ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 4 No. 2 (2024) 35-44
https://publisher.uthm.edu.my/periodicals/index.php/ekst

The Web Application of Tenant Credit Scoring Using Python

Kuan Hiu Yan[1], Siti Suhana Jamaian[1*]

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA*

*Corresponding Author: suhana@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2024.04.02.005

Article Info

Received: 27 December 2023
Accepted: 11 January 2024
Available online: 12 December 2024

Keywords

Tenant Credit Scoring, Credit History, Logistics Regression, Web Application

Abstract

Having a poor credit score or no credit history can restrict options for both housing and employment. This study seeks to mitigate the issue of credit invisibility among the low-income demographic with limited credit history by using tenant credit scoring web application. In this study, a credit scoring model is created using tenants' attributes, monthly rent and rental payment history through the implementation of graphical user interface. A simple logistics regression is applied to compute the credit score of tenants based on their characteristics. Based on the findings of this study, the primary determinants of the tenant's credit score include gender, age, the number of months with late payments, the expense-to-income ratio, and the previous monthly rent. The development of the web application involves the utilization of HTML, CSS and flask library in Python. Finally, the web application is developed and the creditworthiness of a tenant is calculated and displayed in a credit scoring model.

1. Introduction

Housing affordability has become a major concern in many countries including Malaysia due to the significant increase in housing prices which have surpassed income growth rates, resulting in greater challenges for individuals seeking to purchase or rent homes. To measure housing affordability, National Housing Department Malaysia [1] state that a prevalent approach involves analyzing the median household income in a particular locality against the median price of houses or rent. When the ratio of housing cost to income surpasses a certain threshold, this could suggest that the housing is beyond the means of many individuals in the region. The median monthly gross income of households was RM5,209 in the year 2020, which was lower than the previous year's median of RM5,873 [2]. During the Tenth Malaysia Plan period from 2011 to 2015, a total of 102,201 affordable homes were completed through various initiatives that aimed to meet the specific needs of different target groups. The Perumahan Rakyat 1Malaysia (PR1MA), Program Perumahan Rakyat (PPR) and Rumah Mesra Rakyat 1Malaysia were designed by the government of Malaysia to assist the low-income households [3].

The development of credit scoring model has become a crucial concern due to the intense competition in the credit business and issues with bad debt [4]. A credit score is a creditworthiness tool used by bank or other financial institution to determine the loan eligibility of a borrower. The higher the borrowers score in model, the higher the loan will be [5]. In the past, the presence of prior financial actions such as loan repayment is required to assess creditworthiness. The strict standards made it so that individuals who had little to no financial history were unable to establish credit, therefore making them 'credit invisible' [6]. In rural areas, housing loan applications from the B40 group are commonly declined by banks, often due to their classification as 'credit unscored' or 'credit invisible'. As a result, these individuals usually resort to renting properties. Unfortunately, their rental payment history does not play a significant role in housing loan applications.

This research discovered that several factors were responsible for the rejection of home loan applications. These included poor state of the Central Credit Reference Information System (CCRIS), inadequate income, insufficient creditworthiness, a high rate of defaults, concerns about the ability to recover loan and auction operating costs, incapability to repay the loan, absence of evidence of regular income, the lack of collateral, and inability to provide a down payment. As a result, the Malaysian government has implemented several programs over time with the aim of making housing affordable and accessible to its citizens [7]. Additionally, private credit reporting agencies such as Credit Tip-Off Service (CTOS) and RAM Credit Information Sdn Bhd. (RAMCI) provide Malaysians with their reports and credit scores. There are two main credit scoring models which are FICO scoring model which also known as the Fair Isaac Corporation and the VantageScore model was introduced by Equifax, Experian, and TransUnion in United States [8]. Table 1 shows the characteristics of credit scoring models.

Table 1 Characteristics of credit scoring website

	Credit Scoring Model	Depend on Credit History	Must Purchase	Include Advertisement
FICO	/	/	/	/
VantageScore	/	/	/	/
CTOS	/	/	/	/
RAMCI	/	/	/	/
Our proposal model	/			

Individuals in economies with low incomes are unable to obtain a credit or bank account due to their lack of financial history which makes it impossible to assign them a credit score. Therefore, some credit bureaus help to assess the accuracy of the models using non-traditional data instead of traditional data such as credit card usage, loan repayments and mortgage history to anticipate credit risk for new accounts in lieu of credit history [9]. Credit scoring utilizes non-traditional data such as rental and utility payments, delinquency, traffic violations, bankruptcy, legal proceedings, employment data [6].

Various credit scoring models have been created by banks and researchers to address the classification issue based on factors like age, credit limit, income, and marital status. These models include linear discriminant analysis (LDA), logistic regression (LR), multivariate adaptive regression splines (MARS), classification and regression tree (CART), case-based reasoning (CBR), and artificial neural networks (ANNs) [4].

Credit scores have a significant impact on various aspects beyond the loans one can acquire or the interest rates they are charged. Credit scores serve as a financial instrument, but their effectiveness can either be constructive or destructive based on their quality. Therefore, this study aims to raise the creditworthiness of the demographic with modest income and restricted credit history, thereby increasing the possibility of approval of their mortgage application.

2. Methods

The emphasis of this study is on the demographic with modest income and restricted credit history, the debt-to-income ratio factor is substituted with the expense-to-income ratio [10]. Hence, gender, age, monthly income, household income group, expense-to-income ratio, number of dependents, previous monthly rent and number of months late payment are the factors that are contemplated to affect the tenant credit score. The analysis also includes examining how the considered factors affect the tenant's credit score, and if any factor is found to have minimal significance in the model, the logistic regression parameters will be recalculated after removing that factor. In the meantime, a credit scoring model for the tenant is proposed, utilizing the generated probability of the tenant defaulting. Hence, a proposed tenant's credit scoring model using graphical user interface is evaluated.

2.1.1 Simple logistics regression

In a linear regression model, we assume that the outcome variable can be represented or described by

$$g(x) = \beta_0 + \beta_1 x_1 + \varepsilon\,, \tag{1}$$

where x is independent variable, β is the logistic regression parameter. The ε is distributed with a mean of zero and a constant variance in linear regression. The logistics regression model is [11]

Enhanced Knowledge in Sciences and Technology Vol. 4 No. 2 (2024) p. 35-44

$$\pi(x) = \frac{1}{1+e^{-g(x)}}, \tag{2}$$

where $\pi(x)$ is the conditional probability of tenant defaulting and $g(x)$ is the simple logistics regression.

2.1.2 Maximum likelihood estimation for logistic regression

The natural log of the likelihood function, $L(\beta)$ can be defined as

$$L(\beta) = \sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1-y_i)\ln[1-\pi(x_i)]\}, \tag{3}$$

since the dependent variable in this study is a binary outcome [11].

When the training data is separated, solving the separation problem requires the utilization of a new objective function. This objective function incorporates penalty ridge regression on $L(\beta)$ as depicted in Equation (4) [12].

$$\text{minimize } F(\beta) = -L(\beta) + \lambda R(\beta) \tag{4}$$
$$R(\beta) = \sum_{j=1}^{p} \beta_j^2,$$

where λ is a positive regularization strength, $R(\beta)$ is ridge regression and p is the number of independent variables considered. The closer the λ is close to zero, the smaller the effect of the ridge penalty term. Consequently, a value of 0.1 is chosen over 0.01 as the optimal λ in this study [5].

2.1.3 Credit scoring

Based on the analysis of logistic coefficient ($\lambda = 0.1$), the logit of the model is written in Equation (5) [5].

$$g(x) = -7.5989 + 1.2831x_1 + 0.3265x_2 - 0.2327x_3 - 0.1406x_4 \tag{5}$$
$$+1.0339x_5 - 0.2294x_6 - 0.9136x_7 + 2.3615x_8,$$

where x_1 = gender, x_2 = age, x_3 = monthly income, x_4 = household income, x_5 = expense-to-income ratio, x_6 = number of dependents, x_7 = previous monthly rent, and x_8 = number of months late payment.

2.1.4 Tenant credit scoring model

In this research, we presented a credit assessment framework that includes a credit rating scale ranging from zero as the lowest score to 100 as the highest score. In the suggested model, a tenant with a reduced probability of defaulting will be assigned a higher credit score. The proposed credit score of the tenant is computed as Equation (6)[5].

$$\text{Credit score } = 100\big(1-\pi(x)\big). \tag{6}$$

2.1.5 Web application

This study presents a web application that offers a user-friendly graphical interface, designed specifically for Malaysian tenants who lack a credit history, to access their credit scores. The application operates online, allowing individuals to conveniently assess their creditworthiness without traditional credit records. The development of the web application involves the utilization of HTML, CSS and flask library in Python. Flask handles the backend logic, including routing, handling requests, interacting with databases, and generating dynamic content. HTML is used to structure and define the content of web pages, specifying how data should be presented to users. The web application might include several tasks such as retrieving credit score data, calculating credit score of tenants, displaying the credit score and automated scoring updates. This combination

of technologies enables the creation of an interactive and dynamic user interface for the application. This research is to enhance the graphical user interface of tenant credit scoring by utilizing Python programming.

3. Results and Discussion

Fig. 1 shows the flow chart of algorithm for generating tenant credit scoring.

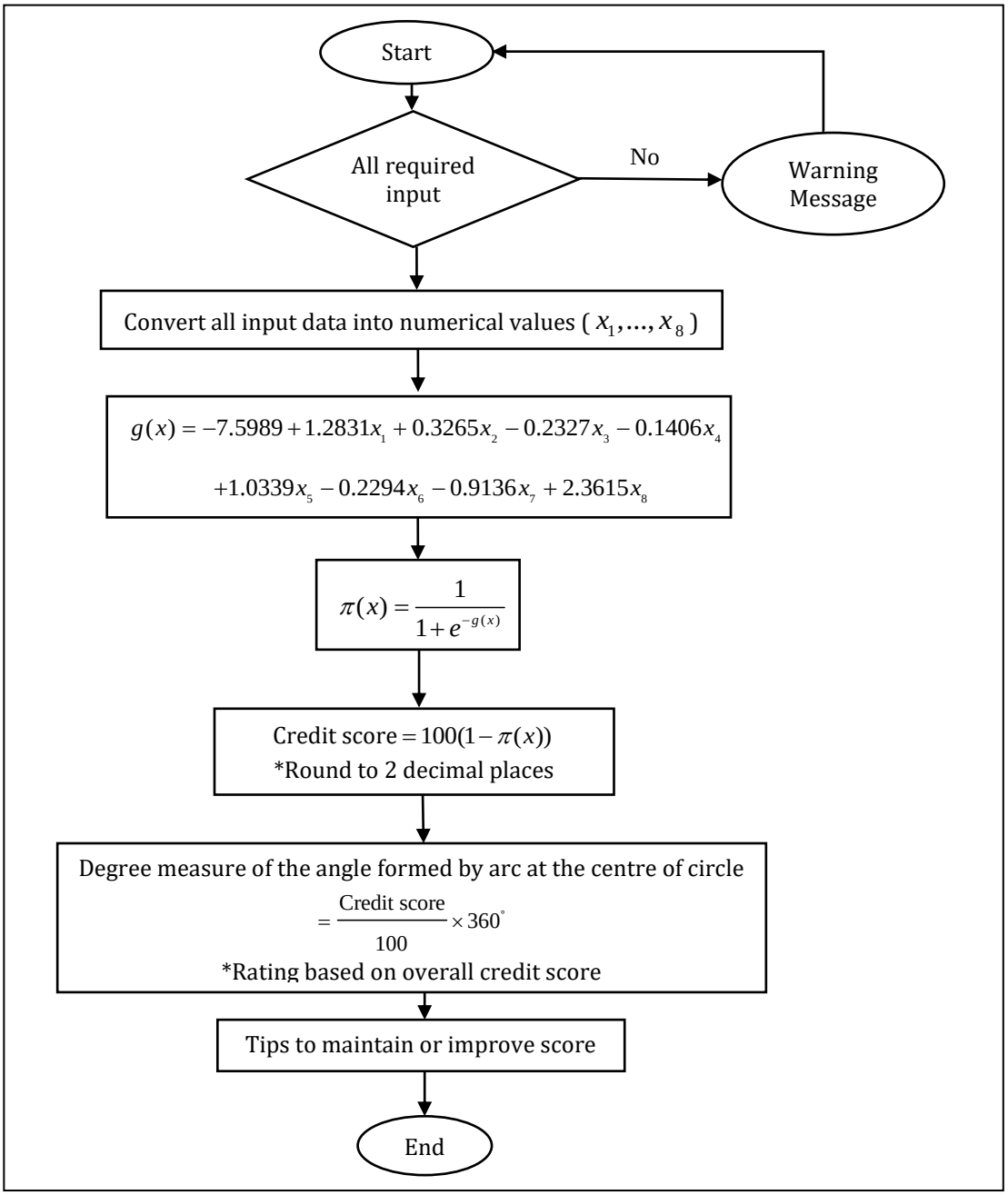


Fig. 1 Flow Chart of Algorithm for Generating Tenant Credit Scoring

The web application is developed using HTML and Flask Library in Python. The home page of the web application contains 3 parts which are personal information, employment information and miscellaneous details, as illustrated in Fig. 2.

Enhanced Knowledge in Sciences and Technology Vol. 4 No. 2 (2024) p. 35-44

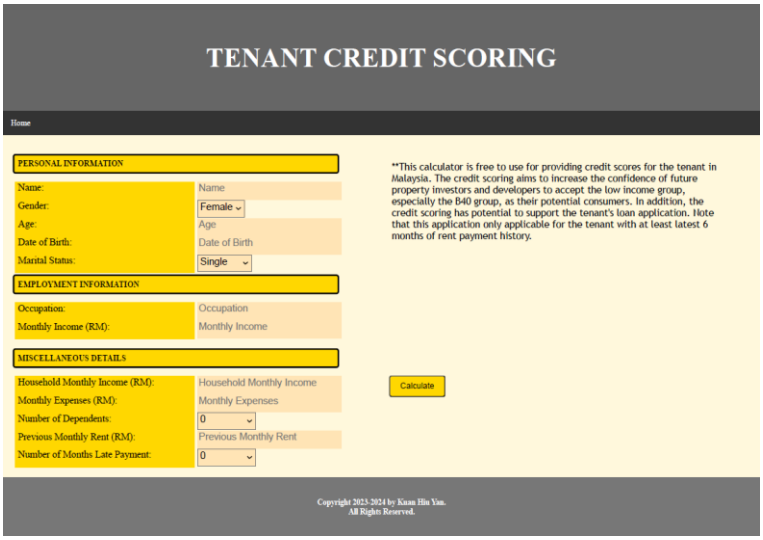
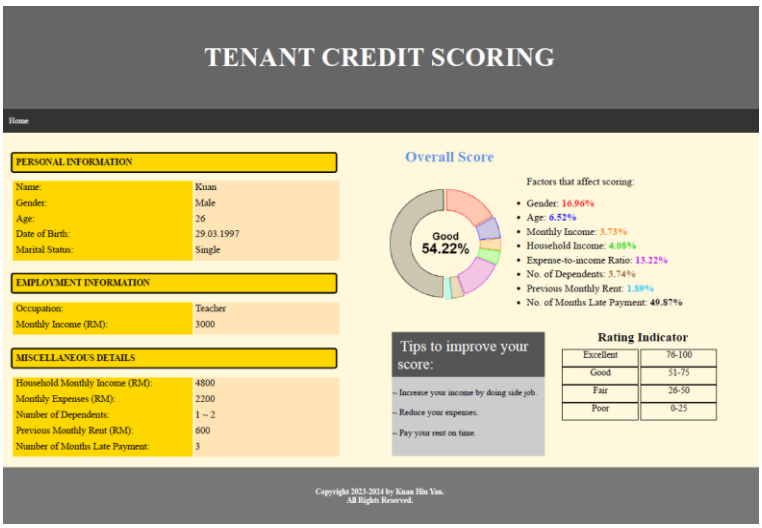


Fig. 2 Home page of web application

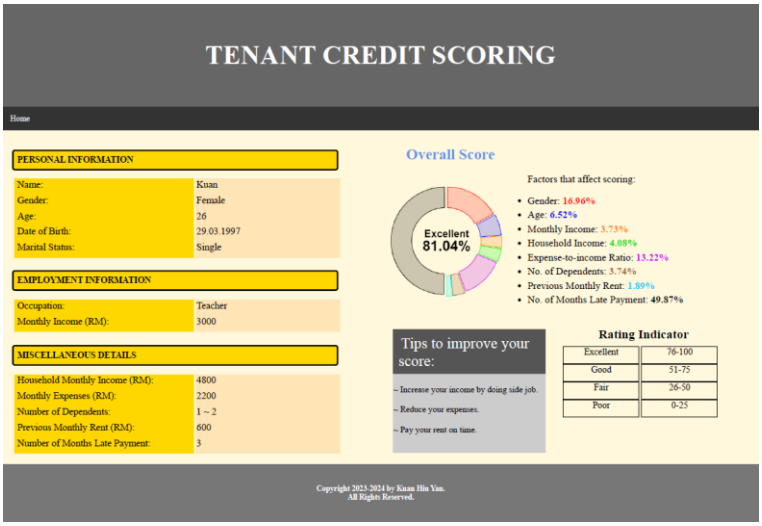
The user is required to fill in the input data and clicks the “Calculate” button. Subsequently, a new page will showcase the user's overall credit score, the rating corresponding to the score, the factors influencing the scoring, and guidelines for maintaining or enhancing the credit score. The user must fill in the age at least 18, monthly income at least RM300, household monthly income at least RM300, monthly expenses and previous monthly rent greater than 0. Conversely, the user will be given warning message to fill in again. The user also has the option to click on the "Home" link located at the top left of the navigation bar to navigate back to the home page.

3.1.1 Tenant credit scoring analysis

Fig. 3 presents a contrast in the credit reports of tenants with different genders. As depicted in the figure, the female tenant exhibits a higher credit score in comparison to the male tenant.



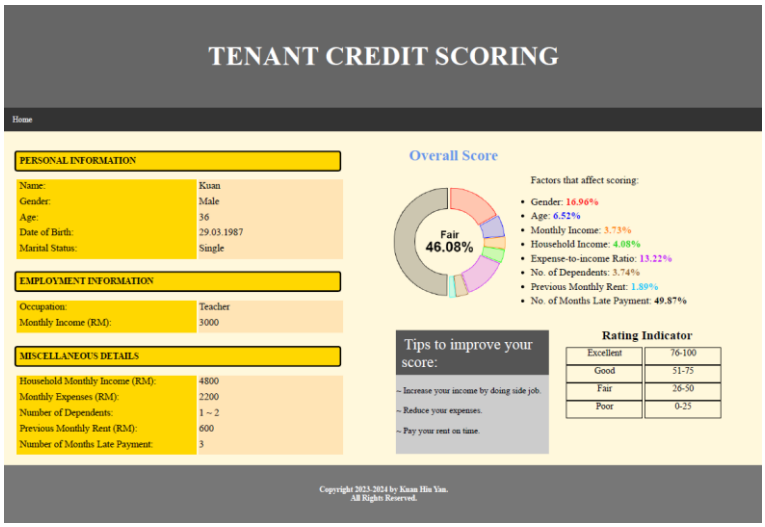
(a)



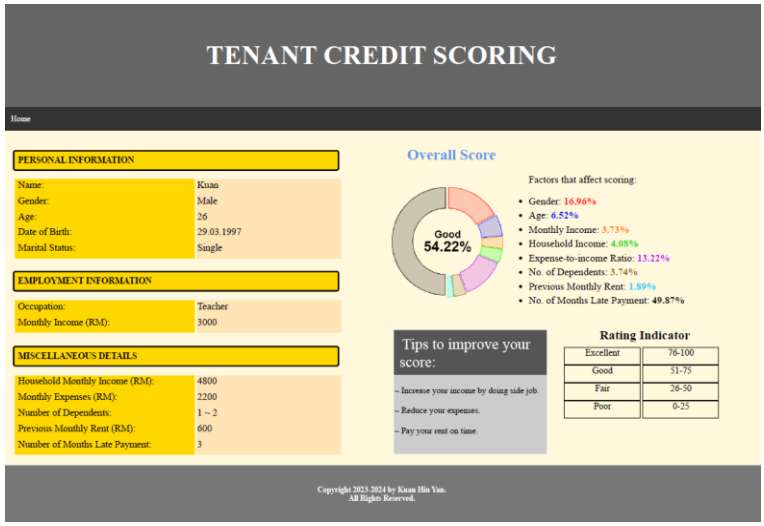
(b)

Fig. 3 Comparison of credit report with different gender (a) Male (b) Female

Fig. 4 illustrates the credit history of tenants across various age groups. A comparison between Fig. 4(a) and Fig. 4(b) reveals that the older tenant possesses a credit score that is lower than that of the younger tenant.



(a)

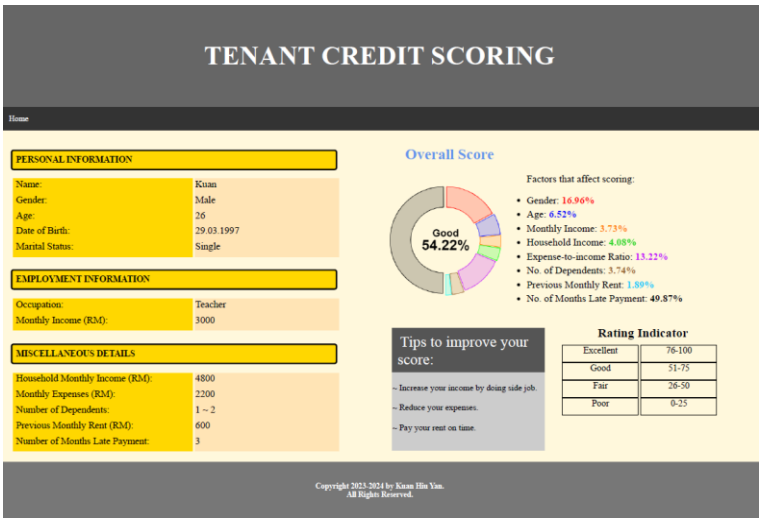


(b)

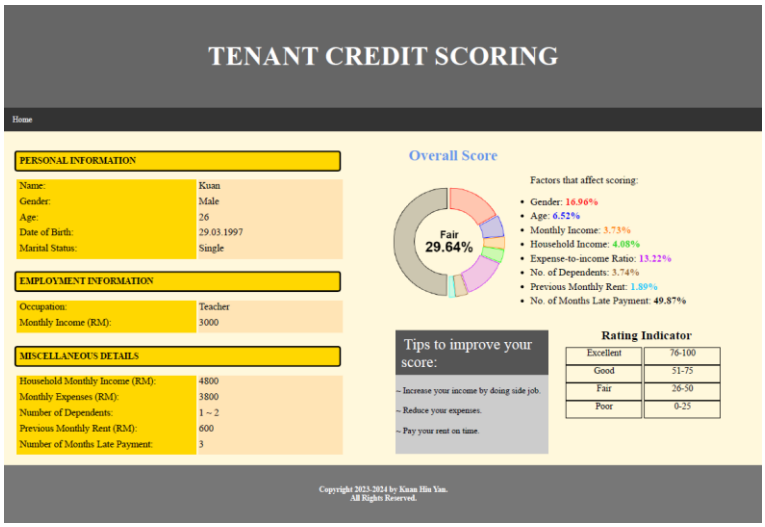
Fig. 4 *Comparison of credit report with different age (a) Older tenant (b) Younger tenant*

Enhanced Knowledge in Sciences and Technology Vol. 4 No. 2 (2024) p. 35-44

Fig. 5 depicts the credit history of a tenant under varying expense-to-income ratios. In Fig. 5(a), the tenant maintains an expense-to-income ratio of 0.73, while in Fig. 5(b), the ratio is 1.27. This suggests that the tenant with the lower expense-to-income ratio enjoys a higher credit score compared to the tenant with the higher expense-to-income ratio.



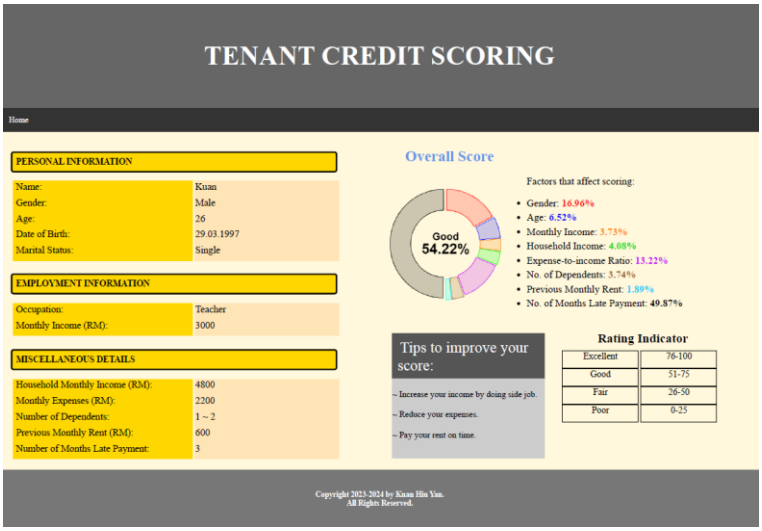
(a)



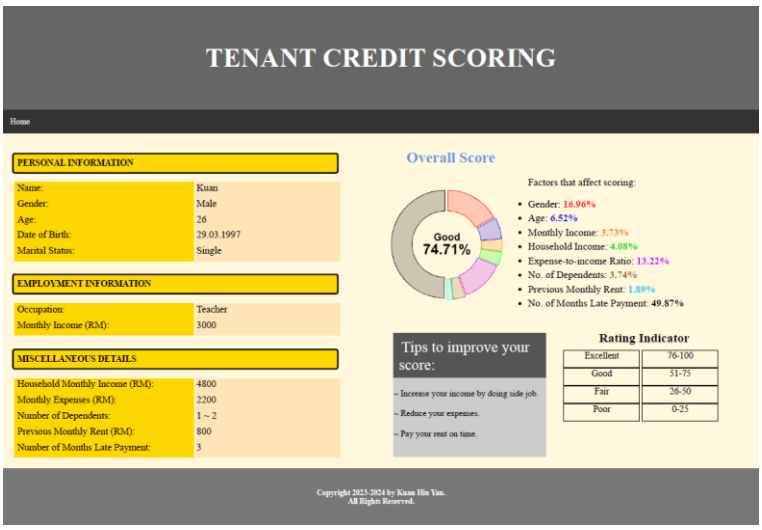
(b)

Fig. 5 Comparison of credit report with different expense-to-income-ratio (a) Lower expense-to-income ratio (b) Higher expense-to-income ratio

Fig. 6 illustrates the credit report of the tenant with different previous monthly rents. According to Fig. 6, the higher previous monthly rent tenant has a higher credit score than the lower previous monthly rent tenant.



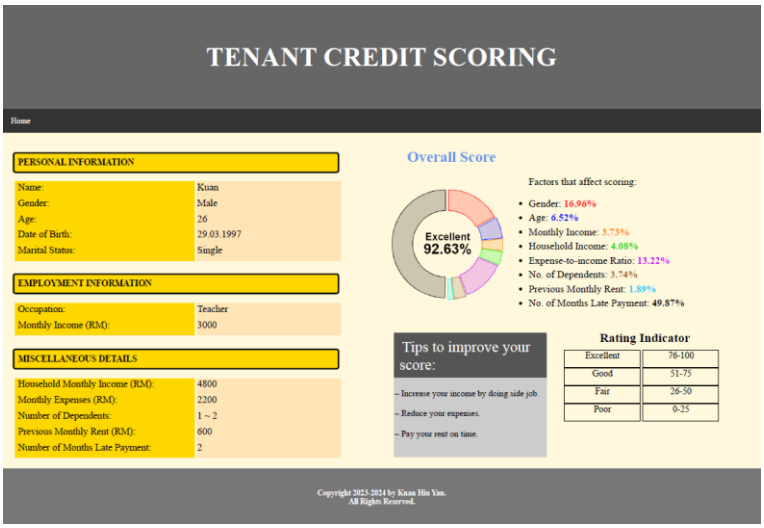
(a)



(b)

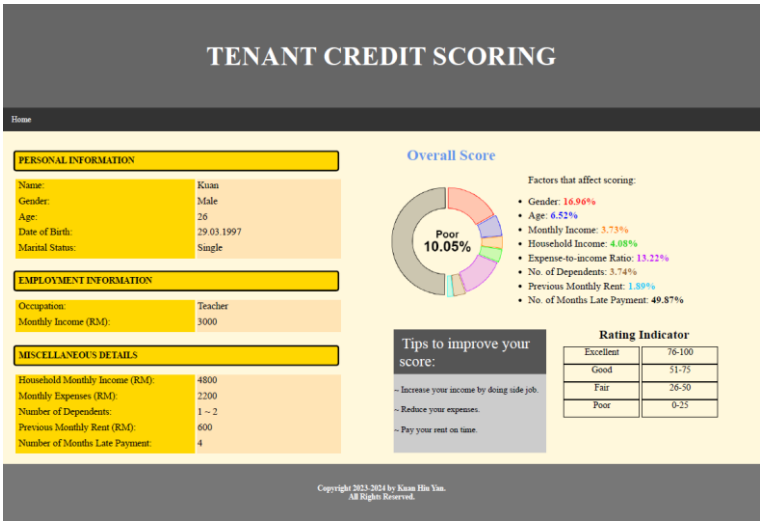
Fig. 6 Comparison of credit report with different previous monthly rent (a) Lower previous monthly rent (b) Higher previous monthly rent

Fig. 7 displays the credit history of a tenant with different instances of late payments over various months. As per the data in Fig. 7, the tenant with a higher number of months with late payments has a credit score lower than that of the tenant with fewer months of late payments.



(a)

Enhanced Knowledge in Sciences and Technology Vol. 4 No. 2 (2024) p. 35-44



(b)

Fig. 7 Comparison of credit report with different number of months late payment (a) Fewer months late payment (b) More months late payment

Table 2 shows the rankings of all factors when λ is 0.1. It can be inferred that the key determinants of the tenant's credit score include the number of months with late payments, gender, expense-to-income ratio, previous monthly rent, and age.

Table 2 Analysis of logistics coefficient [5]

Factor, x	Coefficient, β			Odds ratio (increases by one unit)
	Value	Sign	Ranking	
Gender	1.2831	+	2	3.6078
Age	0.3265	+	5	1.3861
Monthly income	-0.2327	-	6	0.7924
Household income	-0.1406	-	8	0.8688
Expense-to-income ratio	1.0339	+	3	2.8120
Number of dependents	-0.2294	-	7	0.7950
Previous monthly rent	-0.9136	-	4	0.4011
Number of months late payment	2.3615	+	1	10.6068

The percentage of factors that affect the credit score is computed using Equation (7).

$$\text{Percentage} = \frac{\text{Odds ratio of respective factor}}{\text{Sum of odds ratio}}. \tag{7}$$

The higher the percentage, the more affecting the credit score. Note that the negative sign of the logistic coefficient implies that as the value increases, the probability of default decreases, and conversely. Consequently, females are less prone to default compared to males. Additionally, a greater number of months of late payments, a higher expense-to-income ratio, or an older age correspond to an increased probability of default. Conversely, a higher previous monthly rent for the tenant, more dependents, belonging to a higher household income group, or having a higher monthly income are associated with a decreased probability of default. The deduction that the likelihood of default diminishes as monthly rent increases may stem from the tenant's inclination to choose a rental amount aligned with their financial capacity. As tenants age, there is an elevated likelihood of forgetting or missing rent payments, consequently raising the probability of default. Conversely, tenants with a higher number of dependents tend to exhibit a heightened sense of responsibility toward rent payment. Consequently, the likelihood of default decreases with an increase in the number of dependents for the tenant.

4. Conclusion

In conclusion, the development of a tenant credit scoring web application using Python, HTML, and CSS offers a tailored solution for Malaysian tenants lacking a traditional credit history. The utilization of flask, a Python web framework, facilitates the seamless integration of back-end processes, ensuring efficient routing and handling of user requests. Flask typically uses HTML templates to generate dynamic web pages by combining static HTML with placeholders that are filled in with data at runtime. HTML and CSS play a crucial role in designing an intuitive and visually appealing front-end, contributing to the accessibility of the application for users with varying levels of technological proficiency. HTML provides the structure and content of a web page by using a set of predefined elements and attributes. HTML elements are used to define the structure of a document, such as headings, paragraphs, lists, style and more. Overall, this research contributes to the advancement of credit scoring practices, especially in regions where traditional credit histories may be limited. The use of Python, HTML, and CSS in this context showcases the versatility and effectiveness of these technologies in developing practical solutions for real-world challenges.

Acknowledgement

The authors would thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia for its support.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Kuan Hiu Yan, Siti Suhana Jamaian;* ***analysis and interpretation of results:*** *Kuan Hiu Yan;* ***validation of results:*** *Siti Suhana Jamaian;* ***draft manuscript preparation:*** *Kuan Hiu Yan. All authors reviewed the results and approved the final version of the manuscript. They have agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.*

References

[1] National Housing Department Malaysia. (2019). National Affordable Housing Policy. Putrajaya: National Housing Department Malaysia.
[2] Department of Statistics Malaysia. (2021). *Household Income Estimates and Incidence of Poverty Report, Malaysia, 2020*.
[3] Liu, J., & Ong, H. T. (2021). Can Malaysia's National Affordable Housing Policy Guarantee Housing Affordability of Low-Income Households? *Sustainability, 13*(16), 8841. https://doi.org/10.3390/su13168841
[4] Chuang, C., & Lin, R. (2009). Constructing a reassigning credit scoring model. *Expert Systems With Applications*, *36*(2), 1685–1694. https://doi.org/10.1016/j.eswa.2007.11.067
[5] Ling, K. S., Jamaian, S. S., & Mansur, S. (2023). Credit scoring model for tenants using logistic regression. In *Springer proceedings in physics* (pp. 213–227). https://doi.org/10.1007/978-981-99-2850-7_19
[6] Njuguna, R. G., & Sowon, K. (2021). Poster: *A Scoping Review of Alternative Credit Scoring Literature. In Acm sigcas conference on computing and sustainable societies* (pp. 437–444). https://doi.org/10.1145/3460112.3471972
[7] Ebekozien, A., Abdul-Aziz, A.-R., & Jaafar, M. (2019). Housing finance inaccessibility for low-income earners in Malaysia: Factors and solutions. *Habitat International, 87*, 27-35. https://doi.org/10.1016/j.habitatint.2019.03.009
[8] Albanese, J. (2021). Big Data & Big Errors. Washington: Student Borrower Protection Center.
[9] Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M. K. (2021). Enhancing credit scoring with alternative data. *Expert Systems With Applications, 163*, 113766. https://doi.org/10.1016/j.eswa.2020.113766
[10] Ling, K. S., Jamaian, S. S., Mansur, S., & Liew, A. K. H. (2023). Modelling tenant's credit scoring using logistic regression, *SAGE Open*, *13*(3). https://doi.org/10.1177/21582440231189693
[11] Bolton, C. (2010). Logistic regression and its application in credit scoring (Unpublished doctoral dissertation). University of Pretoria.
[12] Duffy, D. M., & Santner, T. J. (1989). On the small sample properties of norm-restricted maximum likelihood estimators for logistic regression models. Communications in Statistics, *18(3)*, 959–980. https://doi.org/10.1080/03610928908829944