Journal of Engineering Education Transformations ,
Volume 36 , No. 4 , April 2023 , ISSN 2349-2473, eISSN 2394-1707

10.16920/jeet/2023/v36i4/23124

Best peer active learning strategies for problem solving using C programming

Saraswathi Meena R[1], Saranya R[2], Neeraja U[3], Senthil Shanmugam K[4]
[1,2,3,4] Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering Madurai, India
[1]saraswathimeenacse@gmail.com
[2]saranyarajagopal2002@gmail.com
[3]neerajausuja@gmail.com
[4]senthil.alagesh@gmail.com

Abstract : Peer learning is a process where one or more students teach other students and provides support throughout the learning process. Active learning is a process of learning concepts through thinking, discussing, investigating, and creating. In order to improve the learning of students in the concept of "problem solving using C programming," peer and active learning approaches are implemented by forming groups. The general four methods for peer active learning, such as think-pair-share, the zig-zag method, the coding test, and mini project, are chosen as activities for the students. An unsupervised machine learning algorithm is used to create clusters of students based on pretest and posttest scores. With the help of the K-Means clustering technique, the increased change in student performance after peer active learning approaches is clearly visible.

Keywords : Peer Learning, Active learning, K-Means Clustering.

Saraswathi Meena R
Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering Madurai, India
saraswathimeenacse@gmail.com

1. Introduction

With the enormous array of technology and gadgets attracting the attention of this new generation of students, specific strategies are essential to efficiently educate the kids in the various skills and assist them in building their problem-solving abilities. Students must be consistently engaged and curious to learn new things in order to enhance their skills and abilities, and as it is known, students are rarely interested in learning new things.

To attract students' attention, teaching and learning methods must elicit a high level of curiosity that can project itself over the extremely addictive nature of social media applications, video games, and attention-getting movies. The students must be so engrossed in the active learning methods that they do not readily become distracted by other factors. Active learning approaches that are very engaging, such as holding a coding competition or requiring students to build a small-scale project, can easily pique students' interest in learning the topics on their own.

Peer learning also helps students enhance their communication skills and appropriately communicate their grasp of concepts to others. This development in communication is shown in their increased self-assurance and capacity to acquire a new topic or skill with the assistance of their peers. Active learning also aids students in making better judgments in general

Based on the learning accomplished via the traditional technique, a pretest was administered.

Based on the learning accomplished utilising four peer-active learning methodologies, a pretest was administered.

because they gain a deeper understanding of the things they are learning. Students gain hands-on experience with tools, increasing their confidence in their ability to apply them in larger, more complex applications in the future. using a C programming course to demonstrate the effectiveness of active learning and peer learning in problem resolution. Unsupervised machine learning algorithms were employed. The unsupervised machine learning approach primarily functions on unlabeled datasets where each piece of information has a specific function.

Clustering techniques come in a variety of forms, including Partitioning clustering, K-Mean clustering, and Density Based clustering. But in this case, K-Means Clustering technique was taken into account to demonstrate that employing clusters, peer learning, and active learning improved students' performance.

2. Study Context

Programming is an important subject for students of all disciplines, especially those studying computer science. It is critical to instill a passion for programming skills in their students in order for them to succeed in their careers. A real-time problem's solution is usually translated into an algorithm, which is then taught to a computer using a programming language. As a student, one can communicate with a computer using programming languages such as C, C++, Java, and Python, among others. Students have been learning the C programming language for the past three decades. If they thoroughly mastered the C programming language, they would be well prepared to study additional programming languages such as C++, Java, and Python.

To assist students in learning the C programming language, this study employs peer learning and some active peer learning methodologies. In this peer learning scenario, 40 students are divided into 20 equal-sized groups to learn. The students were given the work and subsequently evaluated it using peer-active learning methodologies. Think Pair Share, Zig Zag Method, Coding Contest, and Mini Project are some of the peer active learning methodologies that are utilized for evaluation.

After introducing the fundamental concepts in C, such as looping control, decision control, and case control, in a traditional lecture style, each of the 40 students was tested individually. This test result was unexpected. Then, in order to boost the student's learning ability, peer learning was used as a tool. Some active learning methodologies were employed for evaluation in order to analyze peer learning performance.

When it comes to peer learning, grouping is an essential component. Students in this scenario are allowed to establish their own groups, with a maximum group size of two. In fact, this leads to the formation of the most effective study groups. As a result, the following individuals have been added to the group: 1. Gender-specific groups 2. social circles with close friends 3. Groups of people with similar learning abilities.

3. Methods

A. Think Pair and Share

A cooperative learning strategy in which the entire group is given a single question or prompts to discuss is known as Think-Pair-Share. Students think about the subject for a few minutes on their own before forming pairs to share their replies. Finally, they share their ideas with the rest of the group. Even if they have little personal interest in the issue, pupils are motivated to participate. Students not only think about

the issue, but they also practise communication and problem-solving skills. Preparation is straightforward, and implementation has resulted in increased student involvement and improved learning results in a variety of settings. The questions are usually about higher-order cognitive skills. Think Pair Share can be a 5-minute activity or a 30-minute or longer activity.

The implementation of Think Pair Share is simple. Since, the grouping task is already implemented that is a group consists of a pair of students. Then, there is discussion on the given topic. Here the topic chosen was Case Control Statements, Iterative Statements and Decision Control Statements in the subject C programming language. These topics are the building blocks of any programming language especially in the basic programming language C. At the end this process, the students are given a set of multiple choice questions that are highly challenging and align with respect to the goal of the day or week in the learning process. Finally, the scores are calculated for the team.

B. Zig-Zag Method:

Zig-Zag method is one of the most common and modern methods used in the active learning process. With the zig-zag strategy, activities are planned in a way that makes pupils rely on one another to succeed. It divides classes into groups that individually put together a portion of an assignment and, after done, synthesise their work. The zig-zag approach is a cooperative learning strategy that encourages both individual responsibility and team goal attainment.

The method's name is derived from the zig-zag puzzle since it entails fitting the assignment's components together to create a whole image. The class is divided into the same number of groups as the assignment's components, and the assignment is separated into parts. Each of these groups is assigned a unique subject to study and is permitted to do so. To create new groups made up of individuals from each group, these groups are shuffled.

The method divides courses into mixed-gender groups to focus on minor issues that the group then combines into a final result. An in-class assignment, for instance, is broken up into themes. After that, the class is divided into groups with a member allocated to each subject. Each student researches their subject independently before presenting it to the class. Students then assemble into topic-specific groups. Each participant addresses the subject group once more. Students in topic-related groups balance opposing viewpoints and compile data. They produce a summary report. The original groups then get back together and hear everyone's speeches.

The final presentations enable each group member to comprehend both their own content and the conclusions from topic-specific group discussions.

C. Coding contest:

A programming contest, often known as a coding contest, entails writing a code or, more commonly, typing the program for a set of problem statements. Time complexity, space complexity, comparing the solution to the actual program, and the number of problem statements successfully addressed by the candidate in a given amount of time are used to evaluate these types of contests. The popularity of coding contests or programming contests stems from the fact that they improve a student's skill in a programming language and shift the student's problem-solving level from writing a simple solution with fewer constraints to writing the best solution for a problem with more constraints.

Competitive programming's goal is to generate source code for computer programs that can tackle certain issues. The great majority of programming contest problems are mathematical or logical in character. Here, using peer learning concepts, the students solve a problem in pairs. Students can use the programming contest to help them choose the best notion to solve a problem from the Case Control Statements, Iterative Statements, and Decision Control Statements topics. As a result, the student learns how to construct a valid C program from the ground up. In a programming competition, students engage in active learning as well as peer learning.

D. Mini Project:

A mini project is an assignment in which the student attempts to apply all of the theoretical concepts in a small but effective application. This mini project helps the students to boost their skills and widen their thinking and learning processes.

It also helps them gain deeper knowledge and understanding of the fundamental concepts. When a mini project is implemented by having a pair of students. Hence, active learning and peer learning do

take place.

Since the mini project helps the students to solve higher complexity problems, it is taken up that the students teach and help their partners to solve them, indulging all the constraints. Students build extra skills necessary for their future success, such as time management and critical thinking, while also allowing them to work together or take charge of their own learning. Perhaps more crucially, it allows students to produce authentic projects that are both personal and meaningful to them. Students have the opportunity to explore their own interests, which provides wonderful learning opportunities for both students and teachers.

Mini Project also teaches kids important technology skills that they will need in the future, as well as skills that they may apply in other classrooms and in their future. Students get more control over their learning and growth by learning to produce multimedia presentations, problem solve, think critically, quickly access resources, and interact with others. Mini project helps the student to integrate all the concepts that were chosen in the C programming language.

4. Data Analytics – Implementation

To analyze the data for peer learning and active learning, a series of steps like collecting the data, preprocessing the data collected, visualizing the processed data, and applying the K-Means unsupervised learning algorithm. As it establishes the natural grouping among the available unlabeled data, clustering is crucial. Clustering is a simple method for carrying out several surface-level analyses that can provide you with immediate benefits in a number of disciplines. As it establishes the natural grouping among the available unlabeled data, clustering is crucial.

A. Data Description and Data Pre-processing

The data collected is grouped since the students worked together in pairs throughout the learning process. The pretest and posttest scores are grouped according to the group members by using the mean aggregation method. Hence, since each attribute is in different ranges, the data is normalized to a range of 0 to 10. So, every attribute will be given equal weightage.

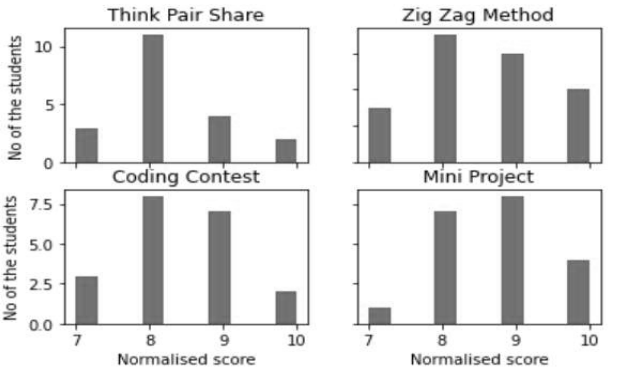

Fig. 1: Distribution of all four activities.

The frequency graph depicts the frequency of students for each normalized score in the activities of think pair share, zig zag method, coding contest, and mini project. For example, in the coding contest method, there are three groups of students having a normalized score of 7, but in the mini project method there is only one group having a normalized score of 7. Other peer groups have a greater score than 7. This normalized score is the result of the peer learning methods. The x axis depicts the range of normalized scores, whereas the y axis depicts the frequency of the student's pair or group.

The effectiveness of the pupils' learning progress is not only visible in the clusters that have been produced and plotted with various ranges. As a result, multiple measurements from the clusters were used to demonstrate that students had improved their problem-solving skills using the C programming language. The first metric used here is the mean, which is the most popular and significant metric for determining whether or not the values in a dataset are evenly distributed. The mean is the sum of all the data points in a particular dataset divided by the total number of data points in that dataset.

$$Mean = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1)$$

The median is the second measure utilized. Another key position in the dataset is the median, which is where the data is arranged from least to biggest. When the mean metric is impacted by extremely small and large data values, the median is preferable. To summarize, whenever a dataset comprises extreme values, the median is frequently used as the preferable measure of central position.

$$Median\ (X) = \begin{Bmatrix} X\left[\frac{n+1}{2}\right], n\ is\ odd \\ \frac{X\left[\frac{n}{2}\right]+X\left[\frac{n}{2}+1\right]}{2}, n\ is\ even \end{Bmatrix} \quad (2)$$

Variance is a data-based measurement of variability. The ratio of variation from the mean to the number of data points in a dataset is known as variance. The difference between each dataset data point and the matching dataset mean is calculated. When comparing the variability of two or more variables, the variance is useful. When the variables are compared, the one with the highest variance has the most variability.

$$Variance = \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n} \quad (3)$$

The final metric used is the standard deviation. The positive square root of variance is measured by standard deviation. The standard deviation is simpler to interpret than the variance because the standard deviation is measured within the same units as the data.

$$Standard\ Deviation = \sqrt{\sigma^2} \quad (4)$$

Skewness is a measure of how far a probability distribution deviates from the symmetrical normal distribution (bell curve) in a given collection of data. A symmetric distribution, or data collection, has an identical appearance to the left and right of the centre point. A normal distribution has zero skewness, and any symmetric data should have a skewness close to zero. Negative skewness numbers show data which is skewed left, whereas positive skewness values suggest data which is slanted right. The term "skewed left" indicates the length of the left tail in comparison to the right tail. Skewed right indicates that the right tail is longer than the left tail. If the data is multi-modal, the sign of the skewness may be affected.

$$Skewness \quad \frac{(\sum_{i=1}^{n}(x_i - \bar{x})^3/n}{\sigma^3} \quad (5)$$

Figure distribution of pretest and pretest compares the pretest and posttest using a graphical display. The mean score has risen from the pretest to the posttest by 1.14. Pretest scores ranged from 5.5 to 10, whereas posttest ratings remained between 8 and 10. The important thing to remember is that whereas the density was 0.4 in the pretest, it was 1.35 in the posttest. This graph shows how student performance and knowledge improved with the use of peer active learning techniques.

From the distribution graph of all four activities (Think pair share, zig-zag, coding contest and Mini project). There are four graphs in this illustration. Each graph shows a different action that was carried out. The distribution graph shows all the necessary statistics for all the four activities (Think pair sharing, Zig-Zag Method, Coding Contest, Mini Project) that were used, including mean, median, skewness, kurtosis, and standard deviation. Even if each activity is independent of the others, the picture may be used to compare student performance in a given activity. The bottom and upper limits of the students' scores are represented in the frequency distribution graph. As can be seen from the aforementioned measurements, the best active and peer learning methodologies significantly increase students' problem-solving skills

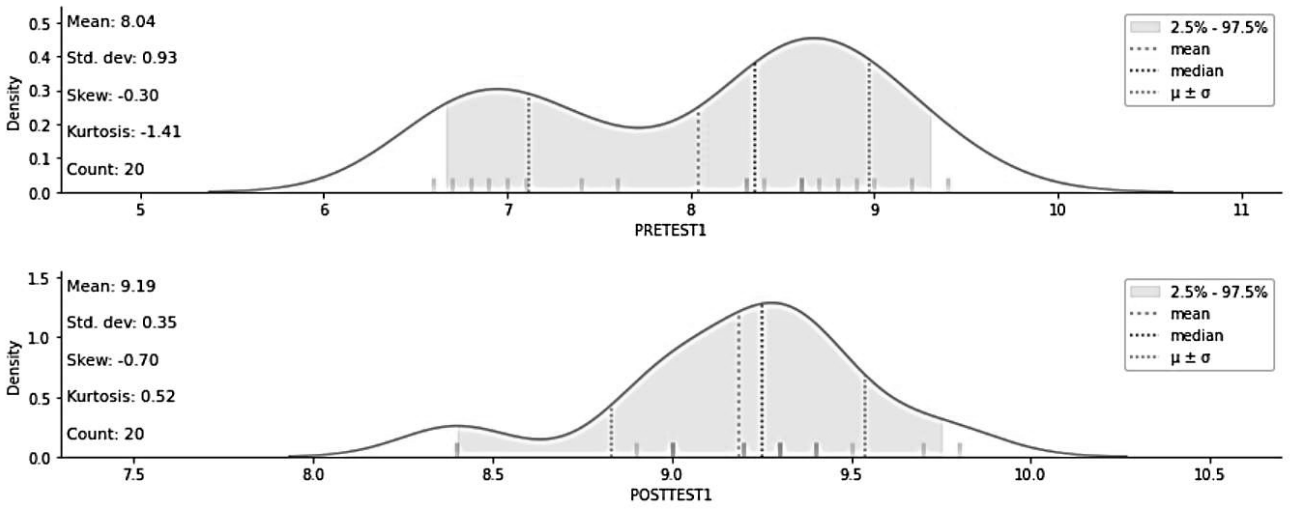

Fig. 2. : Distribution of pretest and Posttest.

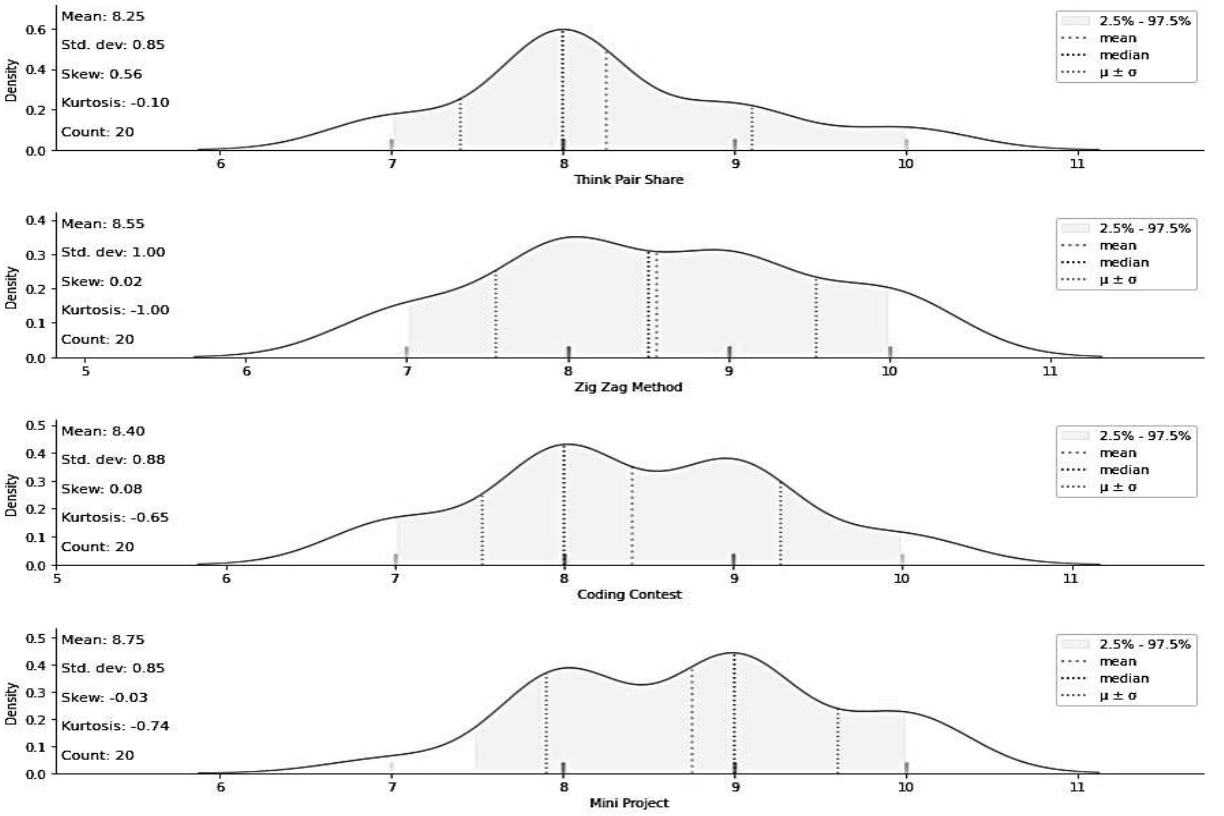

Fig. 3: Distribution of all four activities.

for the C programming language in a short amount of time.

Kurtosis is a measure of how heavy or light the data is compared to a standard distribution. Data sets having a high kurtosis are mostly contain heavy tails, or outliers. Light tails or an absence of outliers are common in data sets with low kurtosis. The most extreme instance would be a uniform distribution. A large kurtosis suggests a high level of risk in an investment since it shows a high probability of extraordinarily large and extremely tiny returns. A minor kurtosis, on the other hand, indicates a moderate amount of risk because the possibilities for maximum returns are minimal.

$$Kurtosis = \frac{\left(\sum_{i=1}^{n}(x_i-\bar{x})^4\right)/n}{\sigma^3} \quad (6)$$

B. Correlation and Covariance

Correlation is defined as how one attribute is related to other attributes in the data. Correlation: a statistical technique which determines how one variable moves or changes in relation to the other variable. It gives us an idea of the degree of the relationship between the two variables. Here, correlation is defined for the post-test parameter as how the posttest gets affected by other factors or if it is an independent variable.

$$r = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\ \sum(y_i-\bar{y})^2}} \quad (7)$$

As a feature selection algorithm, we employ Person's Correlation Coefficient. When two traits are highly correlated, they are linearly dependent and, as a result, have approximately equal effects on the dependent variable. r is the range of correlation. If r = 0, they are highly correlated; if r = 0, they are uncorrelated. Values are independent if the correlation is zero. The use of the correlation coefficient defines whether the unlabeled attributes are linearly dependent on other attributes or independent of other attributes. The formula below is used to calculate a person's correlation coefficient.

From the correlation, the posttest depends on the four activities of peer active learning strategies since they are positively correlated leading to positive relationship with posttest. This gives a clear cut view how these four activities improve the performance of

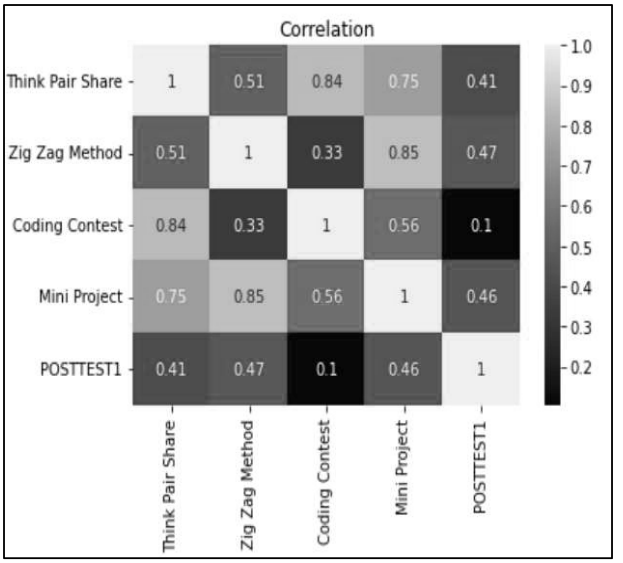

Fig. 4 : Correlation depiction.

the student knowledge in the posttest. Not only correlation graph conveys that activities are positively correlated towards posttest performance, but also the covariance graph does convey that activities are associated with posttest positively. Hence, it is the first clear evidence that student's performance was improved by implementing the peer active learning strategies

Covariance is defined as the measure of how changes in one variable are associated with another variable. Here, covariance measures the degree to which two variables are linearly separated. The sign of the covariance result identifies the tendency in the linear relationship between two variables.

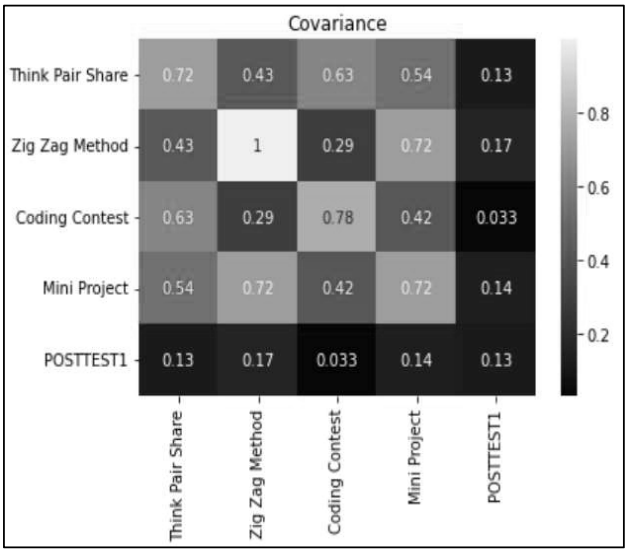

Fig. 5 : Covariance depiction.

$$cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N-1} \quad (8)$$

C. K-Means Clustering

Finally, data analytics is employed to look at the information from numerous perspectives. Clustering is an exploratory data analytics technique for gaining an intuitive understanding of the information structure. Discovering different and interesting patterns in data using clustering algorithms as a result, clustering is spoken as "knowledge discovery" or "pattern discovery."

The K-means clustering algorithm calculates the centroids and repeats the process until the best centroid is discovered. It presumes that the quantity of clusters is known beforehand. Additionally known as the flat clustering algorithm. The "K" in K-means stands for the algorithm's count of clusters that were extracted from the data.

Here we know the k value as 3, as we are grouping the teams under 3 classes average, fair, excellent. So to identify what class does a team belong to we use k-means approach. Since the goal of this approach is to cluster the data points in a way that minimizes the sum of the squared distances between the data points and the centroid. It is important to realize that less variations within the clusters will result in more similar data points within the same cluster. Here we group to compare and see the performance of teams in pretest and post test scores before and after implementing peer active learning strategies and note down the increase or decrease in the posttest scores.

K-Means Clustering is an iterative technique that splits an unlabeled dataset into k different clusters, with each dataset belonging to just one of those groups. It allows us to cluster data into different groups and provides a straight forward technique to work out the categories of groups in an unlabeled dataset with none training. It is a centroid-based approach, which implies that each cluster has its very own centroid. the most goal of this method is to reduce the sum of the distances between data points and thus the clusters that they belong to. The pie chart explains how students improved in three clusters of excellent, average, and fair percentages. In the cluster "Excellent," the percentage has been drastically improved from 20% to 50% in the case of pretest to posttest. Therefore, the cluster "Fair" percentage was reduced by 30% such that the student performance

was shifted to the clusters "Excellent" and "Average". But the pie chart with "assignment and posttest" shows the improvement of student performance in the cluster "Excellent" as well as "Average". It is because the cluster average plays an important role with respect to all the four activities conducted as well as the posttest scores. The main point to be noted is that the cluster average's percentage was not decreased but the range for the average cluster was changed, leading to having a constant value of 40% of students who belong to that cluster.

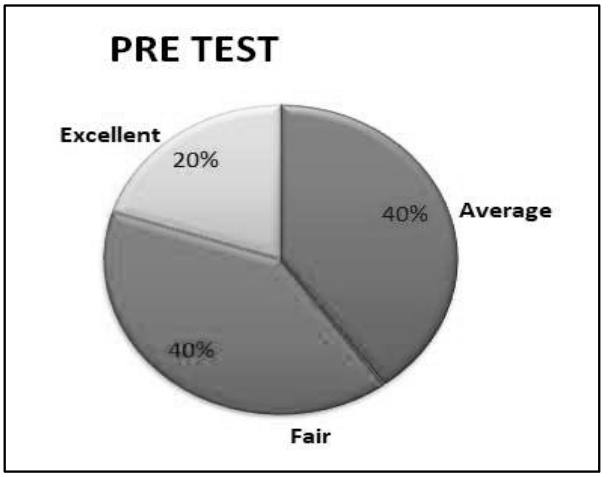

Fig. 6: Performance results of Pretest.

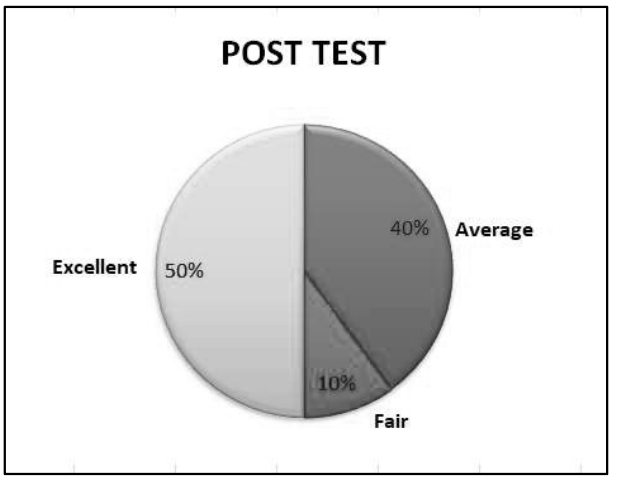

Fig. 7: Performance results of posttest.

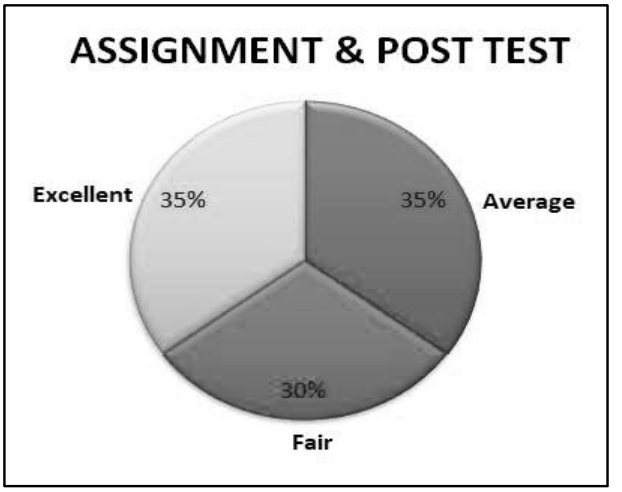

Fig.8:Performance results of posttest and assignments.

The pretest graph depicts the cluster of students in the range between 32 and 46. The posttest graph depicts the cluster of students in the range between 42 and 49. The main difference can be seen in the range of these graphs. Clearly, the posttest scores after the peer-active learning methods are very high, up to 49, whereas the pretest scores have only a high score of 46. And also, the other difference is that the lower cluster in the posttest has only 2 groups of students at a score of 42, whereas in the pretest it has 8 groups in the range of 32-38. Similarly, the other two clusters in both graphs also differ where the number of groups with the scores in the posttest is higher than the pretest clusters and the scores. So, here it is obvious that the posttest score range is high and also the number of students with high scores proves that peer-active learning methodology has improved their efficiency.

The post-test result's much improved compared to the pretest after implementing all the activities in peer learning and active learning. Using K-Means clustering, the student's performance is clustered into three clusters into classes: cluster0, cluster1 and cluster2. K-Means clustering results are visualized as pie charts for better acknowledgement, and the

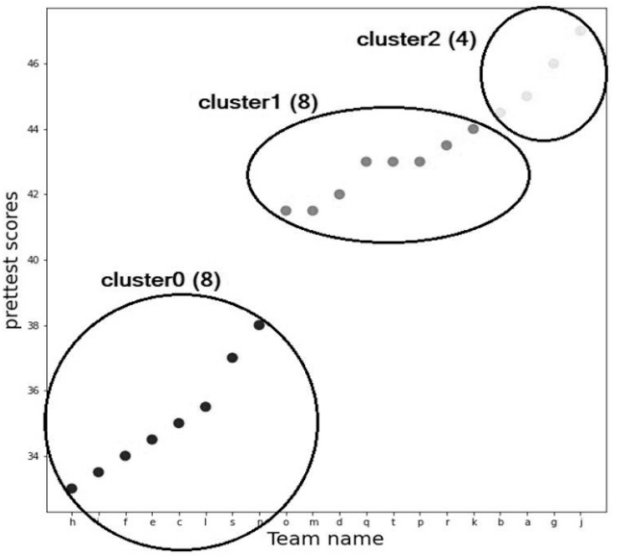

Fig. 9: Cluster's w.r.t pretest scores of students.

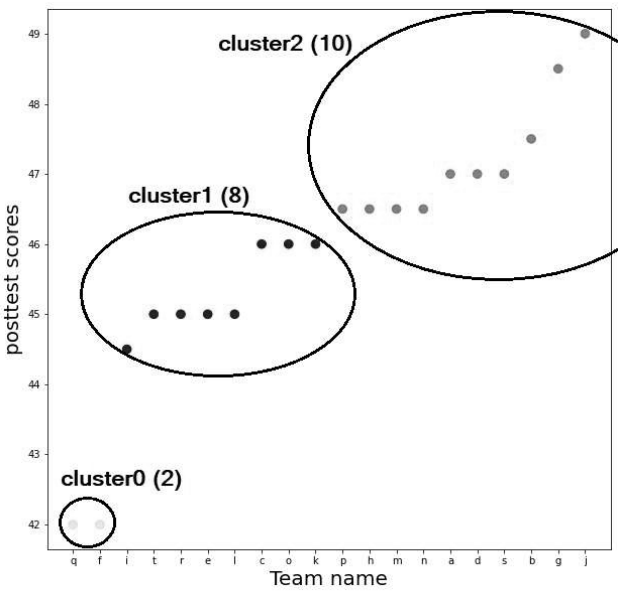

Fig. 10: Cluster's w.r.t posttest scores of students.

Table 1: Summary of k-means clustering algorithm

PARAMETERS	CLUSTER0 (FAIR)	CLUSTER 1 (GOOD)	CLUSTER2 (EXCELLENT)
Pretest scores	8	8	4
Considering all the 4 active and peer learning methods as well as posttest scores	5	7	8
Posttest scores	2	8	10

clusters are named as "Excellent", "Fair", and "Average" based on the test scores. The next part is comparing the student improvement from pretest to posttest using their scores and clusters formed. The scatter plot also defines the clusters such as “Excellent”, “Fair” and “Average”. But, here the scores ranges has also been changed. The change is nothing but the scores had been increased in such way the ranges moved from 34 to 42 in the lower boundary and in the upper boundary.

5. Discussion On Result

One method for enhancing pupils' subject-matter knowledge is peer-active learning. The clustering result makes it clear that students' understanding of problem solving in C programming has greatly increased. As a consequence, the students may learn the ideas through peer-active learning techniques including think-pair-share, the zig-zag approach, coding competitions, and mini-projects. Therefore, peer active learning techniques yield superior outcomes and speed up learning in the area of C programming problem solving.

The table shows a definite improvement in that the number of students in the cluster outstanding has progressively climbed from 4 to 10, while the number of students in the cluster fair has fallen from 8 to 2. The cluster excellent indicates that the students' knowledge had been greatly enhanced, thus even if the count remained constant, their scores had increased when the four activities were taken into account together with the posttest as well as when the posttest was taken in isolation. This is the outcome of the K-Means clustering algorithm for the best peer active learning strategies for problem solving using C programming.

6. Conclusion And Future Enhancement

This paper presents a better way of learning to solve problems using the programming language C. The process is implemented by conducting a pretest and a posttest, where the posttest plays the major role. It is because the posttest scores are high as four activities of active peer learning were conducted. These activities include think pair share, zig zag method, coding contest, and mini project.

Here, in order to prove that there is improved leaning towards the subject, K-Means Clustering algorithm is used. K-Means Clustering algorithm is unsupervised machine learning in nature, so this algorithm proves it by clustering the pretest and posttest scores. Finally, by concluding that although traditional methods play a vital role in learning the concepts, these methods of active peer learning encourage the students to acquire more knowledge about problem solving using C programming language faster and easily.

These best peer active learning techniques may be used to a variety of disciplines and topics to help students learn as much as they can in a short amount of time. Again, the student's performance for the topic studied utilizing clustering methods shows a comparable improvement session. This active peer learning approach is also applicable to subjects including theory and practice. This offers the benefit of enabling students' performance and knowledge to advance more quickly. Thus, this represents the finest peer active strategy's future development.

The course on "problem solving using C-language" aims to introduce the student to the fundamental ideas of the C programming language and also develops the skill to understand the mathematical concept behind the problems, analyze the logic and develop solutions error-free. In accordance to the "problem solving using C-language", the course outcomes in the subject which was considered plays the major role with the subject. The course outcomes can be compared with peer activities and a comparative study can be built further.

References

[1] Eickholt, J., Johnson, M. R., & Seeling, P. (2020). Practical active learning stations to transform existing learning environments into flexible, active learning classrooms. IEEE Transactions on Education, 64(2), 95-102.

[2] Kavitha, D., & Anitha, D. (2018). Flipped Classroom Using ICT Tools to Improve Outcome for the Course'Soft Computing'-A Case Study. Journal of Engineering Education Transformations, 32(2), 2349-2473.

[3] Chandrasekaran, J., & Anitha, D. (2021). Enhancing student learning and engagement in the course on computer networks. Journal of Engineering Education Transformations, 34, 454-463.

[4] Liu, C. Y., & Chen, H. L. (2020). Effects of peer learning on learning performance, motivation, and attitude. International Journal of Education Economics and Development, 11(4), 420-443.

[5] Wang, T. D., & Warshaw, L. Peer Learning in Data Analytics: Engaging Students and Developing Professional Skills.

[6] Leyk, T., McInvale, R., & Chen, L. (2017). Structured Peer Learning Program-An Innovative Approach to Computer Science Education. arXiv preprint arXiv:1703.04174.

[7] Lawson, W. G., Secules, S., Bhattacharyya, S., & Gupta, A. (2016, June). An Application-based Learning Approach to C Programming Concepts and Methods for Engineers. In 2016 ASEE Annual Conference & Exposition.

[8] Altintas, T., Gunes, A., & Sayan, H. (2016). A peer-assisted learning experience in computer programming language learning and developing computer programming skills. Innovations in Education and Teaching International, 53(3), 329-337.

[9] Sabarinath, R., & Quek, C. L. G. (2020). A case study investigating programming students' peer review of codes and their perceptions of the online learning environment. Education and Information Technologies, 25(5), 3553-3575.

[10] Canaleta, X., Vernet, D., Vicent, L., &Montero, J.A. (2014). Master in teacher training: A real implementation of active learning. Computers in human behavior, 31, 651-658.