

Leveraging Machine Learning to Uncover Key Factors Influencing Satisfaction Among Free Tertiary Education Recipients in the Philippines

John Raymund B. Baragas*, Lea D. Austero*, Jennifer L. Llovido*, Lany L. Maceda*, Mideth B. Abisado**

*Computer Science and Information Technology, College of Science, Bicol University, Albay, Philippines

** College of Computing and Information Technologies, National University, Manila, Philippines

jrbbarajas@bicol-u.edu.ph, ldaustero@bicol-u.edu.ph, jllovido@bicol-u.edu.ph, llmaceda@bicol-u.edu.ph, mbabisado@national-u.edu-ph

Abstract— In spite of the broad implementation of the Universal Access to Quality Tertiary Education Act (UAQTE) – a groundbreaking legislation benefitting over 2 million students since its enactment in 2017 – a comprehensive evaluation of its outcomes has been notably absent. To bridge this gap, an extensive survey was undertaken among graduating tertiary students in selected regions of the Philippines. This strategically designed survey aimed to pinpoint overlooked aspects of UAQTE and capture firsthand insights from its recipients. The methodology employed to create this survey included focus-group discussions with various stakeholders (i.e., students, parents, faculty) and a pilot test reflecting the target demographic. To facilitate analysis of the results of 1462 responses, five regression machine learning algorithms were then employed to analyze questionnaire data. The decision tree regressor with a root-mean-squared-error of 0.6881 was found to be the best performing model describing the collected questionnaire data. Shapley explanations of the best performing model highlighted the desire of the recipient to pursue international employment as the top predictor of satisfaction in UAQTE among its recipients. Furthermore, insights from employed topic modeling among the open-ended questions in the deployed survey suggested potential inadequacy of UAQTE subsidies, specifically to recipients whose pursued degrees are in the science, technology, engineering, and mathematics courses. This substantial finding promises valuable insights into the effectiveness of the legislation and may inform future policy adjustments to better address the diverse needs of tertiary education in the Philippines. Overall, this research provides a robust framework for assessing the impact of UAQTE and showcases a methodologically sound approach in integrating machine learning and qualitative analysis.

Keywords— Machine Learning, Tertiary Education, Satisfaction, Free Education, Philippines

I. INTRODUCTION

In the dynamic tapestry of Philippine legislation, the Universal Access to Quality Tertiary Education Act (UAQTE), signed into law in 2017, stands as an enduring monument to the commitment of the Philippine government to foster an inclusive and empowered society. This groundbreaking legislation has not only redefined the educational landscape in

the country but it also emerged as a beacon of hope which extended the promise of higher learning to at least 2 million eager minds. With this achievement, the UAQTE then represented a pivotal chapter in the ongoing narrative of Philippine education and this signaled a profound shift towards democratizing access to tertiary education especially in a third-world country like the Philippines.

With the goal of enshrining the principle that education is a fundamental right, this newly implemented policy indeed dismantled barriers that once hindered the pursuit of knowledge for many aspiring students. By championing universal access to tertiary education, the legislation has become a catalyst for social mobility that has already started a future where education is not a privilege but essentially a birthright.

However, amidst its widespread implementation, a comprehensive evaluation of the outcomes and impacts of UAQTE has remained conspicuously absent. This void in understanding prompted the initiation of an extensive survey among but not limited to graduating tertiary students in selected regions of the Philippines, with the overarching goal of shedding light on the nuanced facets of UAQTE positioned to capture firsthand insights from its beneficiaries.

In an attempt to achieve this primary goal, this work employed a methodology that was both strategic and inclusive. Through focus-group discussions involving various stakeholders, including students, parents, and faculty, the deployed survey of this study was meticulously designed to identify potential oversights in the UAQTE framework. A subsequent pilot test, mirroring the demographic profile of the target respondents, ensured the relevance and efficacy of the survey instrument. To rigorously analyze the wealth of data gathered from the 1462 collected survey responses, five regression machine learning algorithms were employed. Notably, the decision tree regressor emerged as the best-performing model and achieved the lowest root-mean-squared-error across all tested models. Shapley explanations derived from this model spotlighted as well that the baccalaureate degree was the paramount predictor of satisfaction among UAQTE recipients.

Beyond quantitative analysis, this work also delved into qualitative insights through topic modeling of open-ended survey questions. This exploration unveiled potential inadequacies in UAQTE subsidies, particularly for students pursuing degrees in science, technology, engineering, and mathematics (STEM) fields. It could also be inferred from the results of this work that the significance of these findings extended beyond the academic realm. As a result, this findings from this work also offered invaluable insights into the effectiveness of UAQTE as a policy, potentially guiding future policy adjustments to better align with the diverse needs of tertiary education in the Philippines. In essence, this study not only fill a critical void in the evaluation of UAQTE but also presented a robust framework that could potentially integrate quantitative and qualitative analysis as a precedent for future assessments of educational policies.

II. METHODOLOGY

This paper employed explainable machine learning to pinpoint the factors that best influence the satisfaction of the recipients of UAQTE. To add more context to these findings, topic modelling, which was used in earlier studies [1]-[3], was also utilized to understand the gaps that UAQTE may have. To achieve these objectives, the following regression models were employed: (1) k-Nearest Neighbors (kNN), (2) linear (LR), (3) decision tree (DT), (4) random forest (RF), and (5) gradient boosting machines (GBM). Explanations on how the best performing regression model predicts satisfaction of respondents with UAQTE was facilitated through the use of shap values as described in [4]-[7]. Moreover, topic modelling was done through the use of the open-source package BERTopic [7].

A. Development of survey instrument

A comprehensive consultation process involved approximately 10 pairs of parents, 8 faculty members, 40 tertiary students, and representatives from local universities and colleges (LUCs) and private schools. This inclusive approach comprised three rounds of focus-group discussions, ensuring diverse perspectives from key stakeholders, including implementers and school administrators, to enrich the research findings. The goal was to identify pertinent questions for evaluating the impact and effectiveness of the UAQTE implementation. Following the conclusion of these focused dialogues, a total of 30 questions were thoughtfully crafted based on the rich engagement with the stakeholders. Of these 24 questions, half of these were dedicated to identifying the demographics of the target respondents while the remaining questions were strategically formulated to delve into the assessment of the impact and efficacy of UAQTE on its recipients. In addition, within this framework, a Likert scale question was ultimately incorporated to prompt respondents to rate their satisfaction on a scale from 1 to 5 (1 being the lowest and 5 the highest). This addition aimed to provide a nuanced understanding of the sentiment among recipients regarding their satisfaction with the implementation of UAQTE.

B. Pilot-testing of developed survey

Before its full deployment, the developed survey underwent a pilot-testing phase to ensure its validity in capturing the sentiments of the intended respondents regarding the implementation of UAQTE. For this pilot test, 35 volunteer graduating tertiary students, all beneficiaries of UAQTE, were enlisted to evaluate the effectiveness of the developed survey. Feedback from the volunteers primarily centered on enhancing the clarity of the survey questions. Additionally, participants in the pilot test recommended the inclusion of open-ended questions aimed at eliciting suggestions on how to broaden the scope of UAQTE to better support struggling students. Following a process of iterative feedback collection from the pilot-test participants and subsequent revisions to the survey, the survey deployment proceeded once the final version met the necessary recommendations and was deemed satisfactory. The survey was conducted through the use of BOSESKO, a digital citizen participatory toolkit.

C. Model Training and Selection Phase

Five machine learning regressions models were used to develop a model that would extract patterns in the survey data (data frame with 1462 records containing 13 features, inclusive of the satisfaction ratings given by the respondents) that would be useful in assessing the impact and effectiveness of the implementation of UAQTE. Particularly, the kNN, LR, DT, RF, and GBM regressors, as detailed and implemented in [8]-[9] at default parameters, were adopted in this work. For the creation and training of these models, 75% of the survey data was designated as the training and validation set while the remaining 25% utilized were used as the holdout set. Lastly, the evaluation of model performance was based on the calculation of root-mean-squared errors for each model.

D. Model Interpretation Phase

To extract insights from the best performing regression model, this study used Shapley values. These values were instrumental in comprehending the influence of individual features within the trained model on the satisfaction levels of UAQTE recipients following the implementation framework as outlined in [4]-[7].

III. RESULT AND DISCUSSIONS

A. Development of survey instrument

This study focused on three key regions in the Philippines—specifically, CAR, NCR, and Region V—that are recognized to have the highest concentration of UAQTE recipients. Regrettably, owing to political unrest in the Visayas and Mindanao regions due to the barangay elections, this study narrowed its geographical scope to the Luzon archipelago alone to ensure the safety of the survey deployment and collection team. The demographic profile of the respondents is summarized in Table I. Overall, a total of 1462 respondents were surveyed, with 60% of the respondents being women. Notably, the predominant baccalaureate degree among respondents was identified as Bachelor of Science in

Nursing which falls under the science, technology, engineering, and mathematics (STEM) specialization.

TABLE 1. DEMOGRAPHICS OF RESPONDENTS

Demographics	Frequency
Average Age (years)	21.8 ± 1.17
Gender	
Man	408
Woman	897
LGBTQIA+	79
Prefer not to say	78
Course*	
BS Secondary Education	118
BS Nursing	112
BS Entrepreneurship	64
BS Information Technology	58
BS Accountancy	56
BS Agribusiness	54
BS Psychology	50
BS Business Administration (Management)	46
BS Business Administration (Financial)	45
BS Biology	42

*Due to limited space, only the top 10 courses are shown.

Table 2 presents the root-mean-squared-error of the trained regression models. Comparison of the training, validation, and holdout scores revealed that the decision tree regressor at a max depth of 4 gave the lowest holdout RMSE at 0.6881. Remarkably, the decision tree regressor consistently outperformed all other models including robust alternatives like the random forest and gradient boosting machines regressors across all scoring metrics. These compelling findings unequivocally establish the decision tree regressor as the optimal performer on the dataset. With these results, the decision tree regressor was chosen as the best performing model to be interpreted using Shapley values.

TABLE 2. PERFORMANCE OF TRAINED MODELS

Model	Training RMSE	Validation RMSE	Holdout RMSE
k-Nearest Neighbors	0.9398	0.9112	0.9447
Linear	0.9374	0.9304	0.9555
Random Forest	0.9644	0.9321	0.9332
Decision Tree	0.7230	0.7123	0.6881
Gradient Boosting Machine	0.9714	0.9455	0.9543

In the context of machine learning models, lower RMSE values indicate better predictive performance. Based on the provided values, the Decision Tree model appears to be the most effective, as it consistently achieves the lowest RMSE across all three phases (training, validation, and holdout), making it the optimal performer in this study.

B. Interpretation of Best Performing Model

The visual representation of the impact of various features on the prediction of the best-performing model towards respondent satisfaction with UAQTE is shown in Figure 1 and Figure 2 (through bar and beeswarm plots). Specifically, it was found that the primary determinant of satisfaction with UAQTE implementation was the inclination of the UAQTE recipient towards engaging in international employment post-graduation. This discernible trend in the results indicated that the best performing model excels in predicting a satisfied recipient if there is a lack of desire to pursue international employment.

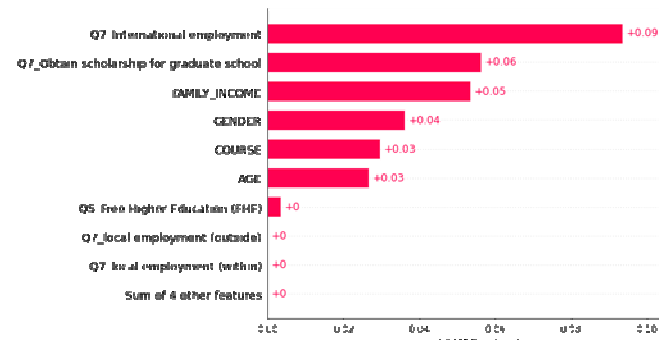


Figure 1. A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

Conversely, dissatisfaction is predicted with higher accuracy when the respondent expresses a desire for overseas work. While this work acknowledges the need for further validation to establish this prospect causality, one plausible speculation is that UAQTE recipients may perceive the educational subsidy provided by the Philippine government as insufficient to confer a competitive advantage in securing employment opportunities abroad. For instance, the popularity of the BS Nursing course in the Philippines, known for its high employment prospects overseas, highly underscores this point.

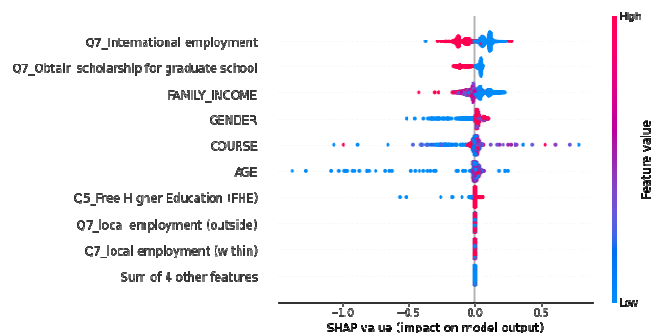


Figure 2. A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

Given that UAQTE does not cover overhead costs associated with employment abroad, students aspiring to work internationally may seek additional subsidies to sustain themselves during the waiting period for overseas employment and to cover costs for processing their work visas.

This perceived situation suggests a nuanced interplay between the perceived adequacy of the subsidy and the specific career aspirations of UAQTE recipients especially to those with ambitions for international employment. Further exploration and validation of these nuanced dynamics could deepen our understanding of recipient satisfaction within the context of UAQTE implementation.

C. Results of Topic Modelling

The thematic analysis of the Universal Access to Quality Tertiary Education Act, depicted in Figure 3, highlighted four key themes from student feedback. In conjunction with Figures 1 and 2 which leveraged Shapley values for predictive modeling, it is evident that post-graduation employment preferences are pivotal to student satisfaction. Delving into these themes, “Financial Transparency and Concerns” (Topic 0) addresses the balance of gratitude for fee elimination and worry over hidden costs. “Infrastructure and Resource Adequacy” (Topic 1) reflects anxieties about the physical quality of educational facilities. “Emotive and Temporal Reactions” (Topic 2) capture the immediate feelings of students towards the implementation of the legislation, and “Beneficiary Experience and Challenges” (Topic 3) speaks to the personal hurdles students face despite being primary beneficiaries. Each theme identified contributed to a comprehensive understanding that while the Act is beneficial, it also surfaced areas for policy enhancement to align with preferred career trajectories of students with a noticeable emphasis on global employment aspirations.

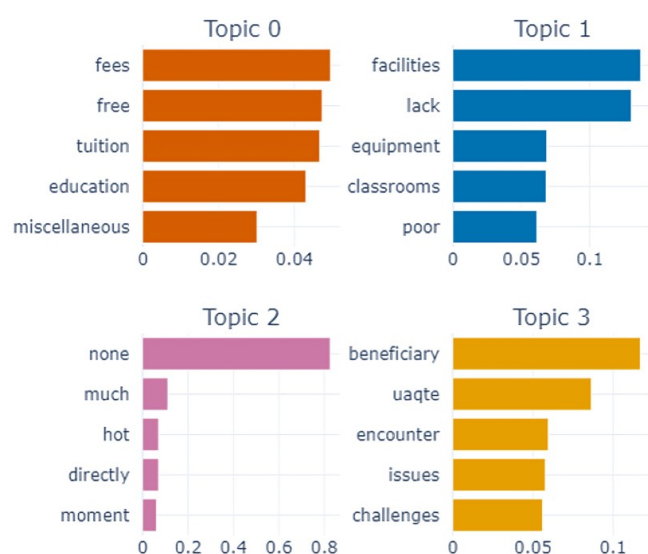


Figure 3. Topic Modelling Scores of the Top 4 Topics Observed.

D. Practical Implications of Findings

The revelation of this study that the inclination towards international employment significantly influences satisfaction of UAQTE recipients has noteworthy practical implications. Educational institutions and policymakers could use these insights to tailor support mechanisms, allocate resources more efficiently, and enhance career guidance services. For example,

consider a scenario where a university that is based on the insights from this study, identifies a high proportion of engineering students within their UAQTE program expressing a desire for international employment post-graduation. Recognizing this trend, the university could tailor its career services to provide specialized guidance workshops for these engineering students that would offer (1) information on job markets abroad, (2) assistance in preparing for international job interviews, and (3) guidance on navigating the complexities of work visas. Simultaneously, the university might also allocate additional resources to create targeted networking events with global employers or industry professionals who can share insights about international career paths in this aspect. This approach not only addresses the specific needs of the students but also optimizes resource allocation to enhance their global employability and overall satisfaction within the UAQTE program. Policymakers may also consider adjustments to the UAQTE program to better align with the career goals of students aspiring to work abroad – specifically, to address the perceived inadequacy of educational subsidies for global competitiveness. Furthermore, the findings of this work highlight the importance of continuous research to ensure adaptive and responsive educational policies that support students in their diverse career aspirations which would ultimately contribute to their satisfaction and success in both local and international professional contexts.

IV. CONCLUSIONS

While the UAQTE Act is recognized as beneficial, this study underscores the need for policy enhancements to better align with students' preferred career trajectories, especially those aspiring to work globally. The findings emphasize the significance of understanding and addressing nuanced dynamics related to perceived subsidy adequacy and career aspirations to enhance recipient satisfaction within the UAQTE program. This research provides actionable insights for educational institutions and policymakers to refine support mechanisms and optimize resource allocation, fostering overall satisfaction and success for UAQTE recipients in both local and international professional contexts.

V. FUTURE WORK

A crucial avenue for future work would involve further validation of the identified correlation between satisfaction and the desire for international employment through longitudinal studies or qualitative assessments. Additionally, exploring the specific challenges faced by UAQTE recipients aspiring to work abroad, such as financial constraints or bureaucratic hurdles, would also provide a more nuanced understanding of their needs of the UAQTE recipient. Further investigations could also delve into the effectiveness of potential UAQTE policy adjustments to specifically examine whether tailored support mechanisms and enhanced career guidance would have a measurable impact on the satisfaction and success of students who have global career aspirations. As the educational landscape continues to evolve, continuous

research efforts would be essential to ensure that policies such as UAQTE remain responsive to the dynamic needs of students to foster an environment that optimally supports diverse career goals within their respective specializations.

ACKNOWLEDGMENT

The authors gratefully acknowledge the generous support provided by Commission on Higher Education (CHED) through the LAKAS program. The authors would also like to extend their gratitude to Bicol University for the administrative support given during the conduct of this study. Moreover, the opinions expressed in this manuscript are solely those of the authors and do not reflect the position of CHED, Bicol University, and their respective offices.

REFERENCES

- [1] Vayansky, Ike, and Sathish AP Kumar. "A review of topic modeling methods." *Information Systems* 94 (2020): 101582.
- [2] Grootendorst, Maarten. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv preprint arXiv:2203.05794* (2022).
- [3] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." *arXiv*, 2017. doi: 10.48550/ARXIV.1705.07874.
- [4] L. D. Austero, C. Y. Sy and M. J. P. Canon, "Discovering Themes from Online News Articles on the 2018 Mt. Mayon Eruption," 2018 International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 2018, pp. 242-245, doi: 10.1109/IS3C.2018.00068.
- [5] S. M. Lundberg et al., "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1. Springer Science and Business Media LLC, pp. 56–67, Jan. 17, 2020. doi: 10.1038/s42256-019-0138-9.
- [6] R. Mitchell, E. Frank, and G. Holmes, "GPUtreeShap: Massively Parallel Exact Calculation of SHAP Scores for Tree Ensembles." *arXiv*, 2020. doi: 10.48550/ARXIV.2010.13972.
- [7] S. M. Lundberg et al., "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," *Nature Biomedical Engineering*, vol. 2, no. 10. Springer Science and Business Media LLC, pp. 749–760, Oct. 10, 2018. doi: 10.1038/s41551-018-0304-0.
- [8] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011. [Online]. Available: <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>.
- [9] P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, no. 3. Springer Science and Business Media LLC, pp. 261–272, Feb. 03, 2020. doi: 10.1038/s41592-019-0686-2.
- [10] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv*, 2022. doi: 10.48550/ARXIV.2203.05794.



wastewater treatment research.

John Raymund B. Barajas has earned a Master of Science in Chemical Engineering from De La Salle University - Manila, as well as a Master of Science in Data Science from the Asian Institute of Management. His expertise spans a range of interdisciplinary fields, including data science, data analytics, and natural language processing. Furthermore, he possesses specialized knowledge in material science, with a particular focus on water and



2018 Mount Mayon Eruption; and Solving course timetabling problem using Whale Optimization Algorithm.

Lea D. Austero is an ASSISTANT PROFESSOR at Bicol University, Legazpi City, Albay, Philippines. She received her Doctor in Information Technology at the Technological Institute of the Philippines Quezon City just last May 2023. Her published works include Determining resource capacity in disaster assistance using a model-driven decision support system;

Discovering themes from internet news articles on the 2018 Mount Mayon Eruption; and Solving course timetabling problem using Whale Optimization Algorithm.



Dr. Jennifer L. Llovido is a faculty member of the Computer Science and Information Technology Department at Bicol University College of Science, Legazpi City, Philippines, with an academic rank of Associate Professor V. She completed her Doctor in Information Technology (DIT) at the University of the Cordilleras, Baguio City, Philippines. Her published research works are centered on the fields of natural language processing, data mining, and system design and development. She can be reached at jllovido@bicol-u.edu.ph.



Lany L. Maceda earned her Doctorate in Information Technology from University of the Cordilleras, Baguio City, Philippines, in 2020. She is a faculty member of the Department of Computer Science and Information Technology, holding an academic rank of Associate Professor V at Bicol University. Moreover, she also serves as the Director of the Research, Development and Management Division at the same institution. She has been actively promoting data-driven policy-making through her research papers published in reputable international journals and conferences with research interests on machine learning particularly on natural language processing and data mining. She can be reached at llmaceda@bicol-u.edu.ph.



Mideth B. Abisado is an Associate Member of the National Research Council of the Philippines and a Board Member of the Computing Society of the Philippines Special Interest Group for Women in Computing. She is the Director of the CCIT Graduate Programs. She completed her Doctor in Information Technology (DIT) at the Technological Institute of the Philippines. Her research focuses on Emphatic Computing, Social Computing, Human-Computer Interaction, and Human Language Technology. She can be reached at mbabisado@national-u.edu.ph.