后，达到最优 Q 值的收敛判据表示为：

（4-20）

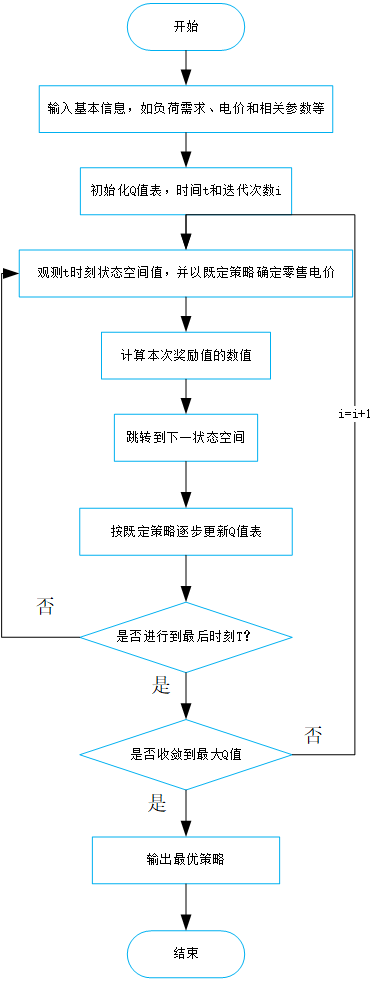


图4.2 动态定价算法流程图

其中为接近于 0 的一个很小的常数值。可认为此时。当收敛条件满足时，迭代停止，并输出此时最佳的动态定价策略。

综上所述，基于强化学习的私有企业用户负荷需求响应动态定价模型求解步骤如下：

步骤 1：输入算法所需基本参数信息，主要包括：居民基本电能需求、居民舒适度参数信息、批发市场电价信息*,*，零售电价边界系数等；

步骤 2：初始化 Q 值表为 0，时间间隙为 1，迭代次数为 1；

步骤 3：观测状态空间值，并在动作空间中以 ε-greedy 策略选择零售电价；

步骤 4：计算奖励值大小，并跳转到状态空间, 按式（4-19）更新Q值表

步骤 5：判断是否更新至最后时间状态 T，若是，则跳转至步骤6，否则跳转至步骤3；

步骤 6：判断是否收敛到最大的Q值，若是，则输出最优定价策略；否则跳转至步骤2。

流程图如图4-2所示。

**4.3 算例及结果分析**

1. **总结和展望**
2. **参考文献**
3. 孙云翔.电力系统负荷预测研究综述与发展方向的探讨[J].电子乐园, 2019(19):1.
4. Du J , Cheng Y , Zhou Q ,et al.Power Load Forecasting Using BiLSTM-Attention[C]//IOP Publishing Ltd.IOP Publishing Ltd,2020:032115 (11pp).DOI:10.1088/1755-1315/440/3/032115.
5. Tang D,Li C,Ji X, et al. Power Load Forecasting Using a Refined LSTM[P]. Machine Learning and Computing,2019.
6. Óscar G,Vanessa G,Manel M. Forecast-informed power load profiling: A novel approach[J]. Engineering Applications of Artificial Intelligence,2020,96.
7. Guo X, Jiang Y, Li L, et al. Short-term power load forecasting based on DQN-LSTM[C]//2022 34th Chinese Control and Decision Conference (CCDC). IEEE, 2022: 855-860.
8. Kong X, Wang Z, Xiao F, et al. Power load forecasting method based on demand response deviation correction[J]. International Journal of Electrical Power & Energy Systems, 2023, 148: 109013.
9. Lu M, Wu X, Chen Y, et al. A novel method for the short-term power load forecasting[C]//2022 IEEE 9th International Conference on Power Electronics Systems and Applications (PESA). IEEE, 2022: 1-5.
10. Fan L,Guang J. Research on power energy load forecasting method based on KNN[J]. International Journal of Ambient Energy,2019.
11. Cui C, He M, Di F, et al. Research on power load forecasting method based on LSTM model[C]//2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC). IEEE, 2020: 1657-1660.
12. Di S. Power system short term load forecasting based on weather factors[C]//2020 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM). IEEE, 2020: 694-698.
13. Wen L, Zhou K, Feng W, et al. Demand side management in smart grid: A dynamic-price-based demand response model[J]. IEEE Transactions on Engineering Management, 2022.
14. Hassan M A S, Chen M, Lin H, et al. Optimization modeling for dynamic price based demand response in microgrids[J]. Journal of cleaner production, 2019, 222: 231-241.
15. Hassan M U, Rehmani M H, Chen J. Differentially private dynamic pricing for efficient demand response in smart grid[C]//ICC 2020-2020 IEEE International Conference on Communications (ICC). IEEE, 2020: 1-6.
16. Tang R, Wang S, Li H. Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids[J]. Applied Energy, 2019, 250: 118-130.
17. Zhong S, Wang X, Zhao J, et al. Deep reinforcement learning framework for dynamic pricing demand response of regenerative electric heating[J]. Applied Energy, 2021, 288: 116623.
18. Zeng H, Shao B, Dai H, et al. Natural gas demand response strategy considering user satisfaction and load volatility under dynamic pricing[J]. Energy, 2023, 277: 127725.
19. Hassan M A S, Assad U, Farooq U, et al. Dynamic price-based demand response through linear regression for microgrids with renewable energy resources[J]. Energies, 2022, 15(4): 1385.
20. Zhang K, Shi Y, Liu Y, et al. Power demand response incentive pricing model[C]//2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019: 2605-2614.
21. Radoszynski A M, Dvorkin V, Pinson P. Accommodating bounded rationality in pricing demand response[C]//2019 IEEE Milan PowerTech. IEEE, 2019: 1-6.
22. Tsiligkaridis A, Paschalidis I C, Coskun A. Data center demand response pricing using inverse optimization[C]//Proceedings of the Tenth ACM International Conference on Future Energy Systems. 2019: 400-402.
23. Basnet A, Zhong J. Pricing mechanism for demand response based on penalty paradigm[C]//2019 IEEE Milan PowerTech. IEEE, 2019: 1-5.
24. Nainar K, Pillai J R, Bak-Jensen B. Incentive price-based demand response in active distribution grids[J]. Applied Sciences, 2020, 11(1): 180.
25. Wen L, Zhou K, Feng W, et al. Demand side management in smart grid: A dynamic-price-based demand response model[J]. IEEE Transactions on Engineering Management, 2022.
26. 孙毅,刘迪,李彬等.深度强化学习在需求响应中的应用[J].电力系统自动化,2019,43(05):183-191.
27. 邵明明,刘友波,陈婧婷等.分散式储能自趋优经济运行的强化学习算法[J].电网技术,2020,44(05):1696-1705.DOI:10.13335/j.1000-3673.pst.2019.2108.
28. 郝旭东,孙伟,程定一等.基于Q强化学习的综合能源服务商现货市场申报策略研究[J].电力建设,2020,41(09):132-138.