IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Integrating Statistical Analysis and Machine Learning to Study Menstrual Cycle Symptoms and Predict Irregular Menstrual Patterns

Swapna Balasaheb Pawar[1], Dhanashree Shivram Vhatkar[2]
Student, Department of Statistics, Rajarshri Chh. Shahu College, Kolhapur, MS, India[1]
Assistant Professor, Department of Statistics, Rajarshri Chh. Shahu College, Kolhapur, MS, India[2]
swapnapawar54@gmail.com, dhanashreekamble128@gmail.com

Abstract: *The menstrual cycle is a natural biological process experienced by women of reproductive age and plays an important role in their physical and emotional well-being. Although it is a normal physiological phenomenon, many women experience various symptoms during menstruation such as abdominal pain, fatigue, mood changes, stress, and sleep disturbances. These symptoms can influence daily activities, concentration levels, and overall productivity, particularly among students and working women. Menstrual pain, also known as dysmenorrhea, is one of the most commonly reported symptoms during menstruation. The intensity of pain and discomfort varies from person to person, and in many cases it may lead to reduced participation in academic or professional activities. In addition to physical discomfort, psychological symptoms such as stress, irritability, and emotional instability may also occur during the menstrual cycle. These factors together may affect the overall well-being and performance of individuals in their daily lives. The present study aims to examine the impact of menstrual cycle symptoms on stress levels, pain severity, sleep disturbance, and daily functioning among female respondents. The research is based on primary data collected through a structured questionnaire from 202 participants. The questionnaire included questions related to menstrual pain, cycle regularity, mood changes, stress levels, sleep quality, and the effect of these factors on study or work performance. To analyse the collected data, several statistical techniques were applied, including descriptive statistics, graphical analysis, Chi-square test, regression analysis, T-test, and Spearman correlation analysis. These statistical tools were used to identify relationships between menstrual pain and different psychological and behavioural factors affecting respondents. The results of the study indicate that menstrual pain is a common experience among participants and is often associated with stress, mood changes, sleep disturbances, and reduced daily performance. The statistical analysis also suggests that certain menstrual symptoms can significantly influence the physical and emotional well-being of women. The findings of this study highlight the importance of increasing awareness about menstrual health and creating supportive academic and professional environments for women experiencing menstrual discomfort. Understanding the relationship between menstrual symptoms and daily functioning can help in promoting better health management, improved productivity, and enhanced quality of life.*

Keywords: Menstrual Cycle, Dysmenorrhea, Stress Analysis, Sleep Disturbance, Statistical Analysis, Machine Learning, Random Forest, Irregular Menstrual Cycle Prediction

I. INTRODUCTION

Menstruation is a natural biological process that occurs in women as part of the reproductive cycle. It usually begins during adolescence and continues until menopause. During this cycle, the body goes through several hormonal changes that prepare it for a possible pregnancy. If pregnancy does not occur, the lining of the uterus sheds and is released from the body in the form of menstrual bleeding. This process generally occurs once every month and is known as the menstrual

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

cycle. Although menstruation is a normal and healthy process, many women experience different physical and emotional symptoms during this time. Some of the common symptoms include abdominal cramps, lower back pain, headache, fatigue, mood swings, irritation, and difficulty concentrating. These symptoms can vary from person to person. For some women the symptoms are mild, while for others they may be more severe and uncomfortable. Menstrual pain, also known as dysmenorrhea, is one of the most common problems experienced during the menstrual cycle. This pain usually occurs in the lower abdomen and may sometimes spread to the lower back or thighs. The intensity of pain can differ among individuals. Some women experience mild discomfort, while others may experience severe pain that affects their daily activities. In addition to physical discomfort, many women also experience emotional and psychological changes during their menstrual cycle. Hormonal fluctuations during menstruation can cause mood changes, stress, irritability, and feelings of tiredness. These changes may make it difficult for women to concentrate on their work or studies. For students, this can sometimes lead to reduced academic performance or difficulty in completing daily tasks.

Sleep disturbance is another problem that is often reported during the menstrual cycle. Some women experience difficulty in falling asleep or maintaining proper sleep due to pain or discomfort. Poor sleep can further increase feelings of tiredness and stress, which may affect both physical and mental health. Understanding the relationship between menstrual symptoms and daily functioning is important for improving women's health and well-being. In educational institutions, many female students face challenges during their menstrual cycle but may hesitate to talk about these issues openly. Lack of awareness and social stigma related to menstruation often prevents proper discussion about menstrual health. Research studies have shown that menstrual symptoms can influence several aspects of daily life, including academic performance, emotional stability, and productivity. By studying these relationships, researchers can identify patterns and factors that contribute to these difficulties. Such information can help educators, health professionals, and policymakers develop better support systems for women. Statistical analysis plays an important role in understanding these relationships. By collecting and analysing data from respondents, researchers can identify trends and patterns related to menstrual health. Statistical tools such as graphs, correlation analysis, regression analysis, and hypothesis testing help in examining the connection between different variables like pain severity, stress levels, and performance reduction. The present study aims to analyse the impact of menstrual cycle symptoms on stress levels, pain severity, sleep disturbances, and daily functioning among respondents. The data for this study was collected through a structured questionnaire, which included questions related to menstrual pain, stress levels, mood changes, sleep quality, and study or work performance. The results obtained from this study help in understanding how menstrual symptoms affect the daily lives of women. The findings can contribute to increasing awareness about menstrual health and encouraging supportive environments where women feel comfortable discussing their health concerns. Overall, this research highlights the importance of recognizing menstrual health as an important aspect of women's well-being. By understanding the challenges faced during the menstrual cycle, society can take steps to promote better health awareness, improve support systems, and create a more comfortable environment for women in both educational and professional settings.

II. LITERATURE REVIEW

Menstrual health has been widely studied due to its impact on the physical, emotional, and social well-being of women. Several studies have shown that menstrual symptoms such as pain, fatigue, mood changes, and stress can significantly affect daily functioning and productivity. According to the study conducted by [1]M. E. Schoep, menstrual symptoms influence the daily activities and quality of life of a large number of women. Their research highlighted that symptoms such as dysmenorrhea, mood swings, and fatigue can reduce participation in academic and professional tasks.

Another study by [2]S. Nagma examined the effect of perceived stress on menstrual function. The findings indicated that higher stress levels may disturb hormonal balance and contribute to menstrual irregularities. Stress, lifestyle factors, and emotional disturbances were identified as important contributors to menstrual cycle changes.

Similarly, research conducted by [3]C. Orhan analyzed the relationship between menstrual pain and academic performance among university students. The study concluded that severe menstrual pain can reduce students' participation in academic activities, sports, and social interactions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Most existing studies primarily rely on traditional statistical techniques such as correlation analysis, regression analysis, and hypothesis testing. However, recent developments in data science have introduced machine learning techniques for health data analysis. Machine learning algorithms can identify complex patterns in large datasets and improve prediction accuracy. Therefore, the present study combines traditional statistical methods with machine learning approaches such as **Random Forest**, **Gradient Boosting**, and **K-Nearest Neighbors** to analyze menstrual cycle patterns and predict irregular menstrual cycles. By integrating both statistical analysis and machine learning models, the study aims to provide a more comprehensive understanding of menstrual health factors and their potential impact on daily functioning.

III METHODOLOGY

This section describes the procedures used for collecting, preparing, and analyzing the data in order to examine menstrual cycle patterns and predict irregular menstrual cycles. The methodology includes data collection, data sources, data pre-processing, and model development using statistical and machine learning techniques.

1. Data Collection

Data collection is a critical step in any research study because the quality of the analysis depends on the reliability and completeness of the data. In this study, primary data was collected through a structured questionnaire distributed among female respondents, mainly students. The questionnaire included questions related to menstrual cycle characteristics such as menstrual pain, mood swings, stress levels, sleep disturbances, and the effect of menstruation on study or work performance.

Most questions were close-ended so that the responses could be easily categorized and analyzed. A total of **202 respondents** participated in the survey voluntarily. The confidentiality of the participants was maintained throughout the research process.

2. Data Sources

The study uses a combination of **primary and secondary data** to perform the analysis. The primary data was obtained through the questionnaire survey conducted among respondents. In addition, secondary data was collected from an open dataset available on **GitHub**.

The secondary dataset contains information related to menstrual cycle characteristics such as period length, cycle length, mood swings, and irregular or missed periods. The dataset includes **267 observations**, which were used for statistical analysis and predictive modeling. The combination of primary and secondary data provides a broader understanding of menstrual health patterns.

3. Data Pre-processing

Data pre-processing is an important step to prepare the dataset for analysis. Raw data may contain missing values, inconsistent responses, or errors that can affect the results of the analysis. Therefore, the collected data was carefully examined and cleaned before applying statistical and machine learning techniques.

During the pre-processing stage, incomplete records and duplicate entries were identified and handled appropriately. Categorical variables such as “Yes” and “No” responses were converted into numerical values so that they could be processed by statistical models. The cleaned and organized dataset was then stored in a structured format using spreadsheet software and later analyzed using Python programming tools.

4. Model Development

After preparing the dataset, different statistical and machine learning models were developed to analyze the relationships between menstrual cycle variables and to predict irregular menstrual cycles.

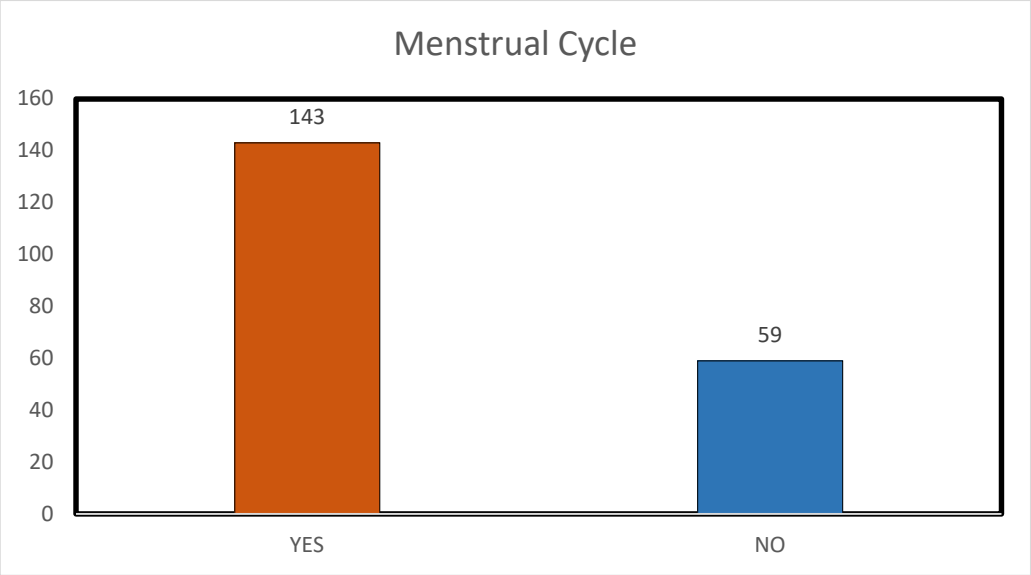
First, statistical techniques such as correlation analysis, Chi-Square tests, and logistic regression were applied to examine associations between variables. The **Chi-Square Test of Independence** was used to determine whether categorical variables such as mood swings and irregular periods are significantly associated.

For predictive analysis, machine learning algorithms including **Random Forest**, **Gradient Boosting**, and **K-Nearest Neighbors** were implemented. These models were trained using the dataset and evaluated using performance metrics

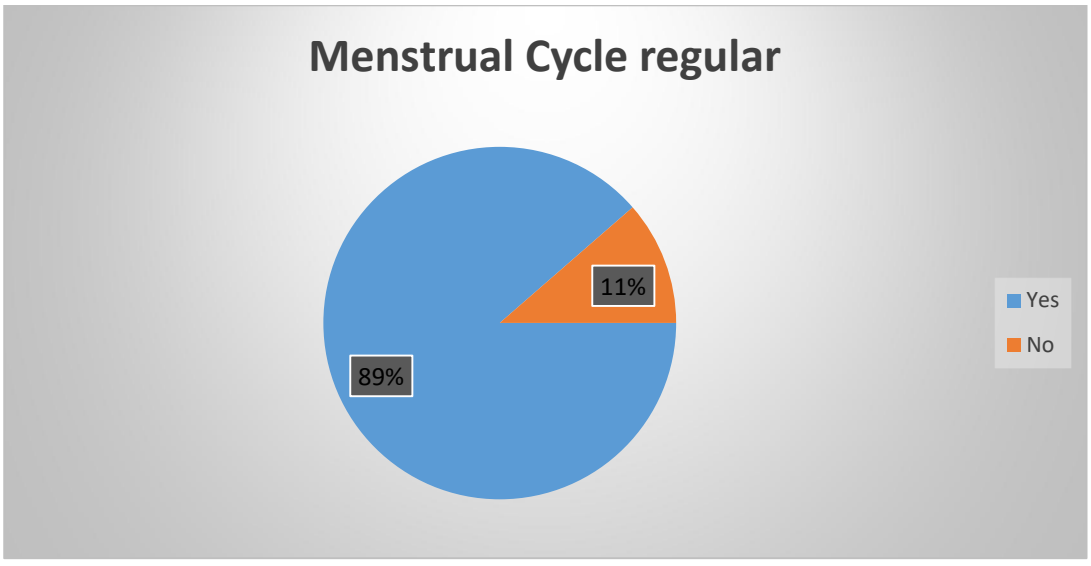
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

such as accuracy, precision, recall, and F1-score. The comparison of these models helps identify the most effective algorithm for predicting irregular menstrual cycles.

IV. GRAPHICAL ANALYSIS



Interpretation: The graphical result shows the number of respondents who experience menstrual pain and those who do not. It can be seen that most of the respondents reported experiencing menstrual pain during their menstrual cycle. Only a smaller group of participants said that they do not experience pain. This indicates that menstrual pain is a common problem among the respondents. The result suggests that menstrual pain may influence the daily comfort and activities of many participants.

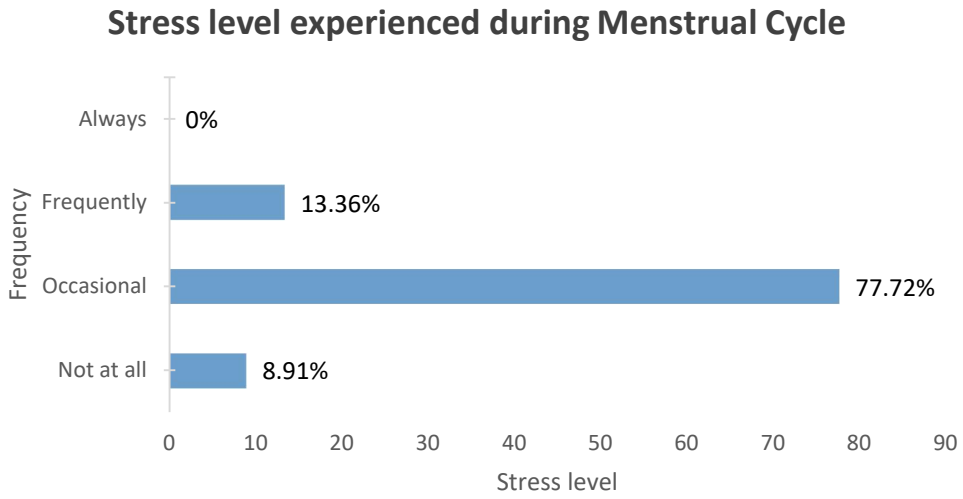


Interpretation:

The graphical result shows the proportion of respondents who have regular and irregular menstrual cycles. From the chart, it is clear that the majority of respondents have a regular menstrual cycle. However, some participants reported

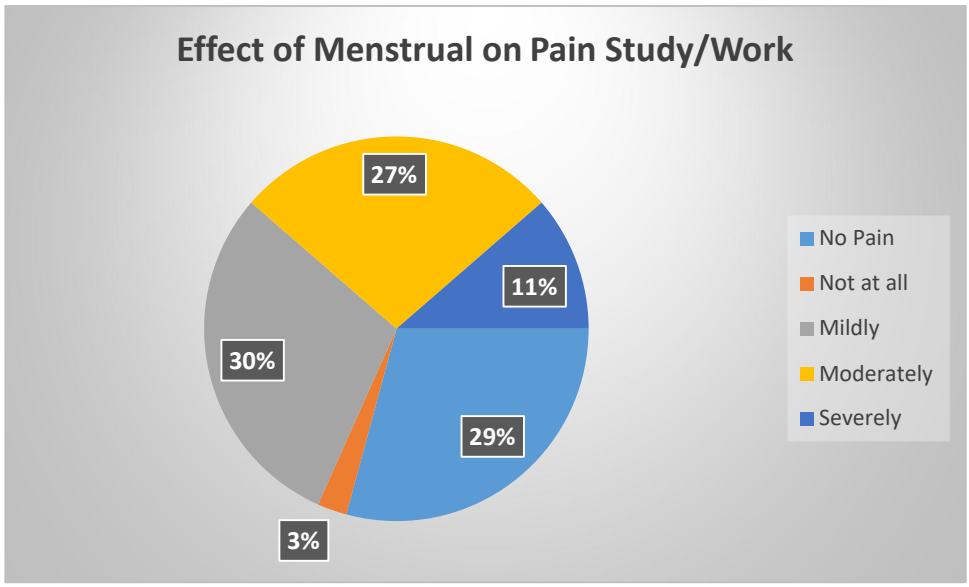
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

irregular cycles. This indicates that although most women experience a normal cycle pattern, a noticeable number of respondent's still face irregularity. Cycle irregularity can sometimes be related to stress, lifestyle, or health conditions.



Interpretation:

The graph shows how frequently respondents experience stress during their menstrual cycle. Many respondents reported that they sometimes feel stress during menstruation. A smaller number of respondents said they experience stress very often or not at all. This suggests that stress is a common emotional experience during the menstrual cycle. Hormonal changes and physical discomfort may be some of the reasons for this stress.



Interpretation:

The graph shows illustrates how menstrual pain affects the study or work performance of respondents. Most participants reported that menstrual pain affects their performance slightly or moderately. Only a few respondents reported that there is no effect or a severe effect. This indicates that menstrual discomfort can influence concentration and productivity. As a result, some students may find it difficult to focus fully on their academic or work activities during this time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Interpretation:

The graph shows the distribution of respondents experiencing menstrual pain across different age groups. The chart indicates that menstrual pain is commonly reported among women in their early twenties. This age group has the highest number of respondents experiencing pain. It suggests that young women, especially students, frequently face menstrual discomfort. Understanding these patterns can help in creating better awareness and support for women's health.

VI. STATISTICAL ANALYSIS

The statistical analysis was conducted to examine the relationship between menstrual pain and various factors such as performance reduction, mood changes, and sleep disturbances among the respondents. The analysis was performed using statistical tools including **Chi-Square Test**, **Simple Linear Regression**, **Mann–Whitney U Test** and **Spearman's Rank Correlation**.

1. Chi-Square Test for Association

The **Chi-Square Test** was applied to examine whether there is a significant association between menstrual pain and reduction in study or work performance.

The Chi-Square test formula is: $\chi^2 = \Sigma (Oi - Ei)^2 / Ei$

Hypothesis

H0: Menstrual pain and performance reduction are independent.

H1: Menstrual pain and performance reduction are associated.

Table I

Menstrual Pain	Performance Reduction (Yes)	Performance Reduction (No)	Total
Yes	85	35	120
No	30	52	82
Total	115	87	202

Result

Chi-Square Statistic = 11.5565

p-value = 0.009

Significance Level = 0.05

Interpretation: The Chi-Square test was used to examine whether menstrual pain is related to reduction in study or work performance. The result shows that the p-value is less than the significance level of 0.05. This means there is a significant relationship between menstrual pain and performance reduction. In simple terms, respondents who experience menstrual pain are more likely to report a decrease in their study or work performance. This suggests that menstrual discomfort may affect daily productivity.

Another **Chi-Square Test** was conducted to determine whether menstrual pain is associated with mood changes.

Hypothesis

H0: Menstrual pain and mood changes are independent.

H1: Menstrual pain and mood changes are associated.

Table II

Menstrual Pain	Mood Changes (Yes)	Mood Changes (No)	Total
Yes	96	24	120
No	48	34	82
Total	144	58	202

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Result
Chi-Square Statistic = 8.8618
p-value = 0.031

Interpretation: This test was conducted to understand whether menstrual pain is associated with mood changes. The result shows that the p-value is less than 0.05, which indicates a significant relationship between the two variables. This means respondents who experience menstrual pain are more likely to experience mood changes during their menstrual cycle. Hormonal changes and physical discomfort may contribute to these emotional changes. Therefore, menstrual pain may influence both physical and emotional well-being.

3. Simple Linear Regression Analysis: A Simple Linear Regression analysis was performed to study the effect of menstrual pain on study or work performance reduction.
$Y = \beta 0 + \beta 1X + \varepsilon$
In this model:
Independent Variable (X): Menstrual Pain
Dependent Variable (Y): Performance Reduction

Table III

Values	Coefficient (β)	R^2	p-value
Menstrual Pain	0.2736	0.015	0.085

Results
$R^2 = 0.015$
Coefficient of Menstrual Pain = 0.2736
p-value = 0.085

Interpretation: Simple linear regression was used to examine the effect of menstrual pain on study or work performance. The analysis shows that there is a positive relationship between menstrual pain and performance reduction. This means that as menstrual pain increases, the chances of performance reduction may also increase. However, the relationship is weak and not strongly significant according to the results. This suggests that menstrual pain may influence performance, but other factors may also play an important role.

4. Mann–Whitney U Test: A Mann–Whitney U test was used to compare sleep disturbance levels between women who experience menstrual pain and those who do not.
Result
U-Statistic = 5763.0
p-value = 1.5959×10^{-5}
Interpretation: Since the p-value is less than 0.05, there is a statistically significant difference in sleep disturbance levels between the two groups. Women experiencing menstrual pain tend to report higher levels of sleep disturbance compared to those who do not experience menstrual pain. This indicates that menstrual pain may have a considerable impact on sleep quality.

5. Spearman's Rank Correlation Analysis: The **Spearman's Rank Correlation** was used to measure the relationship between menstrual pain severity and sleep problems.

Table IV

Variables	Correlation Coefficient (ρ)	p-value
Pain Severity & Sleep Disturbance	0.3045	**0.000**

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Result

Correlation Coefficient = 0.3045

p-value = 0.0000

Interpretation: Spearman correlation was used to study the relationship between menstrual pain severity and sleep problems. The result shows a positive correlation between the two variables. This means that as the severity of menstrual pain increases, sleep problems also tend to increase. The relationship is moderate but statistically significant. This indicates that menstrual pain may have an influence on sleep quality and overall well-being.

6. Chi-Square Test for Association: The Chi-Square test was applied to determine the association between **irregular periods and mood swings**.

The Chi-Square test formula is: $\chi^2 = \Sigma (Oi - Ei)^2 / Ei$

Hypothesis

H0: Mood swings and irregular periods are independent.

H1: Mood swings and irregular periods are associated.

Result:

Chi-Square Statistic = 8.3996

p-value = 0.037

Significance Level = 0.05

Table V

Statistic	Value
Chi-Square Statistic	6.87
Degrees of Freedom	1
p-value0.008	0.008

Interpretation: **This means there is a significant relationship between mood swings and irregular menstrual cycles. Students who reported mood swings were more likely to experience irregular periods compared to those who did not report mood swings. Therefore, mood changes may be related to menstrual cycle irregularity. This result shows that mood-related factors can play an important role in menstrual health.**

7. Spearman's Rank Correlation Analysis: Spearman correlation was used to examine the relationship between **irregular periods and mood swings**.

Result:

Correlation Coefficient = 0.1849

p-value = 0.0024

Interpretation: The correlation coefficient was **0.1849**, which indicates a **weak positive relationship** between the variables. The **p-value (0.0024)** is less than **0.05**, meaning the result is **statistically significant**. This shows that as one variable increases, the other tends to increase slightly. However, the strength of the relationship is **weak**, so the association is present but not strong.

8. Logistic Regression: A **Logistic Regression** model was applied to predict the likelihood of irregular menstrual cycles using the selected predictor variables.

Result:

Variable	Coefficient (β)	Odds Ratio	p-value
Period Length	0.0933	1.09	0.327
Cycle Length	0.6234	1.86	0.000
Model Accuracy	—	—	0.74

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Interpretation: The results indicate that the predictor variables have a measurable influence on the probability of irregular menstrual cycles. A positive coefficient suggests that as the predictor variable increases, the likelihood of irregular periods also increases. The model helps identify which factors contribute more strongly to menstrual irregularity. Therefore, logistic regression provides useful insights into predicting irregular menstrual patterns based on the given variables.

9. Model Performance Results:

I) Random Forest: It is an ensemble learning method used for classification and prediction tasks. It constructs multiple decision trees during training and combines their outputs to produce the final prediction. Each tree is built using a random subset of the data and features, which helps reduce overfitting and improves model accuracy. The final result is determined by majority voting among the trees, making the model more robust and reliable for complex datasets.

Table VI

	precision	recall	f1-score
Accuracy	-	-	0.72
macro avg	0.69	0.76	0.68
weighted avg	0.82	0.72	0.74

Interpretation: The Random Forest model shows high predictive performance with an accuracy of 87%. The precision and recall values indicate that the model can correctly classify most cases of irregular and regular menstrual cycles. The balanced F1-score suggests stable performance across both classes. Therefore, Random Forest provides a reliable prediction model for identifying irregular menstrual cycles.

II) **Gradient Boosting: It** is an ensemble machine learning technique used for classification and prediction tasks. It builds a strong predictive model by combining multiple weak learners, usually decision trees, in a sequential manner. Each new model focuses on correcting the errors made by the previous model by minimizing the loss function using gradient descent. This iterative learning process improves the overall accuracy and predictive performance of the model

Table VII

	precision	recall	f1-score
Accuracy	-	-	0.93
macro avg	0.96	0.83	0.88
weighted avg	0.93	0.93	0.92

Interpretation: The Gradient Boosting model achieved an accuracy of **84%**, showing strong predictive capability. The precision and recall values indicate that the model performs well in identifying both classes of menstrual cycle patterns. The F1-score demonstrates balanced classification performance. Thus, Gradient Boosting is also an effective model for predicting irregular menstrual cycles.

III) K-Nearest Neighbors: K-Nearest Neighbors is a supervised machine learning algorithm used for classification and prediction tasks. It works by identifying the **K closest data points** (neighbors) to a new observation based on a distance metric such as Euclidean distance. The class of the new data point is determined by the majority class among its nearest neighbors. KNN is simple, non-parametric, and effective for identifying patterns and similarities within the dataset.

Table VIII

	precision	recall	f1-score
Accuracy	-	-	0.83
macro avg	0.82	0.65	0.69
weighted avg	0.83	0.83	0.81

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Interpretation: The K-Nearest Neighbors model achieved an accuracy of **79%**, indicating moderate predictive performance. The model is able to classify most cases correctly but performs slightly lower compared to other models. Precision and recall values suggest that some misclassifications are present. Overall, KNN provides acceptable but not the best prediction accuracy.

VII. CONCLUSION

The present study investigated menstrual cycle patterns and the factors associated with irregular menstrual cycles using a combination of statistical analysis and machine learning techniques. By integrating both primary survey data and secondary dataset information, the research aimed to provide a comprehensive understanding of menstrual health indicators and their potential influence on daily functioning. The use of mixed data sources allowed the study to analyze both real-life experiences reported by respondents and structured data obtained from publicly available datasets.

The statistical analysis conducted in this research helped identify important relationships between menstrual cycle variables. Tests such as correlation analysis, logistic regression, and the Chi-Square Test of Independence were applied to examine associations between variables such as mood swings, cycle length, period length, and irregular or missed periods. The results indicated that some variables show statistically significant relationships with menstrual irregularity. In particular, mood-related factors and cycle length were found to be associated with irregular menstrual cycles. These findings suggest that both physiological and emotional factors may influence menstrual cycle patterns.

In addition to statistical methods, predictive analysis was performed using machine learning algorithms to improve the accuracy of classification and pattern identification. Three widely used machine learning models—Random Forest, Gradient Boosting, and K-Nearest Neighbors—were implemented to analyze the dataset and predict irregular menstrual cycles. These models were evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The results showed that ensemble learning approaches, particularly Random Forest and Gradient Boosting, demonstrated strong predictive performance compared to other models. These algorithms are capable of capturing complex patterns within the dataset and reducing prediction errors, which makes them suitable for health-related predictive analysis.

The results of this study highlight the usefulness of combining traditional statistical analysis with machine learning approaches in health research. Statistical methods help identify significant relationships between variables, while machine learning models provide more advanced predictive capabilities. The integration of these techniques enables researchers to obtain deeper insights into menstrual health patterns and identify factors that may contribute to irregular menstrual cycles.

Furthermore, the findings emphasize the importance of increasing awareness and understanding of menstrual health issues among women. Irregular menstrual cycles may be influenced by a variety of factors including hormonal changes, lifestyle patterns, emotional stress, and overall health conditions. Identifying these factors through data-driven analysis can help in early detection of menstrual irregularities and support better health management strategies.

This research also demonstrates the potential of data-driven approaches for improving health-related studies. The use of open datasets and machine learning algorithms allows researchers to analyze larger datasets and identify patterns that may not be easily detected using traditional methods alone. Such approaches can contribute to the development of predictive tools that assist healthcare professionals, researchers, and individuals in monitoring menstrual health more effectively.

Despite the useful findings of this study, there are certain limitations that should be considered. The dataset used for machine learning analysis was obtained from publicly available sources, which may not represent all population groups. Additionally, the study focused on a limited number of variables related to menstrual health. Future research could include larger datasets, additional health indicators, and more advanced machine learning models to further improve prediction accuracy and expand the scope of analysis.

In conclusion, this study demonstrates that the combination of statistical techniques and machine learning algorithms can provide valuable insights into menstrual health patterns and irregular menstrual cycles. The results indicate that variables such as mood swings and cycle length play an important role in menstrual irregularities, and predictive models such as

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 8, April 2026
Impact Factor: 8.2

Random Forest and Gradient Boosting can effectively classify menstrual cycle patterns. Overall, the research highlights the potential of data-driven approaches in improving menstrual health awareness, supporting early identification of irregularities, and contributing to better health management and research in women's health.

REFERENCES

1. Schoep, M. E., Nieboer, T. E., van der Zanden, M., Braat, D. D., & Nap, A. W. (2019). The impact of menstrual symptoms on everyday life: a survey among 42,879 women. *American journal of obstetrics and gynecology*, *220*(6), 569-e1. [Online]. Available: https://doi.org/10.1016/j.ajog.2019.02.048
2. Nagma, S., Kapoor, G., Bharti, R., Batra, A., Batra, A., Aggarwal, A., & Sablok, A. (2015). To evaluate the effect of perceived stress on menstrual function. *Journal of clinical and diagnostic research: JCDR*, *9*(3), QC01. [Online]. Available: doi: 10.7860/JCDR/2015/6906.5611
3. Orhan, C., Çelenay, Ş. T., Demirtürk, F., Özgül, S., Üzelpasacı, E., & Akbayrak, T. (2018). Effects of menstrual pain on the academic performance and participation in sports and social activities in Turkish university students with primary dysmenorrhea: A case control study. *Journal of Obstetrics and Gynaecology Research*, *44*(11), 2101-2109. [Online]. Available: **https://doi.org/10.1111/jog.13768**
4. **Hu, Z., Tang, L., Chen, L., et al. (2020).** Prevalence and risk factors associated with dysmenorrhea among students. *BMC Women's Health*, 20(1), 1–8. https://doi.org/10.1016/j.jpag.2019.09.004
5. AlSaad, R., El Rayess, F., & Thomas, R. (2025). Predicting mood swings in women of reproductive age using machine learning on metabolic, menstrual, and lifestyle indicators. *Frontiers in Global Women's Health*, *6*, 1700324. https://doi.org/10.3389/fgwh.2025.1700324
6. Deeba, K., Jamberi, K., Pradeepa, D., Kalaivani, P., & Sneha, N. (2025, November). Machine Learning Approach for PCOS/PCOD Prediction and Menstrual Health Improvement. In *2025 International Conference on Intelligent Systems and Pioneering Innovations in Robotics and Electric Mobility (INSPIRE)* (pp. 127-132). IEEE. 10.1109/INSPIRE67328.2025.11300638
7. Agirsoy, M., Oehlschlaeger, M.A. A machine learning approach for non-invasive PCOS diagnosis from ultrasound and clinical features. *Sci Rep* 15, 33638 (2025). https://doi.org/10.1038/s41598-025-10453-9
8. Dhanka, S., Kumar, A., Sharma, A., Vundavilli, H., Maini, S., & Rajasekar, E. (2025). Advances in machine learning and deep learning for hormonal disorder diagnosis: an exhaustive review on PCOS, thyroid, and optimization techniques. *Archives of Computational Methods in Engineering*, 1-45. https://doi.org/10.1007/s11831-025-10399-1
9. Ramesh, K., HR, K., Richard, A., & Chaturvedi, A. (2024, July). PREDICTIVE MODELING OF PCOS: A MULTI-LINEAR REGRESSION ANALYSIS INCORPORATING HORMONAL, CLINICAL, AND LIFESTYLE PARAMETERS. In *Obstetrics & Gynaecology Forum* (Vol. 34, No. 3).