

# Prediction of Teacher Enrollment for Pakistani Schools by Using SVM

Samina Kausar, Xu Huahu, Muhammad Shahid Iqbal, Xiangmeng Wang, Muhammad Yasir Shabir, and Tamoor Khan

**Abstract**—Civilization distinguishes human being from other creatures. Civilization is adorned with education. Education uplifts not only the standard but also the historical records of nations. Teachers, students, public, government and curriculum are the main components of education. This study analyzes the number of school teachers and predicts the future annual enrollment of teachers using the SVM model (support vector machine). Time series data of 46 years (1971-2017) have been taken from the Handbook of Statistics on Pakistan Economy. Our results depict that school teachers must be further enrolled at all levels. Our method leads to good precision.

**Index Terms**—Enrollment polices, future prediction, school teachers, and teacher enrollment.

## I. INTRODUCTION

Human being is a masterpiece of creature. As the needs and demands of human being began to increase more and more, it started to discover the solutions of its problems. The human population went on growing more rapidly than small civilizations came into being. These civilizations were adorned with different gifts. The man acquired many qualities and values which categorized him into different standards of civilization. The main gift of nature which was bestowed upon him was to know what was right and what was wrong. What were the things which are beneficial to him? Who were his foes and friends? To answer all these questions a boon and blessing of nature knocked at his door. This knocker was a great gift in the form of education. No doubt education was a beauty for him. It was more than a golden ornament for him. It directed him towards the right path in the darkest clouds of ignorance and wilderness. Education grew with the passage of time for the betterment of humanity. Education developed and conquered civilization of the world.

Manuscript received March 20, 2019; revised July 5, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61572434 and Grant 91630206, and in part by the Shanghai Science and Technology Committee under Grant 16DZ2293600.

Samina Kausar, Xu Huahu, and Xiangmeng Wang are with the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444, China (e-mail: saminamalik7@yahoo.com, huahuxu@shu.edu.cn, xiangmeng\_wang@163.com).

Muhammad Shahid Iqbal is with the School of Computer Science and Technology, Anhui University, Hefei, China (e-mail: nawabishahid@yahoo.com).

Muhammad Yasir Shabir is with the Department of Computer science and Information Technology, University of Kotli, AJ & K, Pakistan (e-mail: yasir.shabir14@gmail.com).

Tamoor Khan is with the School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, People's Republic of China (e-mail: tamoorkhan525.iiui@yahoo.com).

The study of Athens and Sparta which were old civilizations of the Greeks revealed that the secret of their beautification was the education. Undoubtedly teachers are super and supreme who guided the human beings. The focus point and the pivot around which the whole system of education revolved was the child. As a coordinator between the teacher and the tot (student) was the educational system but the logistics which were necessary for all the three afore-mentioned pillars were curricula. All these four pillars were dependable and correlator to each other. None of them could be denounced. All of them carry and bear history which was predated to be updated. In Pakistan education is divided into three tiers and all these tiers depend upon teachers. In order to standardize education, the quality of teachers is to be enhanced but in Pakistan quantity of the teachers is also enumerated along with quality. The analysis of 2005 and 2006 has shown that there was a great dropout percentage of the students due to many reasons and discrepancies. They may be on behalf of text books, teachers and department of education [1]. Pakistan has two types of education system, public and private, public institutions are under Pakistan's government. Public institutions in Pakistan cannot meet the quality of education due to the rapidly growing population with growth rate of 2.7% [2]. Hence, the government cannot fulfill the responsibility of providing education to the rapidly growing population [3]. In this study a statistical linear regression analysis and SVM model is applied in the education system to predict future enrollment of teachers in Pakistan. This method can be applied to all third world countries where educational institutions are less than their growing population. The third world countries desperately need new policies to increase enrollment of teachers for large population like Pakistan.

This research mainly focuses on predicting teacher enrollment at all levels, we use linear regression model and SVM to predict future teacher enrollment in Pakistan, time series data is analyzed over the time period of 46 years. As the data size is small so we use linear regression model and SVM accuracy having accuracy rate 61% and 73% respectively. The dataset contains three types of schools containing separate schools for boys, girls and coeducational schools. As there are separate male and female schools in Pakistan and both types of schools are less than the population so the growing population needs more male and female schools. A comprehensive sketch of the state of Pakistani schools enrolled teachers growing at all levels (primary schools, middle schools, high schools) is provided to contextualize the results and endorsements presented in this article and predicts how school teachers enrolled are increasing each year.

## II. METHOD

Most traditional initialization methods involve time-consuming and laborious manual intervention [4]-[10]. This work introduces an automatic framework to eliminate artificial participation and improve the prediction accuracy. Fig. 1 shows the overall flowchart of the proposed initialization method. The process consists of two major steps. The first step is to use static methods to divide the pre-process data into several stages. As the SVM algorithm allows annually enrolled teachers to have predictive relationships with multiple schools with varying degrees of membership based methods, therefore it is more reasonable in applications for educational institutions. Fig. 1 shows the yearly flow of teacher enrollment.

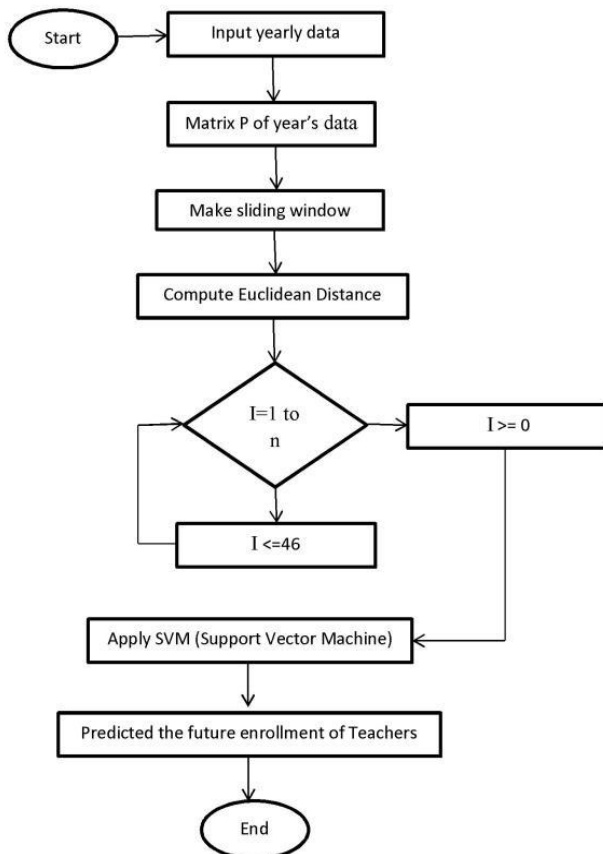


Fig. 1. Main steps of the algorithm for teachers enrolled per year in Pakistani schools.

### A. Data Collection

TABLE I: DESCRIPTION OF TEACHER ENROLLMENT PAKISTAN ECONOMY STATISTICS MANUAL

Types of Schools	Range of Teacher's Enrollment	Sum of Teacher's Enrollment	Range of Teacher's Enrollment	Sum of Predicted Teacher's Enrollment
Primary	1971-2017	14,831,523	1971-2035	25,721,290.97
Middle	1971-2017	8,051,403	1971-2035	16,177,612.4
High	1971-2017	10,756,303	1971-2035	20,703,397.92

Time sequences data from 1971 to 2017 are used annually

to inspect the association between teacher enrollment and education (kinds & levels). The data dimension is represented by 3 columns and 46 rows (tuples) because there are 3 different types of teacher's levels in Pakistan. The variables used in this study are the number of levels in teacher enrollment, primary school output (0), middle school output (0), and secondary school output (0). Detailed description of the data is given in Table I.

## III. RESULTS

In this section, inclusive experiment is presented using data set designed based on teachers enrolled per year, acquired from Pakistan Economy Statistics Handbook as described in data collection. The dataset consists of the record of enrolled teachers in Pakistan over 46 years (1971-2017). Therefore we have a feature matrix of 46 rows over 3 columns and the accuracy is 61 percent and 73 percent. Fig. 2 shows the predicted data and original data, two methods of LR and SVM (support vector machine) are used in the predicted data. Primary education is seen as one of the main vehicles for promoting economic growth and improving living standards in developing countries [11], but it is likely to improve the quality of education if primary education involves new knowledge and adequate technical training at the primary level. Expanding access to primary education is a widely accepted priority in the fight against poverty. It is the first stage of compulsory education that establishes the academic foundation of the students and is considered as a fundamental right of all people. Primary education is highly linked to public [12] and private [13]-[15] institutions and the Express Tribune [16]-[19]. In this context, we took the data (Manual of Pakistan Economy Statistics) from the elementary education sectors from 1971 to 2017. In elementary schools enrolled teachers were 105,700 in 1971, 140,900 in 1980, 280,900 in 1990, 408,871 in 2000 and 441,689 in 2010. It continually increased and reached at 486,903 in 2017. The predicted teachers' enrollment in 1971 is 104.885.0936 while 178, 773.3532 in 1980, 271,133. 6778 in 1990 and 369, 375. 2373 in 2000, 455,854.3269 in 2010 and finally the predicted value of enrolled teachers in 2035 is 680,573.9234. The teachers enrolled in primary schools and their future predictions are shown in Fig. 2.

Middle schools follow grades 6 to 8; rural areas still favor single - sex education in middle schools. Middle school subjects include Urdu, English, arts, Islamic studies, mathematics, science, social studies, and computer science with accessible equipment. We took data from the Pakistani Economics Statistics Handbook, from 1971 to 2017 from middle schools. The middle schools enrolled teachers in 1971 were 36000, 51400 in 1980 and 79900 in 1990, 209,691 in 2000 and 331,464 in 2010, it continually increased and reached 445,455 in 2017. The predicted teachers enrolled in 1971, 1980, 1990, 2000, 2010 are 3771.0782, 66361.01, 144598.4247, 228088.9464, 301073.2542 respectively and finally the predicted value of enrolled teachers in 2035 is 491131.5828. Pakistan population is increasing rapidly as compared to middle schools. Fig. 2 shows the teachers enrolled in middle schools per year and their future prediction.

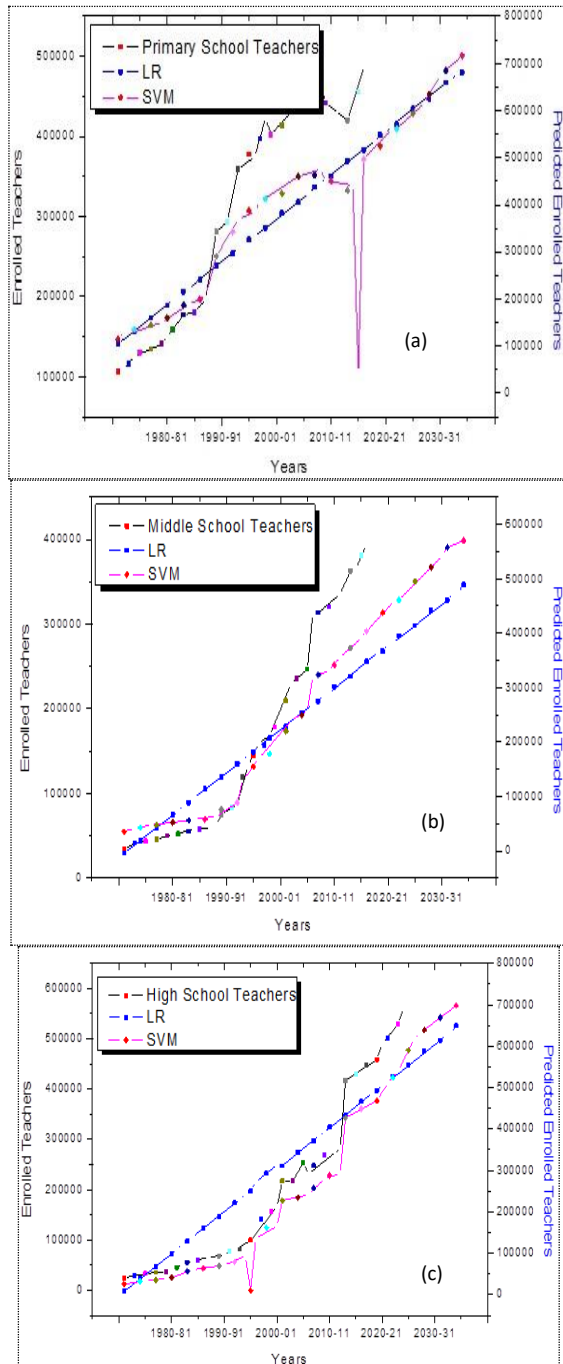


Fig. 2. This figure shows the three types of schools in Pakistan and comparison of two methods LR and SVM, and the prediction accuracy of SVM is higher than LR. Fig. 2(a) shows the primary schools and their prediction LR and SVM, Fig. 2(b) shows the middle schools and their prediction LR and SVM, Fig. 2(c) shows the high schools and their prediction LR and SVM.

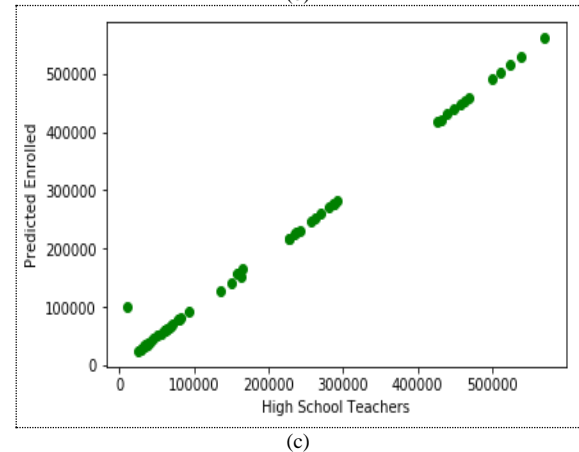
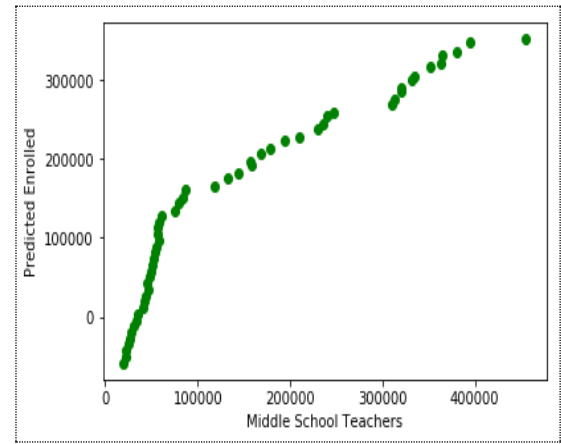
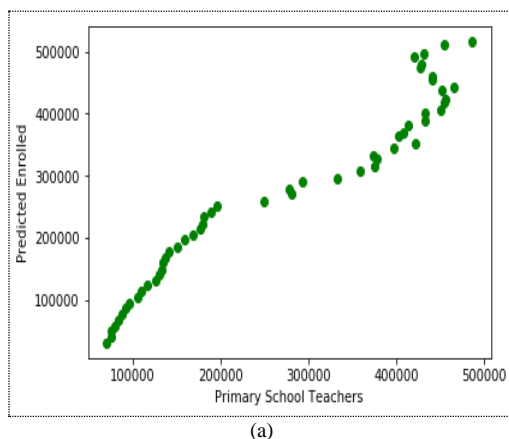


Fig. 3. This figure shows the three types of schools in Pakistan and SVM predicted data. Fig. 3 (a) shows primary school predicted data and Fig. 3 (b) shows middle school predicted data and Fig. 3(c) shows high school predicted data.

With annual examinations, high schools cover grades ninth to tenth. The students receive a secondary school diploma upon completion of grade 10. If they have the desire to continue inter-college further education, they must subsequently pass the high school examination. For high school forecasts, we took data from 1971 to 2017. The high schools enrolled teachers in 1971 were 37,900, 63,800 in 1980, 140,400 in 1990, 260,293 in 2000, 447,130 in 2010, it continually increased and reached 560,642 in 2017.

TABLE II: PREDICTED TEACHERS ENROLLMENT ACCURACY

METHODS	ACCURACY
LR	61%
SVM	73%

The predicted teacher enrolled value in 1971 is 8101.267835, in 1980 is 88879.79633, in 1990 is 18985.957, in 2000 is 313299.614, in 2010 is 404841.3103 and finally the predicted value of enrolled teacher in 2035 is 648952.5006. Fig. 2 shows the total number of teachers enrolled in primary, middle and secondary schools from 1971 to 2017 and the prediction of enrolled teachers in primary, middle and high schools from 1971 to 2035. The teachers enrolled in high schools and their future predictions are shown in Fig. 2 per year. In Fig. 3, the SVM (Support Vector Method) and Table II predicted data show the accuracy of both LR (Linear Regression) and SVM methods and the accuracy of both methods is 61% and 73% respectively. It proves that SVM presents the better prediction on this data

set. We used two different methods LR and SVM for prediction and their respective accuracy are shown in Table II.

#### IV. CONCLUSION

The stakeholders of education are the teachers, students, educational system and the curricula. The development of a country mostly depends of its educational standards. In order to establish standard in education we need competent, efficient, and proficient and devoted teachers. The teachers are considered to be guides, motivators, revolutioners, reformers, the resource persons, facilitators, apostles, messengers, prophets, directors, educators, game changers, fate changers and character builders. In order to develop education for the children, we need female teachers due to having motherly nature. It is an admitted fact that women can perform vital role in the development of education of kids. The number of teachers in Pakistan is lower than the growing population. The predicted results show that the number of registered teachers is less than the population rate in 2035. Pakistan needs to develop new teacher enrolment policies and provide free basic education for every child. The results recommend that more female teachers should be enrolled in schools at all levels in the future, as in Pakistan the female population is 52 percent of the total population, while the female ratio is less than male teachers. Depending on the female population, additional female teachers are required. Recent progress in technology has altered the learning behaviors of students; besides giving a new impulse it reshapes the education itself. It can be said easily that the improvements in technologies empower students to learn more efficiently, effectively and contentedly.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### AUTHOR CONTRIBUTIONS

Samina Kausar and Xuhuahu designed the study, implemented the analysis, and prepared the draft. Samina Kausar also wrote the paper along with literature. Xu Huahu, Shahid Iqbal and rest of the authors contributed in finalizing the manuscript. All authors have read and approved the final manuscript.

#### ACKNOWLEDGMENT

The authors are grateful to the School of Computer Engineering and Science, Shanghai University, Shanghai, 200444, China for their support and cooperation.

#### REFERENCES

- [1] Government of Pakistan, *Draft National Education Policy 2009 Ministry of Education*, 2009.
- [2] Government of Pakistan, Finance Division, *Economic Survey of Pakistan*, Islamabad: Ministry of Education, 2011.
- [3] Government of Pakistan, *National Education Policy (1998-2010)*, Islamabad: Ministry of Education, 1998.
- [4] S. Kausar *et al.*, "Integration of data mining clustering approach in the personalized e-learning system," *IEEE Access*, vol. 6, pp. 72724-72734, 2018.
- [5] Y. He, Y. Li, and X. Zhao "Analysis of energy saving potential of power industry in China based on multiple liner regression model," in

- Proc. 2012 8th International Conference on Computing Technology and Information Management (ICCM)*, 2012, vol. 1, pp. 206-209.
- [6] M. A. Arafa, M. I. Awad, and F. A. Tolbah "Proportional myoelectric prosthetic hand control using multi-regression model estimator with pattern classifier selector," *Advanced Robotics and its Social Impacts (ARSO)*, 2017, pp. 1-5.
- [7] Y. Yabuuchi, "The difference between the formulations of possibilistic robust regression model," *Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS), 2017 Joint 17th World Congress of International*, 2017, pp. 1-6.
- [8] R. Julianjatonso *et al.*, "High-resolution automated fugl-meyer assessment using sensor data and regression model," in *Proc. 2017 3rd International Conference on Science and Technology-Computer (ICST)*, 2017, pp. 28-32.
- [9] X. Feng *et al.*, "Contact temperature prediction of high voltage switchgear based on multiple linear regression model," in *Proc. 2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2017, pp. 277-280.
- [10] F. Li *et al.*, "Constrained interval-valued linear regression model," in *Proc. 2017 20th International Conference on Information Fusion (Fusion)*, 2017, pp. 1-8.
- [11] D. Suryadarma, A. Suryahadi, S. Sumarto, and F. H. Rogers, "Improving student performance in public primary schools in developing countries: Evidence from Indonesia," *Education Economics*, vol. 14, no. 4, pp. 401-429, 2006.
- [12] H. Alderman, P. F. Orazem, and E. M. Paterno, "School quality, School cost, and the public/private school choices of low-income households in Pakistan," *Journal of Human Resource*, vol. 36, pp. 304-326, 2001.
- [13] A. Shaikat. (2014). Education: Bill to regulate private schools yet to be tabled. *Express Tribune*. [Online]. Available: <http://tribune.com.pk/story/692703/education-bill-to-regulate-private-schools-yet-to-be-tabled/>
- [14] N. Siddiqui. (2017). Socio-economic segregation of disadvantaged children between schools in Pakistan: Comparing the state and private sector. *Educational Studies*. [Online]. Available: <http://dx.doi.org/10.1080/03055698.2016.1277139>
- [15] P. Srivastava, "Neither voice nor loyalty: School choice and the low-fee private sector in India," *Occasional Paper*, vol. 134, 2007.
- [16] The Express Tribune. (2015). Vacation fees: LHC seeks details of action taken against private schools. [Online]. Available: <http://tribune.com.pk/story/898641/vacation-fees-lhc-seeks-details-of-action-taken-against-private-schools/>
- [17] J. Tooley *et al.*, "The relative quality and cost-effectiveness of private and public schools for low-income families: A case study in a developing country," *School Effectiveness and School Improvement*, vol. 21, no. 2, pp. 117-144, 2010.
- [18] B. Zeitlyn *et al.*, "Inside private secondary schools in Malawi: Access or exclusion?" *International Journal of Educational Development*, vol. 43, pp. 109-117, 2015.
- [19] Iqbal, M. Shahid, and B. Luo, "Prediction of educational institution using predictive analytic techniques," *Education and Information Technologies*, 2018, pp. 1-15.

Copyright © 2019 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



**Samina Kausar** received the M.S. degree in computer science from the International Islamic University Islamabad, Pakistan, in 2007. She is currently working as a researcher and also a Ph.D. scholar with the School of Computer Engineering and Science, Shanghai University, China. She has been working as an assistant professor with the University of Kotli, Azad Jammu and Kashmir, Pakistan. Her research interests are in the fields of big data, bioinformatics, computer networks, cloud computing, e-learning data mining, and machine learning algorithms.



**Xu Huahu** is currently a doctoral supervisor and a professor with the School of Computer Engineering and Science, Shanghai University, where he is also the director of the Information Office of Shanghai University. He is the chairman of the Shanghai Security and Technology Association. His research interests include multimedia technology, CIMS, and computer network technology.



**Muhammad Shahid Iqbal** is currently working towards his Ph.D. at the School of Computer Science and Application Technology at Anhui University Hefei China. His research interest includes computer vision, image processing, machine learning, automated software engineering and software testing.



**Wang Xiangmeng** is doing her master's degree with the Department of Materials Genome Institute, Shanghai University. She is working in machine learning and data mining region and her major is computer science and technology.



**Muhammad Yasir Shabir** was born in District Kotli Azad Jammu & Kashmir, Pakistan. He has done the MS Computer Sciences from International Islamic University Islamabad, Pakistan and BSIT from University of Azad Jammu and Kashmir. Currently, he is working as a lecturer in University of Kotli Azad Jammu & Kashmir, AJ&K, and Pakistan. His major research area is computer networks & security, cloud computing, big data, machine learning algorithms and IoT.



**Tamoor Khan** is doing the Ph.D. with the School of Economic Information Engineering in Southwestern University of Finance and Economics China. His doctoral research investigates to improve auditing process with artificial intelligence, social media with AI and machine learning with finance.