

关联规则数据挖掘技术在保险行业的运用分析

张宪雷

对外经济贸易大学统计学院, 中国·北京 100029

【摘要】通过相应的算法企业可以对客户的数据信息进行分析从而得知各个数据之间的关联程度。根据得到的结论企业就可以对销售方式和销售产品的内容进行调整。而在保险行业, 保险企业将客户的数据信息以及曾经的保险购买记录进行分析。就可以得知客户与各个险种之间的联系, 保险企业就可以为客户提供更合适的保险服务, 进而提升保险企业的销售率以及经济效益。

【关键词】关键规则; 数字挖掘技术; 保险行业

Analysis of the Application of Data Mining Technology in the Insurance Industry

Zhang Xianlei

School of Statistics, University of International Business and Economics, Beijing 100029, China

[Abstract] Through the corresponding algorithm, enterprises can analyze the customer data information to know the degree of correlation between each data. According to the conclusion, the enterprise can adjust the sales method and the content of the sales products. In the insurance industry, insurance companies analyze customers' data and information and former insurance purchase records. We can know the connection between the customers and each type of insurance, and the insurance enterprises can provide the customers with more appropriate insurance services, and thus improve the sales rate and economic benefits of the insurance enterprises.

[Key words] Key rules; digital mining technology; the insurance industry

前言

随着当前社会的不断进步人民的消费观念和生活观念都与曾经有了很大的不同, 随着个人财富的不断积累, 各个消费者对于保险投保的意识逐渐加强, 对于参保的态度也有了很大的进步。而保险行业作为我国增长速度较快的几个行业之一, 对国民经济的增长有着极为重要的促进作用。而面对复杂的社会环境与市场竞争关系, 如何在市场占有一席之地就是各个保险公司所要面临的首要问题。在当前的大数据分析时代在当前的数据分析, 保险企业由于历史数据的不断积累导致信息爆炸, 在进行业务办理时却缺少有用的信息。这就导致保险行业无法为消费者提供有效的保险帮助。而通过计算机技术及相应先进的信息互联手段就可以对保险公司中的大量客户进行信息采集与分析, 由此就可以帮助公司进行相应的业务调整以及企业路线的规划。本文将通过采取数据挖掘技术对, 保险公司提供的续保客户数据中的相关保险产品进行分析, 再通过关联规则的算法分析各不续保客户购买不同保险产品之间的关联性, 并根据数值对客户购买的保险产品进行推测。促进保险公司满足不同保险客户的需求, 公司应配合相应数据的分析, 对所推出的业务进行调整, 以提升消费者对于保险产品的购买率, 最终提升保险公司的经济效益。

1 保险险种的关联销售分析

1.1 保险产品关联销售的概念

关联销售的本质是指通过科技手段进行特定的数据挖掘技术来筛选客户购买保险的信息, 并根据筛查出的数据来挖掘各个保险产品与客户特征所具有的关联性, 保险公司可以通过相应的数据分析调整保险产品的销售方式与企业规划。保险公司利润收入的主要来源之一就是关联销售的方式, 而关联销售的方式

除了可以依靠传统的保险销售经验外, 主要还是通过对过往的保险销售数据进行挖掘分析, 最终得到新的保险产品或者保险产品销售组合, 在根据不同的人群进行推荐, 尤其是客户特征与保险信息具有一定关联性的客户群体, 通过此种方式最终实现提升保险销售效率的目的。

1.2 数据挖掘关联规则方法应用

本文主要是通过采取关联规则的算法这种算法的主要特征是通过算法反映事物之间的相互联系以及相互依赖等关系。其中采用的算法是非常著名的Apriori算法, 这种算法对数据处理的速度很快, 并且对规则的数量没有明确的要求或者限制。另外在这种算法中会给出特定的两个重要的参考标准和阈值, 即最小支持度和最小置信度。而数据挖掘的目的就是从已经给定的数据库中发掘出满足两项关联规则。本次数据研究使用的数据挖掘软件为SPSSModeler, 方法关联规则算法。

1.3 保险数据准备

数据准备工作就是对初始的数据信息进行一定的处理, 处理的主要流程包括数据的采集, 筛选以及预处理功能等。进行数据处理是为了在一定程度上提高模型构建的准确性和科学性。保险数据的准备流程包括最保险数据进行采集, 然后对采集到的数据进行整理以及清洗。对于整理清洗后的数据要进行汇总, 变换以及规约。根据相关保险数据的准备要求, 首先对保险公司参与续保客户数据进行处理, 首先将脱敏数据, 无关变量数据进行筛选与剔除。例如, 如果保险公司客户的原始数据中存在客户的投保时间, 受教育程度, 以及家庭人数等信息。而由于这三类数据都属于脱敏数据, 并通过相应的分析可以得知这三类数据都属于无效数据, 所以就要对这类数据进行剔除。对于参与续保的客户, 有一些客户有工作, 但在过去

的三年中及平均年收入水平低于 10 万的人群不属于特殊数据, 所以将这类数据视为脏数据。对于脏数据也要进行剔除。另外就是有一些客户有职业但是他的平均年收入为零元。为了提升相关数据的准确性以及有效性, 所以要对这些年收入为零元的客户进行数据分析和剔除。根据 EXCEL 表格对数据进行处理后, 保险公司就可以更加系统并直观的了解到保险客户的相关信息。

1.4 关联规则模型建立

进行数据挖掘所采用的关联算法中的Apriori算法来客户的历史保险数据信息。并分析续保险与保险之间的关键性。首先通过使用EXCE对数据进行处理, 得到对应整理过的客户续保数据信息, 然后在 SPSSModeler 空白界面中创建流, 随后导入选择从 EXCEL 文件中读取相应的输入源。接下来从相应的字段选项中添加类型数据节点通过读取数据值并未其所需的纸标设置所输入字段的角色类型或格式。然后, 在其中在加入相应的设置标志节点并选择创建一些险种, 然后为标志字段以及客户号进行汇总。接下来, 在建立模型之前可以对添加的图形选项进行设置, 可以使保险企业更加直观的感受的各个保险之间的关联性以及关联程度。然后, 再次添加相应的类型节点, 为所需要分析的数据目标设置相应的角色, 来满足接下来建模所需要的要求。最后, 建立数据的模型节点, 通过采用Apriori的算法, 然后选取最低条件的支持度与最小规则的置信度所需要的一般参数。最低条件支持度的一般参数为十最小规则支配度的一般参数为80%。在实际的研究中要根据不同的数据要求进行调整。接下来就是要建立数据模型, 数据模型的建立需要选择专家模式, 并自行设置相应的数据参数, 通过这些数据建立起相应的模型并对其进行分析。

1.5 模型结果分析

根据所建立出的数据模型可以得知各种保险之间的关联性, 根据建立出的模型, 企业可以及时的对保险产品进行调整。根据所建立出的模型可以得知模型中的线条粗细程度可以直观的体现出不同险种之间所存在的关联。这样就可以分析出哪种保险之间的关联性最强, 哪种保险之间不存在联系。而通过模型进行分析得到的结果可以看出输出的规则主要包括五个方面。分别是支持度, 自信度, 规则前项与规则后项以及实力记录数。根据这些数值在对保险消费者进请推荐时, 可以增加保险各种种类间的组合销售, 以此提升保险的销售效率, 为企业获取更多的经济效益。通过对保险数据规则信息的分析, 就可以了解到客户买不同保险的概率。业务人员在对客户进行推销时就可以侧重于概率高的保险种类。并且根据对建模数据的分析可以看出续保客户购买的不同保险之间的确存在着联系。续保险种之间存在着一定的关联性, 业务人员可以根据消费客户的历史购买数据以及险种信息推测出其中的关联程度, 对于关联程度较强的险种进行有针对性的销售, 以满足不同客户对保险种类的个性需求。通过对以上一系列数据的分析可以看出, 保险险种的关联交叉销售有助于提升保险企业的销售率, 以提升保险企业的经济效益。并有助于提高客户忠诚度, 降低企业的销售成本, 实现对于企业经营资源的合理分配。

2 研究结论与建议

2.1 研究结论

根据上述一系列分析, 可以得知, 保险公司在进行日常的

保险销售过程中, 除了要注意分析不同客户的保险需求以及购买倾向做好客户流失的预警工作。保险企业还要善于从客户曾经购买记录中挖掘出客户未来的购买欲望, 并针对不同的客户群体推出有针对性的保险产品或者产品组合项目。总而言之, 对客户群体进行数据挖掘主要是为了企业能够更好的为客户提供服务, 并通过数据挖掘工作为企业提供了明确的发展方向, 企业可以通过数据支持进行相应的业务调整和企业规划, 提升了企业的经济效益收入。本文通过数字挖掘技术探讨了保险公司客户群体特征与不同保险险种之间的联系, 是基于关联规则的关联销售行为。从保险产品的角度出发, 了解不同产品与不同客户之间的联系分析客户购买产品时的习惯与特征, 并运用著名的Aproti算法进行数据处理, 虽然存在一定的局限性, 但并不影响该研究方式的使用。

2.2 相关建议

无论采取哪种数据建模方式对数据进行分析, 都需要保险公司具有丰富的数据库, 所以保险公司需要加强对于自身数据库的建设工作, 为未来保险公司通过数据分析进行产品销售打下一个坚实的基础。除此之外, 通过大数据挖掘技术可以为企业带来可靠的数据分析信息提取, 企业根据所知道的信息就可以进行销售调整。而且通过数据挖掘技术在各行各业都可以实现精准的客户定位以及客户流失预警, 管控风险评估及相关预测, 并分析预测客户未来的需要。企业通过数据挖掘技术可以被客户量身定制个性化的产品需要, 以满足客户的个性化要求。因此洗衣液要积极引进先进的科学技术手段进行创新工作。现如今的企业发展要以客户为中心, 所以企业需要从客户的角度进行分析精确客户群体的需求, 提升客户的满意程度以及可接受程度, 并与其他数据分析技术方法进行结合, 构建一个完善系统的客户需求分析体系, 为保险营销人员后续工作提供一个科学的依据, 并为企业增加经济效益。

3 结语

随着当前科学技术的发展, 信息数据处于大爆炸时代。这也表示数据分析对于企业发展有着至关重要的作用, 企业通过建立一个客户端的数据库, 在利用先进的科学技术手段对数据进行分析处理。通过数据挖掘技术就可以了解到客户所需要的服务信息, 企业针对相关的数据信息就可以进行服务调整和规划调整。除了在保险行业外, 现如今各行各业都已经与数据挖掘技术有了一定的接触。大数据数据挖掘技术是企业了解客户需求的一个重要方法, 企业根据了解到的信息可以为客户提供更加满意服务。根据所掌握的数据信息企业可以改善自身的发展方式并提升自身的竞争力, 在激烈的市场环境中站稳脚跟, 是企业实现既定目的的必要途径。

参考文献:

- [1] 袁杨钦, 杨杉. 关联规则数据挖掘技术在保险行业的运用研究[J]. 电子世界, 2020, 000(8): 2.
- [2] 叶飞, 金先好. 基于关联规则数据挖掘算法的推荐系统研究与实践[J]. 信息与电脑, 2020, 32(7): 3.

作者简介:

张克雷(1994-)男, 汉, 山东菏泽人, 大学本科, 对外经济贸易大学统计学院在职人员高级课程研修班学员。研究方向: 统计学。