

**How to Cite:**

Vigneswaran, L., & Nagadeepa, N. (2022). Prediction of research studies status on astrological data using machine learning (ML) classification techniques. *International Journal of Health Sciences*, 6(S6), 9696–9710.  
<https://doi.org/10.53730/ijhs.v6nS6.12531>

## **Prediction of research studies status on astrological data using machine learning (ML) classification techniques**

**L. Vigneswaran**

Research Scholar (Part-Time), Bharathiar University, Coimbatore

\*Corresponding author email: [vigneswaran.research@gmail.com](mailto:vigneswaran.research@gmail.com)

**Dr. N. Nagadeepa**

Principal, Sri Sarada Niketen College for Women, Karur, Tamilnadu, India

Email: [nagadeepa1012@gmail.com](mailto:nagadeepa1012@gmail.com)

**Abstract**--Astrology specifies both past action as well as the future prediction which act as the subset of astronomy that have been represented in the Galaxy system. Horoscope chart is an astrological representation which involves the movement of nine planets located in twelve houses with periodic positional movements. Based on the horoscope representation, some of the houses are empty and others might contain one or more planets. Specifically, this research considered the 9<sup>th</sup> house of a horoscope which focuses on individual education and higher studies including research etc. Generally, 4<sup>th</sup> house signifies the education of the individual person but 9<sup>th</sup> house is suggested for higher studies and it has been considered for the prosperous and blessed life. Initially, this research involved statistical Analysis of Mercury, 9<sup>th</sup> place, 9<sup>th</sup> place planet and 9<sup>th</sup> place planet from its 9<sup>th</sup> place (M999) planets which assist to obtain scores of mercury planet and those three 9<sup>th</sup> place scores. However, the acquired scores from the horoscope of 201 persons have helped to justify manually and provide recommendation but no validation. Therefore, this paper has focused on validating the M999 planets score using ML methods and compared with existing Kala Purusha Thathuva” (KPT). Furthermore, the Weighted Synthetic Minority Oversampling Technique (WSMOTE) is taken into account for the goal of oversampling unbalanced data for accomplishing a balanced dataset. Thus, the prediction of research studies using M999 planets score can be justified through random forest method with higher accuracy of 92.31%.

**Keywords**--astrology, machine learning (ML), 9<sup>th</sup> place planet, M999 planets score, WSMOTE, random forest, research studies.

## Introduction

Astrology is the ancient discipline of calculating the planet positions to forecast good and bad events that occur in the world. Astrology is a subject that provokes a great deal of interest among people which assist to discover scientific hypotheses and their connections to astrology, adjustments in the way astrology is seen must be made. The astrological forecast of many events in human life is established on the resemblance of zodiac signs, planet positions and aspects amongst them. To make a forecast for a person, the astrologer analyze and compare the prior horoscopes. In addition, the astrologers have employed similarity analysis and diagnosis of person's prior horoscopes to foretell various occurrences in person's life in advance about him or her. In astrology, a person's career can be decided by examining their birth chart, whereas the principles employed by astrologers for prediction are not standardized, and there is little scientific proof accessible. Thus, it is necessary for scientists and astrologers to collaborate and combine their efforts for creating standard norms and scientific validity for forecast. [1].

The nine planets are known as "Navagraha," which involves the Moon, Sun, Mercury, Mars, Jupiter, Saturn, Venus, Rahu (North node), and Kethu, (South node) have been utilized for astrological forecasts. The astrological chart is the primary document for making predictions which represent the "Rasi," a group of twelve dwellings with equal partitions. However, they are named as Taurus, Aries, Cancer, Gemini, Virgo, Libra, Leo, Sagittarius, Scorpio, Aquarius, Capricorn, and Pisces. Hence, all twenty seven stars and nine planets are placed in any one of the twelve houses [2]. The relevance of Data Mining (DM) in information analysis can't be overstated. DM is the process of extracting information and hidden truths from large data collections. There are two types DM namely descriptive DM and predictive DM.

Association rules, clustering or sequence finding approaches are the major method utilized in descriptive mining which tends to focuses on the basic data condition. In the case of predictive mining, the recent dataset has been utilized for prediction using regression, time series analysis or classification techniques. Several difficulties faced in various fields like business, security, education, astrology, medicine etc., are resolved using DM process. This paper focus on DM technique which undergone various classifications and predictions technique for minimizing the effort involved in determining the scientific or astrology basis. Moreover, the data analysis is carried out by design instantaneously utilizing these methods. Therefore, the huge quantum of information prevailing for a person's planetary position is analyzed by virtue of incorporating these large amounts of data with learning methods. Thus, it assist in capable of forecasting the numerous facets of a person's life could be made or established.

Numerous classification methods utilized in this work have a lot of similarities to the manner in which the astrology functions. Carrying out different classification systems of artificial intelligence which has been treated as black box for the previously prevailing methods that yielded respectable results in different sectors in classifying data is the first step for forecasting. Discovering the attributes have assist in enhancing the astrology's dynamic forecasting and then apprising the

system's knowledge as several archives comes into play a role is the subsequent step. Classification techniques have helped in generating its out comes within the form of rules. Such rules produced by classification techniques are subjected to backtracking and they are confirmed with respect to the rules utilized by astrologers. The astrology's scientific basis provided are coincidence amongst the rules given by the astrology and the rules produced by classification techniques [3]. Thus, the reason it should come as no surprise that the problem of class imbalance has gathered considerable research interest. There are two main methods that addresses the imbalance data problems namely data-level and algorithm-level method.

In the case of algorithm level, the traditional classification algorithms are adjusted for enhancing the learning task based on the minority class and further regulate the class imbalance ratio for accomplishing the balanced distribution among the classes. In the data-level method, the adaptability and simplicity have played a major role by either conducting major under sampling instance or minority oversampling are utilized as default. Moreover, the avoidance of eliminating major instance is done through oversampling algorithm. Therefore, the Synthetic Minority Oversampling Technique (SMOTE) algorithm is preferred for performance on minority class and generalization capacity [4]. Thus, the novelty of SMOTE represent its synthetic instance creation mechanism that tried to cope with challenges arisen from overfitting the minority class when performing random oversampling, the state-of-the-art method at that time.

This research involve sastronomy prediction of the research studies with the help of scores obtained by mercury planet, 9<sup>th</sup> place, 9<sup>th</sup> place planet and 9<sup>th</sup> place planet which is placed form its 9<sup>th</sup> place. In addition, the M999 planets recommended procedure is evaluated with ML technique to acquire the accuracy of astrology prediction and due to under sampling instances, the accuracy generated is very less. Therefore, this study focus on avoiding the eliminating the oversampling that is done through WSMOTE algorithm that assist in improving the accuracy astrology prediction to justify the research scholar available in the dataset to understand the research scholars about their research studies status.

The paper is organized as follows, the section 2 discusses the related work of the ML technique used in the astrology for the prediction in different aspects and WSMOTE algorithm to avoid under sampling instance and minority oversampling. The proposed M999 planets procedure with WSMOTE is introduced as analyzing factors of the astrology to justify the status of research studies from their horoscope is discussed in section 3. The evaluation results of proposed M999 planet with WSMOTE is compared with existing KPT using various ML technique are discussed in section 4. Finally, the section 5 deal with concluding the proposed M999 planets with WSMOTE accuracy score is higher than existing KPT model has assist to predict the research studies from their horoscope.

## **Literature Review**

This literature review assist in identifying the data mining application that plays the major role for predicting horoscope is a new way. In addition, the SMOTE algorithm assist in improving the accuracy of ML algorithm by modifying the

imbalance instance to balance instance that have been discussed below Dr. K. Kumaravel et al has proposed a study on Predictive based on Knowledge Discovery Decision Tree (KDDT- DSADS) using Machine Intelligence Learning. They used different artificial intelligence algorithm for the classification of horoscope astrology automatically. The algorithms K-Nearest Neighbour, Decision Stump Algorithm Decision Stump (DSADN), KDDT, Bayesian, ANN are applied on the astrology birth chart. The result found that all the artificial intelligent classification algorithms works effectively to predict the future of a person automatically [5].

Helgertz et al has proposed a method to validate the astrological predictions on marriage and divorce. They have analyzed the Swedish record data. They have examined the data for the prediction using longitudinal analysis. They have examined how contrasts in prophetic favorability affect assistant option in marriage including the isolated danger amid hitched couples utilizing Sweden's longitudinal individual level information in between the ages of 1968-2001. As per the outcomes given by this method, their outcome has been unsuccessful in providing and substantial and dependable proof to strengthen the belief and thought that astrologically several steady couples are one or the other overrepresented amid witnessed matrimonial combinations or subjected to a minor risk of getting divorced [6]. O.P. Rishi et al has developed a common method for the astrology prediction based on persons birth chart. They have collected the birth chart from persons from different professions. Common patterns are identified from the person's horoscopy using standard pattern matching algorithm. Relation between the events and planets with their common factors are mapped to form astrology prediction rule. The Case Based Reasoning (CBR) with nearest neighbour hood algorithms used to store and find the previously experienced prediction pattern [7].

One of the more recent algorithms that have also preserved the interpolation mechanism of the original algorithm is Weighted-SMOTE [8], which computes a weight matrix in defining the synthetic information quantity to be produced for every minority occurrence. By utilizing the Euclidean distance of a given minority occurrence in relation to all other residual occurrences of same class, the weights are found. The smaller the distance, the greater the amount of synthetic instances is generated for that minority data sample. Amongst minority examples, this suggestion is presented so as to address sparse regions in a convenient manner. A ratherak in interpolation method is utilized in the Sigma Nearest Oversampling based Convex Combination (SNOCC) [9] oversampling technique, however the synthetic instance is computed as a convex blend of two randomly selected minority instances. This technique aims to generate instances that naturally model the distribution of original samples.

SMOTE is incorporated to present synthetic samples along the line sections joining some or every k-minority class nearest neighbors, taking in view of every minority class instance in information in the case of oversampling scenario. The collection of k nearest neighbors is carried out arbitrarily grounded on the quantity of oversampling required. In order to attain better sampling distribution, SMOTE algorithm has frequently testified successes. Nevertheless, it may yield

suboptimal outcomes or could be counterproductive as well in numerous cases, when it is incorporated in its original form [10].

Adisaptura discusses the fraud prevention in e-commerce using ML and this method is developed but imbalancing data get balanced by use of SMOTE algorithm. Moreover, the SMOTE has improve the various ML algorithm unbalanced data to balance. Hence, utilizing confusion matrix that attains the uppermost accuracy of 96% in Neural Network, the outcome gets assessed. Thus, the SMOTE has improved the F1-score from 67.9% to 94.5% [11]. A novel algorithm is recommended for producing biomedical information's precise classification in this work. The high dimensionality issue and the skewed data distribution are undertaken by this. Established on farther sample, the method is devised by conjoining the FD\_SMOTE and PCA. Various classifiers based on FD\_SMOTE and PCA is utilized from the quantitative and qualitative analysis and it discloses that the novel method upsurges the performance of Area Under Curve (AUC) metrics that is utilized on several information of biomedical sector. The combined technique as displayed by the current analysis seems to be the most effective as compared to the other prevailing approaches like Information Gains (InfoGs) and Correlation based Feature (CFs) [12].

This session provided a study of the different techniques for predicting horoscope astrology and balancing the oversampling scenarios. This research focus in developing the imbalancing data that get balanced by use of various SMOTE algorithm. Moreover, the SMOTE has improve the various ML algorithm for producing unbalanced data to balance. This paper discusses various techniques for developing the imbalanced data to the balanced to avoid underfitting as well as overfitting scenarios by different researchers. The predictive method of astrology is evaluated through ML technique to validate the astrological predictions.

### **Research Methodology**

One of the most ambitious phase faced in this study is to provoke horoscope analysis of the person with their higher studies. Most of the persons wished to choose their research studies as well but they may not be aware about the completion of research studies. However, the astrology of the each person is analyzed by their individual horoscope that has an efficient influence on choosing their career. Hence, it has been found that the horoscope of each person is centered on twelve (12) houses wherein the nine planets are positioned amongst any one of these twelve houses. Thus, key factor of horoscope is lagnam which is the major factor of time that the in fan gets connected with this universe. Moreover, it defines the rasi name that rise in sky during the time at which the baby is born. Therefore, Lagnam plays a significant role in identifying the twelve Bhavas and each Bhavas has it individual characteristics.

### **Data collection**

The collection of data consists of both aspired students who wished to enroll the research studies as well as students who have already enrolled research scholar of various majors from different university and also different geo-location of

research scholar's horoscope. The respective horoscope chart is redefined with 14 attributes inclusive of Ascendant (Asc) with respect to global planet named "Kala Purusha Thathuva" (KPT). The collected detail is preprocessed for any missing imputation and accomplished with 201 scholar's data in the collected dataset and the sample is shown in Figure 1.

	NAME	POB	TOB	DOB	Asc	Sun	Moon	Mars	Mercury	Jupiter	Venus	Saturn	Rahu	Ketu
0	record_Completed_1	Pudukottai	4.00AM	16.07.1981	3	3	9	3	3	6	4	6	4	10
1	record_Completed_2	Mannargudi	5.00AM	15.06.1979	2	2	11	1	3	4	2	5	5	11
2	record_Completed_3	Tenkasi	9.30AM	30.07.1982	6	4	8	7	4	7	3	6	3	9
3	record_Completed_4	Tanjore	7.45PM	30.10.1986	2	7	6	10	8	11	7	8	12	6
4	record_Completed_5	Nagercoil	11.40AM	21.06.1977	5	3	4	1	2	2	1	4	6	12

Figure 1. Data collection of research scholar horoscope

The recommendation of Mercury, 9<sup>th</sup> place, 9<sup>th</sup> place planet and 9<sup>th</sup> place planet from its 9<sup>th</sup> place (M999) planets using Statistical Analysis (SA) has assist to obtain scores for M999 planet are shown in Figure 2. This collection of data involves higher studies completed scholars and similarly data have been collected from aspiring scholars as well as pursuing for research studies persons.

	Name	POB	TOB	DOB	asc_zodiac_sign	mercury_score	p9_score	p9_bm_score	p9_distance_score	Status
0	Dr. Carmel Prabha Mary	Pudukottai	4.00AM	16.07.1981	Gemini	0.833333	0.833333	1.0	0.5	completed
1	Dr. Geetha	Mannargudi	5.00AM	15.06.1979	Taurus	0.666667	0.500000	1.0	0.5	completed
2	Dr. Athi Narayanan	Tenkasi	9.30AM	30.07.1982	Virgo	0.333333	0.833333	1.0	1.0	completed
3	Dr. Bhuvana	Tanjore	7.45PM	30.10.1986	Taurus	0.666667	0.666667	0.0	1.0	completed
4	Dr. B.T.Geetha	Nagercoil	11.40AM	21.06.1977	Leo	0.833333	0.833333	0.5	1.0	completed

Figure 2. Data collection of M999 Planets score

The performance accuracy of the classification method involves in this research acquire less because of imbalanced data. Thus, WSMOTE has been preferred for balancing the imbalance learning in the ML technique.

### Working of WSMOTE

The classifier performance over minority class can be enhanced using WSMOTE which assist in improving the capacity of generalization and excluding the possibility of overfitting occurrence. The purpose of involving mechanism in synthetic instance creation is based on recent minority instances that are not regenerated with existing records but it is placed in the neighborhood. However, the each minority instance 'y' inX, at first the k-neighborhood has been defined with several metric functions that may be implemented using Euclidean distance method. Hence, the creation of 'n' synthetic instance is a formation of convex combination shown in equation 1.

$$S_i = y + D_i^u(0,1) * (\hat{y}^{r(i)} - y), \quad i = 1, 2, \dots, n \quad (1)$$

Where,

$\hat{y}^{r(i)}$  = Random neighbor from the k-neighborhood

$D_i^u(0,1)$  = Exact placement of random number produced by uniform distribution

Thus, the creation of instance is developed on the line among the minority instance  $y$  and the random neighbor of the k-neighborhood. The WSMOTE algorithm is illustrated below.

### WSMOTE Algorithm

Input: Minority set as  $X$ , Number of synthetic data as  $n$  and amount of nearest neighbor as  $k$ .

- Step 1: Let the synthetic data set is considered as  $S=0$  initially.
- Step 2: When  $\forall y \in X$ , then the k-neighborhood has been determined as  $N_k(y)$  of  $y$  as an instance of synthetic creation.
- Step 3: The creation of  $n$  synthetic instances are defined with for condition as  $i = 1:n$  that randomly select the neighbor from the k-neighborhood  $\hat{y}^{r(i)} \in N_k(y)$ .
- Step 4: The convex combination of synthetic instance creation is provided to avoid overfitting of the sample available in the data as  $S = S \cup s_i$
- Step 5: Repeat the step 3 and step 4 till the overfitting avoided.
- Step 6: Return

The WSMOTE algorithm has enhance the development for avoiding the imbalance ML data by the mechanism of synthetic instance creation. However, the creation of synthetic instance along the line segment contiguous with chosen class neighbors. Therefore, the algorithm has been incorporated the synthetic instance creation by integrating the compatibility of default description shown in equation 2.

$$S_i = y + D_i^u(0,1) \odot (\hat{y}^{r(i)} - y), \quad i = 1, 2, \dots, n \quad (2)$$

Where:

$\odot$  = Hadamard product

Based on the WSMOTE algorithm influence, the complete issue of ML from the imbalanced data have avoided the default interpolation mechanism that integrate and modify the other certain procedures.

```
In [29]: # After balancing using WSMOTE
         np.unique(y_res, return_counts = True)

Out[29]: (array([0, 1, 2]), array([67, 67, 67], dtype=int64))
```

Figure 3. Balancing of horoscope samples using WSMOTE

Figure 3 illustrates the implementation of interpolation mechanism that generate a range restriction with the structure of the nearest neighborhood. This assist in affecting the location of a newly created synthetic instance. One of the earliest and well-established WSMOTE-based extensions involved in this type of interpolation is WSMOTE that has strengthened the outline minority by providing the synthetic creative instance in random uniform distribution. Therefore, the category of completed, pursuing and aspiring of research studies scholars are 67, 67 and 67. Once the samples are modified and split based on train as well as test dataset with 75% and 25% respectively.

### **Algorithm for M999 planet with WSMOTE**

Input: Metadata of completed, pursuing and aspiring

Output: scores of mercury and 9<sup>th</sup> planet

- Step 1: Before collection of metadata, data processing is initiated based on 12 houses and 9 planets. “Asc” is always placed the first position and followed by 9 planets.
- Step 2: After data processing, metadata are collected and followed by planet weightage calculation.
- Step 3: Based on the recommended rules, weightage of mercury is calculated by its mean score of mercury as mercury\_score. Similarly 9<sup>th</sup> planet scores are computed and mean is calculated and stored in p9\_score field.
- Step 4: Benefic\_malfic\_dict is created as per recommended astrology rule and mean score of benefic and malfic is determined and maintained in the p9\_bm\_score.
- Step 5: Confident 9<sup>th</sup> position of 9<sup>th</sup> planet is computed for determining the place of higher studies in the house of horoscope and handled in the field of p9\_distance\_score.
- Step 6: Based on the asc\_zodoac\_sign, the step 3, step 4 and step 5 are progressed and maintained as record for the respective records.
- Step 7: If imbalancing of data causes over-fitting then WSMOTE algorithm is progressed else balanced data with no over-fitting returns and end if.

The fitting predictive model is generated to avoid over-fitting and produce balanced data to improve the accuracy of the astrology predictive model to define the exact information to the scholar of pursuing and aspiring about the research studies status. Thus, the proposed M999 planet with WSMOTE is evaluated through confusion matrix parameter with existing KPT to justify that proposed method shows better accuracy in astrology prediction than existing KPT.

### **Results and Discussion**

This research focus on understanding the power of lagnam that supports for research studies with the help of recommended M999 planets procedure. The procedure guided to define the weightage of mercury planet, 9<sup>th</sup> place score, 9<sup>th</sup> place planet score and 9<sup>th</sup> place planet placed from its position. However, the experiment initially applied ML concept with Kala Purusha Thathuva (KPT) method but it doesn't perform high accuracy. Hence, the recommended astrology



rules have been applied and generated the procedure of M999 planet with the score of mercury, and those three 9<sup>th</sup> place scores. Thus, the accomplished scores are interpreted with the status of respective completed, pursuing and aspiring status are evaluated through various ML techniques. After synthetic instance balancing dataset there are 201 records are available in which 151 records as 75% are considered for train dataset and 50 records as 25% are considered for test dataset.

The various ML techniques involved in the evaluation are Support Vector Classifier (SVC), Naïve Bayes (NB), Decision Tree Classifier (DTC), Multinomial Logistic Regression (MLR), K-Nearest Neighbor (KNN) and Random Forest Classifier (RFC). The performance of M999 planets procedure analysis by WSMOTE for research studies using ML technique and the existing KPT with ML has been determined by accuracy score. The accuracy scores obtained from the confusion matrix parameters of several ML techniques that are obtained multi-class classification for both KPT and M999 planet with WSMOTE method. The existing KPT of binary classified accuracy score for various ML techniques are listed in Table 1.

Table 1  
Accuracy score of ML techniques from existing KPT for astrology prediction in research studies

ML Technique	Accuracy score for Existing KPT
Decision Tree Classifier (DTC)	0.3823
Support Vector Classifier (SVC)	0.5000
Random Forest Classifier (RFC)	0.6176
Multinomial Logistic Regression (MLR)	0.5294
Naïve Bayes (NB)	0.5588
K-Nearest Neighbor (KNN)	0.5294

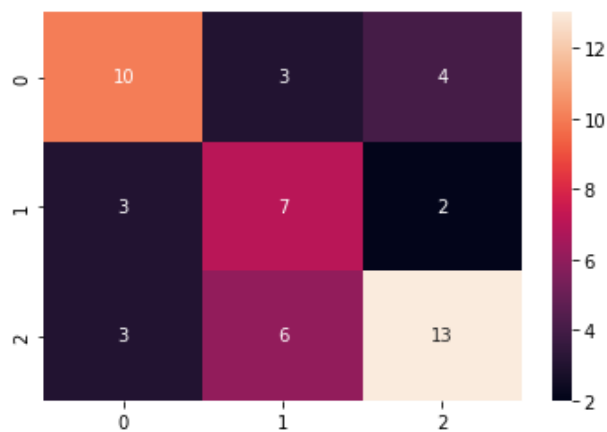


Figure 4. Confusion Matrix for multi-class classification of Decision Tree Classifier (DTC)

Figure 4 illustrates that three type of classification issues can be determined through this kind of categorization. A similar as binary classification, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) plays the major role but it not as typical as binary classification. For example, TP, TN, FP and FN value for DTC are computed and array '0' as completed, '1' as Pursuing and '2' as Aspiring.

### For Completed

$$\begin{aligned} TP &= 10 \\ TN &= (7+13+2+6) = 28 \\ FP &= 3+3 = 6 \\ FN &= 3+4 = 7 \end{aligned}$$

Similarly, the **pursuing** and **aspiring** is calculated but multi-label binary problems need to be manipulated for generating Micro F1-score which can be utilized in assessing multi-label binary quality. Micro F1-score is the aggregated contribution of all classes namely completed, pursuing and aspiring. As per micro F1-score metrics calculated globally are measured to be equal is shown in equation 1.

$$\text{Precision} = \text{Recall} = \text{Micro F1 - Score} = \text{Accuracy} \quad (1)$$

The micro precision and micro recall can be calculated using total TP, total FP and total FN whereas the three classes are accumulated. The same instance of DTC is shown in Table 2.

Table 2  
Multi-class classification of DTC with binary classification

Classifier Model Description	TP	TN	FP	FN
DTC_Completed	10	28	6	7
DTC_Pursuing	7	30	9	5
DTC_Aspiring	13	23	6	9
	30		21	21

### Micro Precision

Formulae is as like as general precision formula of binary classification but instead of TP and FP, total TP and total FP is used.

$$\begin{aligned} \text{Micro precision} &= \frac{\text{Total TP}}{(\text{Total TP} + \text{Total FP})} \\ \text{Micro precision} &= \frac{30}{(30 + 21)} = 0.5882 \end{aligned}$$

### Micro Recall

Formulae is as like as general recall formula of binary classification but instead of TP and FN, total TP and total FN is used.

$$\text{Micro Recall} = \frac{\text{Total TP}}{(\text{Total TP} + \text{Total FN})}$$

$$\text{Micro Recall} = \frac{30}{(30 + 21)} = 0.5882$$

Hence, as per equation 1 Micro F1-score is 0.5882 which is said to be accuracy of the DTC model. Thus, the accuracy of DTC model is 0.5882 (58.82%). Similarly, there are various classifier methods are calculated and shown in Table 3.

Table 3  
Accuracy of various classifier model with proposed M999 planet by WSMOTE technique

Classifier Model	Accuracy (%)
DTC	58.82
Support Vector Classifier (SVC)	23.53
Random Forest Classifier (RFC)	92.31
Multinomial Logistic Regression (MLR)	35.29
Naïve Bayes	49.02
K-Nearest Neighbor (KNN)	54.90

Table 4 is provided with comparative accuracy score for proposed M999 planet with WSMOTE and existing KPT. There are various ML techniques are used to evaluate the accuracy in predicting the astrology score for research studies.

Table 4  
Accuracy comparison of various classifier model with existing KPT and M999 planet by WSMOTE technique

ML Technique	Accuracy Score	
	M999 planet by WSMOTE	Existing KPT
DTC	0.5882	0.3823
SVC	0.2353	0.5000
RFC	0.9231	0.6176
MLR	0.3529	0.5294
Naïve Bayes	0.4902	0.5588
KNN	0.5490	0.5294

Figure 5 illustrates the accuracy score for M999 planet with WSMOTE has high accuracy of 0.9231 in RFC followed by DTC with 0.5882 and KNN as 0.5490 correspondingly while compared to existing KPT. In the case of Naïve Bayes, MLR and SVC, the accuracy score is 0.4902, 0.3529 and 0.2353 respectively are lesser while compared to existing KPT. However, the overall comparison in accuracy

score for RFC has high score of 0.9231 that is higher than all other ML techniques. Hence, the astrology prediction of research studies is done better in RFC for M999 planet with SMOTE.

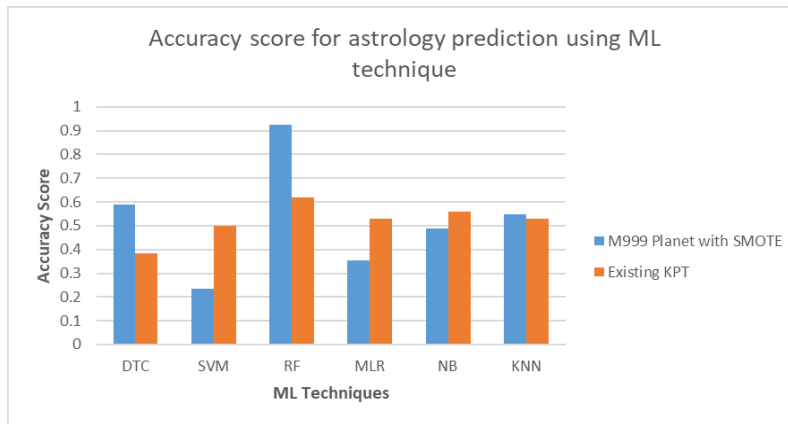


Figure 5. Accuracy score for astrology prediction using ML technique

Moreover, the astrology prediction using M999 planets with SMOTE may assist in understanding the scholar's horoscope about their status like completed, pursuing and aspiring. Hence, the recommended rules of bhavas may give confidence to the person who believe in horoscope for their research studies. The CM parameters like precision, recall and F1-Score for each category in the proposed M999 planet with WSMOTE have been evaluated to the respective ML techniques are calculated and considered as macro average F1-score. Similarly, the weighted score is obtained by considering weight for each category are made to sum up the total number of samples. The macro average and weighted average of precision, recall and F1-Score with respect to ML techniques are shown in Table 5.

Table 5  
Macro Avg. and weighted Avg. of CM parameters for M999 with SMOTE

ML Technique	Macro Average precision score	Macro Average Recall score	Macro Average F1-score value	Weighted Average Precision Score	Weighted Average Recall Score	Weighted Average F1-Score value
valeDTC	0.58	0.59	0.58	0.61	0.59	0.59
SVC	0.08	0.33	0.13	0.06	0.24	0.09
RFC	0.76	0.77	0.75	0.81	0.76	0.77
MLR	0.36	0.36	0.35	0.37	0.35	0.36
Naïve Bayes	0.47	0.47	0.47	0.51	0.49	0.50
KNN	0.52	0.52	0.51	0.57	0.55	0.54

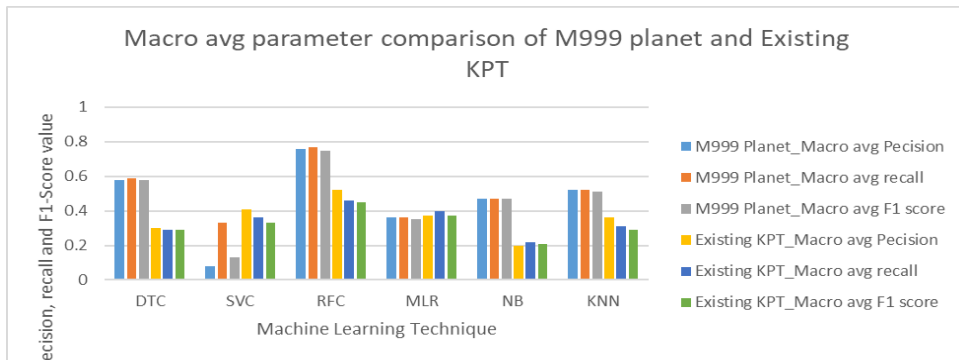


Figure 6. Macro AvgPrecision, Recall and F1-Score value for astrology prediction using ML techniques

Figure 6 illustrates the macro avg precision, recall and F1-Score for various ML techniques that shows RFC method performance of macro avg precision, recall and F1-Score is higher than other ML technique as well as existing KPT. The macro avg precision of RFC is 0.76, macro avg recall for RFC is 0.75 and macro avg F1-Score is 0.75 whereas for existing KPT is 0.52, 0.46 and 0.45 respectively. Simultaneously, the RFC method has higher CM parameter score while compared to other ML techniques. The FP, FN and TP is considered for determining the prediction of astrology to research studies with M999 planet with WSMOTE is high.

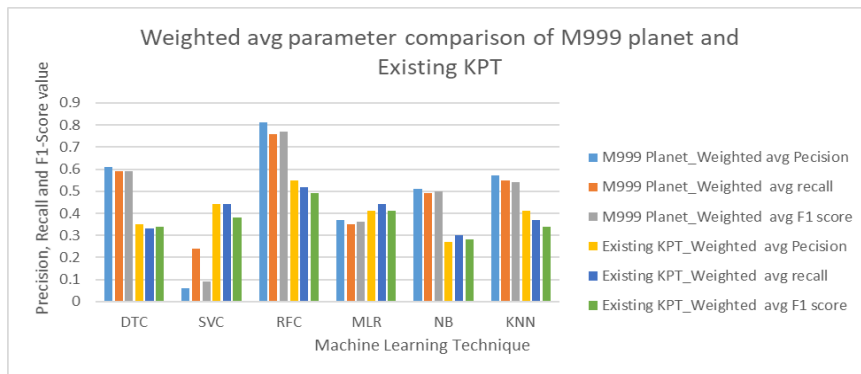


Figure 7. Weighted Avg Precision, Recall and F1-Score value for astrology prediction using ML techniques

Figure 7 illustrates the weighted avg precision, recall and F1-Score for various ML techniques that shows RFC method performance of weighted avg precision, recall and F1-Score is higher than other ML technique as well as existing KPT. The weighted avg precision of RFC is 0.81, weighted avg recall for RFC is 0.76 and weighted avg F1-Score is 0.77 whereas for existing KPT is 0.55, 0.52 and 0.49 respectively. Simultaneously, the RFC method has higher CM parameter score while compared to other ML techniques. The FP, FN and TP is considered along with the individual categories sample counts weightage for determining the prediction of astrology to research studies with M999 planet with WSMOTE is high.

## Conclusion

Astrology is a well-known conventional science that is utilized all over the world and it assist in making decisions in various circumstances namely life partner selection, job selection, business selection, land purchasing, starting new house construction, etc. However, the astrology is an unavoidable and also pursue certain traditional prediction methods which keep away from modern technologies because of not accepting the astrology as provoked science in modern science. Hence, this work shows that DM can be proficiently made applicable to astrological prediction. By analyzing its sample and applying the different ML algorithms, the higher studies of the person's horoscope can be classified and predicted. This type of application of DM has assisted the traditional systems that extremely improve the credibility of astrology. This, the M999 planets with WSMOTE has better prediction results with RFC method as 92.31% while comparing with other ML technique as well as for existing KPT method. In future work, the accuracy scores will be increase by recommending certain more rules as well as increasing the horoscope data. Moreover, recommended rule based procedure is evaluated through deep learning methods to provide high accuracy in predicting the research studies present in the horoscope of the person.

## References

1. A.Saputra, Suharjito, "Fraud detection using machine learning in e-commerce", *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 9, 2019.
2. Chaplot N, Dhyani P, Rishi O. P, "Predictive approach of case base reasoning in artificial intelligence: in case of astrological predictions about famous personalities", *ACM International Conference Proceeding Series*, ISBN:9781450339629, 2016.
3. Chaplot, N., Dhyani, P. and Rishi, O.P., 2015, May. Astrological prediction for profession using classification techniques of artificial intelligence. In *Computing, Communication & Automation (ICCCA)*, 2015 International Conference on (pp. 233-236). IEEE.
4. Dr.K.KumaravelandMr.S.Dineshkumar , " Analysis study on Predictive based on Knowledge Discovery Decision Tree (KDDT- DSADS) using Machine Intelligence Learning ", *IJCRT | Volume 6, Issue 2 April 2018 | ISSN: 2320-2882*.
5. Fernández, S. García, F. Herrera, N. V. Chawla, "SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary", *Journal of Artificial Intelligence Research*, Vol. 61, 2018, pp. 863-905.
6. Hadi, E. H., Lafta, H. A., & Al-Rashid, S. Z. (2022). Survey for Lung diseases using machine learning methods. *International Journal of Health Sciences*, 6(S3), 4510–4523. <https://doi.org/10.53730/ijhs.v6nS3.6887>
7. Helgertz, Jonas, and Kirk Scott. "The validity of astrological predictions on marriage and divorce: a longitudinal analysis of Swedish register data." *Genus* 76.1 (2020): 1-18.
8. J.A. Sáez, J. Luengo, J. Stefanowski, F. Herrera, Smote-ipc: addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering, *Inf. Sci.* 291(2015), 184–203.

9. M. R. Prusty, T. Jayanthi, K. Velusamy, “WeightedSMOTE: A modification to SMOTE for event classification in sodium cooled fast reactors”, *Progress in Nuclear Energy*, Vol. 100, 2017, pp. 355-364.
10. N. Mustafa, J. Li, R. A. Memon and M. Z. Omer, “A classification model for imbalanced medical data based on PCA and farther distance based synthetic minority oversampling technique,” *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 1, pp. 61–67, 2017.
11. O.P. Rishi and NeelamChaplot, “Archetype of astrological prediction system about profession of any persons' using case based reasoning” , *International Conference on Communication and Computational Intelligence*, Dec 2010, PP. 373 – 377,Dec 2010.
12. Pérez, A. V., Gámez, M. R., Briones, V. F. . V., Viteri, C. G. V., & Molina, L. A. V. (2018). Sustainable development seen from environmental training in university linkage. *International Journal of Life Sciences*, 2(1), 12–20. <https://doi.org/10.29332/ijls.v2n1.75>
13. Shajan R, Gladston R. S, “Horoscope analysis and astrological prediction using biased logistic regression (BLR)”, *International Journal of Innovative Technology and Exploring Engineering*, ISSN:22783075, 2019.
14. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). The COVID-19 pandemic. *International Journal of Health Sciences*, 5(2), vi-ix. <https://doi.org/10.53730/ijhs.v5n2.2937>
15. Z. Zheng, Y. Cai, Y. Li, “Oversampling method for imbalanced classification”, *Computing and Informatics*, Vol. 34, No. 5, 2016, pp. 1017-1037.