

Cracking the Code: Using Decision Tree to Forecast 100% Chance of Teacher Education Licensure Exam Success

*Diwag, Glen S.
College of Teacher Education
University of Eastern Pangasinan*

ABSTRACT

Forecasting the chances of licensure examination success by analyzing big data and various factors including academic performance, comprehensive examination scores, and pre-licensure assessments is an important area of research in a Teacher Education Institution (TEI). A TEI can use these insights to enhance the curriculum and prepare graduates for licensure examination success. Using around 800 data entries from 2014 to 2023, this study aims to identify the primary predictor of passing and failing the Licensure Examination for Professional Teachers, to determine the value of an outcome variable that predicts the one-hundred percent chance of passing and failing the Licensure Examination for Professional Teachers (LEPT), and to determine the target score that should be achieved to predict the chances of one-hundred percent passing the LEPT. Using Chi-squared automatic interaction detection (CHAID) as the growing method of classification tree analysis, the result shows that the Majorship rating is the primary predictor of the result of the LEPT. The predictor for a one-hundred percent Failed LEPT result and the predictor for a one-hundred percent chance of passing LEPT were both divided into three cases. The correctness of the model to predict a one-hundred percent chance of passing the LEPT is 94.3%. The conduct of competency appraisal is suggested, a score in the majorship/specialization of more than eighty-two percent of the one-hundred-fifty items, which is at least one hundred twenty-four points over one-hundred fifty is a predictor of a one-hundred percent chance of passing the LEPT.

Keywords: CHAID, Classification Tree Analysis, LEPT Predictor, Forecasting, Licensure Exam

INTRODUCTION

The quality of learning is the top priority for Teacher Education Institutions (TEIs) in addressing Sustainable Development Goal (SDG) No. 4, which aims to increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries. According to the Commission on Higher Education (CHED) Memorandum Order No. 80, series of 2017, TEIs are required to implement selective policies for Teacher Education Programs to ensure high-quality instruction for pre-service teachers.

The University of Eastern Pangasinan (UEP), a TEI, was established following a proposal by Dr. Ramon V. Guico III, the former Vice Mayor of Binalonan and the incumbent Governor of Pangasinan. This proposal was presented to former President Gloria Macapagal Arroyo, former Speaker of the House of Representatives Jose de Venecia, former Mayor Ramon N. Guico Jr., and local officials. This initiative culminated

in the approval of Ordinance No. 2005-007 on December 9, 2005, which established the University of Eastern Pangasinan. The College of Teacher Education (CTE) began its journey in 2009 with just two students enrolled in the Bachelor of Elementary Education (BEd) program, followed by the introduction of the Bachelor of Secondary Education (BSEd) in 2011. Today, the college has grown to over 2,400 students, while the entire university serves approximately 15,000 students, many of whom come from marginalized families lacking the resources to attend private institutions or larger universities outside Pangasinan.

In line with this mission, Senate Bill 2495, known as the "Bawat Pamilya May College Gradweyt Act," aims to ensure that at least one member of every eligible Filipino family receives assistance to pursue and complete a college education. This aligns with the university's commitment to providing opportunities for all students to enhance their qualifications as job applicants and secure better careers with higher earning potential. At the College of Teacher Education, it is essential to elevate the qualifications of pre-service teachers to meet the required standards and competencies. One effective strategy for achieving this is the implementation of review classes and mock board examinations in general education, professional education, and specialized fields, all designed to prepare students for the Licensure Examination for Professional Teachers (LEPT), a critical milestone for aspiring educators.

Predicting the performance of graduates in the licensure examination is a crucial area of research. UEP as a TEI can use these insights to enhance the curriculum and better prepare graduates for licensure examination success. Few studies have been conducted on predicting and forecasting the success of the licensure examination for teachers. Significant findings show that there is a correlation between academic performance/ pre-licensure examination and board examination (Amanonce et al 2020; Balinario et. al. 2023, Valle and Brobo 2022). The study suggests that high academic performance can result in high chances of success in the licensure examination. On the other hand, Fiscal and Roman (2022) made a linear model for the prediction of performance in LEPT of Bachelor of Elementary Education (BEd) and Bachelor of Secondary Education (BSEd) graduates based on the performance in the pre-LEPT. Using regression analysis, it was found that for general education and professional education components of licensure examination for BEd, the pre-LET scores of education graduates are significant predictors. Similarly, the LET performance of BSEd graduates can be predicted using their general education, professional education, and field of specialization components. The linear model can now be utilized in predicting the LEPT ratings of BEd or BSEd graduates using their pre-LET scores. Another study conducted by Delos Angeles (2020) determined predictors of the teacher licensure examination performance through regression analysis. The study showed that the strongest predictors of LET performance identified are professors' mastery of the subject matter and the quality of the content of the general education courses. This study presents how important qualifications of the teaching force and its general education curriculum. The level of accreditation of a school or the quality of a university is one big factor in the success of the licensure examination. As stated by Maramag et. Al (2020), relative to school profile, graduates of TEI with higher accreditation status

have significantly better LET performance than those enrolled in TEI with lower accreditation status. Most recently, Velasco et al (2023) made a study about predicting licensure examination performance using data mining techniques. The study compared 4 data mining techniques based on their accuracy rates namely J48, Random Forest, Logistic Regression, and Naïve Bayes. The study revealed that J48 performed best in terms of accuracy percentage. In a follow-up research, Velasco and Victoriano (2024), in their study of predicting board performances using classification algorithm and time series analysis, concluded that decision tree algorithms excel in predicting student success, particularly in identifying key attributes.

In general, most of these studies conducted associations or correlations of two or more variables to forecast or predict performances in the licensure examination for teachers. Pearson Correlation Coefficient, Regression Analysis, and Logistic regression are the common statistical tools utilized in treating two or more associated variables. Data mining tools were also compared to check which algorithms perform better with accuracy and objectivity. The present study would like to address the analysis of data coming only from the Professional Regulation Commission (PRC), with around 800 data entries of LEPT from 2014 – 2022 of the University of Eastern Pangasinan. The research gap shows that no specific model was generated to address a specific concern of an institution. Chi-squared automatic interaction (CHAID) as the growing method of classification tree analysis was utilized in this study. According to Yang (2023), CHAID can analyze data very well and reliably detect significantly correlated factors. To better understand predictions of the result of the licensure examination for teachers, the researcher focused on the following objectives:

1. To determine the predicting variables of the LEPT result.
2. To determine the value of an outcome variable that predicts the chance of 100% failing the LEPT.
3. To determine the value of an outcome variable that predicts the chance of 100% passing the LEPT.
4. To determine the target score that should be achieved to predict the chance of 100% passing the LEPT
5. To identify the overall correctness of the model.

Scope and Delimitation

This study aims to identify and analyze the predictors of success in the Licensure Examination for Professional Teachers (LEPT) among teacher education graduates in Pangasinan, specifically targeting alumni from the University of Eastern Pangasinan. The research examines data from the Professional Regulation Commission (PRC) covering the years 2014 to 2023. Key variables of interest include the Majorship rating, Professional Education rating, General Education rating, year of graduation, type of examinee (first-time or retaker), number of retakes, the year the licensure examination was taken, the duration in months before taking the exam, institutional passing rate, and national passing rate. A quantitative research design was employed, utilizing archival

data analysis to explore the relationships between these variables and licensure exam success.

To refine the focus of the research, several delimitations have been established. The study is limited to graduates of the University of Eastern Pangasinan who participated in the licensure examination for professional teachers during the specified period from 2014 to 2022. Additionally, the analysis is confined to the variables extracted from the PRC documentation, excluding any other profiles not specified in the provided data.

Literature Review

The licensure examination for professional teachers (LEPT) is a critical assessment that evaluates the readiness of graduates from tertiary education institutions to enter the teaching profession. Identifying the predictors of success in this examination is essential for enhancing teacher education programs and improving educational outcomes. This literature review synthesizes research findings on various factors that influence the success of graduates in the LEPT.

A significant body of research highlights the role of academic achievement as a predictor of success in licensure examinations. Studies consistently demonstrate that Grade Point Average (GPA) during teacher education programs correlates positively with licensure exam performance. For instance, a study by Langcao, Toquero and Tusoy (2023) found that candidates with higher GPAs were more likely to pass the licensure exam on their first attempt. Similarly, Amanonce and Maramag (2020) and Apare and Arcilla Jr (2018) reported that the academic achievement of teacher education graduates had a positive and significant relationship to their performance in the LET.

Balinario, Ofqueria, and Arca (2024) identified key predictors influencing the performance of candidates in the Licensure Examination for Teachers (LET). Their research indicated that candidates often struggle with the specialization sections of the exam. Furthermore, graduates of Bachelor of Secondary Education (BSEd) tend to perform better in three specific areas of the licensure examination compared to their non-BSEd counterparts. The study also found that higher academic achievements and satisfactory admission scores correlate with improved performance in the Licensure Examination. These insights contribute to enhancing the educational quality of TEIs by updating practices such as seminars, review classes, and policies related to admissions and student retention.

Mulay and Potdar (2024) highlighted the potential of data mining algorithms to significantly enhance educational quality in Higher Educational Institutions by predicting student academic performance. Alturki et al. (2021) observed that Educational Data Mining (EDM) has emerged as a vibrant research domain, adept at extracting meaningful insights from educational databases for a variety of applications. Baek and

Doleck (2022) further analyzed the growing emphasis within EDM research on utilizing these techniques to devise innovative strategies for improving educational outcomes, particularly through the development of predictive models for evaluating student performance. Bilici and Özdemir (2021) noted an increasing interest in data mining studies focused on academic performance, concluding that these technologies primarily aim to predict student success, classify students for specific purposes, and assess their effects on both students and educators. Among the various data mining techniques, Song and Lu (2015) identified decision trees as particularly effective with the most commonly employed algorithms for developing decision trees being CART, C4.5, CHAID, and QUEST. Furthermore, Velasco and Victoriano (2024) discovered that the Decision Tree algorithm outperformed other methods in terms of classification efficiency, with Logistic Regression and Naïve Bayes showing relatively lower effectiveness in predicting board performance success. Additionally, Jijo and Abdulazeez (2021) reported that the decision tree algorithm achieved its highest accuracy of 99.93% when utilizing a machine learning repository as the dataset.

Liu and Lewis (2018) describe machine learning as a branch of artificial intelligence (AI) that enables systems to learn and enhance their performance based on experience, all without the need for explicit programming. According to Demirovic et al. (2021), decision tree learning is a popular method in machine learning, particularly valued for its ability to produce clear and interpretable models. The decision tree process involves constructing a classification model in a tree format. This model categorizes instances into groups or forecasts the values of a dependent (target) variable based on the values of independent (predictor) variables. As noted by Abraham Iorkaa et al. (2021), the fundamental concept is to divide the sample responses into new sub-samples that are as uniform as possible while being distinct from one another. This process of splitting continues until it creates a framework for decision-making. The entire dataset serves as the root node, with the resulting sub-samples referred to as nodes.

Jijo and Abdulazeez (2021) reported several types of decision tree algorithms namely: Iterative Dichotomies 3 (ID3), Successor of ID3 (C4.5), Classification And Regression Tree (CART), CHi-squared Automatic Interaction Detector (CHAID), Multivariate Adaptive Regression Splines (MARS), Generalized, Unbiased, Interaction Detection and Estimation (GUIDE), Conditional Inference Trees (CTREE), Classification Rule with Unbiased Interaction Selection and Estimation (CRUISE), and Quick, Unbiased and Efficient Statistical Tree (QUEST). Four of these algorithms namely CRT, CHAID, Exhaustive CHAID, and QUEST are available in IBM Statistical Package for Social Science (SPSS). The other algorithms can be simulated using other software such as Waikato Environment for Knowledge Analysis (WEKA), R (rpart and party packages), Scikit-learn, RapidMiner, MATLAB, SAS Enterprise Miner, and others.

Rahman and Muktedir (2021) identified the SPSS as a crucial tool for quantitative data analysis, allowing researchers to conduct complex analyses and create a variety of charts and graphs within the software. Its intuitive interface, along with online tutorials and extensive documentation provided by the parent company and user communities, has contributed to its reputation as one of the most accessible statistical analysis programs on the market.

METHODOLOGY

This study employed a quantitative research design utilizing data mining techniques and a decision tree algorithm to predict success in licensure examinations. The research focused on analyzing historical data from candidates who have previously taken the licensure examination to identify patterns and factors that contribute to success.

As pointed out by Matzavela and Alepis (2021) there are four stages in data mining and decision tree learning namely data collection, classification, predictive model, and evaluation. In the data collection, data on the 780 licensure examination results was collected from official documents provided by the Professional Regulation Commission (PRC). Available data includes all UEP graduates who took the LEPT from 2014 to 2023. In the classification, the dependent variable was the LEPT result, (passed/failed) treated as categorical data. The independent variables were classified as either categorical or continuous. Independent variables include Majorship rating, Professional Education rating, General Education rating, year of graduation, type of examinee (first-time or retaker), number of retakes, the year the licensure examination was taken, the number of months between graduation and date of examination, institutional passing rate, and national passing rate. In the creation of the predictive model, IBM SPSS's classification tree analysis was utilized. The growing method used is CHAID with a minimum parent node of 60 and minimum child node of 30. The maximum tree depth is 3. A fixed value of 10 for the intervals (bins) of all independent variables was chosen. For the evaluation, 10-fold cross-validation was utilized.