

建模背景

在金融服务业中，评估客户的贷款违约风险是信贷决策过程中的核心环节。为了提升风险评估的准确性与科学性，构建一个非线性预测模型，用于估计客户发生贷款违约的概率具有重要意义。该模型可广泛应用于信用评分、风险定价以及贷款审批策略的制定。

模型输入变量涵盖客户的关键财务与背景信息，包括年收入、负债收入比、信用评分、贷款期限和工作年限。通过将这些变量映射至一个非线性函数中，模型能够输出客户违约的概率值，从而辅助金融机构进行更精细化的风险管理与客户筛选。

建模公式

模型采用逻辑回归风格的非线性函数形式，其核心思想是将多个输入变量经过非线性变换后，组合成一个综合风险评分，并通过S型函数将其映射至[0,1]区间，表示违约概率。

建模公式如下：

$$\text{fun}(\text{income}, \text{debt_ratio}, \text{credit_score}, \text{loan_term}, \text{employment_years}) = \frac{1}{1 + e^{-(a \cdot \text{income} - b \cdot \text{debt_ratio} + c \cdot \ln(\text{credit_score}) + d \cdot \text{loan_term} + e \cdot \text{employment_years})}}$$

其中各项变量含义如下：

- `income`：客户年收入，反映其还款能力；
- `debt_ratio`：负债收入比，衡量客户当前债务压力；
- `credit_score`：信用评分，体现客户历史信用行为；

- `loan_term`: 贷款期限, 反映贷款时间风险;
- `employment_years`: 工作年限, 用于评估职业稳定性。

函数中各项系数 a, b, c, d, e 为模型参数, 分别表示各变量对违约概率的影响权重。通过非线性组合, 该模型能够有效捕捉变量间的复杂关系, 提供更准确的风险预测结果。