

需提交的资料包括以下三部分，请分别压缩后在“会员中心”统一提交：

1、论文正文（pdf 格式），并压缩成“**论文正文.zip**”，提交时间：**2016 年 5 月 15 日 24:00** 前；

2、论文正文（doc 格式）、源数据（组委会提供的源数据外）、过程数据、程序及模型文件，压缩成“**附件资料.zip**”，提交时间：**2016 年 5 月 16 日 24:00** 前；

3、参赛承诺书打印后签名，并扫描生成 pdf 文件，压缩成“**参赛承诺书.zip**”，提交时间：**2016 年 5 月 15 日 24:00** 前。

所选题目：

综合评定成绩： _____

评委评语：

评委签名：

以下模板仅供参考，参赛者可根据实际情况进行调整！

论文报告标题（请替换）

摘 要：简要描述论文摘要

关键词：提取论文关键词

The thesis title

Abstract: A brief description of the abstract

Key words: The keyword extraction

目 录

1. 挖掘目标.....	5
2. 分析方法与过程.....	5
2.1. 总体流程	5
2.2. 具体步骤	5
2.3. 结果分析	6
3. 结论	6
4. 参考文献.....	6

1. 挖掘目标

请简要描述本次数据挖掘建模要达到的目标。

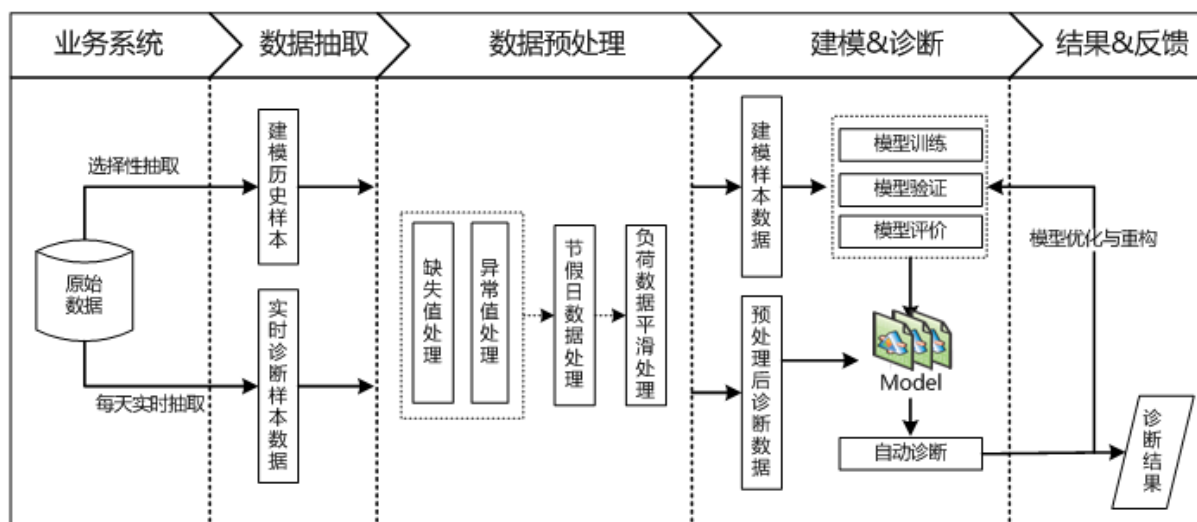
例如：本次建模目标是利用电力系统中积累下来的海量真实数据，采用数据挖掘技术，分析各类电力数据直接的相互关系、发现事件之间的内部关联，构建反映用户窃漏电规律的评价指标体系、专家样本库和窃漏电评价模型，实现对新采集用户计量数据的窃漏电自动评价。从而可以为防止某些用户在用电上作假和偷电提供有效的实时监控。

2. 分析方法与过程

2.1. 总体流程

用一个总体流程图描述建模方法及过程，并对各部分进行简要说明。

例如：



本用例主要包括如下步骤：

步骤一：...

步骤二：...

步骤三：...

2.2. 具体步骤

结合总体流程图，对每一步骤做详细说明。

例如：

对本例自适应窃漏电诊断建模法进行详细说明。

步骤 1: 样本数据抽取 由**错误! 未找到引用源。**知, 窃漏电评价影响因子主要有电量类指标、负荷类指标、线损类指标和报警类指标等, 故进行窃漏电诊断时需根据建模要求有选择性地从营销系统及计量系统中抽取电量、负荷、报警及线损等数据。

为了尽可能全面覆盖各种窃漏电方式, 建模样本包含近三年来所有的窃漏电用户及部分正常用户。

步骤 2: 样本数据预处理 数据预处理主要包括缺失值处理、异常值处理、节假日数据处理、负荷数据平滑处理等。

✧ 缺失值处理

在原始计量数据, 特别是实时负荷数据抽取过程中, 发现存在缺失的现象, 为确保建模数据的有效性, 将这些值抛弃掉, 计算平均值时, 不列入其中。

✧ 异常值处理

在原始样本数据中, 存在大量异常值的情况, 如-10000, 以及一些超出指标阈值范围的数据, 对于这类异常数据, 当成缺失值处理。

✧ 节假日数据修正

节假日的用电量及实时负荷数据同工作日比起来, 会明显偏低, 为了确保不同日期类型的计量数据在一定时间周期内具有可比性和连贯性, 需要对这些特殊类型日期的数据进行修正处理, 处理规则是: 去除星期六星期的数据。

.....

2.3. 结果分析

对数据挖掘建模过程中产生的图表结果进行解释分析。

3. 结论

结合研究目标和实现效果, 对本次研究下一个结论性的评语。

4. 参考文献

列举在本次研究中所参考的文献.

例如:

[1]刘涛, 张良均. 大规模智能用电系统海量数据处理与数据挖掘技术研究及应用[Z]. 博士后重大专项