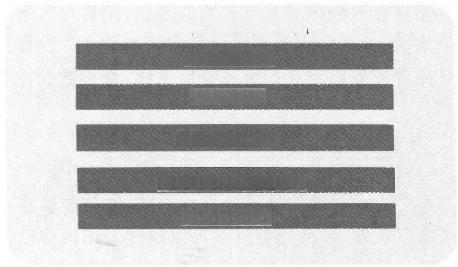
**3.5** **数据挖掘**

数据挖掘是从海量数据中挖掘潜在的信息与知识，挖掘结果的应用往往可以 带来产品的创新以及业务的创新。B2B 、B2C 等电子商务平台在每天的运营过程 中，信息系统都会产生大量数据信息，数据挖掘技术能够把历史积累的大量数据进 行分析以及模型化的挖掘和处理，从中发现隐藏的规律或模式，提取辅助商业决策 的关键性数据，为决策提供支持。MIC 挖掘价值体现见图3.17所示。



MIC挖掘价值体现

**产品深度支持**

**精确化销售**

客户精细化服务

**精细化推广与市场培养**

**客户深度认识**

**图3.17** **MIC** **挖掘价值体现**

**3.5.1** **价值体现**

**1)产品深度支持**

新产品设计：对于竞争激烈的互联网企业，好的产品就是企业的生命。随着信 息化的发展，如何能够设计出一个用户满意的产品，已经不再完全凭借个人的直觉 和行业经验，更多的还是需要数据挖掘和分析的支持，例如产品的结构设计、产品 如何定价，以及该产品适用人群分析等。

旧产品优化：每一产品都有其生命周期， 一个过去很好的产品现在可能已经过 时了，所以对于老产品的效果分析与挖掘是必不可少的，怎样正确的对当前产品的



**商** **业** **智** **能** **在** **电** **子** **商** **务** **中** **的** **实** **践** **与** **应** **用**

效果作出客观评价，同时对产品将来的发展趋势作出准确的预测都需要数据挖掘 技术的支持。

**2)客户深度认识**

服务客户：俗话说知己知彼，百战不殆。电子商务对于用户来讲提供的也是一 种服务，如何能够让用户对我们的服务认可，首先就需要对客户有所了解，这样我 们才能够有针对性地为客户提供服务。例如客户细分、流失客户挖掘等都是通过 数据挖掘的技术使我们对客户能有一个更深入地了解。

客户风险控制：对客户深度认识除了能够为客户更好的服务以外，还可以控制 由客户带来的风险，尤其是互联网这样一个特殊的行业，欺诈等现象时有发生，影 响非常严重，可是少量的异常客户隐藏在海量的客户当中我们很难区分，但有了数 据挖掘技术的支持，我们可以对客户有更深度地认识，就可以把异常客户区分出 来，并对异常客户进行监控，从而对客户可能带来的风险进行控制。

**3)精确化营销与服务**

精确化营销与服务是基于对客户深度认识的基础上，针对于不同的客户采用 不同的营销策略和服务策略，提高营销的成功率和服务效果。要做到精确化单靠 人工去完成成本非常大，或者说根本无法实现，但数据挖掘技术可以精确地发现用 户的需求并且挖掘出最合适的产品与营销服务策略，销售与客服人员可以通过数 据挖掘结果的数据支持，自动的实现精确化的营销与服务。例如现在大型网站都 非常流行的个性化推荐服务以及客户维系挽留系统，都是数据挖掘技术在精确化 营销与服务的应用。

32

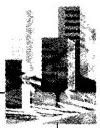
**4)精确化推广与市场培养**

精确化推广：网络推广是电子商务网站得到迅猛发展的重要手段之一，网络推 广最大的难题就是推广成本与推广效果的平衡问题，如何能够在有限的推广成本 前提下取得最好的推广效果，数据挖掘技术可以通过公司内部以往推广的一些历 史数据挖掘出最有效的推广方式，提供辅助精确化推广的决策支持。

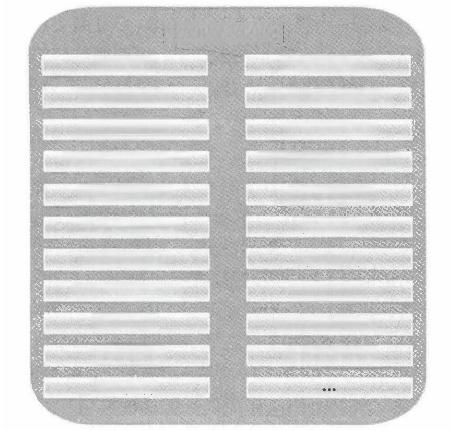
精确化市场培养：市场是公司赖以生存的外部环境，数据挖掘技术能够对目前 市场情况作出准确评估，并对市场将来的可能的变化趋势作出预测，使决策层根据 市场分析结果作出正确的市场决策，抓住市场机会，规避市场风险。

**3.5.2** **主要应用**

上文中我们在讨论数据挖掘价值体现的时候实际上已经提及了一些数据挖掘 的应用，例如客户细分、客户流失预警等。就是通过这些方方面面的数据挖掘应 用，数据挖掘的价值才能在电子商务行业中体现出来。经过总结，数据挖掘在电子 商务行业主要有如图3.18所示的应用。



**第** **三** **章** **商** **业** **智** **能** **在** **电** **子** **商** **务** **的** **解** **决** **方** **案**



数据挖掘应用

客户流失预警

产品搜索效果评估

关联产品推荐

搜索利益最大化挖掘

点击流挖掘 市场评估 市场预测

风险预警

买家个性化推荐 销售工作预警

客户分群 潜在客户挖掘

客户价值提升 客户维系挽留设计

客户价值评估 客户忠诚度评估 客户需求预测 客户满意度评估

产品设计

产品推荐

产品定价

**图3.18** **数据挖掘的应用**

**3.5.3** **应用实现要求**

**1)完善的开发规划**

33

应用之间的关系要求：数据挖掘的多项应用并不是独立的，绝大部分应用需要 联合起来才能发挥价值，例如客户维系挽留系统一般就需要以客户流失预警系统 为前提，客户细分是大部分其他应用的基础，这种依赖关系存在但是又不绝对，所 以导致不同应用之间的关系非常复杂。必须制订完善的开发计划。

应用的价值要求：对于一般的应用开发计划肯定是价值大、难度低的先开发， 但是数据挖掘的应用有时候价值的大小很难直接评估，例如客户细分从短期内看 价值不大，但是它是很多挖掘应用的基础，长期来看价值非常大。因此应用的价值 和长远的开发计划关系非常密切，只有详细的开发计划出来了，才能准确地评判什 么样的应用应该先做，什么样的应用应该后做。

数据要求：数据挖掘一般对历史数据都有要求，虽然现在的电子商务都有自己 的数据仓库，储存了大量的历史数据，但是具体到某一个实际的应用，我们还是经 常会发现数据不足的问题，原因就是历史数据往往是业务系统实现现有业务而产 生的，不一定会考虑到后续的数据分析与挖掘，缺少一些关键的分析指标，因此经 常为了能够收集到这些数据而将一些很重要的数据挖掘应用延后，如果我们有完 整的数据挖掘应用开发规划，就可以根据计划提前准备必要的数据。

**2)与时俱进的挖掘技术**

虽然多数的数据挖掘应用只用到了最基础的挖掘算法，基础的挖掘算法非常

**商** **业** **智** **能** **在** **电** **子** **商** **务** **中** **的** **实** **践** **与** **应** **用**

有用，但是挖掘技术毕竟是一门难度非常高的技术， 一旦需要用到比较高端的挖掘 技术时才来学习是来不及的，因此在平时的工作当中，保持挖掘团队技术的先进性 是非常必要的，这些技术是数据挖掘应用能够稳定实现的有力保证。

**3.5.4** **常见问题**

**1)执行效率**

复杂算法的执行效率：电子商务行业是业内公认的海量数据的行业之一，在这 里有很多效果非常好的算法因为效率的问题而搁浅，特别是有些实际的数据挖掘 应用不能用抽样的方式解决效率问题。因此在电子商务行业有时候并不是结果最 准确的挖掘算法才是最好的。 一定要找到成本与效果的平衡点，这样的挖掘算法 才是最好。

业务系统实时要求：很多挖掘算法不是把挖掘结果分析出来就算结束，这些结 果还要提供给业务系统，供业务系统实时处理事务时使用。这类需求要求实时性 要求非常高， 一种比较实用的方法是在BI 系统中离线训练挖掘模型，等模型规则 评估通过以后，将规则封装到业务系统中进行实时处理。另一种方法就是尽量提 升挖掘算法的效率，BI 实时处理数据并实时同步到业务系统当中。

**2)数据质量**

34

数据缺失：电子商务行业大部分用户是虚拟的，这就导致了用户产生的数据经 常是部分信息丢失，例如用户如果不登录就在网站上进行浏览操作，那么这些操作 的主体信息就会缺失，间接地与主体相关的信息也会缺失。

数据失真：电子商务行业并不是完全需要实名制，因此很多信息用户自愿填 写，填写只要符合规则就可以，对填写的真实性一般是不做验证的，这就导致了大 量数据失真。

做电子商务行业的数据挖掘的时候，由于数据的缺失与失真导致数据的预处 理一般要比其他行业工作量大很多，挖掘的难度也大很多。在这一点上在做电子 商务行业数据挖掘项目的时候需要充分评估项目的难度与实现周期，提前做好 准备。

**3)云计算**

近年来云计算越来越流行，特别是发展迅速，数据量暴增的电子商务行业，云 计算的应用取得了长足的进步；相对于传统的数据挖掘技术，能不能更有效的和云 计算相结合，开发出基于云计算技术的数据挖掘应用，是电子商务行业未来数据挖 掘新的发展方向。焦点科技目前做的云计算平台下的个性化推荐系统就是基于云 计算平台的数据挖掘应用。