22956 39050

Consumer Credit analysis.docx

Assignment

Class

University

Document Details

Submission ID

trn:oid:::1:2929183600

Submission Date

May 23, 2024, 10:56 PM UTC

Download Date

May 23, 2024, 10:56 PM UTC

File Name

AgADwRcAAvFbgFI

File Size

41.2 KB

7 Pages

766 Words

4,740 Characters

How much of this submission has been generated by AI?

0%

of qualifying text in this submission has been determined to be generated by AI.

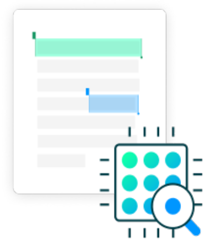
Caution: Percentage may not indicate academic misconduct. Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

What does the percentage mean?

The percentage shown in the AI writing detection indicator and in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was generated by AI.



Our testing has found that there is a higher incidence of false positives when the percentage is less than 20. In order to reduce the likelihood of misinterpretation, the AI indicator will display an asterisk for percentages less than 20 to call attention to the fact that the score is less reliable.

However, the final decision on whether any misconduct has occurred rests with the reviewer/instructor. They should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in greater detail according to their school's policies.

How does Turnitin's indicator address false positives?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be AI-generated will be highlighted blue on the submission text.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

What does 'qualifying text' mean?

Sometimes false positives (incorrectly flagging human-written text as AI-generated), can include lists without a lot of structural variation, text that literally repeats itself, or text that has been paraphrased without developing new ideas. If our indicator shows a higher amount of AI writing in such text, we advise you to take that into consideration when looking at the percentage indicated.

In a longer document with a mix of authentic writing and AI generated text, it can be difficult to exactly determine where the AI writing begins and original writing ends, but our model should give you a reliable guide to start conversations with the submitting student.

Consumer Credit analysis

Question a

Discriminant Analysis (DFA) analysis is mainly performed on a set of data to identify the linear functions existing in the data. In this case, Canonical Discriminant Analysis (DFA) was performed on the dataset "Ass2Credit" to identify a linear function of the variables that best discriminates between individuals who pay off their debt (TARGET = 0) and those who do not (TARGET = 1).

Table 1.0

Canonical Discriminant Analysis

The DISCRIM Procedure
Canonical Discriminant Analysis

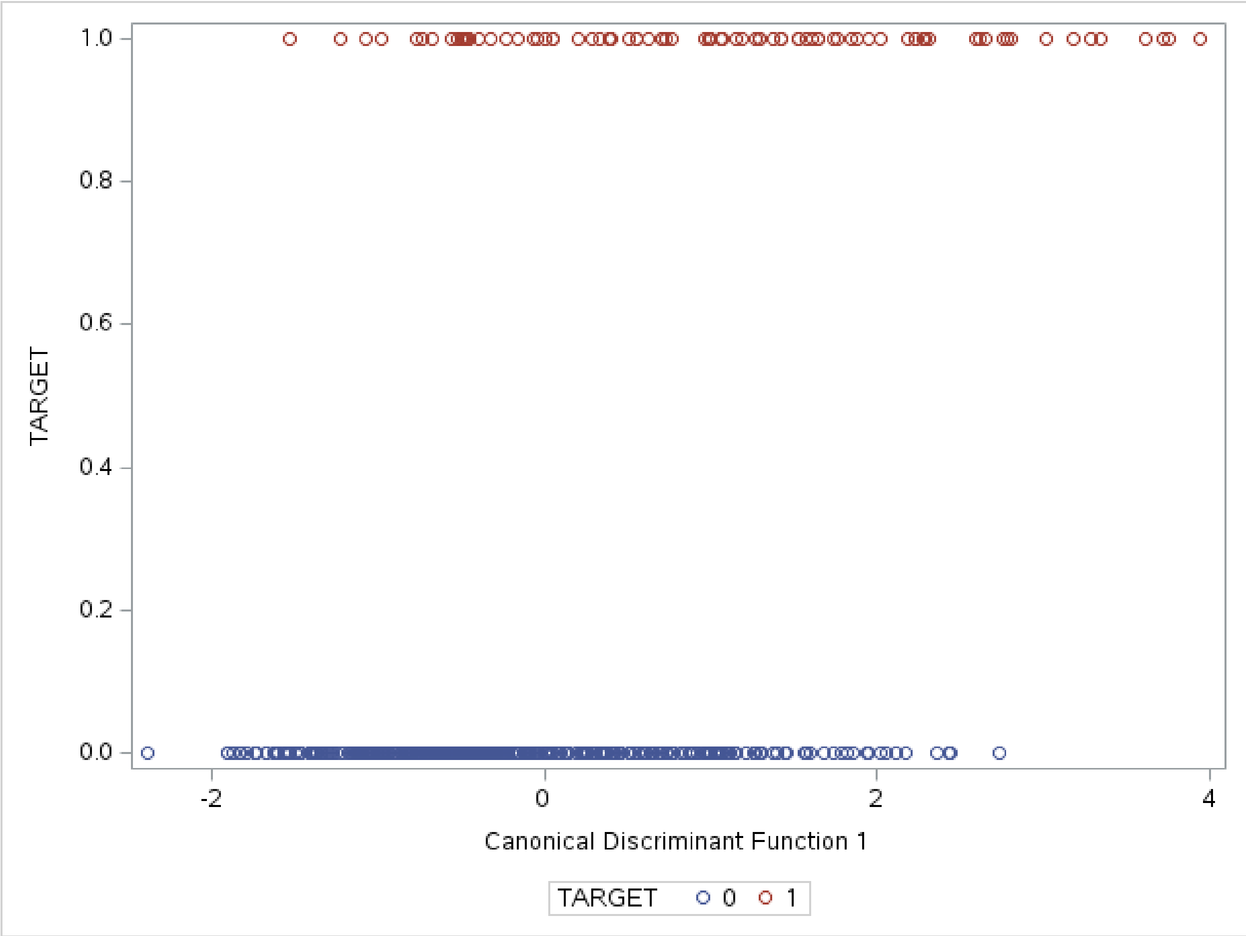
	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation	Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)				Test of H0: The canonical correlations in the current row and all that follow are zero				
					Eigenvalue	Difference	Proportion	Cumulative	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.450851	0.436162	0.036790	0.203267	0.2551		1.0000	1.0000	0.79673341	13.04	9	460	<.0001

1) Only one discriminant function was identified in the data. The Canonical Discriminant Analysis table above provides different statistics and also identifies the number of linear functions in the data.

The below plot justifies the use of a single discriminant function in separating the two individual groups.

Figure 1.0

Plot of Canonical Discriminant (Function) Analysis



2) Identifying the most important variables that discriminate between the different classes in the data, can be achieved by critically analyzing the "Total-Sample Standardized Canonical Coefficients" and "Pooled Within-Class Standardized Canonical Coefficients" tables from the output of the PROC DISCRIM procedure allowing to determine which variables contribute the most to the discrimination between classes. The variables with the highest absolute coefficient are deemed the most important.

Table 2.0

Total-Sample Standardized Canonical Coefficients

Total-Sample Standardized Canonical Coefficients		
Variable	**Label**	**Can1**
CollectCnt	Number Collections	0.0499636785
InqFinanceCnt24	Number Finance Inquires 24 Months	0.5466558736
InqTimeLast	Time Since Last Inquiry	0.1033763125
TLTimeFirst	Time Since First Trade Line	-.1687640127
TLBalHCPct	Percent Trade Line Balance to High Credit	0.3080052121
TLSatPct	Percent Satisfactory to Total Trade Lines	-.7163946321
TLSum	Total Balance All Trade Lines	0.0752990652
TLOpenPct	Percent Trade Lines Open	0.3295953728
TLDel60Cnt24	Number Trade Lines 60 Days or Worse 24 Months	0.3712633783

Table 3.0

Pooled Within Canonical Structure Coefficients

Pooled Within Canonical Structure		
Variable	**Label**	**Can1**
CollectCnt	Number Collections	0.172287
InqFinanceCnt24	Number Finance Inquires 24 Months	0.486341
InqTimeLast	Time Since Last Inquiry	-0.083938

Pooled Within Canonical Structure		
Variable	**Label**	**Can1**
TLTimeFirst	Time Since First Trade Line	-0.181288
TLBalHCPct	Percent Trade Line Balance to High Credit	0.401838
TLSatPct	Percent Satisfactory to Total Trade Lines	-0.618060
TLSum	Total Balance All Trade Lines	-0.029255
TLOpenPct	Percent Trade Lines Open	-0.135203
TLDel60Cnt24	Number Trade Lines 60 Days or Worse 24 Months	0.599692

The most important variables from the data are TLSatPct, TLDel60Cnt24 and InqFinanceCnt24. These variables are the most critical in distinguishing between individuals who pay off their debt and those who do not.

3. I would consider reducing the variables in the data and focus on the most important variables as identified. Reducing the variables and focusing on the most important variables in the data will ensure the DFA methods sustains higher discriminative Power, Simplicity and increase classification accuracy.

Question b

Fisher Discriminant Analysis

Fisher Discriminant Analysis also know as Linear Discriminant Analysis (LDA) is a mathematical method used to classify data points based on their characteristics that separates the data distinctively. To classify the individuals, Fisher Discriminant Analysis method was applied. The method produced an overall classification accuracy of 68.40%. The method correctly classified 75.85% of class 0 and 67.44% of class 1 and had an error rate 28.36%.

Table 4.0

Number of Observations and Percent Classified into TARGET

Number of Observations and Percent Classified into TARGET			
From TARGET	**0**	**1**	**Total**
0	314 75.85	100 24.15	414 100.00
1	28 32.56	58 67.44	86 100.00
Total	342 68.40	158 31.60	500 100.00
Priors	0.5	0.5	

Table 5.0

Error Count Estimates for TARGET

Error Count Estimates for TARGET			
	0	**1**	**Total**
Rate	0.2415	0.3256	0.2836
Priors	0.5000	0.5000	

linear discriminant analysis is more preferable than the quadratic discriminant analysis method since the LDA results satisfies the assumption of equality of the covariance matrices of the two classes. The assumption of linearity holds for the data, this implies that the method finds a linear combination of the predictor variables that best separates the classes. Based on the visual inspection, the assumption of linearity is reasonable for the most important variables. In practical applications, especially in financial and consumer credit analysis, simpler models that are easy to explain to stakeholders are often preferred. The performance difference between the two methods is not substantial enough to warrant the additional complexity of QDA.

References

Discriminant Function Analysis | SAS Data Analysis Examples. (2018). https://stats.oarc.ucla.edu/sas/dae/discriminant-function-analysis/

Fisher's linear discriminant functions. (2018, April 6). https://communities.sas.com/t5/Statistical-Procedures/Fisher-s-linear-discriminant-functions/td-p/451752