

Predicting the Relationship between Parent Teaching Activities and Emergent Literacy in Preschool Children of Oxford Louise Academy of Dasmarinas using a Correlation and Clustering Analysis

^[1]Jennifer E. Galarosa, ^[2]Ma. Micah R. Encarnacion, ^[3]Dr. Maryli F. Rosas, ^[4]Dr. Paulino H. Gatpandan,

^[1] ^[2] ^[3] ^[4]De La Salle University-Dasmarinas

^[1]gje0381@dlsud.edu.ph, ^[2]emr2373@dlsud.edu.ph, ^[3]mfrosas@dlsud.edu.ph, ^[4]phgatpandan@dlsud.edu.ph

Article Info

Volume 81

Page Number: 5226 - 5235

Publication Issue:

November-December 2019

Abstract

Literacy development is the fundamental factor in the language development of a child. To enhance the children's language skills, the preschool teachers are challenged to constantly take part in the individual children's interactions. Since parents are considered the first teachers of their child, it is important to take into consideration how parents can assist on improving the language and acquiring important literacy skills of their children before their children attend to formal schooling.

The main goal of this research is to discover the relationship of the parent's intervention on the growing preschool children student of a private school in Cavite and the teaching activities. This was done through performing a survey assessment for random samples.

Further, the study on the basis of the evaluation of teachers before the child learning capabilities would aim to forecast students who needs to have intervention classes. This research focuses mainly on the application of the data mining techniques in the academic setting.

Keywords: correlation, clustering, preschool, emergent literacy

Article History

Article Received: 5 March 2019

Revised: 18 May 2019

Accepted: 24 September 2019

Publication: 25 December 2019

I. INTRODUCTION

Perceptively reading and writing are the vital requirements of many schools for the primary or preschooler before transitioning them to grade school level. Usually, at the age of 5 to 6, children in most cases starts to learn to read. However, many kids still do not acquire this important skill.

The significant development of a child greatly happens during pre-school years wherein the child's brain rapidly and fully

developing. Further, during this stage, the self-esteem and moral foundations of the child started to establish. Hence, teacher as well as parents should be guided in facilitating these for the development of the child.

According to a survey, Filipinos age ten and above are literate. They also mentioned that nine out of ten Filipinos whose age are from ten to sixty four were functionally literate[1].

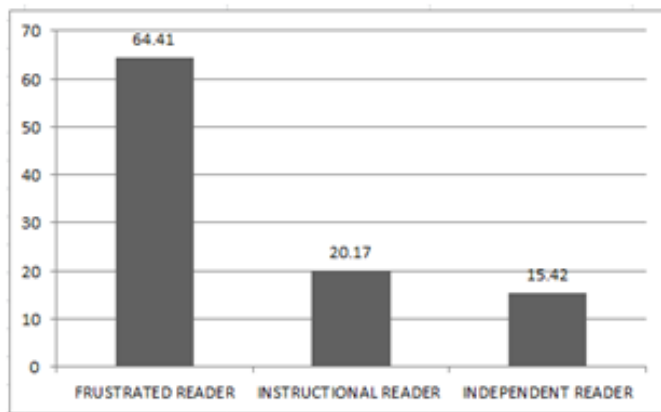


Fig. 1. Reading levels of grade 1 students in the city schools division of Manila.[2]

Literacy starts at home and is crucial to the development of the child capabilities. Learning capabilities of early childhood would involve reading of books, playing educational toys which will stimulate their brains and could have clear grasp of readiness and engagement. Parents interaction to their child would also affect the child's positive learning experience.

Children has to explore their environment in order to have better understanding of their surroundings, create and discover new learnings based on actual play-based activities. These will eventually leads them to becoming emergent literates and, helps them to naturally acquire the proficienciesneeded in school[3].

According to the article of PIA which mentioned on the UNESCO Report, being literate is the ability to read and write[1].While the ability of understanding a plain message in any language is basic literacy. Further, functional literacy is attained when a person has attained basic literacy with numerical skills[1].

A study reported by the U.S. Department of Education, kindergartners coming from low income family has lowest scores in reading and math. While kindergartners coming from high income households got highest scores in math and reading. Further, in terms of math and reading scores, children coming from Asian

countries had higher scores than children coming from other races or ethnicities[4].

Differences in child's skill levels often affects later their overall development in terms of cognitive, language, and academics[5][6].

The goal of this study wereto discoverthe relationofparent's activities to emergent literacy of the child's class performance in academics; and to predict who among the preschool students need to undergo summer class due to lack of reading and writing capabilities in Oxford Louise Academy of Dasmarinas Inc.

Oxford Louise Academy of Dasmarinas Inc., (OLA)is a private non-secretariat, co-educational institutional of learning offers early childhood programs in the City of Dasmariñas, Philippines.

II. REVIEW OF RELATED LITERATURE

The academic performance of students can be attributed to different factors. Heredity and peer pressure contribute to this. In this research, the researchers uncover the students' academic performance attributed to parental interventions.

According to social studies researches, a parent's support and intervention is important to children who are in their preschool years. The children's early reading and writing experiences should be done at school and at home, with a parent's help [7]. In a research by Yonson (2016), Filipino parents have willingness to support their children, either in elementary or secondary, in their academics, extracurricular aspects, and as members of the Parent-Teacher Association [9]. Besides a parent's direct involvement, their permission of their children watching cartoons or child TV shows.

In the academic setting, data Mining enables data-driven decision-making for the improvement of educational practice[10]

Data mining has been used for many aspects of education over the years. Following are researches which used analytics and varying techniques of mining students' academic performance.

Venkatesan and Selvaragini (2017) compared the use of two common algorithms in predicting students' academic performance: clustering and classification. For classification, they used the algorithms C4.5, k-Nearest Neighbor, and Naïve Bayes. On the other hand, they used Expectation Maximization (EM) and k-Means for clustering. The data they used are from undergraduate and post-graduate students of four private Arts and Sciences colleges in Tamilnadu, India. To predict a student's academic performance in end semester exams, they used the student details and results in their internal exams. The processes of compartmentalization and bunch grouping are used to process the data in MATLAB and WEKA. The clustering algorithms both got high accuracy – EM: 90% and k-Means: 99% [11].

In the study by Govindasamy and Velmurugan (2017), have used undergraduate and post-graduate students' details. The authors compared the use of clustering and classification algorithms for data mining in the education industry. The data set is wider than the previous research which now consists of the student's family details, her environment, and grade results for all academic activities. The researchers decided on four (4) clusters to group the students to distinct, first, second, and fail. EM has 74% accuracy, while k-Means remain with a higher accuracy of 81%. Overall, clustering averaged 82% in accuracy, while classification got 62.7% accuracy [12].

Another research of Yassein, Helali, and Mohomad whose main goal was to find out data patterns to predict a student's academic future. This research was also conducted to let

academic stakeholders know what actions to take to improve a student's academic performance. The researchers used SPSS and Clementine as data mining tools, while course and student details are used for data set. The research introduced the use of 2-step clustering which is the division of data based on similar aspects to get the number of clusters dynamically. They concluded that the most affecting factor in a student's academic performance are attendance in class, final and mid exam grades. Other affecting factors are personal, socio-economic, and other environmental variable factors [13].

Moucary, Khair, and Zakhem (2011) tries to improve a student's performance in foreign-language based higher education courses with the use of clustering technique and neural networks. The researchers used a hybrid algorithm of clustering and neural networks to predict the General Point Average (GPA) of students. The result can give a careful advising during registration to maintain acceptable General Point Average and high retention rate. Lebanon use French and Lebanese for their day-to-day conversations. However, on higher educations and due to globalization, English is used for many courses. Students might suffer when they are drastically changed to use English as medium of instruction, instead of using their native language. Neural networks are used by the researchers to predict a student's GPA based on their performance in English-taught courses and English-related courses.

Gangurde and Sonar (2014) mentioned in their article that there are several techniques of data mining such as classification, clustering, association, prediction and sequential patterns which are used for knowledge discovery from databases [14].

III. METHODOLOGY

This study used quantitative research. The researchers utilized several statistical methods to analyze the relationships between the participants with respect to learning levels of the preschool students. Furthermore, participant demographic information was obtained and incorporated into the analysis process. For purposes of this study, obtaining a sampling of preschool students in OLA, the researchers used convenience sampling as a method of obtaining participants.

The research methods and procedures used in this study are discussed including how data was collected, analyzed and synthesized. A quantitative non-experimental survey design was used for the research design.

As shown in Fig. 2, the Data model that the researchers used is a well-proven and robust methodology called Cross-Industry Process for data mining.

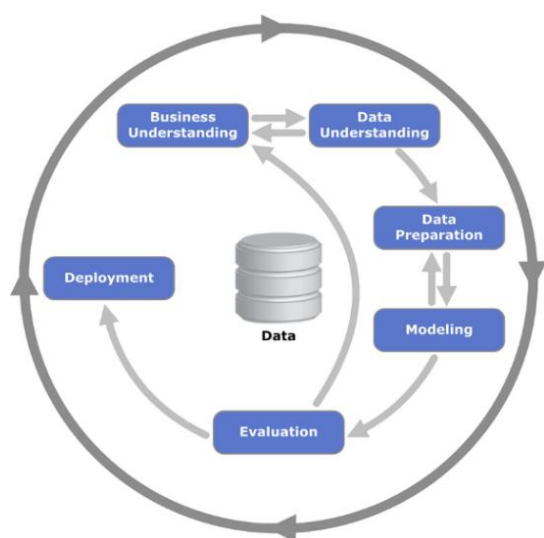


Fig. 2 Cross-Industry Process Diagram for Data mining

Business and Data understanding phase is very important at the start of this methodology. All pertinent data and business rules and requirements are initially planned and loaded.

Data preparation phase is the next stage wherein the researchers decide on what are relevant data needed in the data analysis. This

also includes data cleansing in which all noise data were removed.

Following this stage is the Modelling phase. The researchers then select the appropriate modelling techniques such as clustering and classification.

Results from mining the data are then evaluated if they adhere to business objects.

Data Mining Method

Data mining procedures are used to build model. In building such model, algorithm should be implemented. There are two basic groups of algorithms: supervised and unsupervised.

The researchers used supervised learning wherein it makes use of training data containing the independent variable and dependent variable. By using supervised learning, the target variable/s can be predicted based on the predictor variables [15].

Data Mining Task

Data mining tasks are very diverse because of the nature of the large database where there could be many patterns to discover. These tasks can be classified as summarization, association, classification, trend analysis and clustering. [16]

Clustering

Clustering is a process of identifying same groups from unstructured data [17]. A cluster is a subset of similar objects.

Correlation

Correlation is identifying relationships between variables [18].

Data Mining Tool

A data mining tool used by the researchers is the IBM Statistical Package for the Social Sciences Statistics. This is a software package used for analyze statistical data [19].

The researchers used survey forms from the respondents to gather and extract necessary data that are related to the purpose or goal of the survey.

Fig. 3 is the researchers' sample survey form for parent's intervention on the child emergent literacy while Fig. 4 shows the Sample survey form the evaluation of the teacher's prior to child's learning capabilities.

The researchers have used closed-ended questions surveys which is more efficient in terms of preparation and interpretation. For data gathering and analysis, the sample consist of total 31 preschoolers were selected out of 54. Among them are 13 beginners and 18 kindergarten students.

Student Name: _____
Parent Name: _____

Questionnaires Data Entry – SPSS

Kindly fill up the questions below.

1. What is your child's gender?
☐ Male
☐ Female

2. What is your child's school level?
☐ Beginner
☐ Kindergarten

3. What is your child's age?
Child's age
Choose an item.
Choose an item.
2 years old
3 years old
4 years old
5 years old
6 years old
7 years old

4. What teaching methods used?
☐ Traditional approach (ex. Precise reading books, counting numbers etc.)
☐ Interactive approach using educational apps or games or videos
☐ Enrolling the child to a learning center (ex. Kumon)
☐ None

Fig.3 Parent's intervention on child emergent literacy survey form.

Questionnaires Data Entry – SPSS

Evaluator:
Student name:

Kindly fill up the questions below.

1. What is your student's gender?
☐ Boy
☐ Girl

2. What is your student's school level?
☐ Beginner
☐ Kindergarten

3. How does the student's reading skills?
☐ Very Poor ☐ Poor ☐ Fair ☐ Good ☒ Very Good

4. How does the student's reading skills?
☐ Very Poor ☐ Poor ☐ Fair ☐ Good ☐ Very Good

5. How does the student's reading skills?
☐ Very Poor ☐ Poor ☐ Fair ☐ Good ☐ Very Good

6. How does the student's numerical skills?

Fig.4 Sample survey form for the evaluation of the teacher's prior to child's learning capabilities.

The researchers then identify the type of variables that were needed for the analysis. The

different learning approaches are considered independent variables. While dependent variables are the child's learning capabilities.

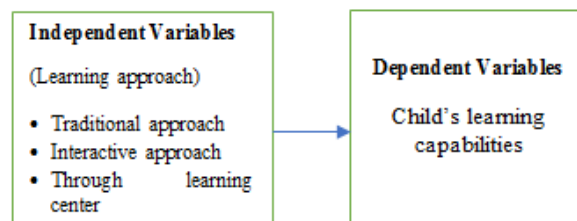


Fig.5. Research Paradigm

In Fig. 5 is the conceptual framework of the study is presented to give of the work to be done. The process is the strategy to be used as generating data using survey or questionnaires which will be distributed in several participants. The possible output or the dependent variable will be based on the participant answers that the researchers evaluated as well.

IV. RESULT AND DISCUSSION

The research was conducted to identify the relationship of having parent intervention prior to emergent literacy of their child regarding of learning capabilities of the students which being evaluated. The researchers selected a sample population of 31 respondents' parents of students and 2 advisory teachers in OLA. This is to find the relation to parent's activities to emergent literacy of the child's class performance in academics; and to predict who among the preschool students need to undergo summer class due to lack of reading and writing capabilities.

Distribution sampling

Type of Predictor (Beginner, Kindergarten levels): Teaching method/learning approach/ Teacher evaluation

To get the Frequency distribution table (FDT), here's the formulas: $(K = 1 + 3.322 \log n)$, $R = \text{highest} - \text{lowest}$, $C = R / K$. The dataset results were done through SPSS software.

Where N = 13 for beginner and N=18 for kindergarten level.

**Table I. What teaching method used?
(Beginner level)**

	Percent	Valid Percent	Cumulative Percent
Valid Traditional approach	38.5	38.5	38.5
Interactive approach	30.8	30.8	69.2
Through learning center	15.4	15.4	84.6
No intervention	15.4	15.4	100.0
Total	100.0	100.0	

Table I. shows that most of the parent applies the traditional approach in a way of teaching their child for early literacy.

Table II. How does the student in reading skills? (Beginner level)

	Percent	Valid Percent	Cumulative Percent
Valid Very poor	7.7	7.7	7.7
Poor	15.4	15.4	23.1
Fair	46.2	46.2	69.2
Good	23.1	23.1	92.3
Very good	7.7	7.7	100.0
Total	100.0	100.0	

In table II.is the reading evaluation of the teachers among 13 beginners.One of them got very poor reading skills.

Table III. How does the student in writing skills? (Beginner level)

Table III shows that 46% got fair in writing skills while 2 among 13 got good ratings.

	Percent	Valid Percent	Cumulative Percent
Valid Very poor	7.7	7.7	7.7
Poor	30.8	30.8	38.5
Fair	46.2	46.2	84.6
Good	15.4	15.4	100.0
Total	100.0	100.0	

Table IV. How does the student in numerical skills? (Beginner level)

In table IV, one out 13 children has very good performance in numerical skills and two of them might be having difficulties.

	Percent	Valid Percent	Cumulative Percent
Valid Very poor	15.4	15.4	15.4
Poor	30.8	30.8	46.2
Fair	23.1	23.1	69.2
Good	23.1	23.1	92.3
Very good	7.7	7.7	100.0
Total	100.0	100.0	

**Table V. What teaching method used?
(Kindergarten level)**

	Percent	Valid Percent	Cumulative Percent
Valid Traditional approach	33.3	33.3	33.3
Interactive approach	38.9	38.9	72.2
Through learning center	11.1	11.1	83.3
No intervention	16.7	16.7	100.0
Total	100.0	100.0	

Table V shows that most of the kindergarten parent applies the interactive approach in a way of teaching their child for early literacy. Three of them have no intervention done for teaching their kids.

Table VI. How does the student in reading skills? (Kindergarten level)

	Percent	Valid Percent	Cumulative Percent
Valid Poor	22.2	22.2	22.2
Fair	22.2	22.2	44.4
Good	38.9	38.9	83.3
Very good	16.7	16.7	100.0
Total	100.0	100.0	

In table VI,three among 18 students being surveyed got very good in reading skills.

Table VII. How does the student in writing skills? (Kindergarten level)

	Percent	Valid Percent	Cumulative Percent
Valid Poor	22.2	22.2	22.2
Fair	22.2	22.2	44.4
Good	38.9	38.9	83.3
Very good	16.7	16.7	100.0
Total	100.0	100.0	

In table VII, shows the highest percentage has on good criterion which means most of them have a good writing skill.

Table VIII. How does the student in numerical skills? (Kindergarten level)

	Percent	Valid Percent	Cumulative Percent
Valid Very poor	5.6	5.6	5.6
Poor	5.6	5.6	11.1
Fair	50.0	50.0	61.1
Good	22.2	22.2	83.3
Very good	16.7	16.7	100.0
Total	100.0	100.0	

In table VIII shows that 3 out 18 students outstand in numerical skills while fair criterion got the highest percentage

Correlation/Clustering matrix result

Table IX. Correlations Matrix

		What type of learning approach does the parent being applied?	What score does the kid got from teachers' evaluation?
What type of learning approach does the parent being applied?	Pearson Correlation	1	.093
	Sig. (2-tailed)		.619
	N	31	31
What score does the kid got from teachers' evaluation?	Pearson Correlation	.093	1
	Sig. (2-tailed)	.619	
	N	31	31

In table IX shows the traditional approach has a strong positive correlation to child's passing score prior to learning evaluation by their teachers.

Fig.6 shows that traditional approach has a high number of being used for emergent literacy of their kids while the lowest is enrolling their kids through learning center.



Fig.6 the FDT learning approach applies

The Fig.7 indicates that 65% of the total 31 of both beginner and kindergarten students sampling has passed the evaluation and 8% need to undergo summer class and 3% of it has failed the evaluation might repeat the school level to complete the requirements to pass in next school level.

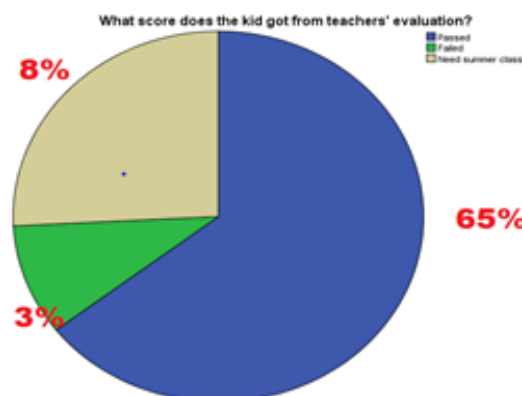


Fig.7 Evaluation of teachers 'prior to child learning capabilities.

Table X. Initial Cluster Centers

	Cluster	
	1	2
What type of learning approach does the parent being applied?	1.0	4.0
Learning capability	5.0	1.0
Gender	1.0	2.0
Preschool level	2.0	1.0

In table X, it has clustered into 2, and the variables selected being analyzed, which learning approach 1 is the traditional approach while 4 is no intervention, learning capability 5 is very good satisfaction rate while 1 is very poor, gender 1 is male and 2 is female, preschool level 1 is beginner and 2 is kindergarten level.

Table XI. Final Cluster Centers

	Cluster	
	1	2
What type of learning approach does the parent being applied?	1.6	3.3
Learning capability	3.9	2.0
Gender	1.5	2.0
Preschool level	1.7	1.5

In table XI, the clusters values were being changed on the final clustering method wherein the value 3 on learning approach is the through learning center. Then 3.9 values indicate fair ratings while 2 is poor.

Table XII. Distances Between Final Cluster Centers

Cluster	1	2
1		2.620
2	2.620	

In table XII shows the two clustering being analyzed and it has same values of 2.620.

Cluster	1	20.000
	2	11.000
Valid		31.000
Missing		1.000

Table XIII. Number of Cases in each Cluster

Table XIII shows the cases in each cluster in total of 31 out 54 sampling. Cluster 1 has 20, while cluster 2 has 11.

V. CONCLUSION AND RECOMMENDATION

In this study, the researchers have used the qualitative methodology to compare the predictive power of clustering algorithm and the correlation analysis. The researchers were able to demonstrate the technique using k-means clustering algorithm. On a data set of Oxford Louise Academy of Dasmarinas Inc., (OLA) 31 out of 54 preschool students were evaluated to produce the numerical interpretation of the results. These models predicted (1) the relation of parent's activities to emergent literacy of the child's class

performance in academics; and (2) who among the preschool students need to undergo summer class due to lack of reading and writing capabilities.

Therefore, clustering algorithm is used to monitor the students' performance in an educational institution. It also enhances the decision making of academic planners in monitoring the behavior of the learner and predicting the learning intervention that the parents can use to improve their child learning skills.

In conclusion, a traditional method of intervention is being used for emergent literacy of their children. This type of intervention strongly affects the students' academic performance positively.

REFERENCES

- [1] Literacy in the age of social media. PIA. Philippine Information Agency. October 10, 2018. <https://pia.gov.ph/news/articles/1013738>.
- [2] Reading levels of grade 1 students in the city schools division of Manila. Source: Philippine-IRI Test, Schools Division of Manila, SY 2003-04
- [3] Department of Education, Retrieve from <http://www.deped.gov.ph/k-to-12/about/k-to-12-basic-education-curriculum/kindergarten/>
- [4] A Matter of Equity: Preschool in America, U.S. Department of Education, April 2015 <https://www2.ed.gov/documents/early-learning/matter-equity-preschool-america.pdf>
- [5] Shields, L., Newman, A., Satz, D. "Equality of Educational Opportunity", Stanford Encyclopedia of Philosophy, First published Wed May 31, 2017,

- <https://plato.stanford.edu/entries/equal-ed-opportunity/>
- [6] Schwab, JF, Lew-Williams, C. "Language learning, socioeconomic status, and child-directed speech", Wiley Interdiscip Rev Cogn Sci. Author manuscript; available in PMC, April 16, 2018. Published in final edited form as: Wiley Interdiscip Rev Cogn Sci. 2016 Jul; 7(4): 264–275. Published online 2016 May 19. doi: 10.1002/wcs.1393
- [7] Bierman, KL., Coie, J., Dodge, K., Greenberg, M., Lochman, J., McMohan, R., Pinderhughes, E., "School Outcomes of Aggressive-Disruptive Children: Prediction From Kindergarten Risk Factors and Impact of the Fast Track Prevention Program", Conduct Problems Prevention Research Group, First published: 05 February 2013, Aggressive Behavior, Volume 39, Issue 2. <https://doi.org/10.1002/ab.21467>
- [8] Semorlan, A. A., & Semorlan, P. (2015). Correlates Of Early Reading And Writing Performance In Filipino Of Prefirst Grade Pupils. Journal of Southeast Asian Studies, 20(1), 173-182. doi:10.22452/jati.vol20no1.11
- [9] Yonson, D. (2016). Level of Parent Involvement in the Elementary and Secondary Levels. The Normal Lights, 10(1). Retrieved from https://www.researchgate.net/publication/309798276_Level_of_Parent_Involvement_in_the_Elementary_and_Secondary_Levels
- [10] Rosas, MF., Ambat, SC, Lagman, AC, "Extraction of Students' Feedback in University Services using Logistic Regression and Chi-Square Automatic Interaction Detection (CHAID) Algorithm", International Journal of Simulation Systems, Science and Technology, April 2019, ISSN: 1473-804x online, 1473-8031 print
- [11] Venkatesan, E., & Selvaragini, S. (2017). Prediction Of Students Academic Performance Using Classification And Clustering Algorithms. International Journal Of Pure And Applied Mathematics, 116(16), 327-334. Retrieved from <https://acadpubl.eu/jsi/2017-116-13-22/articles/16/57.pdf>.
- [12] Govindasamy, K., & Velmurugan, T. (2017). A Study on Classification and Clustering Data Mining Algorithms Based on Students Academic Performance Prediction. International Journal of Control Theory and Applications, 10(23). Retrieved from https://www.researchgate.net/publication/317166866_A_Study_on_Classification_and_Clustering_Data_Mining_Algorithms_based_on_Students_Academic_Performance_Prediction.
- [13] Yassein, N. A., Helali, R. G., & Mohamad, S. B. (2017). Predicting Student Academic Performance in KSA using Data Mining Techniques. Journal of Information Technology & Software Engineering, 07(05). doi:10.4172/2165-7866.1000213
- [14] Gangurde, R. A., Sonar, M. R. (2014). Knowledge Extraction using Data Mining. Department of MCA, KKWagh Institute of Engineering Education and Research, Nashik Maharashtra, India. Spvryan's International Journal of Engineering Sciences Technology (SEST) SEST Issue 1, Volume 1 - Oct. 2014 <http://spvryan.org/Issue1Volume1/12.pdf>
- [15] Avenir, AE. Villota, JD. and Rosas, MF. "Machine Learning: A Means of Diagnosing and Prescribing Treatments

for Tuberculosis Patients “, International Journal of Computer Theory and Engineering, Vol. 7, No. 3, June 2015

- [16] Fu, Y., “Data Mining: Tasks, Techniques, and Applications”, University of Missouri – Rolla, <https://academic.csuohio.edu/fuy/Pub/pot97.pdf>
- [17] S.D.Gheware¹ , A.S.Kejkar² , S.M.Tondare“Data Mining: Task, Tools, Techniques and Applications”, International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 10, October 2014, www.ijarcce.com 8095
- [18] Nikolas, S., “What does it mean if the correlation coefficient is positive, negative, or zero?”, Fundamental Analysis, April 2018, retrieved from <https://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-coefficient-positive-negative-or-zero.asp>
- [19] Quintero, Dino; et al. "Workload Optimized Systems: Tuning POWER7 for Analytics".2012