

12-1-2024

## Toward a globally lunar calendar: a machine learning-driven approach for crescent moon visibility prediction

Samia Loucif  
*Zayed University*

Murad Al-Rajab  
*Abu Dhabi University*

Raed Abu Zitar  
*Liwa College*

Mahmoud Rezk  
*Dubai Electricity And Water Authority*

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Computer Sciences Commons](#)

---

### Recommended Citation

Loucif, Samia; Al-Rajab, Murad; Abu Zitar, Raed; and Rezk, Mahmoud, "Toward a globally lunar calendar: a machine learning-driven approach for crescent moon visibility prediction" (2024). *All Works*. 6827.  
<https://zuscholars.zu.ac.ae/works/6827>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact [scholars@zu.ac.ae](mailto:scholars@zu.ac.ae).

RESEARCH

Open Access



# Toward a globally lunar calendar: a machine learning-driven approach for crescent moon visibility prediction

Samia Loucif<sup>1</sup>, Murad Al-Rajab<sup>2\*</sup>, Raed Abu Zitar<sup>3</sup> and Mahmoud Rezk<sup>4</sup>

\*Correspondence:  
murad.al-rajab@adu.ac.ae

<sup>1</sup> Zayed University, Abu Dhabi, United Arab Emirates

<sup>2</sup> Abu Dhabi University, Abu Dhabi, United Arab Emirates

<sup>3</sup> Liwa College, Abu Dhabi, United Arab Emirates

<sup>4</sup> Dubai Electricity and Water Authority R&D, Dubai, United Arab Emirates

## Abstract

This paper presents a comprehensive approach to harmonizing lunar calendars across different global regions, addressing the long-standing challenge of variations in new crescent Moon sightings that mark the beginning of lunar months. We propose a machine learning (ML)-based framework to predict the visibility of the new crescent Moon, representing a significant advancement toward a globally unified lunar calendar. Our study utilized a dataset covering various countries globally, making it the first to analyze all 12 lunar months over a span of 13 years. We applied a wide array of ML algorithms and techniques. These techniques included feature selection, hyperparameter tuning, ensemble learning, and region-based clustering, all aimed at maximizing the model's performance. The overall results reveal that the gradient boosting (GB) model surpasses all other models, achieving the highest F1 score of 0.882469 and an area under the curve (AUC) of 0.901009. However, with selected features identified through the ANOVA F-test and optimized parameters, the Extra Trees model exhibited the best performance with an F1 score of 0.887872, and an AUC of 0.906242. We expanded our analysis to explore ensemble models, aiming to understand how a combination of models might boost predictive accuracy. The Ensemble Model exhibited a slight improvement, with an F1 score of 0.888058 and an AUC of 0.907482. Additionally, the geographical segmentation of the dataset enhanced predictive performance in certain areas, such as Africa and Asia. In conclusion, ML techniques can provide efficient and reliable tool for predicting the new crescent Moon visibility that would support the decisions of marking the beginning of new lunar months.

**Keywords:** Crescent visibility, Machine learning, Lunar calendar, Prediction, Hyperparameter tuning, Feature selection, Ensemble learning, Astronomy

## Introduction

Approximately 24.1% of the global population, including Muslims [1], adhere to the lunar calendar, which plays a crucial role in defining the start of each month, initiated by the sighting of the new crescent Moon. This method of timekeeping is integral to various religious practices around the world, underscoring the significance of the crescent Moon as a symbol marking religious events. Precise observation of the crescent Moon is

pivotal, as it signifies the onset of new periods in the lunar calendar, influencing numerous religious celebrations.

The importance of the crescent Moon in the lunar calendar can be summarized as follows; determination of months as the lunar calendar is based on the cycles of the Moon, and the sighting of the crescent Moon is used to establish the beginning of each month. The tradition of sighting the crescent Moon is a visual confirmation of the new month. It involves observing the thin crescent Moon shortly after Sunset, and the testimony of reliable witnesses is often used to confirm the sighting. The crescent Moon sighting involves the community, fostering a sense of unity and shared religious observance. In many Muslim-majority countries, committees or religious authorities are tasked with officially announcing the sighting of the crescent Moon. The sighting of the crescent Moon is especially significant during the month of Ramadan. The first sighting of the crescent Moon determines the start of Ramadan, the month of fasting, reflection, and increased devotion for Muslims around the world. The sighting of the crescent Moon also plays a crucial role in determining the end of Ramadan and the celebration of Eid al-Fitr. The month of fasting concludes with the sighting of the crescent Moon, and the next day is celebrated as a day of festivity and thanksgiving. Similarly, the crescent Moon is observed to determine the beginning of the month of Dhu al-Hijjah, during which the Hajj pilgrimage takes place. The 10th day of Dhu al-Hijjah, known as Eid al-Adha, also relies on the sighting of the crescent Moon to mark the end of Hajj and the beginning of the festival of sacrifice. In summary, the crescent Moon holds significant importance in Islam as it is used to determine the start of each lunar month. The practice of crescent Moon sighting is deeply rooted in Islamic tradition and serves as a communal and visual way to mark the passage of time and the beginning of important religious observances.

This paper investigates the use of supervised ML techniques in predicting the visibility of the new crescent Moon. For those ML systems, there are several input features related to geographical aspects such as the longitude and the latitude, and other features such as the atmosphere, Moonset, Sun\_Moon\_lag, age\_of\_Moon, Moon\_altitude, Sun\_altitude, Moon\_azimuth, Sun\_azimuth, elongation, illumination, and conjunction time. The target or the output is the fact that the crescent was seen by the eye or not. Other inputs may be seen by other devices such as the binocular (V\_bino), telescope (V\_tele), Charged Coupled Devices (CCD) camera (V\_ccd). Several ML techniques will be demonstrated in this work, especially the ones that we call static ML systems that do not require deep processing of data or temporal inputs [2]. Some of the models that are used in this paper are: Support Vector Machine (SVM), Decision Tree, AdaBoost, Random Forest, logistic regression, Naïve Bayes, KNN, and Extra Trees. As in most ML problems, the first stage will be feature selection. The training will be applied for the whole 12 months for the prediction of visibility of the new crescent Moon. The training will also be repeated based on classified/clustered continents and regions. Therefore, we may have a predictor for every region rather than having a single predictor for all regions in the dataset. This is an attempt to increase the accuracy by using the same predictor for geographically close reasons that are expected to have the new crescent Moon seen at the same time [3]. Moreover, in order to improve the accuracy, ensembled and hybrid techniques will be used. Meta heuristic optimization such as the Genetic Algorithm (GA) and the Particle Swarm Intelligence (PSO) are used to optimize the hyper parameters of the best

performing ML algorithms [4]. Meta heuristic optimization such as the Genetic Algorithm (GA) and the Particle Swarm Intelligence (PSO) are used to optimize the hyper parameters of the best performing ML algorithms. GA have proven successful in tackling numerous optimization challenges. PSO, on the other hand is known for its simplicity and efficiency in continuous optimization tasks. Both algorithms have shown high performance across a wide range of problems [4–7]. The idea is to, at least, reach fast convergence to the optimum hyperparameters if not much improvement has been achieved in the accuracy. This convergence to the optimum parameters is expected to be faster than the exhaustive grid search method.

Detecting the crescent using our proposed ML methods will not replace the traditional Islamic method of using the naked eye to “see” the new crescent Moon. For religious occasions such as the Holy month of Ramadan and the Eids, the new crescent Moon has to be seen by the naked eye. The ML outcome will definitely support any claims that the new crescent Moon has been seen. The sole target used in training the ML algorithms is  $V_{eye}$  target (i.e., 0 not seen by eye or 1 seen by eye). Those values are usually captured from previous experiences. Therefore, the datasets on those ML methods have high credibility. Therefore, the proposed methods will train the predictors to make decisions based on Islamic law (seen by the eye). In Islamic events such as Ramadan or Eid the crescent must be seen by the naked-eye.

It is worth mentioning that investigations of minimizing the imbalance in the dataset will be conducted using upsampling and downsampling methods [8]. The fact that we treat the problem as a classification problem requires that we should try to cure the imbalance that we have in the original dataset. The imbalanced dataset also implies that the most appropriate performance metrics would be the F1 score, the AUC [Area Under the Receiver Operating Characteristic (ROC) Curve], the precision and the recall. It is not recommended to rely on accuracy metrics for this type of problem. Many studies emphasized the use of AUC for skewed data [9]. Several previous works for predicting the start of lunar months are available in literature. Some authors used classical methods that depend on analytical and numerical calculations [10]. Others used simple supervised ML methods to gain high precision and recall values [11]. Al-Rajab et al. [12] tried several ML methods on their own datasets and obtained encouraging results. Some works used image processing techniques and images caught using telescopes and binoculars [13]. Allawi [14] used a pattern recognizer method to estimate the presence of the new crescent Moon. In this work however, the goal is to achieve the following objectives:

- a) Improve the accuracy of ML techniques for the prediction of the new crescent Moon, investigating a wide spectrum of methods that vary in methodology and approach.
- b) Enhancing the existing methods by ensembling different models and hybridizing the existing ones with hyperparameters optimization.
- c) Dividing the datasets into clusters depending on geography or closeness of the countries.
- d) Using the ( $V_{tele}$ ), ( $V_{bino}$ ), and ( $V_{ccd}$ ) as possible inputs.

It should be noticed here that the target is  $V_{eye}$  values and not any other detection values such as seen by the binocular or seen by CCD camera, or telescope.

The paper is divided into the following sections; “[Background and literature review](#)” section is background and literature review, “[Dataset](#)” section covers the dataset description, “[Methodology and experimentations](#)” section presents the methodology and experimentation, “[Results and discussions](#)” section presents the results and discussions, and finally “[Conclusion and future work](#)” section is the conclusion and future work.

### Background and literature review

There exist several calendar systems used by different civilizations such as Hijri Calendar, Solar Calendar, Gregorian Calendar. For instance, the solar calendar was restricted only to some Asian countries. Gregorian calendar, which is used worldwide [15]. However, Hijri calendar [16] represents an important reference to millions of people, Muslim, and Islamic states, across the world to identify significant religious events. In the Hijri calendar, a year is determined by 12 lunar months, each of which consists of approximately 29.530 days, accommodating the lunar phases. Consequently, the Hijri year encompasses approximately 354.360 days. This lunar-based system deviates slightly from the solar year in the Gregorian calendar, which spans around 365.25 days. Of these months, Muhrram, Rajab, Ramadan, Shawwāl, and Dhū al-Ḥijjah, which bear a unique significance for Muslims.

Before we start covering all research papers which tackled new Moon visibility, let us first introduce the technical terms related to this research. Table 1 summarizes all these terms, the formulas used to compute some features such as Age\_of\_Moon, Sun\_Moon\_Lag, etc. can be found in [12].

The commencement of each Hijri month hinges on the sighting of the new crescent Moon. The Moon’s position relative to the Earth and the Sun causes it to illuminate partially, a phenomenon known as Moon illumination, thereby altering its appearance or phase. Consequently, the Moon cycles through eight distinct phases, initiating from the

**Table 1** Essential vocabulary for this research

Technical terms	Definition
Age_of_Moon	Number of days (time elapsed) since the Moon was last in conjunction with the Sun
Altitude_difference	Difference between the Moon_altitude and the Sun_altitude
Azimuth_difference	Angular difference in azimuth between the Sun and the Moon. The azimuth of an object defines its position on the celestial sphere in the horizontal coordinate system, measured in degrees starting from 0 north and turning in a clockwise direction
Moon_conjunction	Alignment of the Moon between the Earth and the Sun, with its illuminated side facing away from us, making it invisible from Earth
Elongation	Angular distance between the Moon and the Sun
Moon_altitude	Height of the Moon above the horizon
Moon_azimuth	Compass discretion (north = 0°) to the Moon’s position in the sky
Sun_altitude	Height of the Sun above the horizon
Sun_azimuth	Compass direction (north = 0°) to the Sun’s position in the sky
Sun_Moon_Lag	Time difference between the positions of the Sun and the Moon as seen from Earth
V_bino	New Moon was visible through binoculars
V_ccd	New Moon was visible through charge-coupled devices camera
V_eye	New Moon was visible by the eye
V_tele	New Moon was visible through telescope

unseen new Moon, progressing to the slender crescent illuminated by Sunlight, referred to as the new crescent Moon. This is depicted in the initial illustration of Fig. 1. The Moon proceeds through subsequent phases until reaching its final phase, initiating a new cycle. Figure 1, sourced from NASA's website, provides a visual representation of these phases [17].

Numerous studies have addressed the proposal to unify the Hijri calendar worldwide. For instance, lately, Mufid et al. [18] engaged in discussions concerning the unification of the Hijri calendar for the Southeast Asian region, specifically focusing on Brunei, Indonesia, Malaysia, and Singapore. Wahidin [19] presented a literature review that integrates insights from several perspectives: hadith, astronomy, and sociology to investigate potential approaches for the unification of the Hijri calendar. The findings suggest that taking a comprehensive approach could accelerate the process of the Hijri calendar unification.

On the other hand, Maskufa et al. [20] found that the significant challenges when it comes to unifying the Hijri calendar, primarily stemming from a combination of fiqh (Islamic jurisprudence) and astronomical factors. Hafez [21] proposed a mathematical model for the Moon sighting based on the motion of the Moon relative to the position of the observer on the Earth. The model was mainly based on the calculation of the ratio of the thickness of the luminous part of the Moon to its diameter. The author validated the model against data given in the Um-Alqura calendar, Al-Tawqifat Alelhamih book and showed high accuracy of the results.

Another call to the use of science in determining the start of the Hijri month was from Hasan [22] who presented a review highlighting the collaboration between Fiqh and empirical science in determining the start of the Hijri month in Indonesia, emphasizing the role of science in translating fiqh concepts into practical guidelines, the interplay between fiqh and scientific Moon sighting criteria, and fiqh validation of scientific efforts through testimonies. While it focuses on the years from 2010 to 2023, it provides valuable insights into methods and criteria development by the Hisab and Rukyat, with the potential for further



**Fig. 1** The different Moon phases [17]

research to deepen the understanding of this integration and its contribution to advancements in fiqh and empirical science in Indonesia. In his study in 2018, Maskufa [23] placed specific emphasis on the fiqh dimension, with a particular focus on astronomical fiqh. This aspect of fiqh pertains to the regulation to determine when religious activities should take place based on the movements of the Moon and Sun.

Further researchers addressed the visibility of the new crescent Moon by employing image processing methods with the emergence of advanced telescope technologies and satellites such as Kepler, has expanded the possibilities for automating the analysis of observations, prompting the necessity to train computer systems to perform these tasks [24, 25]. For instance, Moshayedi et al. [26] suggested a machine learning prototype system that employs a collection of Moon images to recognize and match the respective lunar phase represented in each image. Fakhar et al. [13] proposed an image processing-based method involving stages like noise elimination with a Gaussian filter, enhancing image quality around the crescent Moon area, and utilizing the circular Hough transform to extract the crescent's features. Sejzei and Jamzad [27] created a Matlab toolbox designed to improve the visibility of a crescent Moon in images, aiding observers in its detection. Utama et al. [28] utilized computer vision algorithms on video data to quickly and accurately identify the young lunar crescent's appearance. The process involved Gaussian Blur, Adaptive Thresholding for capturing the lunar crescent, and the Circular Hough Transform along with the OpenCV package for image extraction, processing, and detection.

Although, several studies have been conducted using AI techniques including ML and DL in several fields [29–33] and in particular the field of astronomy [34–36] but very few to almost no research focused on new crescent Moon prediction. For instance, Tafseer [11] addressed the new crescent Moon visibility problem by employing machine learning algorithms, namely logistic regression (LR), Neural Network (NN), Support Vector Machine (SVM), and Random Forest (RF). In his study, he utilized a dataset collected from the Islamic Crescents Observation Project (ICOP) website [37]. The dataset was augmented with additional features such as the age of the Moon, the Moon's lag time, the altitude difference, DAZ, the Moon phase, and atmospheric conditions. His results revealed that RF achieved the highest precision of 0.88 and a recall of 0.87. On the other hand, Allawi [14], utilized a single ML algorithm, the Artificial Neural Network (ANN) to predict the visibility of the new crescent Moon but confined to one particular region namely, Iraq. The algorithm yielded results with an accuracy level not less than 77%.

In our earlier study [12], we conducted research to predict the start of the Ramadan month, employing classification supervised ML algorithms on a dataset acquired from ICOP and augmented with additional observations and features. Our results revealed that SVM, NN, and RF yielded promising prediction outcomes with an accuracy rate of 91%. In this research, we aim to extend our investigation to establish the Lunar Calendar for the entire Hijri Year of Islamic months, utilizing artificial intelligence.

## Dataset

To begin our research, we started by collecting and building our dataset. We ensured its comprehensiveness and adherence to necessary standards by sourcing raw data from the ICOP, specifically from the "Crescent Observation Results". ICOP is a reliable source for lunar observations, recording data from numerous cities and countries worldwide. Their

extensive network of observers guarantees accurate documentation of new crescent Moon sightings, providing a diverse range of observations from different geographical regions.

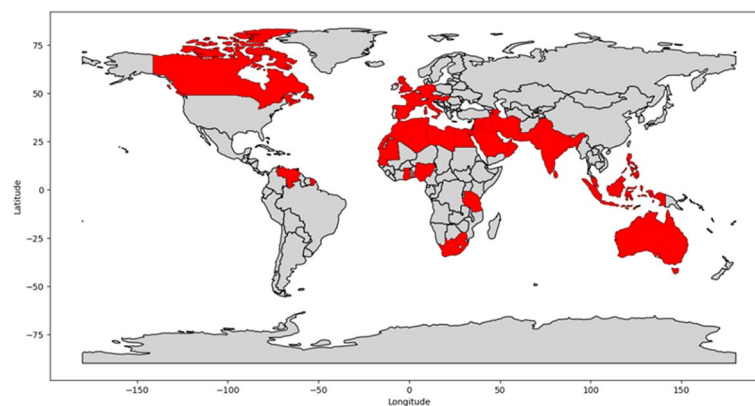
We manually built our own dataset from those observations, focusing on various features as outlined in our previous study [12]. These features include Hijri Day, Conjunction Time, Date, Country, City, State, Atmosphere,  $V_{eye}$ ,  $V_{bino}$ ,  $V_{tele}$ ,  $V_{ccd}$ , Longitude, Latitude, Sunset, Moonset, Sun\_Moon\_Lag, Age\_of\_the\_Moon, Moon\_altitude, Sun\_altitude, Altitude\_difference, Moon\_azimuth, Sun\_azimuth, Azimuth\_difference, Elongation, and Illumination. The formulas for calculating features such as Age\_of\_the\_Moon, Sun\_Moon\_Lag, Altitude\_difference, Azimuth\_difference, and Illumination can also be found in [12].

We processed and translated data such as country and city names from Arabic to English, standardizing spelling variations and handling abbreviations as necessary and explained in the data preparation “Dataset preparation” section. In our previous study [12], we focused exclusively on data for the month of Ramadan. In this work, we have significantly expanded that initial dataset to now include comprehensive data from all months over a 13-year period, from 2010 (1431 Hijri) to 2023 (1444 Hijri).

When deciding which features to include, we used commonly accepted data for this type of study [38] and augmented it using the Accurate Times 5.6 software [37]. By inputting the country, city, and date, the output included extra important features such as sunset, moonset, altitude, and longitude. Our final dataset comprises a total of 2085 observations from 47 different countries, as shown in Fig. 2. A sample of the dataset is presented in Fig. 3.

## Methodology and experimentations

In this research paper, our main contribution is focused on achieving the precision of the new crescent Moon prediction through an exhaustive exploration of diverse machine learning techniques and approaches. Additionally, based on the obtained results, we attempted to combine different models and optimize hyperparameters to achieve better prediction results. Furthermore, we intended to segment datasets into clusters based on geographical or country proximity factors.



**Fig. 2** Map of countries included in Moon observation dataset



index	hijr_year	Month	conjunction	Conjunction_Time	date	hijr_date	country	city	state	atmosphere	v_eye	v_bino	v_tfile	v_ccd	longitude	latitude	sunset	moonset	Sun_Moon_Lag	Age_of_Moon
0	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Algeria	Ghardaia	Ghardaia	1.0	1	1	0	0	48.000	23.000	18:34	19:20	0.766666667	13.7
1	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Australia	Sydney	New South Wales	0.85	1	0	0	0	11.000	-33.57000	7:22	8:47	1.416666667	2.5
2	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Bangladesh	Dhaka	Dhaka	1.0	1	0	0	0	20.160	46.100	12:36	13:22	0.766666667	7.733333333
3	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Ghana	Accra	Greater Region	0.35	1	0	0	0	10.000	36.000	18:15	19:31	1.266666667	13.383333333
4	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Iran	Esfahan	Esfahan	0.35	1	1	0	0	40.450	39.050	15:30	16:11	0.683333333	10.633333333
5	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Iraq	Mosul	Ninawa	0.85	1	0	0	0	08.010	20.160	16:06	16:42	0.6	11.233333333
6	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Oman	Muscat	Muscat	0.35	0	0	1	1	35.360	36.480	14:42	15:33	0.85	9.833333333
7	31	Dhu'l Hijah	06/11/2010 4:52	4:52	06/11/2010	29/11/1431	Saudi Arabia	Jabal Khandamah	Makkah	0.85	1	0	0	0	38.590	59.530	13:55	13:55	0.0	9.05
8	32	Dhu'l Hijah	26/10/2011 19:56	19:56	26/10/2011	29/11/1432	Tanzania	Daresalaam	Tanzania	0.35	1	1	0	0	18.110	-06.49230	15:16	14:57	-0.316666667	-4.666666667
9	32	Dhu'l Hijah	26/10/2011 19:56	19:56	26/10/2011	29/11/1432	United States	Gainesville	Florida	1.0	1	1	0	0	-82.17000	41.000	22:48	22:35	-0.216666667	2.866666667

Fig. 3 Sample of the observation features included in our dataset

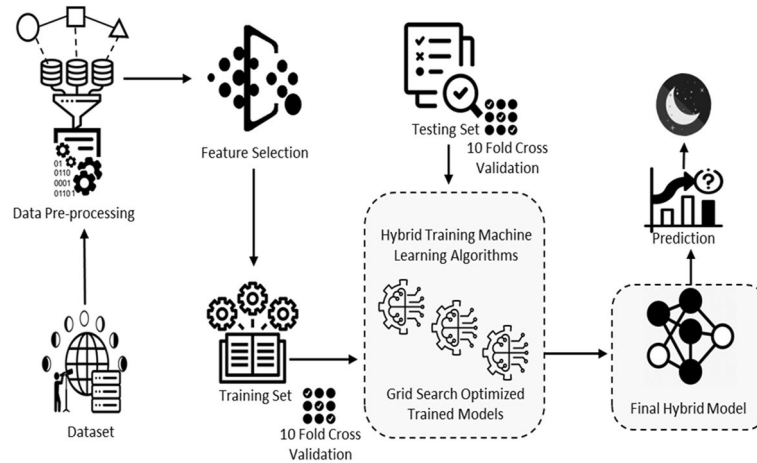


Fig. 4 The proposed framework model

Our overall system architecture is presented in Fig. 4. As shown in the figure, the experiment design goes over different steps, starting from the dataset collection and building, data pre-processing and feature selection, training and validating the data through a combination of ML models and approaches (as it will be explained later).

### Dataset preparation

The data we collected went through a pre-processing phase to guarantee its appropriateness for analysis. Prior to conducting our experiments, we took several pre-processing steps:

- Missing values: all missing or null values were eliminated to ensure data completeness.
- Invalid values: rows with negative Sun\_Moon\_Lag values were eliminated, as these are not valid for new crescent Moon determination.
- Data validity: rows with a Hijri Date of 28 were eliminated (although the observation of the new crescent Moon was recorded on ICOP, including the 28th of Hijri months, it can only be determined on the 29th or 30th of a Hijri month.)
- Data standardization and normalization: we standardized the data using min-max standard scaler normalization, scaling it to values between 0 and 1. This normalization technique is commonly used to ensure that all data is on a consistent scale, making it easier to compare different variables during analysis.

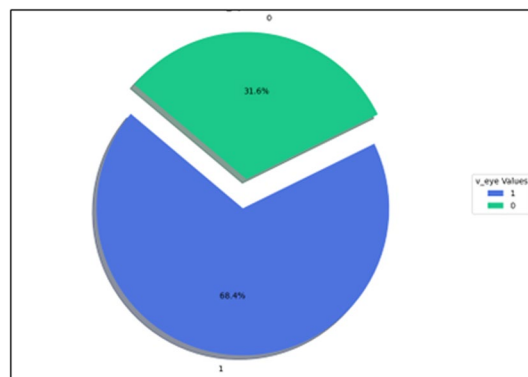
- We translated countries from Arabic to English and standardized spellings to match those used by the Accurate Times 5.6 software (e.g., Mecca to Makkah).
- Abbreviations were expanded to their full names.
- The atmosphere feature was converted to numeric values (Superb = 1, clear = 0.85, hazy = 0.6, Very hazy = 0.5).
- The observation feature, V\_eye, was converted from categorical to numeric to enable better understanding, fitting, evaluation, and extraction of valuable information by the machine learning model.
- Total Number of observations after data processing reduced to 1779 sample points.

To decide on the performance metrics to consider in our study, we had to analyze our data. We are interested in predicting the visibility of the Moon using the naked eye, for that we chose the V\_eye feature as our main prediction label. Plotting the V\_eye values from the 1779 observations, as depicted in Fig. 5, it was found that 68.4%, with 1216 instances out of a total of 1779 of the observations were seen by the naked-eye (V\_eye = 1), and 31.6% with 563 instances out of a total of 1779 of the same observations were unseen (V\_eye = 0). Consequently, to remedy the dataset unbalance, we used the F1-score and precision for the model performance evaluations.

**Feature selection**

After data balancing exploration, feature selection algorithms were explored to increase the possibility of reducing the complexity of the problem by reducing the number of parameters fed in. Different feature selection algorithms were applied [39] where some focus on linear and other focuses on the non-linear characteristics of the data. Each feature selection method was employed with a range of features to identify the optimal subset for training the models. These subsets were then used to train various machine learning algorithms, ensuring that each model had access to the most relevant features for making predictions. A brief description of these feature selection algorithms used in our experiments is provided below.

- *The ANOVA F-test* this method is used to evaluate the significance of each parameter in explaining the variation of a target variable. It is calculated by taking the ratio of



**Fig. 5** The distribution of samples across the two classes (V\_eye=0 and V\_eye=1) in the dataset

variance between the parameters to the variance within each parameter. The result of the ratio (F1 score) is then compared with the P-value to assess the statistical significance of their relationship with the target value, the higher the f-test score the stronger the dependency.

- *The mutual information* this method can capture both linear and non-linear relationships between the features and the target, without making assumptions regarding the distribution. This method relies on the entropy formula to check and evaluate the relationship between the features and target variable. The selected features are based on the higher mutual information score.
- *The recursive feature elimination with logistic regression (RFE)* method combines the principle of back propagation and logistic regression by removing the least important features until it reaches the set number of features. The corresponding coefficient is the controlling metric which arranges the features based on importance.
- *The least absolute shrinkage and selection operator (LASSO)* is used for feature selection and regularization by identifying the most important feature based on a modified linear regression cost function. The cost function assigns zero weights on the less important feature within the iterations.
- *The principal component analysis* is a dimensionality reduction technique in which it transforms high dimensional data to lower dimensional representations. The dimensionality reduction is achieved by finding the linear combinations of the features which would capture the maximum variance.
- *The correlation coefficient* is used to assess the degree of linear association between the target variables and individual features. It is also used to assess the degree of linear association between the features between themselves. The correlation coefficient is calculated and is usually represented as a correlation matrix. It works well when the features are associated with a linear relationship.
- *The gradient boosting importance* it works by applying an ensemble of decision trees, in which the features are split based on the value. The features that are frequently used for splitting the tree are considered of more importance. However, this method works on the predictive power of features not its causation. It can handle class imbalances which is the case in this problem.
- *The linear support vector classifier's importance* aims at finding the best hyperplane which accurately separates the data points that belong to different classes. During the training process each feature is assigned a certain weight which indicates the contribution of each parameter.
- *The random forest importance classifier* create an ensemble machine learning algorithm which can be used for feature selection. It calculates feature importance based on the importance score of each feature in reducing the error of the decision tree. It works best with numerical and categorical features.
- *The extra trees importance* is another ensemble method in which it helps in feature selection based on their importance. It applies further randomness in the tree building process.

Each feature selection method was employed with a range of features to identify the optimal subset for training the models. These subsets were then used to train various

machine learning algorithms, ensuring that each model had access to the most relevant features for making predictions.

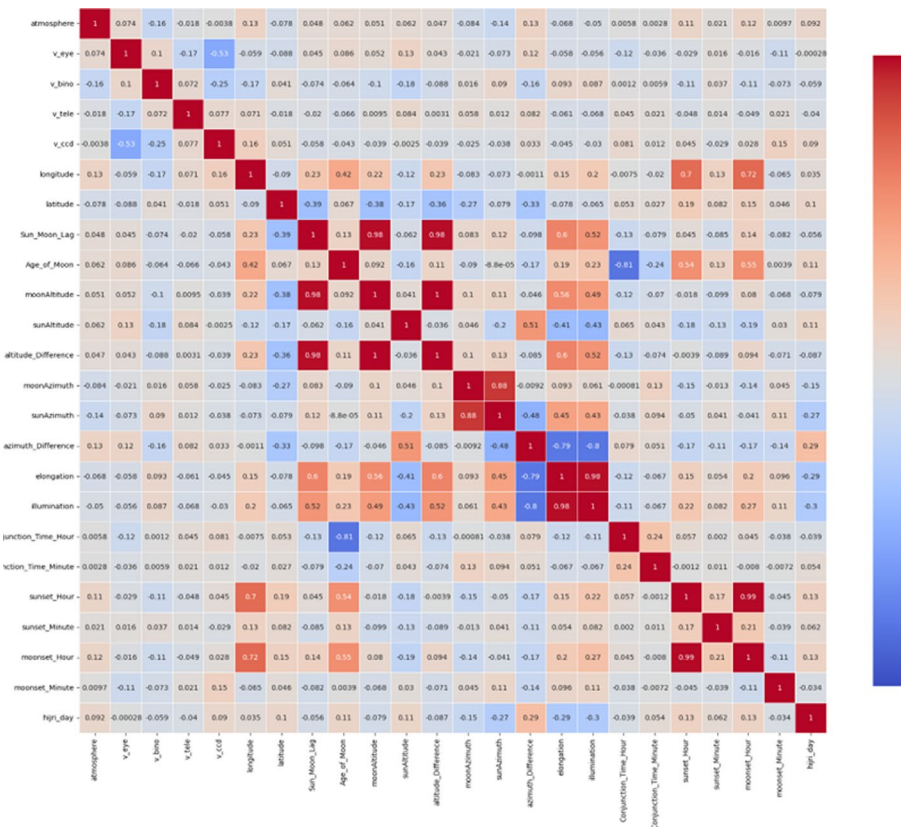
### Implementation details

Our models were implemented and trained using the Collaboratory notebook (Colab), a cloud-based tool provided by Google Research, offering a lightweight and free environment [40]. The programming language used was Python, which incorporates several Python libraries such as Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, among others. Pandas facilitated data manipulation and analysis, while NumPy handled computations. Matplotlib and Seaborn were utilized for data visualization purposes. For the implementation of machine learning algorithms, we relied on the Scikit-learn library.

### Experimental setup

For our experiments, we employed commonly used machine learning algorithms for classification purposes. These included Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), gradient boosting (GB), logistic regression (LR), K-Nearest Neighbor (KNN), Decision Tree (DT), Xgboost, AdaBoost, and Extra Trees. In all our experiments we used tenfold cross validation. The latter exhibits important benefits which include model performance assessment across different subsets of the data, reduction of overfitting risk, provision of robustness against data random variations, and is a widely adopted approach in the field. In the initial phase of our experiments, we conducted a correlation analysis among the features in our dataset. The results were visualized through a correlation heatmap as illustrated in Fig. 6, where important features exhibit high correlation. Particularly, the features such as “illumination” shows significant correlations with “Sun\_Moon\_lag”, “Moon\_altitude”, “Altitude\_difference”, “elongation”, and “Azimuth\_difference”. This indicates a strong linear relationship between these features and the “illumination”. Notably, the most significant correlation is detected between “Sun\_Moon\_lag” and both “Moon\_altitude” and “Altitude\_difference”, each having a correlation coefficient of 0.98. It is worth noting that both “Sunset\_hour” and “Moonset\_hour” are highly dependent on “Age\_of\_Moon” (0.54, 0.55) and “longitude” (0.7, 0.72). Given the results obtained from the heatmap, this suggests the use of the non-linear ML models which are found more suitable for our analysis.

In our research, we conducted a series of experiments aimed at comprehensively analyzing and optimizing our model’s performance. Experiment 1 examined how excluding certain features affected our predictions. Experiment 2 included all features to provide a complete interpretation for comparison. Experiment 3 utilized feature selection and grid search models to enhance the model’s performance by relying on relevant features and fine-tuning hyperparameters. Experiment 4 explored the benefits of ensemble learning, combining multiple models for improved accuracy and robustness. Experiment 5 employed a hybridized approach to efficiently optimize hyperparameters. Experiment 6 investigated a region-based approach to capture geographically specific patterns. These experiments contributed to the iterative refinement and enhancement of our predictive model, advancing our understanding and applicability in the target domain.



**Fig. 6** Dataset correlation Heatmap

**Experiment 1: analysis of full 12-month dataset excluding V\_ccd, V\_tele, and V\_bino features**

In Experiment 1, our approach involved the application of the previously mentioned ML algorithms (“[Experimental setup](#)” section). This analysis was conducted using the default hyperparameters for each of the ML algorithms. It is important to note that the set of input parameters in this experiment excluded three specific features: V\_tele, V\_bino, and V\_ccd. This exclusion was aimed at assessing the predictive capabilities of these models under modified dataset conditions.

**Experiment 2: analysis of full 12-month dataset with all feature included**

In this experiment, we chose to include all the available features of the dataset. This approach was undertaken to gain more comprehensive insights into the dataset characteristics. With all the features present, our objective was to assess whether the overall performance and predictive capabilities of the models improved. The same 10 ML algorithms were applied.

**Experiment 3: feature selection and model optimization using grid search**

In this experiment, the dataset underwent a systematic process involving feature selection and grid search of various ML models which are presented in Fig. 7. This process aimed to identify the most effective model in conjunction with the optimal set of

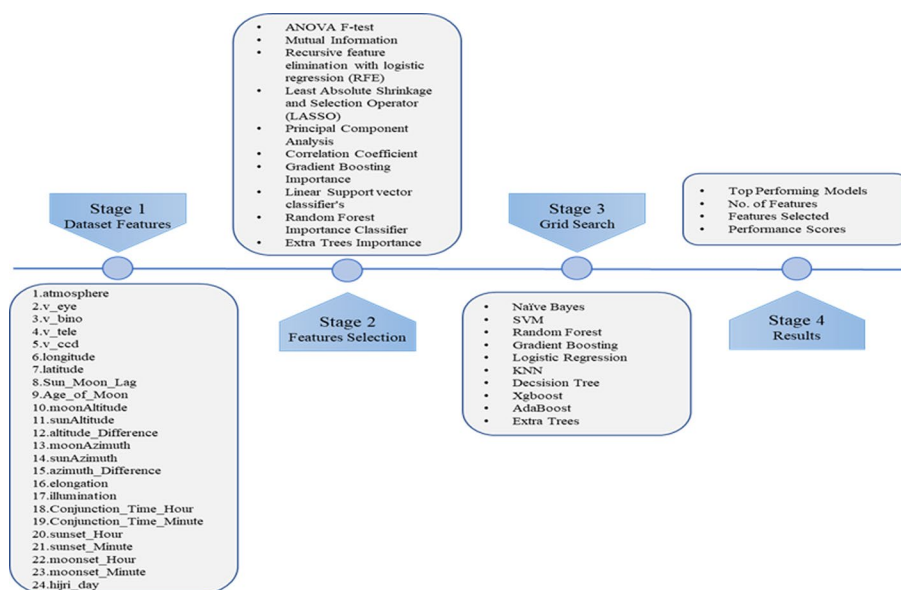
features. The experimental procedure started by defining a number of features,  $x$ , to be selected among the 24 features in our dataset (full dataset). In our case  $x$  was set as 7. Within each iteration of this process, the feature selection algorithm selected a specific number of  $x$  features. These selected features were then input into the grid search technique to fine-tune the hyperparameters of a model.

**Experiment 4: investigating enhanced predictive performance through ensemble modeling**

In this experiment, we aimed to investigate the potential enhancement of predictive results through ensemble technique. This approach integrated the four best-performing models based on the top F1 scores released from experiment 3, each with their respective input features, and passed them through a soft voting classifier, which combined the average of the votes. The objective was to assess whether this ensemble methodology could further augment the performance outcomes. To evaluate the value of the ensemble technique, we benchmarked its results against those of the top 4 performing models obtained in the previous experiment 3. This comparative analysis provided insights into the viability and effectiveness of employing an ensemble of models for improved predictive capabilities in our dataset.

**Experiment 5: exploring hybrid optimization techniques for enhancing ml model performance**

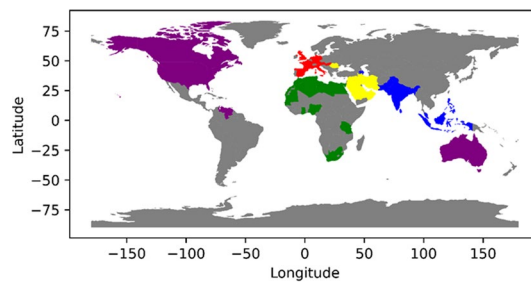
This experimental setup aimed to explore the potential enhancement of the top-performing model, from the results obtained in experiment 3 through hyperparameter optimization using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) techniques. These optimization methods offer advantages in fine-tuning model parameters to maximize performance, thereby helping in a more refined and effective model. To maintain consistency in the experimental conditions, the same set of 14 features, as previously identified, will be utilized for this phase of the study.



**Fig. 7** Feature selection and grid search pipeline

**Table 2** Countries segmented in distinct regions

Region	Longitude range	Latitude range	Countries
Middle East	34° E–60° E	12°–38° N	Saudi Arabia, Iran, Iraq, Oman, United Arab Emirates, Bahrain, Jordan, Lebanon, Syria, Palestine, Qatar
Africa	17° W–51° E	12° S–37° N	Algeria, Libya, Tunisia, Egypt, Morocco, Mauritania, Ghana, Tanzania, Nigeria, Senegal, South Africa, Gambia
Asia	95° E–125° E	1° S–42° N	Bangladesh, India, Pakistan, Sri Lanka, Nepal, Azerbaijan, Indonesia, Philippines, Malaysia, Brunei, Hong Kong
Europe	10° W–31° E	36° N–55° N	Germany, France, United Kingdom, Hungary, Austria, Spain, Netherlands, Bosnia and Herzegovina, Italy
Other	113° E–123° W	10° N–44° S	United States, Canada, Venezuela, Australia



**Fig. 8** Map of countries segmented into regions

**Experiment 6: investigating the influence of geographical factors: a region-based approach**

In this experiment, we explored deeper the influence of geographical features such as longitude and latitude on our models. This exploration was driven by the observation in Experiment 3. We categorized the observations, derived from 47 countries, into five distinct regions: Middle East, Africa, Asia, Europe, and Other. The region segmentation is detailed in Table 2 and illustrated in Fig. 8.

**Results and discussions**

In all our experiments, we used the F1 score and the mean of the area under the curve (AUC) as the main evaluation performance metrics to assess our prediction results.

For experiment 1, the performance of the ML models considered in the analysis is illustrated in Table 3. The GB outperforms all the other models, boasting the highest F1 score of 0.824845 and an AUC of 0.764184, indicating its ability to handle the complexity of the dataset. On the other hand, NB exhibits the lowest performance with an F1 score of 0.608916 and an AUC of 0.617499, showing its struggle with the dataset complexity. In general, ensemble methods like GB, RF, and XGBoost demonstrate robust performance, indicating complex patterns within the data. SVM also shows strong performance with a high F1 score, indicating its capability to manage high-dimensional data and accurately detect true positives.

Experiment 2 involved the comprehensive analysis of the full dataset, inclusive of features V\_ccd, V\_bino, and V\_tele, resulting in a noticeable enhancement in the performance metrics across all applied ML models, as illustrated in Table 4. This

**Table 3** Results of experiment 1 (The bold values in the table indicate the highest performance metrics (Mean F1 Score and Mean AUC) achieved by the machine learning models in experiment 1)

ML algorithm	Mean F1 score	Mean AUC
NB	0.608916	0.617499
SVM	0.820968	0.706106
RF	0.808872	0.719215
<b>GB</b>	<b>0.824845</b>	<b>0.764184</b>
LR	0.800651	0.637309
KNN	0.781338	0.676872
DT	0.740491	0.610156
XGBoost	0.787297	0.717707
AdaBoost	0.814294	0.750441
Extra trees	0.800238	0.683953

**Table 4** Results of experiment 2 (The bold values in the table indicate the highest performance metrics (Mean F1 Score and Mean AUC) achieved by the machine learning model in experiment 2)

Algorithm	Mean F1 score	Mean AUC
NB	0.771363	0.798937
SVM	0.876087	0.874549
RF	0.870211	0.864888
<b>GB</b>	<b>0.882469</b>	<b>0.901009</b>
LR	0.858785	0.833284
KNN	0.866233	0.857882
DT	0.837237	0.752882
XGBoost	0.854933	0.882258
AdaBoost	0.846337	0.807547
Extra trees	0.87041	0.868849

improvement distinctly indicated the significant role these features played in augmenting the predictive F1 score. In particular, GB emerged as the most effective model, achieving an F1 score of 0.882469 and an AUC of 0.901009. This better performance indicated the compatibility of the dataset with sophisticated ensemble techniques, as also evidenced by the high efficiency of models like SVM, RF, and XGBoost. However, the NB, despite its improvement to an F1 score of 0.771363 and AUC of 0.798937, remained the least effective among its counterparts.

Experiment 3 records the outcomes of each iteration, collecting a comprehensive result from the dataset. Upon completion of all iterations, which cover more than 7000 combinations of selected features and model configurations. The most effective four combinations, based on predetermined performance metrics are shown in Table 5. This method enables a thorough examination of the models' effectiveness across diverse feature sets. It also enhances the comprehension of feature significance within the dataset.

Table 5 specifically describes the performance metrics—Mean F1 Score and Mean AUC of the four most performant models which are, the Extra Trees algorithm and the SVM.



The analysis of the dataset, employing the different feature selection techniques, has improved the predictive performance across several ML models. This enhancement indicates the importance of specific features, such as “V\_bino”, “V\_tele”, “V\_ccd”, “longitude”, and “latitude”, among others, in improving the overall model efficiency as depicted by their high respective F1 scores.

It has been found that the Extra Trees model, optimized with parameters of max\_depth: 10 and n\_estimators: 50, and using a set of 14 features selected through ANOVA F-test, stands out with an impressive mean F1 score of 0.887872, and a mean AUC of 0.906242. The marginal differences in F1 scores across models, being less than 0.02%, illustrate the impact of feature selection on model accuracy. On the other hand, the SVM model, employing the Random Forest Importance for feature selection and optimized with parameters of C: 1 and kernel: ‘rbf’, also shows promising results. When applied with 10 selected features, it achieves a comparable mean F1 score of 0.887848, and a mean AUC of 0.888855. Notably, even with a reduced feature set of 8 or 9 features, the SVM model maintains high performance levels.

These results underscore the effectiveness of employing targeted feature selection methods and fine-tuning model parameters. The effectiveness of ensemble methods and the SVM’s kernel trick suggest that the dataset likely exhibits non-linear relationships and interactions among its features, which are more accurately captured by models capable of handling high-dimensional spaces and complex feature interactions. These results highlight how useful it is to carefully choose specific features and adjust model parameters. The success of combining multiple models and using advanced techniques indicates that the dataset may contain complex relationships between its features. Models that can

**Table 5** Results of feature selection models and grid search

Model ranking/type	No. of selected features	Selected features	Mean F1	Mean AUC	Model parameters
Model 1/extra trees	14	Atmosphere, V_bino, V_tele, V_ccd, longitude, latitude, Age_of_Moon, Sun_altitude, Sun_azimuth, Azimuth_difference, elongation, illumination, Conjunction_Time_Hour, Moonset_Minute	0.887872	0.906242	{‘max_depth’: 10, ‘n_estimators’: 50}
Model 2/SVM	10	Latitude, Azimuth_difference, V_tele, longitude, Sun_altitude, V_bino, Conjunction_Time_Minute, Conjunction_Time_Hour, Age_of_Moon, V_ccd	0.887848	0.888855	{‘C’: 1, ‘kernel’: ‘rbf’}
Model 3/SVM	8	Longitude, V_tele, Sun_altitude, V_bino, Conjunction_Time_Hour, Conjunction_Time_Minute, Age_of_Moon, V_ccd	0.887638	0.896323	{‘C’: 1, ‘kernel’: ‘rbf’}
Model 4/SVM	9	Azimuth_Difference, longitude, V_tele, Sun_altitude, V_bino, Conjunction_Time_Minute, Conjunction_Time_Hour, Age_of_Moon, V_ccd	0.887238	0.891651	{‘C’: 1, ‘kernel’: ‘rbf’}

**Table 6** Ensemble model

Model ranking/type	Mean F1	Mean AUC
Ensemble model	0.888058	0.907482

**Table 7** Comparison with hybrid models

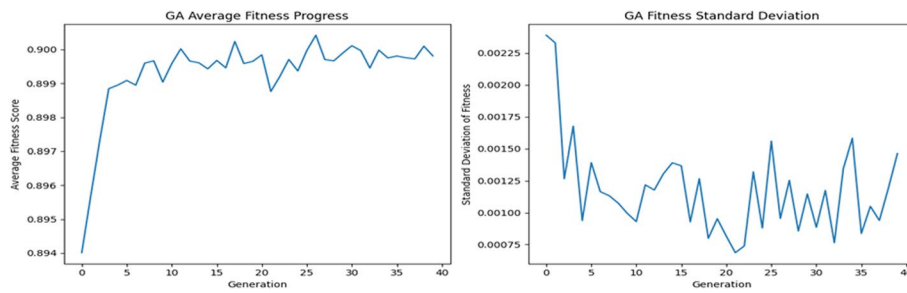
Model ranking/type	Parameters	Mean F1	Mean AUC
Model 1/extra trees	'max_depth': 10, 'n_estimators': 50	0.8878	0.9062
PSO with extra trees	max_depth = 8, n_estimators = 138	0.8811	0.9077
GA with extra trees	max_depth = 10, n_estimators = 199	0.8826	0.9053

process lots of different features and their interactions tend to perform better as it was illustrated in our case.

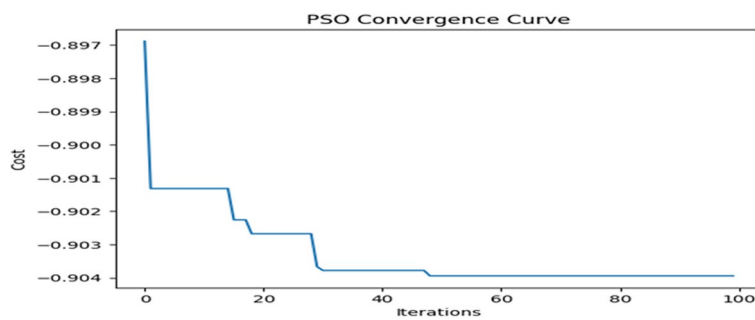
In experiment 4, we extended our analysis to include ensemble models. To evaluate the importance of this technique, we will benchmark its results against those obtained from experiments 3 (Table 5). This comparative analysis aims to reflect on how using a combination of models can enhance the predictive capabilities within our dataset. It is illustrated from Table 6, that the Ensemble Model demonstrates the highest F1 score of 0.888058 coupled with an AUC of 0.907482, compared to the results of (Table 5).

The objective of experiment 5, is to identify the most optimal hyperparameter combination that maximizes model performance. A defined range of hyperparameter values for the Extra trees model, specifically 'n\_estimators' within the range of 10 to 200 and 'max\_depth' within the range of 1 to 50, will be subjected to both GA and PSO optimization processes. This comparative study aims to determine if the GA and PSO-based methods for hyperparameter optimization can outperform the performance standards established by the grid search technique. The outcomes of this optimization will then be compared with the results obtained from Experiment 3, as outlined in Table 7.

The results depicted in Table 7 clarify a crucial aspect of our ML model optimization. The convergence patterns observed from both the GA and PSO optimization indicate that the performance level stays about the same without much improvement [41]. However, the convergence behavior within 40 generations by the GA and the PSO shows the efficiency of those algorithms in reaching the optimal solution without the need for the exhaustive grid search that is usually used to find the optimum hyperparameters values. Figures 9 and 10 depict the convergence behavior of GA and PSO. Since we are applying a tenfold cross validation it would be advisable to use efficient techniques such as the GA and the PSO for parameters hyper tuning optimization. The GA and the PSO were applied on the best performer algorithm which is the "Model 1/Extra trees". The convergence to the optimum hyperparameters was much faster as compared to the exhaustive grid search. Both algorithms used a cost function that is based on the F1 and AUC values. This work leads to propose a possible direction for future work, a more expansive approach to hyperparameter tuning across the spectrum of the 10 ML models employed, this will help us investigate deeply the hyper parameter space, potentially uncovering more effective model configurations.



**Fig. 9** Convergence behavior of the GA over the generations; average fitness (to the left) and standard deviation (to the right)



**Fig. 10** Convergence behavior of the PSO algorithm over the generations

**Table 8** Regional distribution of V<sub>eye</sub> Labels

Region	No. of observations	V <sub>eye</sub> = 1	V <sub>eye</sub> = 0
Middle East	582	313	269
Africa	470	419	51
Asia	307	207	100
Europe	259	133	114
Other	167	139	28

In experiment 6, we divided our dataset into regions. The distribution of V<sub>eye</sub> labels within these regions is presented in Table 8. It is worth noting, we can see that there is a relative balance of datasets in the Middle East and Europe, in contrast to the significant imbalance in the rest of the regions.

To assess the impact of the regional segmentation on the predictive performance, we applied the same methodology adopted in experiment 3. The top performing models for each region are illustrated in Table 9.

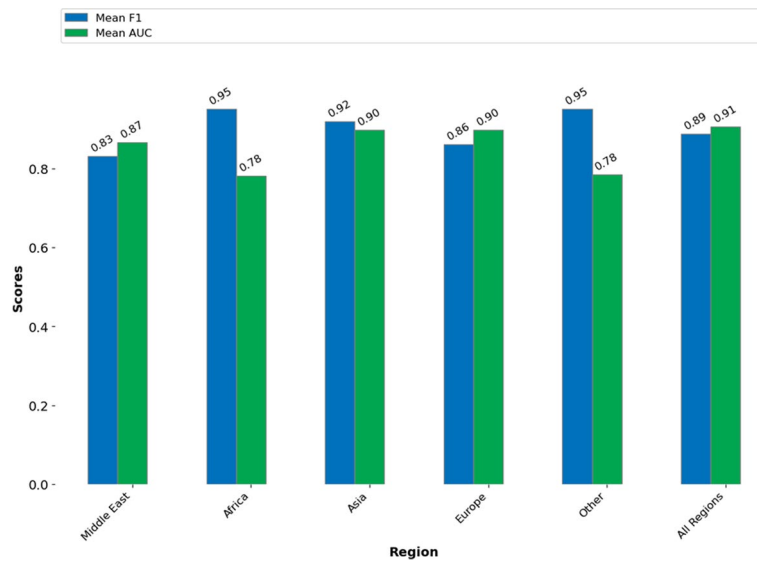
The analysis of the regionalized data, as reflected in Table 9, shows an improvement in the F1 score for some regions such as Africa and Asia, and this compared to the results obtained from experiment 3. However, there is a decline in the AUC scores. This can be due to a reduced number of samples, which can be attributed to the limited false label (“V<sub>eye</sub> = 0”) occurrences in the datasets. When comparing the performance metrics of individual regions to the ‘All Regions’ model, it is visible that

**Table 9** Model evaluation of different regions (The bold and italic values represent the highest Mean F1 Score and Mean AUC, respectively, within each region’s model evaluation. These values highlight the best-performing model for that specific region, demonstrating the model’s ability to most accurately predict the visibility of the new crescent Moon based on the region-specific dataset)

Region	Model	No. of selected features	Selected features	Mean F1	Mean AUC
Middle East	<i>Gradient boosting</i>	15	Atmosphere, V_bino, V_tele, V_ccd, latitude, Age_of_Moon, Sun_altitude, Moon_azimuth, Sun_azimuth, Azimuth_difference, elongation, illumination, Conjunction_Time_Hour, Sunset_Hour, Moonset_Minute	0.8317	0.8672
<b>Africa</b>	<b>KNN</b>	<b>8</b>	<b>V_bino, V_ccd, Sun_Moon_Lag, Age_of_Moon, MoonAltitude, elongation, illumination, Moonset_Hour</b>	<b>0.952</b>	<b>0.782</b>
Asia	Random forest	8	Atmosphere, V_bino, V_ccd, latitude, Sun_altitude, illumination, Sunset_Hour, Moonset_Hour	0.920	0.898
Europe	<i>Gradient boosting</i>	17	Principal component 1, principal component 2, principal component 3, principal component 4, principal component 5, principal component 6, principal component 7, principal component 8, principal component 9, principal component 10, principal component 11, principal component 12, principal component 13, principal component 14, principal component 15, principal component 16, principal component 17	0.8617	0.899
Other	KNN	7	Longitude, Age_of_Moon, Sun_altitude, Conjunction_Time_Hour, Conjunction_Time_Minute, Moonset_Minute	0.9517	0.7846
All regions	Random forest	14	Atmosphere, V_bino, V_tele, V_ccd, longitude, latitude, Age_of_Moon, Sun_altitude, Sun_azimuth, Azimuth_difference, elongation, illumination, Conjunction_Time_Hour, Moonset_Minute	0.888	0.906

while some regions benefit from the regionalization approach in terms of F1 Score, the AUC scores are reduced. This observation indicates the impact of the dataset size on model performance. The number of samples in the different regions was mostly unbalanced.

Figure 11 depicts a comprehensive view of all the scores via the different ML models and regions. The RF, GB, KNN, and Xtree models were the top performers as proven by shown results and the previous results. Probably not much improvement has been noticed by allocating an independent predictor for each one of the 5 regions, however, based on comparisons with the all-regions predictor, some improvement has been noticed for Africa and the “other” region with regard to the F1 score when using the KNN model. It is worth mentioning that the number of samples were still unbalanced despite the region’s division process. In fact, some regions had closely balanced samples such as the “Middle East” but still no improvement was noticed for that region. This leads us to reach some conclusions regarding the training data balance characteristic as will be mentioned in the conclusion.



**Fig. 11** All score values as per region and ML models

This study aimed to establish the Hijri calendar for a span of 13 years, incorporating various countries from diverse regions worldwide. In contrast to our previous research in [12], which solely focused on the month of Ramadan, this study encompasses all months of the Hijri year, utilizing an updated dataset spanning the entire 13-year period. To our knowledge, only two studies have been conducted to predict the visibility of the new crescent Moon. Allawi [14], only one deep learning algorithm, Convolutional Neural Network (CNN), was employed, limited to a single country over a 4-year period. Tafseer [11], four supervised learning algorithms were applied to the dataset. Our study represents the first comprehensive examination of all 12 Hijri months over a 13-year period from x to y. We conducted an extensive analysis employing a wide array of ML algorithms and techniques, including default parameters to grid search, model ensemble, model hyperparameter hybridization, then ending with regional splitting of the data with the main purpose of extracting the best performing model and maximizing its performance.

Overall, the GB and the Extra Trees outperformed the other machine learning models. It is apparent that decision-tree based models (GB and Extra Trees) are more appropriate for this type of problems and due to their ensemble nature, their ability to capture the interactions of complex features, and the ability to manage the imbalanced data effectively. Other machine learning models that have a connectionist nature such as NN (that use weight tuning approach) show less success as sometimes it is more difficult to find a surface that separates between the rough input regions. All these made the GB and the Extra Trees models well suited for predicting the visibility of the new crescent Moon.

The results show that the GB and Extra Trees models outperform the other models considered in this analysis. Each model processes data differently depending on the importance of specific features. In terms of feature selection, the Extra Trees

model uses a broader set of features, which might explain why it performs slightly better than the GB model, which prioritizes features that reduce the overall error. The hyperparameter tuning using the GA and PSO metaheuristic algorithms helps the models learn faster and work more effectively, resulting in the superior performance of the Extra Trees model.

### Conclusion and future work

In this paper, we present ML-based framework that contributes significantly toward the realization of a globally unified lunar calendar. Various ML algorithms were considered, and several experiments were conducted on the dataset, starting with default parameters, proceeding to Grid Search, Ensemble Modeling, model hyperparameter hybridization, and regional data segmentation. The primary objective was to identify and optimize the best-performing predictive model for the new crescent Moon visibility. Results revealed that the GB model outperforms the other models considered in this study in terms of F1 score and AUC. However, with feature selection identified through the ANOVA F-test and optimized parameters, the Extra Trees model turns out to be the best-performing predictive model. The Ensemble model and hybridization, on the other hand, although they exhibit slight improvement compared to the other models, showed faster convergence to the optimum values of the models hyperparameters. Geographical segmentation of the dataset enhanced predictive performance in certain areas, such as Africa and Asia. A noteworthy point is that a balanced dataset is not sufficient to achieve improved performance in the segmented regions when compared to experiments that take all regions in the dataset into account.

For future work, we will consider incorporating new optimization techniques such as Differential Evolution (DE) and Gradient-based Optimization algorithms. Additionally, data synthesis could be used to generate more valuable training and testing data. Collecting more data from different regions will be necessary to further explore the geographical segmentation of the dataset and train more robust ML models. Although the machine learning methods we have employed are advanced, DL models capable of capturing spatial and temporal variations in inputs could also be applied.

To extend our study, we plan to include real-time data for predictions using ongoing observations, particularly during the conjunction period. The input data can be a live stream or stored videos providing continuous observations, similar to real-life scenarios. While this method may require more processing storage and power, it offers a more comprehensive approach by utilizing rich spatial and temporal information to reach a decision.

### Abbreviations

ANN	Artificial Neural Network
ANOVA	Analysis of variance
AUC	Area under the curve
CNN	Convolutional Neural Network
Colab	Collaboratory (Google Research cloud-based tool)
DT	Decision tree
GA	Genetic Algorithm
GB	Gradient boosting
ICOP	Islamic Crescents Observation Project
KNN	K-Nearest Neighbor
LR	Logistic regression

ML	Machine learning
PCA	Principal component analysis
PSO	Particle Swarm Intelligence
RF	Random forest
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine

#### Author contributions

S.L.: contributed to the concept, methodology, analysis, and drafting of the paper. M.A.: contributed to the concept, methodology, analysis, drafting of the paper, development of the code, and supervised the research project. R.A.Z.: contributed to the concept, methodology, analysis, and drafting of the paper. M.R.: contributed to the development of the code used in the research and the validation of the results. S.L., M.A., and R.A.Z.: involved in data collection and interpretation.

#### Funding

The authors acknowledge financial support from Abu Dhabi University's Office of Research and Sponsored Programs (Grant Number: 19300796).

#### Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

#### Declarations

##### Competing interests

The authors declare no competing interests.

Received: 12 April 2024 Accepted: 6 August 2024

Published online: 12 August 2024

#### References

1. Wisevoter. Number of muslims in the world 2023. Wisevoter; 2023. <https://wisevoter.com/country-rankings/number-of-muslims-in-the-world/>. Accessed 2 Jan 2024.
2. Yu L, Sun L, Du B, Liu C, Xiong H, Lv W. Predicting temporal sets with deep neural networks. In: 2020 26th ACM international conference on knowledge discovery & data mining (KDD), CA, USA; 2020. p. 1083–91.
3. Singh A, Thakur N, Sharma A. A review of supervised machine learning algorithms. In: 2016 3rd international conference on computing for sustainable global development (INDIACom), New Delhi, India; 2016. p. 1310–5.
4. Dokeroglu T, Sevinc E, Kucukyilmaz T, Cosar A. A survey on new generation metaheuristic algorithms. *Comput Ind Eng*. 2019;137: 106040. <https://doi.org/10.1016/j.cie.2019.106040>.
5. Pranolo A, Mao Y, Wibawa A, Utama ABP, Dwiyoanto F. Optimized three deep learning models based-PSO hyperparameters for Beijing PM2.5 prediction. *Knowl Eng Data Sci*. 2022;5(1):53–66. <https://doi.org/10.17977/um018v5i12022p53-66>.
6. Zoremsanga C, Hussain J. Particle swarm optimized deep learning models for rainfall prediction: a case study in Aizawl, Mizoram. *IEEE Access*. 2024;12:57172–84. <https://doi.org/10.1109/ACCESS.2024.3390781>.
7. Xue Y, Aouari A, Mansour R, Su S. A hybrid algorithm based on PSO and GA for feature selection. *J Cyber Secur*. 2021;3(2):117–24. <https://doi.org/10.32604/jcs.2021.017018>.
8. Kotsiantis S, Kanellopoulos D, Pintelas P. Handling imbalanced datasets: a review. *GESTS Int Trans Comput Sci Eng*. 2006;30:25–36.
9. Jeni LA, Cohn JF, De La Torre F. Facing imbalanced data—recommendations for the use of performance metrics. In: 2013 Humaine association conference on affective computing and intelligent interaction, Geneva, Switzerland; 2013. p. 245–51. <https://doi.org/10.1109/ACII.2013.47>.
10. Yallop B. Method for predicting the first sighting of the new Crescent Moon. *RGO NAO Technical Note*, vol. 69. 1997.
11. Tafseer A. Predicting the visibility of the first crescent: predicting the visibility of the first crescent. *KIET J Comput Inf Sci*. 2020;3(2):53–61.
12. Al-Rajab M, Loucif S, Al Risheh Y. Predicting new crescent moon visibility applying machine learning algorithms. *Sci Rep*. 2023;13:6674. <https://doi.org/10.1038/s41598-023-32807-x>.
13. Fakhar M, Moalem P, Badri MA. Lunar crescent detection based on image processing algorithms. *Earth Moon Planet*. 2014;114:17–34.
14. Allawi ZT. A pattern-recognizer artificial neural network for the prediction of new crescent visibility in Iraq. *Computation*. 2022;10(10):186.
15. Ohms BG. Computer processing of dates outside the twentieth century. *IBM Syst J*. 1986;25(2):244–51.
16. Farichah F. The java calendar and its relevance with the Islamic calendar. *Al-Hilal J Islam Astron*. 2021;2(2):214–48. <https://doi.org/10.21580/al-hilal.2020.2.2.6725>.
17. Moon phase and libration. 2020. <https://svs.gsfc.nasa.gov/4768>. Accessed 12 Jan 2024.
18. Mufid A, Djamaluddin T. The implementation of new minister of religion of Brunei, Indonesia, Malaysia, and Singapore criteria towards the Hijri calendar unification. *HTS Teologiese Stud/Theol Stud*. 2023;79(1):8774. <https://doi.org/10.4102/hts.v79i1.8774>.

19. Wahidin N. Problem of unification Hijri calendar. *Al-Afaq Jurnal Ilmu Falak Dan Astronomi*. 2022;4(2):275–83. <https://doi.org/10.20414/afaq.v4i2.5761>.
20. Maskufa M, Sopa S, Hidayati S, Damanhuri A. Implementation of the new MABIMS crescent visibility criteria: efforts to unite the Hijriyah calendar in the Southeast Asian region. *Ahkam Jurnal Ilmu Syariah*. 2022. <https://doi.org/10.15408/ajis.v22i1.22275>.
21. Hafez G. Empirical model for moon sighting. *Yanbu J Eng Sci*. 2022;19(2):22–9. <https://doi.org/10.53370/001c.38803>.
22. Hasan M. The interaction of Fiqh and science in the dynamics of determining the beginning of the Hijri month in Indonesia. *J Islam Law*. 2023;4(2):237–57. <https://doi.org/10.24260/jil.v4i2.1433>.
23. Maskufa SH. Global Hijriyah calendar as challenges Fikih astronomy. *Adv Soc Sci Educ Humanit Res*. 2017;162:188–92. <https://doi.org/10.2991/iclj-17.2018.39>.
24. Bhamare AR, Baral A, Agarwal S. Analysis of kepler objects of interest using machine learning for exoplanet identification. In: International conference on intelligent technologies (CONIT); 2021. p. 1–8.
25. Khan MA, Dixit M. Discovering exoplanets in deep space using deep learning algorithms. In: 12th international conference on computational intelligence and communication networks (CICN); 2020. p. 441–7.
26. Moshayedi AJ, Chen ZY, Liao L, Li S. Sunfa Ata Zuyan machine learning models for moon phase detection: algorithm, prototype and performance comparison. *Telkomnika Telecommun Comput Electron Control*. 2022;20(1):129–40.
27. Sejzei AH, Jamzad M. Evaluation of various digital image processing techniques for detecting critical crescent moon and introducing CMD—a tool for critical crescent moon detection. *Optik*. 2016;127(3):1511–25.
28. Utama JA, Zuhudi AR, Prasetyo Y, Rachman A, Sugeng Riadi AR, Nandi, Riza LS. Young lunar crescent detection based on video data with computer vision techniques. *Astron Comput*. 2023;44: 100731. <https://doi.org/10.1016/j.ascom.2023.100731>.
29. Firouzi F, Shiyi J, Krishnendu C, Bahar F, Mahmoud D, Jaeseung S, Kunal M. Fusion of IoT, AI, edge–fog–cloud, and blockchain: challenges, solutions, and a case study in healthcare and medicine. *IEEE Internet Things J*. 2023;10(5):3686–705. <https://doi.org/10.1109/JIOT.2022.3191881>.
30. Virmani N, Singh RK, Agarwal V, Aktas E. Artificial intelligence applications for responsive healthcare supply chains: a decision-making framework. *IEEE Trans Eng Manag*. 2024;71:8591–605. <https://doi.org/10.1109/TEM.2024.3370377>.
31. Wang Y, Xiao J, Wei Z, Zheng Y, Tang K-T, Chang CH. Security and functional safety for AI in embedded automotive system—a tutorial. *IEEE Trans Circuits Syst II Express Briefs*. 2024;71(3):1701–7. <https://doi.org/10.1109/TCSII.2023.3334273>.
32. Mosavi MR, Khishe M, Ghamgosar A. Classification of sonar data set using neural network trained by gray wolf optimization. *Neural Network World*. 2016;26(4):393–415.
33. Khishe M, Mosavi MR. Classification of underwater acoustical dataset using neural network trained by Chimp optimization algorithm. *Appl Acoust*. 2020;157: 107005. <https://doi.org/10.1016/j.apacoust.2019.107005>.
34. Sen S, Agarwal S, Chakraborty P, Singh KP. Astronomical big data processing using machine learning: a comprehensive review. *Exp Astron*. 2022;53:1–43.
35. Bely P. *The design and construction of large optical telescopes*. Berlin: Springer; 2003.
36. Bhavsar R, Kumar JN, Umesh B, Rajesh G, Sudeep T, Gulshan S, Pitsshou B, Ravi S. Classification of potentially hazardous asteroids using supervised quantum machine learning. *IEEE Access*. 2023;11:75829–48. <https://doi.org/10.1109/ACCESS.2023.3297498>.
37. International Astronomical Center. <https://www.astronomycenter.net/>. Accessed 2 Jan 2024.
38. Odeh MS. New criterion for lunar crescent visibility. *J Exp Astron*. 2004;18(1):39–64.
39. Géron A. *Hands-on machine learning with scikit-learn, Keras, and TensorFlow*. 3rd ed. Safari: O'Reilly Media, Incorporated; 2022.
40. Colaboratory. <https://colab.research.google.com/>. Accessed 2 Jan 2024.
41. Shivam M, Shashank A. Feature selection using metaheuristic algorithms on medical datasets. In: *Harmony search and nature inspired optimization algorithms: theory and applications, ICHSA 2018*. Singapore: Springer; 2019. p. 923–37.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.