### 分析流程 数据源： 正向化后数据\_副本(1).xlsx 算法配置： 算法： 偏最小二乘法 变量： 变量Y:{1 try，2 tries，3 tries，4 tries，5 tries，6 tries，7 or more tries (X)}；变量X:{单词是否具有专业性，单词频率，外来词，单词形式，字母重复次数，Orth} 参数： 是否自动确定最大主成分数量:{5} 分析结果： 偏最小二乘法是用于解决多个因变量的回归，请看详细结论进行分析。

### 分析步骤 1. 通过方差解释情况和VIP（累积投影重要性）参考最大主成分数量。 2. 通过成分矩阵表得到主成分的组成情况。 3. 通过因子载荷系数表得到变量的重要性。 4. 最终得到偏最小二乘回归(PLSR)的标准化公式。

### 详细结论

**输出结果1：因子方差解释情况表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 潜在因子 | X方差 | 累计的X方差 | Y方差 | 累计的Y方差(R²) | 调整后的R² |
| 1 | 0.268 | 0.268 | 0.102 | 0.102 | 0.1 |
| 2 | 0.196 | 0.464 | 0.059 | 0.161 | 0.157 |
| 3 | 0.191 | 0.655 | 0.006 | 0.167 | 0.16 |
| 4 | 0.179 | 0.834 | 0.001 | 0.168 | 0.158 |
| 5 | 0.166 | 1 | 0 | 0.168 | 0.156 |

**图表说明：**

上表展示潜在因子的信息综合解释能力。其中，累计的X方差代表对自变量信息的提取，累计的Y²(R²)代表对因变量信息的提取，可以以此为依据确定参数最大主成分数量。

**智能分析：**

因子对方差解释情况表的结果显示，前4个潜在因子就可解释自变量80%的信息，全部潜在因子也不能解释因变量80%的信息。

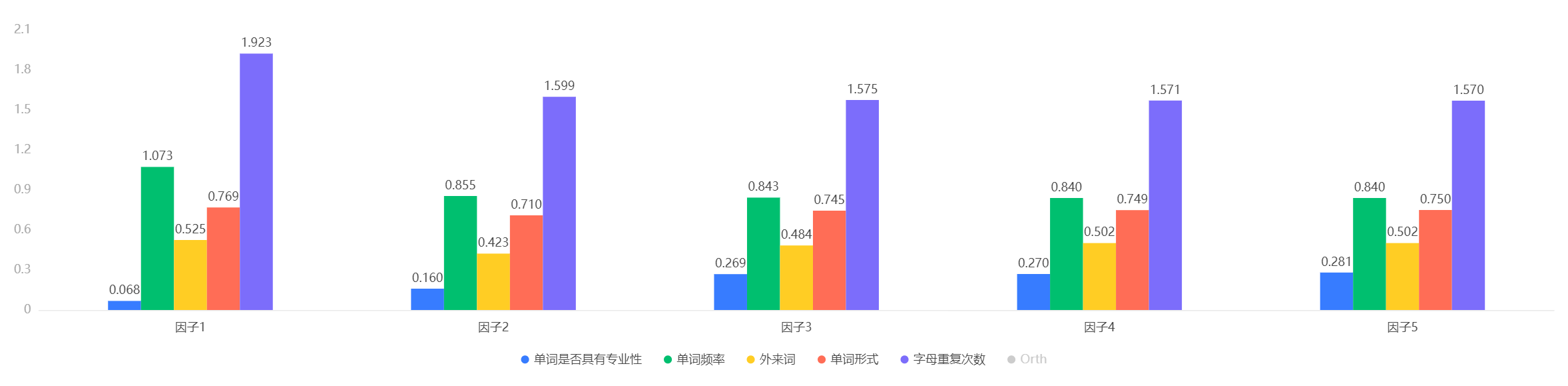
**输出结果2：自变量VIP（累积投影重要性）汇总表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 因子1 | 因子2 | 因子3 | 因子4 | 因子5 |
| 单词是否具有专业性 | 0.068 | 0.16 | 0.269 | 0.27 | 0.281 |
| 单词频率 | 1.073 | 0.855 | 0.843 | 0.84 | 0.84 |
| 外来词 | 0.525 | 0.423 | 0.484 | 0.502 | 0.502 |
| 单词形式 | 0.769 | 0.71 | 0.745 | 0.749 | 0.75 |
| 字母重复次数 | 1.923 | 1.599 | 1.575 | 1.571 | 1.57 |
| Orth | 0.529 | 1.415 | 1.396 | 1.393 | 1.392 |

**图表说明：**

上表展示VIP（累积投影重要性）的情况，它表示成分个数不同时，X对于Y的解释重要性力度，也可以用于参考最大主成分数量。其中对于VIP很大（大于1）的自变量，它在解释潜在因子（从而在解释因变量）时作用相对更大一些。

**输出结果3：自变量VIP（累积投影重要性）图**



**图表说明：**

上图将VIP（累积投影重要性）的情况可视化的展示出来。

**输出结果4：成分矩阵表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 因子1 | 因子2 | 因子3 | 因子4 | 因子5 |
| 单词是否具有专业性 | -0.028 | 0.105 | 0.481 | -0.184 | 0.817 |
| 单词频率 | 0.438 | -0.077 | -0.196 | -0.058 | 0.224 |
| 外来词 | 0.214 | -0.066 | 0.52 | 0.763 | -0.199 |
| 单词形式 | 0.314 | 0.211 | 0.519 | -0.587 | -0.472 |
| 字母重复次数 | 0.785 | -0.394 | -0.31 | -0.05 | 0.158 |
| Orth | 0.216 | 0.889 | -0.345 | 0.192 | 0.049 |
| 1 try | 1.608 | 1.531 | 6.379 | -0.819 | 12.595 |
| 2 tries | 2.567 | 0.255 | 4.351 | 10.659 | -18.166 |
| 3 tries | -2.164 | -1.589 | -9.372 | -14.629 | -2.868 |
| 4 tries | 2.64 | 0.832 | 10.472 | 17.493 | 16.007 |
| 5 tries | -0.324 | -0.705 | 0.625 | -2.987 | -19.544 |
| 6 tries | -2.11 | 0.217 | -4.861 | -9.665 | 8.415 |
| 7 or more tries (X) | 2.195 | 2.016 | 6.649 | 17.331 | 8.088 |

**图表说明：**

由上表可得到主成分分析降维后的成分矩阵表。

**输出结果5：因子载荷系数表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 因子1 | 因子2 | 因子3 | 因子4 | 因子5 |
| 单词是否具有专业性 | 0.043 | 0.162 | 0.535 | -0.126 | 0.845 |
| 单词频率 | 0.417 | -0.026 | -0.149 | -0.083 | 0.211 |
| 外来词 | 0.328 | -0.014 | 0.57 | 0.786 | -0.138 |
| 单词形式 | 0.402 | 0.256 | 0.524 | -0.568 | -0.454 |
| 字母重复次数 | 0.716 | -0.326 | -0.196 | -0.001 | 0.113 |
| Orth | 0.278 | 0.898 | -0.245 | 0.201 | 0.066 |
| 1 try | 0.281 | 0.157 | 0.027 | -0.071 | 0.005 |
| 2 tries | 0.371 | 0.061 | -0.081 | -0.009 | -0.011 |
| 3 tries | 0.361 | -0.189 | -0.093 | -0.014 | 0.016 |
| 4 tries | -0.055 | -0.396 | 0.068 | 0.015 | 0.021 |
| 5 tries | -0.399 | 0.01 | 0.142 | -0.006 | -0.028 |
| 6 tries | -0.254 | 0.264 | 0.015 | 0.005 | -0.004 |
| 7 or more tries (X) | -0.065 | 0.359 | -0.064 | 0.035 | -0.007 |

**图表说明：**

上表为因子载荷系数表，可以分析到每个因子中隐变量的重要性。

**输出结果6：模型系数结果表**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 try | 2 tries | 3 tries | 4 tries | 5 tries | 6 tries | 7 or more tries (X) | 1 try(标准化) | 2 tries(标准化) | 3 tries(标准化) | 4 tries(标准化) | 5 tries(标准化) | 6 tries(标准化) | 7 or more tries (X)(标准化) |
| 常数 | -0.753 | -0.308 | 12.296 | 32.279 | 32.034 | 19.088 | 5.139 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 单词是否具有专业性 | 0.18 | -1.133 | -2.718 | 0.262 | 1.996 | 1.345 | -0.058 | 0.04 | -0.048 | -0.06 | 0.008 | 0.058 | 0.037 | -0.002 |
| 单词频率 | 0.001 | 0.008 | 0.019 | -0.001 | -0.014 | -0.01 | -0.003 | 0.101 | 0.154 | 0.207 | -0.013 | -0.202 | -0.135 | -0.055 |
| 外来词 | 0.038 | 0.692 | 1.433 | 1.914 | -0.395 | -2.341 | -1.151 | 0.008 | 0.027 | 0.029 | 0.056 | -0.01 | -0.059 | -0.044 |
| 单词形式 | 0.268 | 0.786 | 0.395 | -0.878 | -0.379 | -0.208 | 0.039 | 0.174 | 0.098 | 0.026 | -0.084 | -0.033 | -0.017 | 0.005 |
| 字母重复次数 | 0.134 | 1.307 | 3.175 | 0.568 | -2.34 | -2.05 | -0.754 | 0.16 | 0.301 | 0.383 | 0.1 | -0.369 | -0.31 | -0.172 |
| Orth | 0.054 | 0.254 | -0.184 | -0.785 | -0.296 | 0.416 | 0.529 | 0.179 | 0.162 | -0.062 | -0.382 | -0.13 | 0.175 | 0.334 |

**图表说明：**

展示了本次PLS模型结果，主要包括模型的系数，用于分析自变量X对于因变量Y的影响关系情况。

**智能分析：**

模型的标准化公式为：  
1 try = -0.753+0.18 \* 单词是否具有专业性+0.001 \* 单词频率+0.038 \* 外来词+0.268 \* 单词形式+0.134 \* 字母重复次数+0.054 \* Orth  
2 tries = -0.308-1.133 \* 单词是否具有专业性+0.008 \* 单词频率+0.692 \* 外来词+0.786 \* 单词形式+1.307 \* 字母重复次数+0.254 \* Orth  
3 tries = 12.296-2.718 \* 单词是否具有专业性+0.019 \* 单词频率+1.433 \* 外来词+0.395 \* 单词形式+3.175 \* 字母重复次数-0.184 \* Orth  
4 tries = 32.279+0.262 \* 单词是否具有专业性-0.001 \* 单词频率+1.914 \* 外来词-0.878 \* 单词形式+0.568 \* 字母重复次数-0.785 \* Orth  
5 tries = 32.034+1.996 \* 单词是否具有专业性-0.014 \* 单词频率-0.395 \* 外来词-0.379 \* 单词形式-2.34 \* 字母重复次数-0.296 \* Orth  
6 tries = 19.088+1.345 \* 单词是否具有专业性-0.01 \* 单词频率-2.341 \* 外来词-0.208 \* 单词形式-2.05 \* 字母重复次数+0.416 \* Orth  
7 or more tries (X) = 5.139-0.058 \* 单词是否具有专业性-0.003 \* 单词频率-1.151 \* 外来词+0.039 \* 单词形式-0.754 \* 字母重复次数+0.529 \* Orth

### 参考文献 [1] Scientific Platform Serving for Statistics Professional 2021. SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com. [2] 梁北辰,戴景民. 偏最小二乘法在系统故障诊断中的应用[J]. 哈尔滨工业大学学报,2020,52(3):156-164. DOI:10.11918/201805149.