中心探索利用大数据挖掘技术开展主动式运营监测，使用用户行为数据分析预测用户欠费风险。

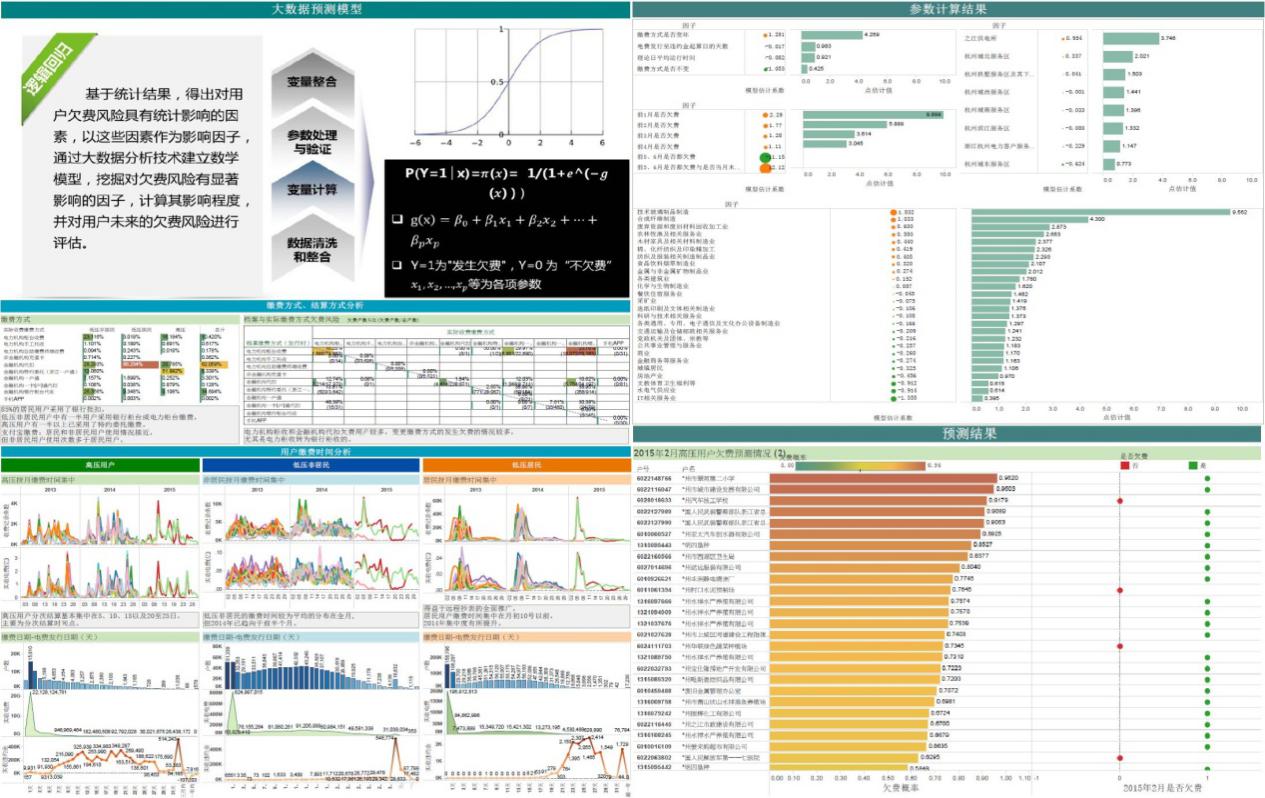


图9-2 主动式“电费回收风险”挖掘分析可视化效果图

通过深入挖掘杭州等地38.34万个客户的档案信息、用电行为、缴费行为、增减容行为、违约窃电欠费行为、95598客服及短信记录、以及行业特征、外部环境等多维度信息共计1900余万条数据，分析识别客户欠费行为与各类客户特征、客户行为、行业趋势、外部因素的关联关系，发现了大量以前未知的业务规律和管理盲点，如欠费时间分布特点、欠费行为与缴费方式变化的关联、违约金起算时间误差、银行批扣电费不及时等问题。通过定量计算各相关因素与欠费行为的关联度，提取关键影响因子,基于逻辑回归算法建立了电费风险预测模型。初步使用该模型对杭州地区大客户的欠费概率进行预测，欠费概率前10名预测准确率达80%，前30名预测准确度达73.33%。

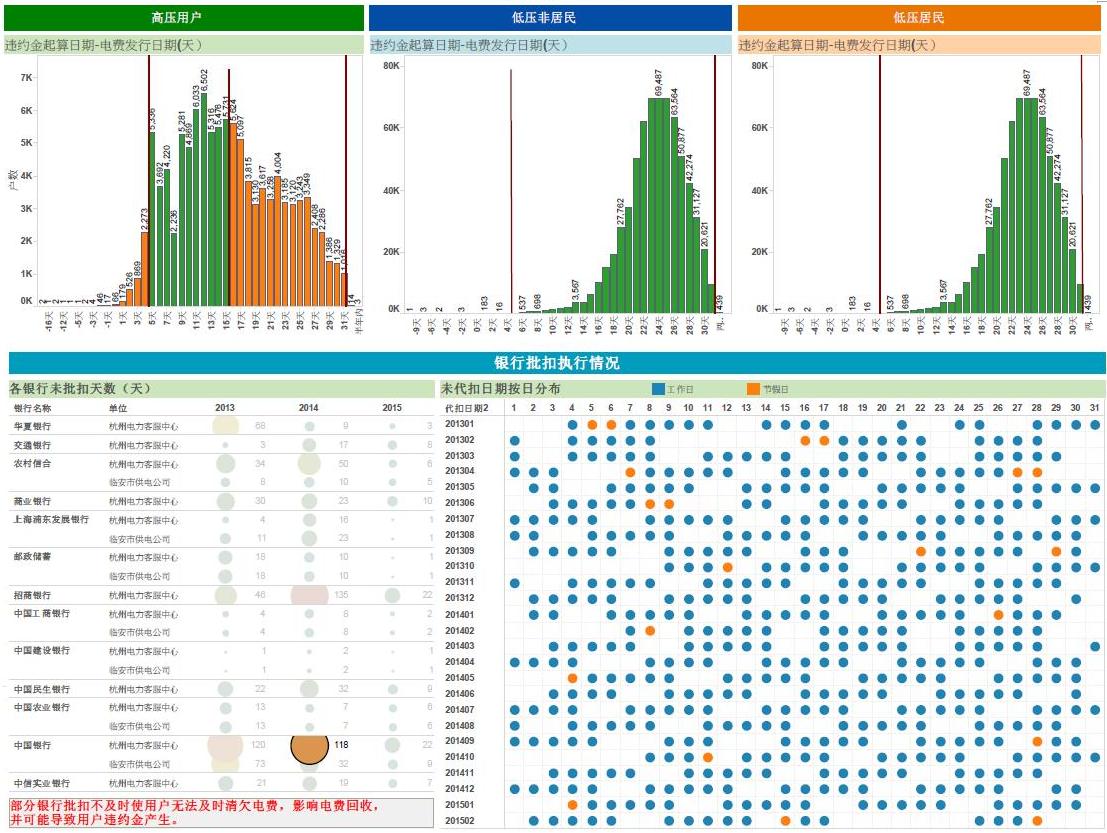


图9-3主动式“电费回收风险”挖掘分析可视化效果图（二）

该项成果可以通过对用户行为数据的在线监测，提早识别高风险的欠费用户，提示营销部门有针对性地采取应对措施，降低电费回收风险。借助大数据技术，初步实现了“主动感知问题，主动预防风险、主动辅助决策”的预期目标。

## 9.2 基于多源数据融合分析的企业风险预警与管控

电网作为全球最大的关键基础设施，面临着严峻的内外部经营风险和挑战。外部风险主要体现在舆情危机，内部风险主要体现在物资管理、廉政建设、法律事务、企业文化等内部经营管理风险。现有的风险监测方法及时性不足，缺乏对关键指标的量化手段和辅助分析工具，过度依赖人工判别，难以支撑企业经营管理风险的精准有效处置。

基于新一代大数据技术，如大数据智能、自主智能等技术，助力网络舆情监测分析，帮助公司及时发现外部经营风险；基于大数据、知识图谱等技术，有效发现物资管理、廉政建设、法律事务、企业文化等内部经营管理制度、流程、规范中存在的逻辑漏洞、落实不当、管理不到位等风险，实现企业内部经营管理风险的常态化监测及告警，提升公司经营管理风险管控水平。国家电网公司（以下简称公司）全面推进了以信息化、自动化、互动化为基本特征的智能电网建设。信息化建设在公司电网发展、生产运行、经营管控、客户服务等各个领域的全面覆盖和不断深入，各类业务数据急剧增长，表现出量大、价值高、结构多样等特征。通过挖掘数据资产潜在价值，使利用数据指导生产和管理企业成为可能。

受到当今外部环境变化的影响，公司内部经营管理的风险复杂性增加。公司的企业经营现状存在数据量大、业务复杂等特点，在很多应用场景下表现出大维度、小样本、非结构化的数据特性，传统的风险分析预警模式已经难以适应当今企业内部风险管控的而需求，而新一代人工智能技术为企业风险管控提供了新的动能。深度学习技术使企业风险管控由经验主义变为理性主义，深度学习能够弱化人为主观分析判断的成分，依托真实数据量化分析，其结果的准确性将大大提升；深度学习技术使企业风险管控由粗放管理变为精细管理，依托深度学习技术构建的风险分析预计体系将能够更精准的定位风险成因并跟踪风险的演化路径，使企业风险管控更加精益化；深度学习技术使企业风险管控由被动变为主动，深度学习技术的应用将实现企业风险的智能化、自动化、在线化分析预警，实现对企业风险的全天候监管，实现风险的提前预判，降低风险的发生率；深度学习技术使企业风险管控由局部变为全局，依托于公司内部经营管理全量数据进行深度学习分析，可以构建企业经营管理的全景分析体系，实现对企业内部多维风险的掌握，依照业务逻辑分析多元风险的交互作用，从而实现对企业风险的全面认知。

**9.2.1基于大数据的电网企业舆情风险管控**

从潜在的舆情风险点来看，电力体制改革进入攻坚阶段，各利益相关方通过媒体表达诉求，观点交锋，一旦当改革无法达到各方预期时，公司有可能成为舆论情绪的宣泄口；新能源发展迅速，电网消纳舆情风险依然突出；供电服务质量预期攀升，一线员工服务水平和意识有待提升；随着国际业务的深入开展和全球能源互联网的不断推进，地缘政治的复杂多变性、业务开展的风险性、政治经济利益的博弈，都可能带来公司海外业务在国内外舆论场的变数和不确定性。

现有风险监测方法及时性不足，缺乏关键指标的量化手段和辅助分析工具，过度依赖人工判断难以支撑企业经营管理风险的精准有效处置。

第三方信源较少，影响复杂问题的传播公信力。国家电网公司直接或间接掌控的第三方信源较少，不太重视负面舆情事件中第三方的作用，尚未形成公司与第三方信源互相协调配合应对舆情事件的工作机制。

议题设置能力有待提升，危机转化能力有待加强。国家电网公司已经形成了较为规范的新闻发布制度，但总体上来看，主动、积极的发布尚需加强，议程设置能力有待提升，容易出现媒体“先声夺人”，选择有利于自身且社会关注的热点事件进行追踪报道。

联动协作障碍较多，跨部门、跨层级配合尚需完善。电网公司内部机构庞大，如果遇到舆情事件，需要调动各部门力量进行配合，沟通及运作的障碍确有不少，尤其是部分县级供电企业，普遍缺乏制度化、系统化的应对机制。此外联动和舆情导控机制不规范、预警处置预案不科学、未建立信息分析研判机制、正面引导机制未形成、与网络媒体合作机制缺失等现象也影响舆情风险及时、有效的处置。

**1．研究****海量Web（微博、论坛、新闻等）多区域、跨渠道外部信息的数据结构化内容提取技术**

**（1）数据噪声过滤**

针对海量数据的国网公司及其关注业务主题、议题的舆情风险管控研究的需求，提出基于自然语言词向量表示学习模型的特征表示方法，解决大规模数据中噪声过滤难题和海量数据分布式并行高效过滤问题，噪声数据类型涵盖了广告、娱乐新闻、体育赛事新闻、心灵鸡汤等噪声数据类型，实现基于有限规模噪声数据样本的高效分布式数据过滤功能。

**（2）信息精准分类**

针对海量数据中的国网公司业务相关主题、议题的精准分类需求，解决自然语言文本中公司业务主题相关的舆情分层分类难题，提出基于文本深度学习模型Deep Graph CNNs的国家电网大数据舆情精准分类，同时利用业务主题标签之间依赖关系的作为正则优化项来提供精准分类准确率和召回率，以及领域内分布式并行训练与分类的优化方法。

**（3）业务内容提取**

针对海量Web（微博、论坛、新闻等）多区域、跨渠道外部信息的数据舆情分析的结构化内容提取的需求，根据公司所涉及业务定义结构化要素，涵盖“时间、地点、人物、公司、职务、关键词、主体、谓词、客体、情绪、类型”等可解释性要素，提出基于LSTM和条件随机场CRF的实体识别模型解决上述要素抽取难题。

**2.研究文档主题生成模型和概率潜在语义分析技术，获取关于公司的主题结构**

**（1）业务领域主题结构检测**

针对与国电公司业务和关注领域网络数据的更细粒度主题检测需求，利用最新融合流式图模型与自然语义词表示学习技术的文档主题检测模型，以及时序稀疏子图的概率潜主题检测的语义分析技术，解决公司关注网民发帖相关主题与潜在主题的检测问题，包括涵盖了关键词、实体、不同关系的异构网络上异常检测的参数选取问题。

**（2）主题结构模型增量更新**

针对文档主题生成模型与概率潜在语义分析均为线上系统的现状，研究包括除了流式图模型检测的数据时序高效增量更新数据和参数外，还有自然语言本身词语义飘逸以及近似增量表示学习瓶颈难题，以及基于估计的增量学习框架，解决文档主题生成模型和概率潜在鱼鱼分析的增量更新。

**3.研究基于自然语言处理的Web 短文本立场判定技术**

**（1）融合舆情热度的情感分类**

针对海量数据中涉及公司的舆情影响力的刻画研究需求，利用与公司业务和关注事件主题类型的情感词库、文本深度学习模型，研究舆情的精准情感极性分类，文本深度学习模型包括卷积神经网络、LSTM以及情感词的Attention模型，解决准确识别与公司有关的正、负面舆情和长期正、负面网络用户和网络媒体。

**（2）面向业务领域的立场判别**

针对海量数据中网民对涉及公司的舆情立场识别的研究需求，利用与公司业务和关注事件主题类型的舆情和评论者，以及立场相关特征词，研究网民对舆情立场的支持与反对的极性判别，文本深度学习模型包括卷积神经网络、LSTM以及情感词的Attention模型，解决与公司业务领域立场的准确识别问题。

**（3）基于深度学习的舆情趋势预测**

针对海量数据中涉及公司的全部舆情都需要热度趋势预测的研究需求，利用积累的历史舆情的发展趋势，包括主题、关键词、参与媒体、参与网民，以及舆情热度值等，研究融合时间序列热度值和关键词序列的舆情趋势预测模型，解决公司关注舆情的趋势长期预测与短期内准确预测难题。

**4.研究基于深度学习的舆情传播网络溯源分析技术**

**（1）构建公司舆情传播异构网络**

针对长期关注公司的传媒、微博用户、论坛用户等网民，研究利用历史舆情数据包括用户传播网络，构建异构的信息传播网络，包括网络节点对公司舆情关注主题词、情感和立场判别信息。

**（2）基于深度学习的演化溯源分析**

针对公司业务相关舆情碎片化现状，以及缺乏统一的时序聚类与描述，研究利用深度学习提取和表示自然语言文本语义特征，实现业务相关舆情的时序聚类，实现语义机演化与溯源计算，同时实现同一事件上演化与溯源序列要素的统一表示。

**（3）融合情感指标趋势的传播评价**

针对传统舆情趋势预测更多关注热度等度量指标缺乏融合网民情感趋势的监控，研究融合网民情感指标随着主题演化和传播完整过程的评价机制，实现融合网民传播行为与情感的舆情传播评价分析功能。

**9.2.2 基于深度学习的企业内部经营管理风险智能识别**

深度学习是机器学习中一种基于对数据进行表征学习的方法，随着人工智能的技术发展和应用，知识图谱以其强大的语义处理能力和开放组织能力，为互联网时代的知识化组织和智能应用奠定了基础，知识图谱作为关键技术之一，已被广泛应用于智能搜索，智能问答，个性化推荐，内容分发等领域。本课题主要涉及的关键技术包括：知识图谱的构建，特征编码技术，基于图的半监督学习，聚类分析，lime算法等。

（1）研究基于知识图谱的不同经营管理主体间多维度信息关联技术；

国网公司基于风险管理体系化、系统化、层级化的结构化思想，进行了风险管理框架的构建，确定了风险管理的层次，明确了各类风险的层级与关联关系。目前国网信息库主要由“通用风险+电网主业风险信息库”“产业特色风险信息库”“金融特色风险信息库”构成，将公司存在的风险分为三级风险，具体包括一级风险编号、一级风险名称、二级风险编号、二级风险名称、三级风险编号、三级风险名称、三级风险描述、三级风险成因等内容。

目前来看，公司在风险框架设计与方面，主要还是依赖人工经验总结与判断的方式对业务进行分类、细分进而设计出框架，虽然从理论上来说，可以基本覆盖公司所面临的业务且满足目前风险管理的需要，但仍存在如下进一步待提升的工作内容：一是目前风险框架只是从业务经验上反映了公司风险分类及层级情况，其完整度、准确度未经科学检验；二是目前风险框架只能反映各类风险的层级关系，但无法反映同级风险的重要性排序以及风险之间的关联关系，不能确定对各项风险的管理优先顺序和有效指导风险管理策略的制定。三是风险框架的更新与运维工作缺少自动化手段。

知识图谱（Knowledge Graph) 旨在描述现实世界中存在的实体以及实体之间的关系,以其强大的语义处理能力和开放组织能力，为互联网时代的知识化组织和智能应用奠定了基础。知识图谱通过描述各种实体或概念及其关系,构成一张巨大的语义网络图，逻辑上可分为模式层与数据层两个层次，数据层主要是由一系列的事实组成，而知识将以事实为单位进行存储。如果用(实体1，关系，实体2)、(实体、属性，属性值)这样的三元组来表达事实，可选择图数据库作为存储介质，模式层构建在数据层之上，是知识图谱的核心，通常采用本体库来管理知识图谱的模式层。本体是结构化知识库的概念模板，通过本体库而形成的知识库不仅层次结构较强，并且冗余程度较小。

企业风险智能识别管理核心是风险，运用共引分析和关键路径网络分析，绘制所有风险的知识图谱，显示了以风险管理为中心的企业风险管理研究的总体情况。根据知识图谱的聚集特征，明确核心研究群体，以及每个群体的代表和研究方向。为了更进一步分析，研究企业风险管理核心研究群体的特征。可以选取某个特征有代表性的研究群体进行样本分析。统计各个风险出现的频次，中心度（表示风险之间的联系，联系越多越紧密，中心度就越高）即表示各个风险的影响力。

通过对知识抽取，知识表示，知识融合，知识推理构建知识图谱之后，利用多维尺度分析的方法来对企业内部风险管理相关的文本信息进行分析，结合风险的相似度权重对分析结果绘制风险框架知识图谱使得更加直观和形象。

(2) **基于深度学习的经营管理风险特征编码技术**

随着大数据时代的到来和数据采集方式的更新换代，在电网企业运营的过程中，会产生大量的历史负面样本和投诉样本。面对这些海量的数据，如何有效的分析并提取出对经营风险有决定意义的特征成为亟需解决的难题。对此，深度学习在特征提取和建模上都有着相较于浅层模型显然的优势，受到了越来越多研究者的关注。

自动编码器则是深度学习中常用的一个主要模型，其在对海量的数据进行特征提取中发挥了重要的作用。它能将具体的特征向量逐渐转化为抽象的特征向量，能很好地满足高维数据空间和低维数据空间双向映射的非线性学习，它采用自适应、多层编码网络将高维原始数据转换成低维抽象数据，并利用类似的解码网络从低维抽象数据中重构原始数据的高维数据表示。

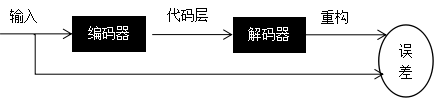


图9-4 自动编码器

一是对无标签数据进行预训练。面对涌现的海量数据，对每个数据进行标注需要耗费大量的人力物力。对此，本课题拟采用自动编码器对文本数据进行无监督学习。根据数据的特点拟从稀疏自编码、栈式自编码、去噪自编码、压缩自编码中选取合适的自编码方式。根据重构误差调整自动编码器的参数。

二是利用有标签数据进行优化模型参数。根据获得的自动编码器模型提取出数据的抽象特征，拟在自动编码器的最顶的编码层添加一个分类器（逻辑回归），然后进行多层有监督训练。分类的目标可以是业务类别（包括一级分类、二级分类、三级分类等），也可以是其它预设的分类标签。通过有监督学习对整个系统的参数进行微调。

根据模型获取有监督数据的抽象特征，结合卷积神经网络构建特征编码模型，即将上述特征输入卷积神经网络进行多层卷积、池化进一步获得特征编码。同理在这一过程中页考虑使用业务类别作为标签或其它预设类别作为训练数据的标签。



图9-5 基于深度学习的企业风险特征编码方法

（3）研究基于半监督学习的经营管理风险识别技术；

国网公司在风险识别工作中，主要是查找企业各业务单元、各项重要经营活动及其重要业务流程中有无风险，有哪些风险。主要采用了生产流程分析法、风险专家调查列举法、分解分析法等方法，核心即是对业务进行分类且细分至与可管理的最小业务单元，最小业务单元即为末级风险。

（4）研究基于特征值扰动的风险成因解释技术

机器学习算法在人工智能、大数据、物联网等热门领域得到广泛应用。但是现在很多前沿的机器学习模型还是一个黑盒，几乎无法去感知它的内部工作状态。这就带来了机器学习模型可信度的问题：该相信那个模型的某个预测结果是正确的吗？或者说该相信那个模型的预测结果普遍都是合理的吗？通俗的来说，如果生产系统正准备替换上一套基于机器学习的系统，它就要确保机器学习模型的工作状态是良好的。

从直观上看，解释每次独立预测背后的基本原理能使我们更容易信任或是不信任预测结果，乃至机器学习分类器本身。即使我们无法理解模型在所有情况下的表现，却有可能理解（大多数情况都是如此）它在某些特定情况下的行为。

## 9.3 电力大数据应用建议

**9.3.1对服务社会与政府部门应用建议**

（1）**社会经济状况分析预测** （电力数据与经济数据结合） （加引用）

电力关系经济发展、社会稳定和群众生活，电力需求变化是经济运行的晴雨表和风向标，能够真实、客观地反映国民经济的发展状况与态势。智能电网中部署的智能电表和用电信息采集系统，可以获取详细的用户用电信息。用电信息采集系统与营销系统所累积的电量数据属于海量数据，需要采用大数据技术，实现多维度统计分析、历史电量数据比对分析、经济数据综合分析等大数据量分析工作。对用户电量数据从行业、区域、电价类别多维度开展用电情况分析，提取全社会用电量及相应社会经济指标，分析用电增长与相应社会经济指标关联关系，归纳总结各指标增长率与全社会用电情况的一般规律。通过对用户用电数据的分析，可为政府了解和预测全社会各行业发展状况和用能状况提供基础，为政府就产业调整、经济调控等做出合理决策提供依据。

（2）**相关政策制定依据和效果分析**（电力数据、交通数据、财政数据）

通过分析行业的典型负荷曲线、用户的典型曲线及行业的参考单位GDP 能耗，可为政府制定新能源补贴、电动汽车补贴、电价激励机制(如分时电价、阶梯电价)、能效补贴等国家和地方政策提供依据，也可为政府优化城市规划、发展智慧城市、合理部署电动汽车充电设施提供重要参考，还可以评估不同地区、不同类型用户的实施效果，分析其合理性，提出改进建议。

**9.3.2 面向电力用户服务类应用领域**

（1）**需求侧管理、需求响应**（气候数据、社交数据、电力数据）

根据不同的气候条件(如潮湿、干燥地带，气温高、低地区)、不同的社会阶层将用户进行分类；对于每一类用户又可绘制不同用电设备的日负荷曲线，分析其主要用电设备的用电特性，包括用电量出现的时间区间，用电量影响因素，以及是否可转移、是否可削减等，对于受天气影响的用电设备，如热水器、空调等，需分析其对天气的敏感性。分析不同用户对电价的敏感性，包括在不同季节、不同时间对电价的敏感性。通过聚合，可得到某一片区域或某一类用户可提供的需求响应总量，再分析哪一部分容量、多少时间段的需求响应量是可靠的，分析结果可为制定需求管理/响应激励机制提供依据。

**（2）用户能效分析和管理**

对用户进行用电效率分析，通过安装的智能电表可以获得较短时间间隔的用电数据，可以通过电表数据，识别用户端的不同类型负荷比例，并与典型数据进行比对得出能效分析结果。

从海量用户的负荷曲线，采用数据挖掘技术，按照特定的函数算法，按行业、季度聚合成行业的典型负荷曲线模型，然后将所有的用户的负荷曲线与行业的典型负荷曲线进行对比，分析出与典型负荷曲线变化趋势不一致的用户，由此对用户的能效给出评价，并提出改进建议。

**（3）业扩报装等营销业务辅助分析**（基建部、营销部、运检部）

业扩报装辅助分析以营配集成为纽带，将用电信息采集系统、营销系统和PMS及SCADA系统的数据相融合，实现对变电站、线路及下挂用户和台区的负荷、电量检测分析，为加快业扩报装速度和提高供电服务水平提供技术支撑。同时极大提高电网设备运行的可靠性，为优化配电网络结构，降低电网生产故障，提高公司用电营销管理精益化水平提供了手段。

**（4）供电服务舆情监测预警**（舆情数据与销售部、运检部）

通过与微博、微信、博客等互联网新媒体的服务对接机制收集海量用电信息、用户信息以及互联网舆论信息与95598电力客服咨询与投诉数据，建设大数据舆情监测分析体系，利用大数据采集、存储、分析、挖掘技术，从互联网海量数据中挖掘、提炼关键信息，建立负面信息关联分析监测模型，及时洞察和响应客户行为，拓展互联网营销服务渠道，提升企业精益营销管理和优质服务水平。

**（5）电动汽车充电设施建设部署**（多领域数据融合）

融合电动汽车用户信息、居民信息、配电网数据、用电信息数据、地理信息系统数据、社会经济数据等，可利用大数据技术预测电动汽车的短中长期保有量、发展规模和趋势、电量需求和最大负荷等情况。参照交通密度、用户出行方式、充电方式偏好等因素，依据城市与交通规划以及输电网规划，建立电动汽车充电设施规划模型和后评估模型，对电动汽车充电设施的部署方案制定和建设后期的效能评估提供依据。

**9.3.3 支持公司运营和发展类应用领域**

（1）**电力系统暂态稳定性分析与控制**

在线暂态稳定分析与控制一直是电力运行人员追求的目标，随着互联电网规模越来越大，“离线决策，在线匹配”和“在线决策，实时匹配”的暂态稳定分析与控制模式已不能满足大电网安全稳定运行要求，因而逐渐向“实时决策，实时控制”方向发展。

（2）**基于电网设备在线监测数据的故障诊断与状态检修**

在实现GIS、PMS、在线监测系统等各类历史数据和实时数据融合的基础上，应用大数据技术进行故障诊断，并为状态检修提供决策，可实现对电网设备关键性能的动态评估与机遇复杂相关关系识别的故障诊断，为解决现有状态维修问题提供技术支撑。

（3）**短期、超短期负荷预测**

分布式能源和微网的并网增加了负荷预测和发电预测的复杂程度。负荷预测也必须考虑到天气的影响以及能源交易状况，包括市场引导下的需求响应等。传统的预测方法无法体现某些因素对负荷的影响，从根本上限制了其应用范围和预测精度。应用大数据技术，建立各类影响因素与负荷预测之间的量化关联关系，有针对性地构建负荷预测模型，可更加精确地预测短期/超短期负荷。

（4）**配电网故障定位**

利用大数据技术，配合故障投诉系统，融合SCADA、EMS、DMS、D-SCADA 等系统中的数据作出最优判断，建立新型配电网故障管理系统，可以快速定位故障，应对故障停电问题，提高供电可靠性。此外，随着分布式电源在系统中比重的逐渐增加，其接入会影响到系统保护的定值及定位判据。对于带分布式电源的配电网故障定位也要根据不同的并网要求选择合适的定位策略。

1. **防窃电管理**

电力公司通过电量差动越限、断相、线损率超标、异常告警信息、电表开盖事件等数据的综合分析，建立窃电行为分析模型，对用户窃电行为进行预警；通过营配系统数据融合，可比较用户负荷曲线、电表电流、电压和功率因数数据和变压器负载，结合电网运行数据，实现具体线路的线损日结算，通过线损管理功能不仅可以知道实施窃电用户所在的具体线路，并且可以定位至某一具体用户，克服目前检查范围广，查处难度大的问题。

（6）**电网设备资产管理**

基于电网设备信息、运行信息、环境信息(气象、气候等)以及历史故障和缺陷信息，从设备或项目的长期利益出发，全面考虑不同种类、不同运行年限设备的规划、设计、制造、购置、安装、调试、运行、维护、改造、更新直至报废的全过程，寻求寿命周期成本最小的一种管理理念和方法。依据交通、路政、市政等可能具备的外部信息，如工程施工、季节特点、树木生长、工程车 GPS

等外部信息，关联电网设备及线路 GPS 坐标，对电网外力破坏故障进行预警分析。

（7）**储能技术应用**

由于储能系统大多是由数量庞大的电池单体组成(动辄以万计)，每个电池单体又包含单体电压、电流、功率、电池荷电状态、平均温度、故障状态等相关信息，汇总起来整个电站监测信息可能达到数十万个点，储能相关数据量十分庞大。利用大数据分析技术，可对储能监控系统相关数据进行有效采集、处理与分析，为储能应用提供依据。

（8）**城市电网规划**

通过实现用户用电数据、用户停电数据、城市电力服务数据(95598 客服电话)、基于 GIS 的城市配电网拓扑结构和设备运行数据、城市供电可靠性数据、气候数据和天气预报数据、电动汽车充电站建设及利用数据、人口数据、城市社会经济数据、城市节能和新能源政策及实施效果数据、分布式能源建设和运行数据等、社交网站数据的整合，识别城市电网薄弱环节，辅助城市电网规划。在上述数据融合的基础之上，利用人口调查信息、用户实时用电信息和地理、气象等信息绘制“电力地图”，可以街区为单位，反映不同时刻的用电量，并将用电量与人均收入、建筑类型等信息进行比照。通过“电力地图”，能以更优的可视化效果反映区域经济状况及各群体的用电行为习惯，为电网规划决策提供直观依据支撑。

**9.3.4 大数据驱动的业务优化与能力提升**



图9-6 企业风险预警与管控流程

一是促进企业信息化水平提升。通过分析电网企业监测关键技术，基于公司运营全业务流程，完善企业管理技术架构，提高信息抓取和在线分析能力，实现异动问题的在线监测，提升公司信息化水平，为实现数据应用统筹、数据资源共享、智能决策支持、运营全景展示打好技术基础，同时为提高企业管理效率提供技术支持。

二是促进企业数据资产优化。通过研究企业原始数据，建立适用于智能化和自动化的数据挖掘关联机制，统一数据质量规范，确定数据之间关联关系及其程度与性质，建立数据间的关联关系网络，为企业制定数据管理模式提供参考；同时通过充分发掘企业数据的价值，有效提升企业数据利用率，使企业管理更加全面，为企业制定相关运营发展策略提供参考和依据。

三是促进主动式智能化闭环监测的实现，降低企业运营风险。通过数据治理技术提升公司运营数据质量，通过自关联挖掘算法在监测指标与数据间建立关联关系，形成关联关系网络，根据监测指标异动向底层数据回溯，定位问题根因，开展智能化预测，实现异动处理的辅助决策和趋势预测，为公司数据化运营提供决策支撑。

2.间接效益

一是促进公司运营监测机制的优化。课题研究引入具备自动化的特征关联架构，构建运监中心与业务部门间稳固的信息和数据保障机制，提高数据在决策中的应用效率，支撑专业管理和战略决策，促进公司运营监测机制的优化，提升公司效率、效益和管理水平。

二是促进企业自我提升及健康发展。通过模拟企业运营中的指标异动，采用推理回溯的分析方法，定位出引起异动的根源问题所在，大大缩小整治范围，有助于优化企业运营策略，提高企业决策效率。同时通过构建异动处理辅助决策模型，对底层数据异动进行分析，基于数据视角提出异动处理建议，及时发现企业运营或管理中的不足。项目成果将有助于降低和规避企业运营的风险，对企业的健康发展起到重要的促进作用，为企业及国家相关部门制定发展战略提供参考。

三是促进业务人员技术水平提升。通过开展指标关联自动识别及监测、指标异动回溯、智能化预测和趋势分析，加深运监人员对主营业务的理解，全面提升业务人员的专业管理能力，提高公司对于经营状况的综合分析能力，保障公司安全、有序、健康、高效运营。

## 9.4 项目规划

**9.4.1 项目近期规划**

电网企业运营水平不断提升，电网信息平台所承载的数据信息量越来越多且难以高效充分利用管理，使得对电网运营监测体系的要求越来越高。电网企业应组织汇聚相关资源与力量，在电网企业范围内建立并完善基于大数据挖掘的主动式运营监测体系，实现“数据源头准确可靠、数据资产规范有序、大数据平台稳定运转、大数据应用丰富实用、成果展现灵活有效、协调控制和保障机制健全有力”，使得数据资产价值得到充分利用，基于大数据平台，应落实业务部门需求，信通部门统一开发交付的信息话统筹管理模式；持续开展应用服务化优化，落实系统优化改革并迁移；广泛开展数据挖掘，创新数据应用模式；实现技术统一、数据统一、管理统一、形成应用灵活的电网企业信息化新架构。

一方面，电网企业应满足横向贯通与精益化管理的监测需求。基于运营监测业务需求，通过横向协同和纵向管控方式，在监测分析和大数据挖掘全过程开展数据管理、异动管理、专项协调、辅助决策等工作，协调相关部门消除异动、解决问题、防范风险。与此同时，应加快其主要建设的两个方向，一是信息化基础平台建设，即建立统一的大数据信息平台，构建企业统一的数据模型和大数据库，通过部署业务应用和采集装置，自动采集业务流程数据和电网数据，完成企业运营行为采样和存储；二是主动式运营监测体系的流畅运转，即通过运营主题库和数据资产关联库的构建应用，采用大数据挖掘方法技术，将企业内部的明细数据梳理清楚，建立数据关联库，形成数据地图和数据关联卡片，并建立适用于电网企业业务流和资源流的分析算法模型，实现主动感知、主动预防和主动参谋，利用内部、外部数据解决内部运营问题，找到企业运营过程中的出血点，并对企业的运营关键点进行预测，进而达到企业精益化管理的目标。

另一方面，电网企业应开展运营检测业务主题库建设，形成主题挖掘、异动监测、问题协调的业务重点环节。建立异动管理机制，实现集异动监测、异动处理、跨业务协同、跟踪反馈、统计分析于一体的异动监测管理。通过统一平台、统一工单、多选工单对象、简化接单人、合并派发、合并工单一次性反馈等促进异动流程的高效流转。实行公司层面问题协调整改的分级管控机制，针对监测分析发现的问题、风险及决策建议，落实整改、防范和决策实施。

**9.4.2 项目远期规划**

通过运营监测（控）中心的统一规划形成主动式运营监测体系，国网浙江电力应遵循电网企业远期发展规划，合理利用近期主动式运营监测所具备的主动感知、主动预防、主动谋划工作成果，有效扩大运营监测视野，将监测成果与电网企业内外部数据资源有机结合，建立协同共享机制，充分沟通、协调、调动并分享各业务数据，延伸大数据技术链、价值链。形成更加多样化的监测主题，有效实现1+1>2的聚合效应。

开展具体业务的大数据挖掘工作是探索运监业务主题库和数据资产关联库的建设思路的目的所在。其远期规划主要业务领域可涵盖有无功电量计算情况专题监测分析及主动式“电费回收风险”挖掘分析等。

无功电量计算需要计量班、营业班和电费复核班相关人员的协同与配合。为提升公司的功率因数调整电费计算精益化管理水平，减少因无功电量计算错误导致的经济损失和法律风险，运营监测（控）中心可采取大数据技术抓取全省100千伏安（千瓦）及以上客户，进而开展无功电量计算专题监测，从营销业务应用系统和用电信息采集系统抽取尽可能多的明细数据，识别无功电量计算错误，并利用大数据技术按错误类别开展穿透分析，查找执行和管理层面的原因，并总结得出相关结论提出更好的建议。

随着经济增速放缓，公司电费回收压力日益增加。运营监测（控）中心探索利用大数据挖掘技术开展主动式运营监测，使用用户行为数据分析预测用户欠费风险。可以通过深入挖掘杭州等地客户档案信息、用电行为、缴费行为、增减容行为、违约窃电欠费行为、95598客服及短信记录、以及行业特征、外部环境等多维度信息及数据，分析识别客户欠费行为与各类客户特征、客户行为、行业趋势、外部因素的关联关系，力争发现以前未知的业务规律和管理盲点，如欠费时间分布特点、欠费行为与缴费方式变化的关联、违约金起算时间误差、银行批扣电费不及时等问题。通过大数据技术定量计算各相关因素与欠费行为的关联度，提取关键影响因子,基于逻辑回归算法建立电费风险预测模型。初步使用该模型对杭州地区大客户的欠费概率进行预测，使得相关部门可以通过对用户行为数据的在线监测，提早识别高风险的欠费用户，提示营销部门有针对性地采取应对措施，降低电费回收风险。借助大数据技术，力争实现并完善“主动感知问题，主动预防风险、主动辅助决策”的预期目标。

## 9.5 未来展望

在信息化深度发展的今天，大数据价值日益凸显，谁掌握了数据，谁就掌握了主动权。国网浙江电网公司应如何应对新经济形势的挑战，将企业运营过程中积累的大量数据真正运用起来，为企业创造商业价值与社会价值，已经成为关系到企业实现跨越式发展的首要问题。目前电网企业的信息化水平在逐步提升，而运用大数据挖掘开展主动式运营监测将是企业管理精益化，决策数字化实现的保障。从浙江电网企业的远景来看，聚焦打造信息化的电网企业，运用大数据全生命周期管理技术重塑电网企业核心价值观，推动电网企业在现代社会的可持续发展是未来的趋势。