Credit Scoring

2024 - Assignment 2

This assignment is to be submitted by 16:00, UK time, on **Thursday April 4.**

You are to submit a single PDF on Gradescope.

Where appropriate, print-out’s of your calculations, manipulations, graphs, data modelling, etc. are to be included in the PDF.

This is an individual assignment.

Therefore, students are expected to work on this assignment independently, and for the work submitted to be their own.

students are expected to express ideas in their own words and not simply re-use phrases from the lecture notes.

There are 14questions. You should attempt all questions.

This assignment contributes 66% of the grade for the course.

**Question 1**  **(5 marks)**

If a bank or other lender uses risk-based pricing as part of the application process for a credit product, the end-to-end application process takes longer than if risk-based pricing is not used.

List three reasons for this. **3**

When this bank identifies a potential new customer who is high quality, they are approved for the credit product they are applying for but are offered it at a reduced price. Why might the bank do this?

**2**

**Question 2 (4 marks)**

An examination of the Lending Club data reveals the following Bad Rates split by attributes of two characteristics, Term and Loan Purpose.

|  |  |  |
| --- | --- | --- |
| Bad Rates | Term | Term |
| Loan Purpose | 36 months | 60 months |
|  |  |  |
| Vacation | 11.3 | 26.4 |
| Debt Consolidation or credit card re-financing | 12.5 | 19.4 |
| Medical Expenses | 13.8 | 15.0 |
| Business | 22.1 | 20.5 |

Table 1

There is a possible interaction here between Term and Loan Purpose.

Explain briefly in terms of these characteristics two ways we might try to model or manage this interaction.

**Question 3 (5 marks)**

In a personal loan business, we have been analysing the performance of cases.

After introducing some reject inference, we have the following “confusion” matrix:

|  |  |  |
| --- | --- | --- |
|  | Predicted Good | Predicted Bad |
| Actual Good | 27177 | 11039 |
| Actual Bad | 1038 | 7204 |

Table 2

a.

Calculate the F1-score for the model being used. **4**

b.

If we swap the definitions of Good and Bad, will the F1-score change? **1**

**Question 4** **(3 marks)**

A lender has two scoring models being considered for use. Consider the graph below showing (ROC) curves for the two credit scoring models. The graph for Model A is in orange (with shorter dashes) and the graph for Model B is in blue (with longer dashes). Both models have a Gini coefficient of 66.75%. The lender here does not have risk-based pricing in place.

Comment on the performance of the models. Which would you recommend for use and why?

Figure 1

**Question 5 (8 marks)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Score | Applications | | Accepts | Rejects | Accept |  |  | | Accept Population |  |
| Band | # | Col % | # | # | Rate % |  | Goods | Bads | | Odds |
| 220-239 | 9802 | 29.59% | 1865 | 7937 | 19.03% |  | 1731 | 134 | | 12.918 |
| 240-259 | 11802 | 35.63% | 10662 | 1140 | 90.34% |  | 9781 | 881 | | 11.102 |
| 260-279 | 10411 | 31.43% | 9804 | 607 | 94.17% |  | 9357 | 447 | | 20.933 |
| 280-299 | 1107 | 3.34% | 1054 | 53 | 95.21% |  | 1029 | 25 | | 41.160 |
|  |  |  |  |  |  |  |  |  | |  |
| Total | 33122 | 103% | 23385 | 9737 | 70.60% |  | 21898 | 1487 | | 14.726 |

Table 4

The scorecard performance reported in the above table produces scores on a range of 80 points, from 220 to 299. We are going to explore this table and the model and output using Reject Inference / Over-ride factors.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Odds | Inferred Goods | Inferred Bads |  | Total Goods | Total Bads | Odds |
| 220-239 |  |  |  |  |  |  |  |
| 240-259 |  |  |  |  |  |  |  |
| 260-279 |  |  |  |  |  |  |  |
| 280-299 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Total |  |  |  |  |  |  |  |

.

Table 5

a.

Calculate the entries for the red boxes in table 5 above, utilising a Reject Inference / Over-ride factor of 0.3.

**5**

b.

In what sense might we think of 0.3 being the “correct” factor? In other words, in what sense does it produce a “very good” scorecard?

**2**

c.

If the factor used is close to 1, what would this say about the application process?

**1**

**Question 6 (2 marks)**

In developing a scorecard for a mortgage product, for example, there can be a debate about what should be the length of the performance window. In developing a scorecard for a PayDay Loan product, however, the developers are not usually concerned about the length of the performance window. Why is this?

**Question 7 (5 marks)**

The profile and performance of cases in a data set for the characteristic Age appears below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | 18-19 | 20 | 21 | 22 | 23 | 24 | 25-27 | 28-29 | 30-34 | 35-42 | 43-47 | 48-52 | 53-56 | 57-64 | 65+ |
| B | 2 | 4 | 3 | 4 | 8 | 15 | 32 | 17 | 92 | 101 | 73 | 50 | 45 | 47 | 7 |
| G | 7 | 23 | 28 | 54 | 55 | 64 | 339 | 331 | 914 | 1417 | 784 | 770 | 534 | 605 | 175 |

Table 3

Use the maximum likelihood monotone classifier approach to produce a coarse classing of the attributes of this characteristic.

**Question 8 (10 marks)**

Earlier in the course, we applied a four-characteristic model to the Irish data set and explored its power.

Here, we are going to add a fifth characteristic.

This additional characteristic has the following scores:

Current Account - Months of Unauthorised Debit Interest L6M

|  |  |
| --- | --- |
| Attribute Grouping | Score |
| 0 | 34 |
| 1 | 21 |
| 2 | 9 |
| 3 | 0 |
| 4 | -10 |
| 5 or 6 | -25 |

Table 6

One way to do this is as follows:

*In the Irish data set Applying the Model tab, insert a new column between Column AM and AN. Give it an appropriate label. Write the Excel command to apply the new characteristic, perhaps using the syntax in columns AJ, AK, or AM as an example, or set up a new Lookup table. Change the commands in the old Column AN to include this new characteristic’s scores.*

*The old scores, i.e. the scores before adding the new characteristic, for the first five records are 142, 164, 122, 180, and 94.*

*Do not change the order of the records.*

a), Derive the new total scores for all of the data. Include the first ten values for the new total score (to confirm that the new characteristic has been added correctly).

**2**

b), Applying this new model to the data, calculate the Kolmogorov-Smirnov statistic.

Also, draw the K-S graph, including appropriate labels on the axes.  **4**

c), The Gini coefficient has a value of approximately 0.61. What do you conclude about whether this new characteristic should be included in the model?

**2**

d), Your results are very likely to be biased. What should we do to eliminate this bias?

**2**

**Question 9** **(2 marks)**

There are several methodologies that can be used to build credit scoring models. One of these is linear regression. It has a weakness that causes many developers to use logistic regression instead. What is this weakness and what is different about logistic regression?

**Question 10** **(5 marks)**

In monitoring an applicant profile, we have the following characteristic analysis and report.

Read the report and then answer the questions below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Expected | Current Sample | Score Weight | Difference | Index |
| Existing Borrowing Customer | 42% | 48.6% | 27 | 6.6% | 1.782 |
| Existing Non-Borrowing Customer | 33% | 27.2% | 17 | -5.8% | -0.986 |
| New Customer | 22% | 15.2% | 6 | -6.8% | -0.408 |
| No Response / Don’t Know | 3% | 9.0% | 15 | 6.0% | 0.9 |
|  |  |  |  |  |  |
|  |  |  |  |  | 1.288 |

Table 7

Report

1. Overall, this characteristic is contributing 1.29 points more to the total score than expected. This is generally good news, although we should be careful that it does not lead to significantly higher acceptance rates than planned.
2. Examining the attributes individually, we can see that we have a higher percentage than expected from existing borrowing customers.
3. We can also see that we have a lower percentage than expected from existing non-borrowing customers.
4. Further, there is a smaller percentage than expected from new customers.
5. There is also a large increase in the percentage of cases where we do not know. This needs urgent investigation.

a,

Why might point 4) be a problem for the business? **1**

b,

Calculate the value of Population Stability Index for this characteristic. **2**

c,

The PSI has some deficiencies. However, it also has one major advantage over the index calculated in the table above. What is this? **2**

**Question 11 (7 marks)**

We are analysing a scorecard being used on the credit product to which the Irish data relate. This scorecard has been developed using the usual PDO scale of 20.

We are looking at the characteristic Loan Purpose. This characteristic is not in the current model.

The attributes have been grouped into three groups:

BOT / REF / RES / TAX

MED / BUS / FAR

All Others

The table of data is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Loan Purpose | BOT / REF / RES / TAX | MED / BUS / FAR | All Others |
| Actual Goods | 1694 |  | 4096 |
| Actual Bads |  |  |  |
| Expected Goods | 1731.1 | 192.4 | 4060.2 |
| Expected Bads | 357.1 | 8.9 | 124.3 |

Table 8

The expected numbers of Goods and Bads come from the scorecard development. (You are given these - you do not need to calculate these.)

Examine the Irish data and complete the table above. **1 1/2**

Then, complete Table 9 below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Attribute | Delta Score | Excess Goods | Possible Attribute Score | χ2 Contribution |
| BOT / REF / RES / TAX |  | -37.1 |  | 4.09 |
| MED / BUS / FAR |  |  |  | 0.43 |
| All Others | 10.19 | 35.8 |  |  |

Table 9 **4 1/2**

Finally, comment on the possible changes to the scorecard. **1**

**Question 12 (2 marks)**

Analysis of lending portfolios for Basel / IFRS9 purposes provided evidence of credit behaviour that might have been predicted – that charging higher levels of interest leads to increased charges and repayments and generally increases the level of defaults.

However, the analysis also revealed a quite different behavioural effect of charging higher levels of interest. What was it?

**Question 13 (2 marks)**

What is the fundamental difference between the way we interpret a K-S statistic?

A in scorecard development, where we might measure the difference between the cumulative distributions of Good and Bad cases

And

B in scorecard monitoring, where we might measure the difference between the cumulative distribution of scores for applications expected, perhaps from the development sample, and the cumulative distribution of scores of applications actually received

**Question 14 (6 marks)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Score** | **Applications** | **Accepts** | **Rejects** |
| **Up to 209** | **8117** | **84** | **8033** |
| **210-214** | **6432** | **134** | **6298** |
| **215-219** | **7897** | **795** | **7102** |
| **220-224** | **6737** | **6535** | **202** |
| **225-229** | **5984** | **5893** | **91** |
| **230-239** | **9791** | **9687** | **104** |
| **240-249** | **8479** | **8403** | **76** |
| **250+** | **16841** | **16770** | **71** |
|  |  |  |  |
| **Total** | **70278** | **48301** | **21977** |

Table 10

The cut-off is currently 220 with the following results:

Scorecard Pass rate = 68.061%

Accept rate = 68.729%

HSO rate = 1.137%

a), What is the LSO rate? **2**

b), if the cut-off had been 215 instead of 220, what would have been the scorecard pass rate?

**1**

c), if the cut-off had been 215 instead of 220, on the basis of the data in the table, what would have been the range of possible values for the Accept rate?

**3**