A data mining model for evaluation of instructors' performance in higher institutions of learning using machine learning algorithms

Aranuwa Felix Ola
School of Science and Engineering,
Malaysia University of Science and Technology,
Selangor, Malaysia.
aranuwa.felix@pg.must.edu.my

Prof. Sellapan Pallaniappan
School of Science and Engineering,
Malaysia University of Science and Technology,
Selangor, Malaysia
sell@must.edu.my

Abstract— **The development of an improved and intelligent model for evaluation of instructors' performance in higher institutions of learning especially in the developing countries can be well motivated from these points of view: the recent national policies on education mandating high stakes evaluation of instructors and learning system which are at the forefront of the education reforms agenda, the reasoning behind the degree of difference in the students' academic outcomes, and the quest for an optimal algorithm suitable for predicting instructors' performance. Worldwide national policies on higher education are giving increasing importance to improve the quality of education offered. Consequently, the evaluation of instructors' performance is especially relevant for the academic institutions as it helps to formulate efficient plans to guarantee quality of instructors and learning process. Effort in this work is directed at modelling an intelligent technique for evaluation of instructors' performance, propose an optimal algorithm and designing a system framework suitable for predicting instructors' performance and as well as recommend necessary action to be taken to aid school administrators in decision making considering the limitations of the classical methodologies. The proposed technique will overcome the limitations of the existing techniques; improve reliability and efficiency of instructors' performance evaluation system, provide basis for performance improvement that will optimize students' academic outcomes and improve standard of education. Consequently, it will contribute to successful achievement of the goals and objectives defined in the vision and mission of the new education reform agenda.**

Keywords- improved model, machine learning, national policy, optimal algorithm, performance evaluation

I. INTRODUCTION

Recent national policies on higher education mandating high stakes evaluation of instructors and the learning system coupled with the quest for an optimal algorithm for evaluation of instructors' performance in higher institutions of learning especially in the developing countries are primary motivation for this work. Additionally, the ever increasing analysis on the outcome being produced by colleges and universities in recent years, has not only been generating questions about the quality and efficiency of their teaching workforce and learning systems, but has continued to occupy a major place on the agenda of higher education leaders and teachers' performance evaluation is becoming a dominant theme in the school reform efforts [1], [2], [3], [4]. With this development, there is a no doubt educational institution both in developing and developed countries have an obligation to deliver moreso that university performance are often judged by the quality and reputation of the awards they provide and the product they produce [5]. According to Andrew, Bankole, and Olatunde (2010)[6], a large mismatch appears to exist between university output and labour market demand in recent times. Their findings show that the performance of recent graduates have clearly deteriorated, primarily because of the operational policies and inadequate level of skilled human resources, especially the quality of university trained work force. Deteriorating quality perception is supported by the results from empirical research of Chiemeke, Longe, and Shaib (2009) [7].

The Obama led administration made state support of rigorous teachers' evaluation systems a pre-condition for competition in "Race to the Top", and has laid out a blueprint for the reauthorization act in which teacher effectiveness defined by evaluation of on-the-job performance is an important facet [8]. Self-proclaimed education reformers such as Bill Gates, Davis Guggenheim and Michelle Rhee in their submission posited that teachers' evaluation should be at the forefront of the education reform agenda and that evaluation results be used as the basis for making decisions about hiring, disciplining, compensating, awarding tenure to and sanctioning ineffective teachers [9]. One of the reasons for this may not be farfetched from the fact that the strength of good education in any educational institution depends on the quality of the academic staff in that institution; and there is no satisfactory substitute for competent staff that possesses sound educational philosophy and dynamic leadership [10].

As the most significant resource in schools, teachers are vital to improve student outcomes and raise education standards [4]. From this perspective, teachers' performance evaluation is a vital step in the drive to improve the effectiveness of learning system and raise educational

standards. According to Denisi and Pritchard (2006) [11], a central reason for the employment of performance evaluation is performance improvement (initially at the level of the individual workforce, and ultimately at the level of the institution). Other fundamental reasons include basis for employment decisions (e.g. promotions, career advancement, performance reward, sanctions, etc). Additionally, performance evaluation can aid in the formulation of criteria and selection of individuals who are best suited to perform required organizational tasks [12]. It can be part of guiding and monitoring employee career development and improvement. Evaluation of teachers' practice and performance evaluation in higher institutions of learning is definitely not a new trend, but what is new is deep interest to enhance ways of evaluating teachers' performance considering the weakness of the classical methodologies and contentious issues on accuracy and dependability [13],[14]. Hence, the need for an improved model that will proffer solution to the limitations of the classical methods, improve the efficiency and accuracy of teachers' evaluation system, and consequently help in no small measure in the new move for educational reform efforts.

II. LITERATURE REVIEW

A. Performance Evaluation

Performance evaluation have been defined as a systematic process of evaluating an individual worker's job performance and effectiveness in relation to certain pre-established criteria and organizational objectives [14],[15]. According to Keifer (2013) [16], evaluation is the process of examining a subject and rating it based on its important features, while evaluation in education can be referred to as the systematic determination of merit, worth, and significance of a learning process by using some criteria against a set of standards or a systematic acquisition and assessment of information to provide useful feedback about some object [17], [18], [19]. Both definitions agree that evaluation is a systematic endeavour and the term 'object' or 'subject' here could be a program, policy, technology, person, need, or activity. However, the later definition emphasizes acquiring and assessing information because all evaluation work involves collecting and sifting through data, making judgments about the validity of the information and of inferences we derive from it.

According to Nakpodia (2011)[20], evaluation is an intervention strategy that has received significant attention in academic, business and political circles for information gathering process, ascertaining the decision to be made, selecting related information, collecting and analysing information in order to report summary data useful to decision makers in selecting among alternatives. Researchers at different levels have proposed and used wide-ranging approaches to evaluating instructors' performance. However, the efficiency and dependability of these classical methods have been controversial [8], [20], [21], [22], [23]. As a result there was no standard method or computerized solution for evaluating instructors' performance that capture the complex nature of the art and science of teaching and learning system in the tertiary institutions [24].

B. Performance Evaluation Methods

Two types of performance evaluation method identified and widely used in the literature are: formative and summative evaluations [4], [25], [26], [27], [28], [29], [30]. Meanwhile, early in the history of teachers' performance evaluation, educators were evaluated based on traits or characteristics which may or may not have been related to performance, and yet no significant body of knowledge confirms to the fact that effective teaching performance is dependent on specific traits. As a result, this form of evaluation was discarded [13], [31]. Formative evaluation refers to a qualitative evaluation on the instructors' teaching assessment, aimed at identifying strengths, weaknesses and providing adequate professional development opportunities for the areas in need of improvement.

Formative evaluation involves the use of classroom observations, student evaluation report etc, as tools to measure the performance and effectiveness of an instructor. The overall intention of this is to provide informative feedback to assist faculty in improving the effectiveness of their teaching performance [32]. While summative evaluation is described as an indispensable source of documentation and recognizable way to evaluate instructors' quality, providing summary statements of a instructor's capabilities through inspection, examination or interviews, in order to measure aptitude and knowledge to ensure that required standards are met, or to promote level of performance for immediate recognition [26]. It is used to determine the worth and career advancement of an instructor, assess that instructors are adopting the actions and best practices which improves student outcomes. Summative evaluation gives crucial information about the instructor being evaluated relatively to what is considered as standards. Hence, summative evaluation is an indispensable source of documentation to hold instructors accountable for their professionalism. In many institutions, it involves the use of annual performance evaluation report (APER) and interview as instruments to measure their quality, performance and effectiveness [20].

Although various research studies supported the usefulness of these tools at different levels; however, there are considerable debates about their dependability when used separately [8], [14], [20]. Guarini and Stacy (2012) [33] claimed any single measure of evaluation will only emphasize one important element at the expense of others. Traditionally, faculty have been sceptical about this and critics have claimed that formative method is too subjective (i.e biased and one-sided) [25], [34]. In the work of Nakpodia (2011)[20], the author comments on the shortcomings of summative method and concludes that formative method alone cannot provide all the relevant information required to evaluate the lecturers. Hence, a supplementary instrument should be used to obtain information dealing with such aspects of professional

development e.g academic advancement, research publications and participation in academic conferences, workshops and seminars. The author however, suggests that if APER is going to be used, ambiguous sections should be eliminated from the instruments. Steele et al., (2010) [35], opined that instructors' evaluation system should employ a diverse set of measures to capture the complex nature of the art and science of teaching and learning system, which is inherently a multidimensional construct.

Hence, the obvious need for an improved and intelligent instructors' performance evaluation system that includes a spread of verifiable and comparable instructors' performance that distinguish instructors' quality and effectiveness, which is the focus of this work. In this work, a data mining based model that are guided by instructors' evaluation principles using machine learning algorithms is presented, focuses at addressing the following: modeling an improved instructors' performance evaluation technique by integrating formative and summative evaluation methods, propose an optimal algorithm and design a framework of instructors' evaluation system suitable for predicting instructors' performance and as well as recommend necessary action to be taken in order to aid school administrators in decision making. Machine learning algorithms have been described a powerful tool concerns the construction and study of systems that can learn from data, and have been used in many applications with suitable results [36], [37].

III. Performance Evaluation Procedure and Metrics

For effective and improved performance evaluation system there is need for adequate procedures for a given objectives, such as the extent of the existence of national standards of education guidelines, or the culture of evaluation in a school. Followed by a development plan (i.e internal regulations and activity plan), at the institution level. According to OECD (2009)[38], proper evaluation of teachers' performance procedures requires the establishment of reference standards and criteria, choice of instruments and sources of information, discussed as follows:

A. Establishment of reference standards

A fair and reliable instructor evaluation model needs reference standards to evaluate relatively. The main reference standard for instructor or teachers' evaluation typically is their academic and professional qualifications (i.e level of education, experience, certification and licensure) [39]. The key element and fundamental precondition of these must be clearly and concisely stated to know what are expected from them at different levels. Instructors' profiles often express levels of performance appropriate to beginning teachers, experienced teachers, and those with higher responsibilities. It is important to note that professional profiles provide the common basis to organize the key elements of the teaching profession such as initial teacher education, teacher certification, teachers' ongoing professional development and career advancement.

B. Establishment of criteria

Another essential basis for good practice in evaluation is the existence of clear and measurable criteria which must be consistently applied by competent (trained and experienced) evaluators [40], [41]. This requires the development of explicit guidelines about what is expected from professional practice. UNESCO's analysis of the European and Latin American teacher evaluation system emphasizes the content knowledge, the pedagogical skills, the abilities to assess instructors and the professional responsibilities vis-à-vis the school and the students as key domains to evaluate teachers. However, one should note that the analysis does not mention the engagement in professional development as a common teaching standard in European systems, with a subsequent risk to undervalue the teacher's engagement and willingness to enhance his or her own practice. Nevertheless, England has recently implemented a framework for professional standards, close to Danielson's one, which includes professional development criteria for the five levels of teaching performance (the award of Qualified Teacher Status, teachers on the main scale, Post Threshold Teachers, Excellent Teachers, and Advanced Skills Teachers).

C. Instruments and Information Sources

Since the way of gathering evidence about a particular subject may influence the evaluation results, the choice of instruments is of chief importance in designing and implementing systems to evaluate instructors' performance. Gathering multiple sources of evidence about instructor meets the need for accuracy and fairness of the evaluation process, taking into account the complexity of what a competent teacher should know and be able to do. A range of instruments and information sources are typically used to evaluate teachers. Some of which are: classroom observation, student evaluation form, inspection and interview, student outcomes, questionnaires and survey, annual performance evaluation report to mention a few.

IV. Validity and Considerations in Measuring Instructors' Performance

Determining what type of instructor evaluation method is best for a given purpose includes taking account of the validity and reliability of the instrument or process being used. According to Millett, Stickler, Payne, and Dwyer, (2007)[42], validity is the most fundamental consideration in assuring the quality of any evaluation. Validity refers to the degree to which an interpretation of a test score, or in this case, a score from a measure of instructors' performance, is supported by evidence. For a measure of instructors' performance and effectiveness to be valid, evidence must support that the measure actually assesses the dimension of instructors' performance and effectiveness it claims to measure and not something else. In addition, evidence that the measure is valid for the purpose for which it will be used is essential. Instruments cannot be valid in and of themselves; an instrument or assessment must be validated for particular purposes [43]. For example, an observation-based score might be validated for professional development purposes but might not be validated for compensation purposes. Determining the

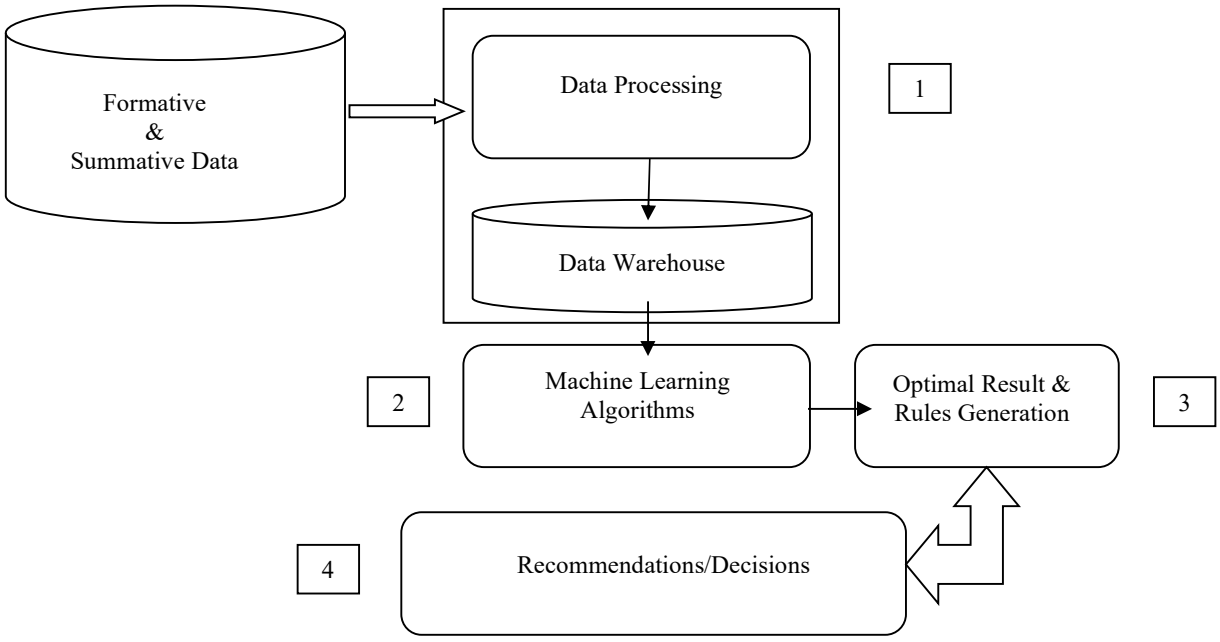

Figure 1: The Architecture of the Proposed Intelligent Evaluation System- (IES**)**

validity of an instrument requires taking account of the evidence regarding what the instrument measures and how the scores are being used. This requires the user of the instrument to be well-informed about these issues and to only make judgments about the degree to which there is sufficient evidence to use a particular instrument for the purpose under consideration. In addition to concerns about validity, there are other measurement concerns. Blanton et al. (2003)[44] identified six criteria that are particularly useful in informing this discussion which are elaborated in [45]. They are:

a) Comprehensiveness: This refers to the degree to which a measure captures all of the various aspects of instructors' performance evaluation. Less comprehensive measures might only capture important elements in one aspect at the expense of others.

b) Generality: refers to how well an instrument captures the full range of contexts in the teaching profession. An instrument is said to have a high level of generality if it measure important contents across contexts.

c) Utility: refers to how useful scores from an instrument are for a specific purpose. For example, scores from an instrument that ignores teaching context may not be useful in identifying contexts that appear to support more effective teaching. The experience of other researchers or a practitioner with an instrument makes it possible to better anticipate its potential uses and limitations.

d) Practicality: refers to the logistical issues associated with a measure. These include the developmental work required to adapt an existing model or measure for one's own purpose.

e) Reliability: refers to the degree to which an instrument measures something consistently. For example, it might be important to know whether scores on an instrument measuring instructor effectiveness vary by time of year, time of day, grade level, or subject matter.

f) Credibility: is a specific type of validity, face validity that is particularly important in measures of instructor performance and effectiveness. A measure is said to be credible if it is viewed as reasonable and appropriate by stakeholders from different groups (e.g., instructors, experts, and administrators). These aspects of measurement: validity, comprehensiveness, generality, utility, practicality, reliability, and credibility must be duly considered in any evaluation system.

V. COMPONENT ANALYSIS OF THE PROPOSED SYSTEM

Figure 1 presents the architecture of the proposed system aggregating four main components of the system model integrating instructors' formative and summative data: The first component is the data acquisition and storage, responsible for storing instructors' information to a data warehouse. The second component is the model building, responsible for obtaining knowledge about the instructors, through appropriates classification models. Specifically neural

networks and decision tree algorithms will be used in search for the best model with high predictive accuracy. The third component is for mapping the pattern in the rules generated with the instructor data to predict performance and the fourth component is the recommendation, responsible for recommending necessary action to be carried out on individual instructor based on the prediction from the evaluation system.

VI. CONCLUSION

The need for an improved and intelligent model for evaluation of instructors' performance in higher institutions of learning especially in the developing countries has become necessary in order to proffer solution to the limitations of the classical methodologies. Using factors and resources obtained from randomly selected stakeholders, a system framework for apt instructors' evaluation system is presented. The framework was designed with some basic components considered by the authors for reliability and efficiency. The proposed system, if fully implemented will aid school administrators in decision making, provide basis for instructors' performance improvement that will optimize students' academic outcomes and improve standard of education. Consequently, this will contribute to successful achievement of the goals and objectives defined in the vision and mission of the new education reform agenda.

REFERENCES

[1] L. Bacow, Performance Assessment Models and Tools from Other Sectors: Lessons for Higher Education *A Seminar Paper by Harvard Graduate School of Education Cambridge, MA 02138,* 2012. Retrieved from http://www.gse.harvard.edu/ppe/programs/highereducation/portfolio/performance-assessment.html

[2] F. Henard and S. Leprince , The Path to Quality Teaching in Higher Education *London School of Economics and Political Science Publication at the OECD* from October 2007 to June2008.

[3] E.S Taylor and J.H. Tyler , The effect of evaluation on teacher performance. *Publication of American Economic Review, 102(7), 3628-3651,* 2012. Available online at http://cepa.stanford.edu/content/effect-evaluation-teacherperformance#sthash. VFt1w51z.dpuf.

[4] J.M. William. Research-Based Options for Education Policymaking-"*Teacher Evaluation". National Education Policy Centre (NEPC)* 2012. Available online at: http://nepc.colorado.edu/publication/options.

[5] S.A. Jusoff , .A. Samah, and P. M. Isa, "Promoting university community's creative citizenry," *in Proceeding of WorldAcademy of Science, 2008, pp. 1-6.*

[6] D. Andrew, O. Bankole and A. Olatunde, Labor Market Prospects of University Graduates in Nigeria. *Publication of* Nigeria University System Innovation Project. November, 2010.

[7] S. Chiemeke, O.B. Longe,, F.A. Longe and I.O. Shaib, Research Outputs from Nigerian Tertiary Institutions: An Empirical Appraisal . *Journal Publication of the Library Philosophy and Practice, 2009.ISSN 1522-0222,*2009.Available online at: *www.webpages.uidaho.edu/~mbolin/chiemeke-longeshaif.htm* .

[8] S. Glazerman, D. Goldhaber, S. Loeb, S. Raudenbush, D. Staiger. and G.J.R. Whitehurst, Evaluating Teachers: *The Important Role of Value-Added.* The Brookings Brown Center Task Group on Teacher Quality Pg 2. November 17, 2010.

[9] M. McFarland-McDaniels , How to Design an Evaluation Instrument for a Teacher's Performance. A *Publication of E-How Education Group* 2012. Available online athttp://www.ehow.com/how_8100468_design-evaluationinstrument-teachers-performance.html#ixzz2XeNE0gs2.

[10] L.W. Anderson and L.A. Van Dyke , *Secondary School Administration*: Houghton Mifflin Company, Boston, U.S.A, 2000

[11] A. Denisi and R. Pritchard, Performance Appraisal, Performance Management, and improving individual performance: A motivational framework. *Management and Organization Review, 2(2),* 253-277, 2006.

[12] K. Manasa, and N. Reddy, Role of Training in Improving Performance. *The IUP Journal of Soft Skills, 3, 72-80,* 2009.

[13] K. Jenrbaun Evaluation of College Instructors' Teaching Competencies Towards Proposed Human Resource Development Program, 2012. Retrieved from http://www.studymode.com/essays/Evaluation-Of-College-Instructors-Teaching-Competencies-1293750.html.

[14] S. Mardikyan, B. Badur, Analyzing Teaching Performance of Instructors Using Data Mining Techniques. *Journal of Informatics in Education, 2011, Vol. 10, No. 2, 245–257 245.*

[15] J.Abu-Doleh and D.Weir, "*Dimensions of Performance Appraisal Systems in Jordanian Private and Public Organizations*" International Journal of Human Resource Management, 18(1), 75-84, 2007.

[16] K. Keifer. A Definition of Evaluation. Colorado State University. Copyright © 1993-2013. Available online at: http://writing.colostate.edu/about/copyright.cfm

[17] P. Bramley, and A. C. NewbyThe Evaluation of Training Part I: Clarifying The Concept. *Journal of European & Industrial Training*, 8,6, 10-16, 1984. Retrieved from: http://www.nwlink.com/~donclark/hrd/sat6.html. 17/6/13.

[18] M. Foxon, "Evaluation of training and development programs": A review of the literature. *Australian Journal of Educational Technology,* 5(2), 89-104, 1989. Retrieved from: http://www.nwlink.com/~donclark/hrd/sat6.html. 17/6/13.

[19] W.M.K. Trochim, Introduction to Evaluation. Web Centre for Social Research Method, 2006. Retrieved 30/6/2013 from http://www.socialresearchmethods.net/kb/intreval.htm.

[20] E.D. Nakpodia, "A Critique of the Methods of Evaluating the Competency of Lecturers in Nigerian Tertiary Institutions"- *African Journal of Education and Technology, volume 1 Number 1, April 2011; pp. 53-59.*

[21] W., McKeachie, Student Ratings; The validity of use. American Psychologist v52, No 11, 1218-1225, 1997.

[22] I. A. Archibong, and M.E. Nja, Towards Improved Teaching Effectiveness in Nigerian Public Universities: Instrument Design and Validation. *Journal of the Higher Education Studies, Canada. Vol. 1, No. 2; December 2011* URL:http://dx.doi.org/10.5539/hes.v1n2p78.

[23] R.L. Rothstein, H.F. Ladd, D. Ravitch, E.L. Baker P.E Barton L. Darling-Hammond, E. Haertel, R.L. Linn, R.J . Shavelson, and L.A. Shepard, Problems with the Use of Student Test Scores to Evaluate Teachers. *A Publication of Economic Policy Institute*, 2010. *Retrieved fromhttp://www.epi.org/publication/bp278/*

[24] O.K. Chaudhari P.G. Khot, and K.C. Deshmukh , Soft Computing Model for Academic Performance of Teachers Using Fuzzy Logic. *British Journal of Applied Science & Technology 2(2): 213-226, 2012*.

[25] K. Bain. What the Best College Teachers Do. Cambridge, MA: Harvard Press, 2004.

[26] M. Isoré, *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review, OECD Education Working Papers, No. 23, OECD Publishing, 2009. doi:10.1787/223283631428.*

[27] J.A. Kulik, Student ratings: validity, utility, and controversy. In: Theall M., Abrami, P.C. Mets, L.A. (Eds.), New Directions for Institutional Research. Jossey-Bass, 109, 9–24, 2001.

[28] H.W. Marsh and L.A Roche, The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American educational research journal, 30(1), 217-251.* 1997.

[29] J. Ory, Teaching Evaluation: past, Present, and Future. *New Directions for teaching and learning, No. 83, Fall ,2000.*

[30] M. Theall. And J. Franklin, Looking for bias in all the wrong places: A search of truth or a witch hunt in student ratings of instruction? *New Directions for Instructional research 109: 45-56,* 2001.

[31] J. Vollmer, Our Schools Need to Change, 1987**:** http://www.studymode.com/essays/Evaluation-Of-College-Instructors-Teaching-Competencies-1293750.html, *http://www.jamievollmer.com/about.html*.

[32] UNESCO, Evaluación del Desempeño y Carrera Professional Docente*: Una panorámica de América y Europa*, Oficina Regional de Educación para américa Latina y el Caribe, UNESCO Santiago, 2007.

[33] C. Guarini. and B. Stacy, Review of Gathering Feedback for Teachers. Boulder, CO: *National Education Policy Center,* 2012*. Retrieved from http://nepc.colorado.edu/files/pb-options-1-teval_1.pdf.*

[34] W. McKeachi, Student Ratings; The validity of use. *American Psychologist v52, No 11, 1218-1225,* 1997.

[35] J.L Steele, L.S. Hamilton, and B.M. Stecher, Incorporating Student Performance Measures into Teacher Evaluation Systems. *Palo Alto, CA: RAND. ISBN: 978-0-8330-5250-6,* 2010. *Retrieved from: http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR917*

[36] M. Mohri, A. Rostamizadeh, A. Talwalkar, "*Foundations of Machine Learning"*, The MIT Press ISBN 9780262018258, 2012

[37] W. Yang, B. Yourganov and M. Strother (2010). Machine Learning in Medical Imaging,*IEEE Signal ProcessingMagazine*, vol. 27, no. 4, July 2010, pp. 25-38

[38] OECD, "Teacher Evaluation: A Conceptual Framework and Examples of Country Practices". *A paper was presented at the OECD-Mexico Workshop Towards a Teacher Evaluation Framework in Mexico: International Practices, Criteria and Mechanisms, held in Mexico City on 1-2 December 2009.*

[39] L, Goe, C, Bell, and O. Little, "Approaches to Evaluating Teacher Effectiveness: A Research Synthesis June 2008. National Comprehensive Center for Teacher Quality", sponsored under government cooperative agreement number S283B050051.

[40] J. M. Drake, "Improving Teacher Performance through Evaluation and Supervision". Paper presented at the annual meeting of the National Association of Secondary School Principals, February 1984. ED 250 782

[41] Training and Development Agency (TDA) for Schools "Professional Standards for Teachers: Why Sit Still in Your Career?", TDA, United Kingdom, 2007.

[42] C. M. Millett, L. M, Stickler, D. G Payne and C. A Dwyer, A Culture of Evidence: Postsecondary Assessments and Learning Outcomes, 2007.

[43] M.T. Kane, "Current Concerns in Validity Theory". Journal of Educational Measurement Volume 38, Issue 4, pages 319–342, 2006.

[44] M.R. Blanton, "The Galaxy Luminosity Function and Luminosity Density". The Astrophysical Journal, 592:819-838, Aug, 2003.The American Astronomical Society. All rights reserved. Printed in U.S.A.

[45] J. G. Coggshall, "Determining Teacher Effectiveness -Center on Innovation and Communication Framework for Measuring Teacher Quality and Effectiveness": Bringing Coherence to the Conversation. Washington, DC, 2007.