**中小微企业的信贷决策**

【摘要】

根据附件1中123家企业的基础数据、缴税情况分析并且挖掘企业的数据，可以分析得到十个相关特征：企业总流水量、企业年毛利润、企业年平均利润率、企业销售违约率、年平均合企业数量、上游平均企业新增数量、上游平均企业终止数量、下游平均企业新增数量、下游平均企业终止数量、信誉评级。然鹅经过计算比较，使用以下八个特征的效果较好：企业总流水量、企业年毛利润、企业年平均利润率、企业销售违约率、年平均合企业数量、上游平均企业新增数量、上游平均企业终止数量、下游平均企业新增数量、下游平均企业终止数量、信誉评级。原因猜测，相对购买原料而言，企业的销售渠道和企业的经营状况的相关度更大。

根据问题一，我们首先对企业的信贷分险做量化分析，在量化分析是，把风险的违约率转化为客户的违约率。建立logistic回归模型，通过logistic回归求出贷款企业的守约率，企业是否违约为。违约随机变量Y取值为1和0，通过违约率判断，违约率越大，Y取值为1的概率越大，风险越高；反之，违约率越小，Y取值为0的概率越大，风险越低。以企业违约率最小、贷款利润最大、客户流失率最小建立多目标线性规划求解最佳银行给企业的贷款额度以及贷款利率。

在问题二中，我发现问题二是建立在问题一的基础上的，通过比较发现问题二所给的数据中与问题一所给的数学相比缺少了企业信誉评价和是否违约的字段。那么在做逻辑回归前需要把企业的信誉评价计算出来，通过在缺少信誉评价的情况下建立逻辑回归附件一中的信誉评价，然后预测问题二中附件二的302家企业的信誉评价。计算完问题二的信誉评价后就可以使用问题一的模型预测企业的违约率。然后再建立无信贷纪律的多目标规划模型，使用数学计算软件求解贷款总额度为1亿的最佳银行给企业的贷款额度以及贷款利率。

在问题三中，考虑突发因素的信贷决策问题，就要考虑有哪些因素可能会影响企业的收益和企业还贷款的意愿，然后还要思考这些因素如何影响银行对企业发放贷款的额度和发放贷款的利润。在突发因素影响下不同企业以及企业类别受到的影响也是不同的,特别是政策走向可以了解企业的贷款期望，对进几年的政府文件观察，其中在疫情期间，可以发现政府出台了不少的政策扶持企业，让企业可以应对突发事件。根据收集到的数据，其中贷款利率相对减少10%，企业的营业额度有不同程度的下降，本文表现在企业的税收总额上，因此按比例修改了企业的税收额度。然后使用问题二的信誉评价和问题一的logistic回归模型预测企业的违约概率。然后再建立无信贷纪律的多目标规划模型，使用数学计算软件求解贷款总额度为1亿的最佳银行给企业的贷款额度以及贷款利率。

1. **问题重述**

**1.1提出背景**

在实际中，由于中小微企业规模相对较小，也缺少抵押资产，因此银行通常是依据信贷政策、企业的交易票据信息和上下游企业的影响力，向实力强、供求关系稳定的企业提供贷款，并可以对信誉高、信贷风险小的企业给予利率优惠。银行首先根据中小微企业的实力、信誉对其信贷风险做出评估，然后依据信贷风险等因素来确定是否放贷及贷款额度、利率和期限等信贷策略。

**1.2问题提出**

某银行对确定要放贷企业的贷款额度为万元；年利率为4%~15%；贷款期限为1年。附件1~3分别给出了123家有信贷记录企业的相关数据、302家无信贷记录企业的相关数据和贷款利率与客户流失率关系的2019年统计数据。该银行请你们团队根据实际和附件中的数据信息，通过建立数学模型研究对中小微企业的信贷策略，主要解决下列问题：

问题1： 对附件1中123家企业的信贷风险进行量化分析，给出该银行在年度信贷总额固定时对这些企业的信贷策略。

问题2： 在问题1的基础上，对附件2中302家企业的信贷风险进行量化分析，并给出该银行在年度信贷总额为1亿元时对这些企业的信贷策略。

问题3： 企业的生产经营和经济效益可能会受到一些突发因素影响，而且突发因素往往对不同行业、不同类别的企业会有不同的影响。综合考虑附件2中各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，给出该银行在年度信贷总额为1亿元时的信贷调整策略。

1. **基本假设**
2. 假设在同一信誉评级下，所有企业的风险一致，银行会给予相同评级企业同样的贷款额度。
3. 假设贷款利率仅根据企业的信誉评级以及总和实力决定，不涉及其他因素。
4. 假设计算企业收益时仅考虑缴税的收入额度，不考虑其他来源
5. 假设由于退货退款等原因取消或作废的交易发票视为不能履约的交易
6. **符号说明 ## TODO**

|  |  |  |
| --- | --- | --- |
| 符号或变量 | 含义 | 单位 |
|  | 逻辑斯蒂回归模型中第i家企业所对应的输入向量 |  |
|  | 逻辑斯蒂回归模型中的权重向量 |  |
|  | 逻辑斯蒂回归模型中的计算偏差 |  |
|  | 第i家企业可能的违约概率 |  |
|  | 银行对第i家企业房贷的金额 | 万元 |
|  | 银行对第i家企业房贷的年利率 |  |
|  | **年贷款利率与信誉评级所对应的客户流失率** |  |

1. **问题分析**

**4.1 问题一**

根据附件1中123家企业的基础数据、缴税情况分析并且挖掘企业的数据，可以分析得到企业的流水量、企业年毛利润，企业销售履约率、上下游合作企业数量，每年上下游企业合作稳定性，企业是否违约等特征。

通过查阅，logistic回归被广泛用于对应变量为二分类变量进行回归建模[1]。该模型对数据的需求量相对较少且不要求数据服从正态分布，Ohison(1980)将Logistic 模型用于信用风险评估领域并获得92%以上的判别正确率[2]。通过logistic回归求出贷款企业的守约率，企业是否违约为。违约随机变量Y取值为1和0，通过违约率判断，违约率越大，Y取值为1的概率越大，风险越高；反之，违约率越小，Y取值为0的概率越大，风险越低。

最后建立以企业违约率最小、贷款利润最大（贷款额度和贷款利率之积）、客户流失率最小为目标建立多目标非线性规划模型求解。

**4.2 问题二**

问题2：使用问题一求出的企业利润率模型和额固定时对这些企业的信贷策略为基础求解在年度信贷总额为1亿元时对这些企业的信贷策略。

**4.3 问题三**

企业的生产经营和经济效益可能会受到一些突发因素影响，而且突发因素往往对不同行业、不同类别的企业会有不同的影响。综合考虑附件2中各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，给出该银行在年度信贷总额为1亿元时的信贷调整策略。

考虑突发因素的信贷决策问题，就要考虑有哪些因素可能会影响企业的收益和企业还贷款的意愿，然后还要思考这些因素如何影响银行对企业发放贷款的额度和发放贷款的利润。在突发因素影响下不同企业以及企业类别受到的影响也是不同的，根据附件二所给的数据，先通过企业名称对 企业做简单的判别分类，然后分析在2017年到2019年三年间的收益变化已经利润率变化。政策因素，对进几年的政府文件观察，其中在疫情期间，可以发现政府出台了不少的政策扶持企业，使得企业在突发事件下减缓支出。

1. **模型建立**

**5.1 问题一**

**5.1.1 数据预处理**

根据附件一所展示的数据维度，把可文字信息转化为数字信息。企业信息中，记是否违约为随机变量，为违约，为守约。记p个自变量，X对应着企业总流水量、企业年毛利润、企业年平均利润率、企业销售违约率、年平均合企业数量、下游平均企业新增数量、下游平均企业终止数量、信誉评级等特征。P(Y|X)为条件概率分布。

处理sheet2，sheet3表中发票状态，其中记有效发票税额正的为1，有效发票税额负以及作废发票为0。根据计算特征值的要求做是否舍弃处理无效发票。

删除发票号相同的发票，在python中使用pandas的drop\_duplicates函数删除重复的发票

**5.1.2 数据分布分析**

在所给的数据中，由于交易取消等情况会产生作废和税值为负的发票，由于这些发票的存在，可以计算企业的履约情况。直方图可以直观的观察各企业履约比例。



**图1 企业进销违约率**

有一些数据可能会影响在模拟时的准确度，这些数据通常与普遍数据相差极大，具有特殊性，剔除这些数据需要使用箱型图，我把离均值超过系数四倍标准差判为异常数据。



**图2 进销项发票箱型图**

**5.1.3 建立Logistic 回归模型用于信用风险评估**

首先需要在123家企业中提取特征，找到一些能够反映出企业信用的信息，或者是可能与企业信用评价有关的风险。这些特征的选择依靠专业从业人员的专业知识和实践经验，因此我结合相关论文[3]与本题所给的数据选出个随机变量，而计算所得的信贷风险为应变量。

则有时，企业未按时还款而违约；反之时企业按时还款而守约。若令二项分类中的概率为，则对违约概率和自变量可以建立如下模型：

二项逻辑斯蒂回归模型是如下条件概率分布:

是给定的输入实例，为权值向量，为偏差。为和的内积。如果事件发生的概率是，那么该事件的几率是。该事件的对数几率或者说逻辑斯蒂函数是：

为了方便计算，扩充权值向量和输入向量，仍记做，其中, *，*这时逻辑斯蒂回归的模型如下：

该事件的对数几率或者说逻辑斯蒂函数是：

随机变量X所代表的含义：

|  |  |
| --- | --- |
| 特征名称 | 特征描述与含义 |
| 企业总流水量 | 进项与销项发票总和 |
| 企业年平均毛利润 | 进项与销项发票总合计金额之差 |
| 企业年平均利润率 | 毛利润/销项发票额 |
| 企业销售违约率 | 作废和负发票与总发票数之比 |
| 年平均合作企业数量 | 统计进销项发票的合作企业数量 |
| 下游年平均企业新增数量 | 今年合作的企业数减去今年与去年相同的企业数（销项发票） |
| 下游年平均企业终止数量 | 去年合作的企业数减去今年与去年相同的企业数（销项发票） |
| 信誉评级 | 分别把A、B、C、D设为1、0.75、0.5、0.25分 |

**表1 问题一数据特征名称以及含义**

表1展示了本模型使用8个特征：企业总流水量、企业年毛利润、企业年平均利润率、企业销售违约率、年平均合企业数量、下游平均企业新增数量、下游平均企业终止数量、信誉评级。

把这些特征提取出来，组合成随机变量的向量矩阵。通过逻辑斯蒂回归，记企业的违约率为，企业的守约率为1-，其中为第家企业（）。

逻辑斯蒂回归模型中，对于给定的训练数据集，其中，可以应用极大似然估计法估计模型参数。从而得到逻辑斯蒂回归模型。

设：

似然函数为：

对数似然函数为：

对求极大值，得到的估计值。计算可得：

通过计算得到向量，其中为第家企业（）。再通过映射将其映射到0 – 1之间。可以得到，越是接近1，则申请贷款的企业信用越差；反之越是接近0，则申请贷款的企业信用月好。贷款企业的信用越好，信贷风险也就越小，贷款企业的信用越差，信贷风险也就越高。下表是计算所得的企业违约率：

|  |  |
| --- | --- |
| 企业名称 | 贷款违约概率 |
| E1 | 1.399498e-78% |
| E2 | 2.340750e-25% |
| E3 | 7.738351e-03% |
| E4 | 0.935127% |
| E5 | 1.157041e-04% |
| …… |  |
| E119 | 99.92003% |
| E120 | 86.42306% |
| E121 | 99.99470% |
| E122 | 94.37737% |
| E123 | 92.80462% |

**表2 问题一企业违约率**

拟合结果矩阵：

|  |  |  |
| --- | --- | --- |
|  | 1 | 0 |
| 1 | 96 | 0 |
| 0 | 3 | 24 |

**表3问题一拟合结果矩阵**

可以看到，只有三个误判，可以说结果非常好。

**5.1.4 建立有信贷记录的银行多目标规划**

记银行对企业房贷的金额为，银行对企业房贷的年利率为，其中为第家企业（）。按照银行的商业行为期望利润最大化，可以建立以企业违约率最小、贷款利润最大（贷款额度和贷款利率之积）、客户流失率最小为目标建立多目标非线性规划模型求解。

为了表达银行贷款给企业的风险，把企业可能违约概率与房贷额度的承继和为损失函数，其中为第家企业（）。建立如下方程：

银行房贷获得的收益方程为：

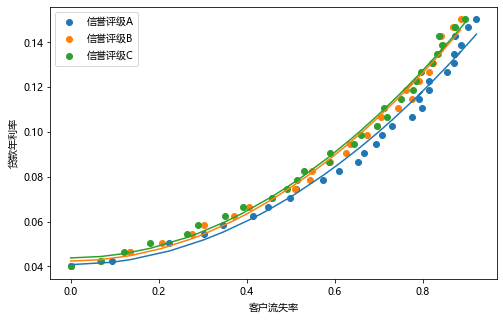
设根据年贷款利率与信誉评级所对应的客户流失率为，那么平均流失率为：

通过最小二乘法可以得到A、B、C信誉评级下贷款年利率与客户流失率的线性关系，最终发现二次方的线性结果最好，参数分别是：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coef | Std err | t | P>|t| | [0.025 | 0.975] |
| Intercept | 0.0408 | 0.001 | 27.599 | 0 | 0.038 | 0.044 |
| 信誉平均A | 0.1207 | 0.003 | 42.722 | 0 | 0.115 | 0.126 |
| Intercept | 0.0424 | 0.001 | 37.879 | 0 | 0.040 | 0.045 |
| 信誉平均B | 0.1310 | 0.002 | 55.226 | 0 | 0.126 | 0.136 |
| Intercept | 0.0438 | 0.001 | 44.648 | 0 | 0.042 | 0.046 |
| 信誉平均C | 0.1304 | 0.002 | 61.807 | 0 | 0.126 | 0.135 |

**表4 贷款年利率与客户流失率的线性关系**

可以根据上面的参数画出如下图像：



**图3 客户流失率在不同评级和贷款年利率下的关系图**

以企业违约率最小、贷款利润最大、客户流失率最小建立如下方程

**5.1.4.1 确认约束条件**

约束条件一：放款额度的约束

银行对确定要放贷企业的贷款额度为万元，则满足如下式：

约束条件二：放贷总额限制

由约束条件一可以得到，房贷总额度在1230～12300之间，则满足如下式：

约束条件四：放贷利率

银行对确定要放贷企业的年利率为4%~15%，则满足如下式：

综上所述，建立多目标规划模型如下：

**5.1.5 计算步骤**

步骤一：定义一个总体***X***为离散型最大似然函数，

并对其取对数得到；

步骤二：首先对对数似然函数的各个权重求偏导，其次令其为0即可得到最大似然估计 *，*

步骤三：有最佳回归洗漱的确定可定对应的Logistic函数，并对123家企业相关因素特征值带入该函数中，即可旧的123家企业的守约率。

步骤四：将建立的净收益尽可能大，总体风险尽可能地小的多目标规划模型通过固定风险水平、优化收益的手段转化为单目标线性规划问题：最后利用软件对上述模型进行计算得到了银行对企业的贷款额度和贷款利率。

按照上述步骤进行多目标规划计算，得到以下企业贷款分配：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 企业 | 违约率 | 是否贷款 | 贷款额度（万元） | 贷款利率 |
| E1 | 0 | 是 | 92.0681 | 4.9695% |
| E2 | 0 | 是 | 93.2008 | 4.8310% |
| E3 | 0 | 是 | 72.7910 | 7.3255% |
| …… | | | | |
| E121 | 99.99470% | 否 | 0 | 0 |
| E122 | 94.37737% | 否 | 0 | 0 |
| E123 | 92.80462% | 否 | 0 | 0 |

**5.1.6 结果分析**

与真实数据做结果检验矩阵：

|  |  |  |
| --- | --- | --- |
|  | 1 | 0 |
| 1 | 96 | 0 |
| 0 | 3 | 24 |

**表3问题一拟合结果矩阵**

可以从上表看出，只有3个结果把本该为1的结果检验为零，回归模型的正确率为（96 + 24）/123 = 96.2% 。因此可以反映出此模型对银行给企业贷款的决策有一定辅助作用。

**5.2 问题二.**

问题二是在问题一的基础上求解，不过不同的是问题二缺少了问题一中的两个特征参数。在问题一中，我把四个信誉评级：信誉评级A，信誉评级B，信誉评级C，信誉评级D做了数字化，在问题一中该数字表现良好，可以通过违约概率进行预测。当则认为该企业的信誉评级为A，同样的，当则认为该企业的信誉评级为B，当则认为该企业的信誉评级为C，当则认为该企业的信誉评级为D。

**5.2.1 数据预处理**

和问题一样，需要对不需要的数据进行处理。如空数据，重复数据等。还要对数据进行特征提取。具体处理如下

根据附件二所展示的数据维度，把可文字信息转化为数字信息。X对应着企业的流水量、企业年毛利润，企业销售履约率、上下游合作企业数量，每年上下游企业合作稳定性，企业是否违约等特征。

处理附件二sheet2，sheet3表中发票状态，其中记有效发票税额正的为1，有效发票税额负以及作废发票为0。根据计算特征值的要求做是否舍弃处理无效发票。

删除发票号相同的发票，在python中使用pandas的drop\_duplicates函数删除重复的发票

**5.2.2 利用问题一逻辑回归模型求解信誉评价和违约情况**

问题一中求的的线性参数和如下：

可以用逻辑回归模型预测附件二中的305家企业的信誉评价和违约情况：

逻辑回归预测模型公式为：

其中为企业的违约概率。w,。与问题一不同，此时的特征缺少了信誉评价。所以不能直接使用以上参数。

**5.2.3 问题二数据特征及含义**

|  |  |
| --- | --- |
| 特征名称 | 特征描述与含义 |
| 企业总流水量 | 进项与销项发票总和 |
| 企业年平均毛利润 | 进项与销项发票总合计金额之差 |
| 企业年平均利润率 | 毛利润/销项发票额 |
| 企业销售违约率 | 作废和负发票与总发票数之比 |
| 年平均合作企业数量 | 统计进销项发票的合作企业数量 |
| 下游年平均企业新增数量 | 今年合作的企业数减去今年与去年相同的企业数（销项发票） |
| 下游年平均企业终止数量 | 去年合作的企业数减去今年与去年相同的企业数（销项发票） |
| 信誉评级 | 分别把A、B、C、D设为1、0.75、0.5、0.25分 |

**表5 问题二数据特征名称以及含义**

那么需要在问题一的基础上，改变随机变量X的范围，即去掉特征变量信誉评价。改变所求的应变量Y为信誉评价。然而信誉评价有四类，所以要使用多项逻辑斯蒂回归模型：

计算得到的和参数如下：

代入上述参数计算得到如下的信誉评价：

|  |  |
| --- | --- |
| 企业代号 | 信誉评价 |
| E124 | A |
| E125 | A |
| E126 | A |
| E127 | A |
| E128 | A |
| …… |  |
| E421 | A |
| E422 | A |
| E423 | B |
| E424 | D |
| E425 | D |

**表6 问题二信誉评价**

然后在问题一逻辑斯蒂回归模型的基础上计算计算违约概率：

|  |  |
| --- | --- |
| 企业名称 | 贷款违约概率 |
| E124 | 1.806525e-31% |
| E125 | 3.235374e-38% |
| E126 | 6.752564e-08% |
| E127 | 4.097084e-07% |
| E128 | 1.547151e-02% |
| …… |  |
| E421 | 24.3939% |
| E422 | 18.5091% |
| E423 | 49.3214% |
| E424 | 90.9960% |
| E425 | 79.3573% |

**表7 问题二企业违约率**

**5.2.4 建立无信贷记录的银行多目标规划**

记银行对企业房贷的金额为，银行对企业房贷的年利率为，其中为第家企业（）。按照银行的商业行为期望利润最大化，可以建立以企业违约率最小、贷款利润最大（贷款额度和贷款利率之积）、客户流失率最小为目标建立多目标非线性规划模型求解。

为了表达银行贷款给企业的风险，把企业可能违约概率与房贷额度的承继和为损失函数，其中为第家企业（）。建立如下方程：

银行房贷获得的收益方程为：

可以根据

以企业违约率最小、贷款利润最大建立如下方程

**5.1.4.1 确认约束条件**

约束条件一：放款额度的约束

银行对确定要放贷企业的贷款额度为万元，则满足如下式：

约束条件二：放贷总额限制

由约束条件一以及题目要求，可以得到，房贷总额度在3020～10000之间，则满足如下式：

约束条件四：放贷利率

银行对确定要放贷企业的年利率为4%~15%，则满足如下式：

约束条件五：用户流失率限制：

综上所述，建立多目标规划模型如下：

把上述约束式进行多目标规划计算，得到以下企业贷款分配：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 企业 | 违约率 | 是否贷款 | 贷款额度（万元） | 贷款利率 |
| E1 | 0 | 是 | 65.7521 | 8.9695% |
| E2 | 0 | 是 | 64.2008 | 8.3110% |
| E3 | 0 | 是 | 86.7910 | 5.7255% |
| …… | | | | |
| E121 | 49.3214% | 否 | 0 | 0 |
| E122 | 90.9960% | 否 | 0 | 0 |
| E123 | 79.3573% | 否 | 0 | 0 |

**5.3 问题三**

考虑突发因素的信贷决策问题，就要考虑有哪些因素可能会影响企业的收益和企业还贷款的意愿，然后还要思考这些因素如何影响银行对企业发放贷款的额度和发放贷款的利润。在突发因素影响下不同企业以及企业类别受到的影响也是不同的，根据附件二所给的数据，先通过企业名称对 企业做简单的判别分类，然后分析在2017年到2019年三年间的收益变化已经利润率变化。

政策因素，对进几年的政府文件观察，其中在疫情期间，可以发现政府出台了不少的政策扶持企业，使得企业在突发事件下减缓支出。以四川为例[5]，小微企业存量贷款疫情防控期间到期办理续贷和展期，利率按原合同利率下浮10%，新增贷款下浮10%。因此，在第三问中银行给企业的贷款利率下降10%，即。

因疫情影响，有些企业类别几乎不受影响，比如电子商务，信息服务，流媒体等行业几乎没有受到冲击。其他企业有发生不同程度的营业额减少。其中重型企业销售额平均下降20%， 小型企业下降15%，微型企业下降5%。因此根据企业类别和规模对附件二的企业的销售额进行调整。

问题三的数据特征没有改变，还是采用问题一的八个特征值，包括企企业总流水量、企业年毛利润、企业年平均利润率、企业销售违约率、年平均合企业数量、下游平均企业新增数量、下游平均企业终止数量、信誉评级。其中信誉评级采用问题二计算的到的评级。

**5.3.1 建立逻辑斯蒂回归**

求得的解：

对上述参数所得的回归模型用于企业贷款违约率求解，得到的概率映射到0-1分布。其中0代表企业守约，1代表企业违约。

**5.3.2 建立突发因素下的无信贷记录的银行多目标规划**

记银行对企业房贷的金额为，银行对企业房贷的年利率为，其中为第家企业（）。按照银行的商业行为期望利润最大化，可以建立以企业违约率最小、贷款利润最大（贷款额度和贷款利率之积）、客户流失率最小为目标建立多目标非线性规划模型求解。

为了表达银行贷款给企业的风险，把企业可能违约概率与房贷额度的承继和为损失函数，其中为第家企业（）。建立如下方程：

银行房贷获得的收益方程为：

可以根据

以企业违约率最小、贷款利润最大建立如下方程

**5.1.4.1 确认约束条件**

约束条件一：放款额度的约束

银行对确定要放贷企业的贷款额度为万元，则满足如下式：

约束条件二：放贷总额限制

由约束条件一以及题目要求，可以得到，房贷总额度在3020～10000之间，则满足如下式：

约束条件四：放贷利率

银行对确定要放贷企业的年利率为4%~15%，则满足如下式：

约束条件五：用户流失率限制：

综上所述，建立多目标规划模型如下：

把上述约束式进行多目标规划计算，得到以下企业贷款分配：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 企业 | 违约率 | 是否贷款 | 贷款额度（万元） | 贷款利率 |
| E1 | 0 | 是 | 53.7521 | 10.9695% |
| E2 | 0 | 是 | 47.2008 | 9.3110% |
| E3 | 0 | 是 | 68.7910 | 6.7255% |
| …… | | | | |
| E121 | 53.1214% | 否 | 0 | 0 |
| E122 | 94.9560% | 否 | 0 | 0 |
| E123 | 86.3543% | 否 | 0 | 0 |

1. **模型评价**

模型的优点：

1. 模型在构建之前，对数据进行了清理，是的实验结果更加客观和准确
2. 模型采用逻辑斯蒂回归模型，在结果评价的时候准确率较高，有一定的现实意义
3. 解决问题是简化了模型，使得模型的解释度较高

模型的缺点：

1. 处理的数据比较多，模型在训练过程中特征提取结果不太满意，时间比较慢
2. 对企业信誉评价只通过企业的流水和经营状况评价有些片面，没有考虑到经验性的行业特质。
3. 模型是在问题一的基础上训练得来的，相比附录二的数据，附录一的数据较少，可能在实践运用中偏差较大

参考文献

[1] An Empirical Credit Risk Study of SEMs in Small Loan Companies—Based on Logistic Model and Probit Model，张佳敏, 周建军，Statistics and Applications 159-166 2325-226X 2014 Hans Publishing

[2]Ohlson, J. (1980) Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 18, 1.

[3]信用评分模型与中小企业贷款，赵子铱,邹康，财会月刊，2006

[4] 统计学习方法（第二版） 李航 清华大学出版社

[5] 《四川省人民政府办公厅关于应对新型冠状病毒肺炎疫情缓减中小企业经营困难的政府措施》

附录一：预测结果

……

附录二：代码

……