



# Unleashing Creativity And AI For Play: Machine Learning Algorithms For Predicting Children's Hobbies And Early Interest Forecasting

<sup>1</sup>Dr. Nazimunissa S, <sup>2</sup>Puneet Shetteppanavar, <sup>3</sup>S Krishna Reddy, <sup>4</sup>Shaik Ahmed Pasha

<sup>1</sup>Associate Professor, <sup>2,3,4</sup>Assistant Professor

<sup>1,2,3</sup>Department of Computer Science and Engineering, <sup>4</sup>Department of Electronics and Communication Engineering

<sup>1,2,3,4</sup>Sree Dattha Institute of Engineering and Science, Hyderabad, India

**Abstract:** A recommendation system plays a vital role in the social media world, and it is used in many applications which provide recommendations based on user preferences. A recommendation is nothing but a choice-making system. The utilization of recommender systems is increasing daily as many choice-based challenges potentially influence users. Now a day's, parents are worried about their kids regarding the future. Most parents need help to identify their kid's interests and force their interests against kids. So, this prediction will be helpful to society. There are many types of recommendation systems with different applications that have been successfully adopted. This work needs to specify the age limits for kids, and it's preferable for below ten years old kids. The dataset is prepared based on different parameters. Fourteen parameters have been identified, and the result is three possible classifications. The first classification is 1) academics, 2) arts, and 3) sports. The academics indicate the kid's interest in science with a mathematics background. A total of 1600 dataset has been used for training and modelling the different algorithm such as Logistic Regression (88.16%), Adaptive Boosting Classifier (87.85%), Decision Tree Classifier (89.41%), Gaussian Naïve Bayes Classifier (90.65%), Random Forest Classifier (91.90%) and Support Vector Classifier (89.41%). We have developed the python-based flask web application to demonstrate the predictions.

**Keywords:** Machine Learning, Children's Hobbies Prediction, Forecasting, Django Framework

## I. INTRODUCTION

Today's technology-driven kids tend to be glued to their gadgets all day. Therefore, it becomes necessary to encourage them to pick up a hobby that may expose them to outdoor activities, bring them closer to nature, and provide a suitable outlet to spike their creativity and interests. Teaching kids new hobbies is an enjoyable and captivating way to spend time. Kids are fond of leisure activities and acquiring new skills. Activities, such as reading, writing, painting, etc., are productive ways to spend time while fostering cognitive and emotional growth. Moreover, participating in hobbies together can reinforce the relationship between parent and child. You can even include your children in your hobbies, such as cooking or gardening. Keep reading to discover interesting indoor and outdoor hobbies for kids that promote their overall development.

A hobby is a fun activity based on a person's interests. Many people practice hobbies during their time off from work or studying. Popular hobbies include drawing, collecting baseball cards, and building Web sites. A person with a hobby is called a hobbyist. Most hobbies involve collecting, creating, observing nature, or using special equipment. Many people collect stamps, coins, seashells, jewelry, or even action figures. Creative hobbyists enjoy arts and crafts—for example, painting, singing, writing, carving, sewing, or building models.

Hobbies for people who enjoy nature include bird-watching and gardening. Photography, ham radio (amateur radio), and computer programming are hobbies that require equipment. Many hobbies are done alone, but people often form groups or read magazines based on their hobby. No matter what the hobby, there is always an opportunity to learn from it. Before the 1900s most people spent nearly all their time working and caring for their families. Only wealthy people could afford to have hobbies. After the Industrial Revolution, modern technology and machines gave people more spare time, so more people developed hobbies. In the late 20th century computers created a whole new class of hobbies. Many people today spend their spare time surfing the Internet and playing video games.

A hobby can shape the personality of a kid positively by helping them become confident, self-reliant, and independent. Having a hobby also enhances the thinking skills of kids. A hobby instils qualities like team spirit, compassion, and patience in them and teaches various life skills that may mould them into healthy adults. Children tend to imitate their elders. So, you can generate his interest by setting a personal example. You can arouse his curiosity by indulging him in your own hobby. He is likelier to join in when he watches you do something fun and creative. And what may start as a fun and light activity can eventually become his hobby. Another way to help your child choose a hobby can be to take him to explore interesting places like museums, collector's studios, art galleries, plays, and concerts. He will not only learn new things which will help broaden his horizon and awareness levels but will also guide him to discover his area of interest. Activities involving enjoyable and exciting ways to improve motor skills and stimulate creative thinking. According to research, these activities foster children's healthy mental health and resilience.

## II. LITERATURE SURVEY

Conducting a literature survey on predicting children's hobbies using machine learning involves exploring various research papers, articles, and studies that examine the intersection of machine learning, child behavior, and hobby or interest prediction. This field is relatively niche but intersects with broader research areas like recommendation systems, machine learning algorithms for predictive modelling, and analyses of children's interests and behaviors.

In future, the people will believe the recommendation of the system rather than the people. The recommendation systems are divided into two types personalized and non-personalized. Then the personalized category is divided into three types 1) Content-based filtering, 2) Collaborative filtering 3) Hybrid filtering [1-4]. Each type uses different algorithms to recommend the item to the user. Collaborative filtering is divided into two types 1) model-based filtering and 2) memory-based filtering [1-4]. Hybrid filtering is a combination of both content-based and collaborative-based [1-4]. The model-based filtering uses popular machine learning algorithms like Clustering, Association, Bayesian, and Neural networks [5]. The memory-based filtering is divided into two types one is user based, and another one is item based. Hybrid model filtering works in both combinations of content-based and collaborative-based. Also, a non-personalized recommendation system works on popularity based.

Riyadi et al. employed the C4.5 and Random Forest algorithms to predict children's talents based on a dataset collected from 1601 parents. They tested eight models using four training/testing splits (50:50, 60:40, 70:30, 80:20) and evaluated performance using 10-fold and 20-fold cross-validation. Performance metrics included accuracy, F-score, precision, and recall. Their study demonstrated that the Random Forest model with a 70:30 split and 20-fold cross-validation achieved the highest accuracy of 91.5%, showcasing the model's robustness in talent prediction.

Amrale et al. implemented machine learning algorithms such as decision trees, Random Forest, support vector machines, and AdaBoost to predict students' career paths. Using Python for development, they emphasized the practical benefits of AI in career counseling. Random Forest stood out, achieving an accuracy of 93%, indicating its potential to guide individuals in making informed career decisions. The study highlighted the transformative role of AI in personalizing career guidance.

Bando et al. proposed a novel approach to estimate hobby categories of Twitter users by analyzing sequential tweets. Their study compared deep learning models (RNN, LSTM, GRU, CNN) and traditional machine learning methods (Random Forest, SVM) with different feature extraction techniques like word embeddings and Bag of Words. The LSTM model performed best, achieving the highest accuracy when analyzing sequences of six tweets, underlining its ability to capture sequential dependencies in text data effectively.

Pandey et al. investigated the use of machine learning and ensemble learning algorithms to predict the career paths of engineering students. Their research focused on identifying factors influencing career choices and how classification techniques can assist students in decision-making. By leveraging these algorithms, the study aimed to improve the accuracy and reliability of career predictions, demonstrating the importance of data-driven support systems for students.

Duraisamy et al. developed a recommender system to predict hobbies in children under ten years old based on a dataset of 1600 samples and 14 parameters. Their classification included three categories: academics, arts, and sports. They employed Azure and auto-sklearn tools to test multiple algorithms. Multiclass Logistic Regression using Azure achieved an accuracy of 93.21%, while Liblinear SVC using auto-sklearn obtained 93%. This study emphasized the potential of machine learning tools to identify interests in young children and guide their development.

### III. PROPOSED SYSTEM

The working of the system starts with the collection of data and selecting the important attributes.

- Then the required data is pre-processed into the required format.
- The data is then divided into two parts training and testing data.
- The algorithms are applied and the model is trained using the training data.
- We have applied various Machine Learning Algorithms namely Logistic Regression, Decision Tree, Random Forest, Boosting Algorithm and Support Vector Machines
- The accuracy of the system is obtained by testing the system using the testing data.

The proposed diagram of kids hobby prediction is given in Figure 1.

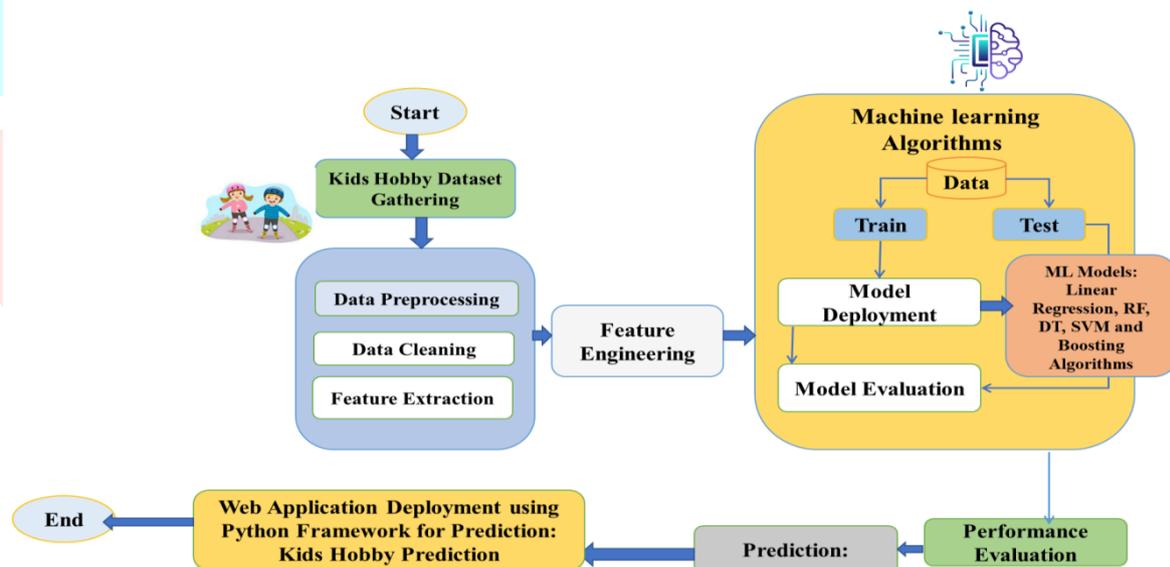


Figure 1. Kids Hobby prediction proposed Diagram

This system is implemented using the following modules.

- Collection of Dataset
- Selection of attributes
- Data Pre-Processing
- Balancing of Data
- Hobby Prediction

### 3.1 Dataset used

This dataset contains 13 attributes and 3 labels, having Categorical, Boolean, and Numerical data. Dataset descriptions is given in Table 1.

Table 1. Dataset descriptions

Features	Meaning
Olympiad_Participation	Has your child participated in any Science/Maths Olympiad?
Scholarship	Has he/she received any scholarship?
School	Love's going to school?
Fav_sub	What is his/her favorite subject?
Projects	Has done any projects under academics before?
Grasp_pow	His/her Grasping power (1-6)
Time_sprt	How much time does he/she spend playing outdoor/indoor games?
Medals	Medals won in Sports?
Career_sprt	Want's to pursue his/her career in sports?
Act_sprt	Regular in his/her sports activities?
Fant_arts	Love creating fantasy paintings?
Won_arts	Won art competitions?
Time_art	Time utilized in Arts?

In our project, implementing a machine learning model involves several stages, from understanding the problem to deploying the model. Here's a high-level overview of the process:

- **Problem Definition:** Understand and define the problem you are trying to solve with machine learning. This involves specifying the task (e.g., classification, regression, clustering) and determining how machine learning can add value.
- **Preparation of the dataset:**
  - Data Collection: Gather the data needed for training and testing the model. This might involve collecting new data, using existing datasets, or augmenting data.
  - Data Cleaning: Clean the data to handle missing values, outliers, and errors.
  - Data Pre-processing: Normalize or standardize the data, and encode categorical variables as necessary.
- **Feature Engineering:** Selecting the most relevant features that can possibly engineer new features to improve model performance need to be considered. This step is crucial because the quality of features significantly impacts model accuracy.
- **Choosing Machine Learning Models:** Select an appropriate machine learning algorithm based on the problem type, data characteristics, and the computational resources available. Common choices include decision trees, support vector machines, neural networks, and ensemble methods.
- **Split the Data into Training, Test and validation:** Split the dataset into training, validation, and test sets. The training set is used to train the model, the validation set to tune the model parameters, and the test set to evaluate its performance.
- **Training different Models:** Train the model using the training dataset. This involves using a machine learning algorithm to learn the relationship between the input features and the target variable. Evaluate the model on the validation set and tune hyperparameter to improve performance. Techniques like grid search, random search, or more advanced methods like Bayesian optimization can be used for hyperparameter tuning.
- **Performance Evaluation of the Model:** Assess the model's performance on the unseen test dataset to gauge its generalization ability. Common evaluation metrics include accuracy, precision, recall, F1 score for classification tasks, and mean squared error, mean absolute error for regression tasks.
- **Deploy the Model using Python Flask Web Application:** After achieving satisfactory model performance, the next step is deployment. This involves integrating the model into an existing application or building a new one around it. Key considerations during deployment include ethical implications, data privacy, and bias mitigation. These aspects depend on the application, tools, data, and problem domain. The project utilized algorithms such as Logistic Regression, AdaBoost, Decision

Tree, Support Vector Machine, Gaussian Naive Bayes, and Random Forest Classifier to achieve optimal results.

### 3.2 Data Preprocessing

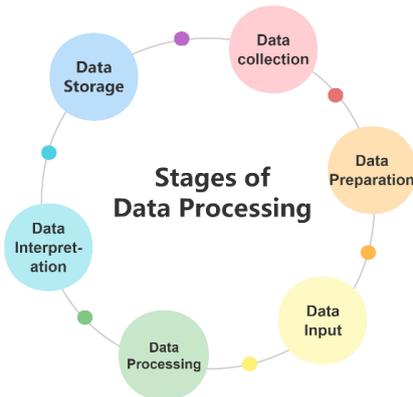


Figure 2. Data Preprocessing

Data preprocessing for a kid's hobby dataset involves addressing null values and converting categorical variables into numerical representations (Figure 2). To ensure data completeness, any instances of missing values are identified and either imputed or removed, depending on the extent of the missing data. This step is essential for facilitating machine learning algorithms to effectively interpret and learn from the data. By performing these preprocessing steps, the dataset becomes more suitable for analysis and model training, ensuring that no null values disrupt the process and that categorical variables are appropriately converted for computational compatibility.

### 3.3 Data Extraction & Feature Engineering

We utilized the diamond dataset from Kaggle for our analysis. The process began with **data preprocessing**, addressing missing values and converting categorical variables into numerical forms to prepare the data for machine learning. Next, **exploratory data analysis (EDA)** was performed to uncover patterns, trends, and relationships within the dataset, offering valuable insights for modeling. The dataset was then partitioned into training and testing sets with an 80-20 split, ensuring sufficient data for training while reserving an independent subset for evaluating model performance. This approach helps assess the model's generalization and predictive accuracy on unseen data.

### 3.4 Prediction

In our analysis of the kid's hobby dataset, we employed various machine learning algorithms, including Linear Regression, Adaboost Classifier Decision Tree, Random Forest Classifier, Support Vector Classifier, GaussianNB to predict kid's hobby. After applying these algorithms, we compared their performance based on key evaluation metrics. To comprehensively evaluate the performance of our chosen model, we employed multiple metrics. The Accuracy score provides a measure of correct predictions.

## IV. SYSTEM DESIGN AND IMPLEMENTATION

System design is a process through which requirements are translated into a representation of software. Design provides us with representation of software development; design is the only phase where requirements are accurately translated into finished or nearly finished software product or system.

System design refers to modelling of a process and approach to create new system. It can be defined as a transition from user's view to programmer's view. System design phase act as a bridge between the requirement specification and the implementation phase. It is the process of defining the architecture, components, modules interface and data for a system to satisfy specified requirements. The design of the system is perhaps the most critical factor affecting the quality of the software; hence we build the system block diagram that will be helpful to understand the behaviour of the system.

The purpose of system design is to create a technical solution that satisfies the functional requirements for the system

The design stages involve two sub stages namely:

- High-level design
- Low-level design

## 4.1 High-Level Design

High-level design involves decomposing system into modules and representing the interface and invocation relationship among modules. It provides an overview of an entire system identifying all its elements at some level of abstraction.

High-level design will usually include a high-level architecture diagram depicting the components; it may also depict work flow/data flow between component systems. This phase is the first step in moving from the problem domain to the solution domain. Here we divide the whole problem into modules, dataflow diagram show flow of data between or among modules.

### 4.1.1 Activities of High-level design

#### Design consideration

It describes many issues, which need to be addressed or resolved before attempting to device a complete design solution. There are many aspects in design; the importance of each should reflect the goals of the software trying to achieve.

Some of the aspects of Design consideration are:

- **Compatibility:** The software is able to operate in the other products that are designed for interoperability with another product.
- **Extensibility:** New capabilities can be added to the software without major changes to the underlying architecture.
- **Fault tolerance:** software resistant to and able to recover from component failure.
- **Maintainability:** the software can be restored to a specified condition within a specified condition within a specified period of time.
- **Modularity:** the resulting software comprises a well-defined, independent component that leads to better maintainability. The components could be then implemented and tested in isolation before being integrated to form a desired software system. This allows division of work in a software development project.
- **Packaging:** printed material such as the box and manuals should match the style designated for the target market.
- **Reliability:** the software is able to perform a required function under stated conditions for a specified period of time.
- **Reusability:** the modular components designed should capture the essence of the functionality expected out of them and no more or less. This single-minded purpose renders the components reusable wherever there are similar needs in other designs.
- **Robustness:** The software is able to operate under stress or tolerate unpredictable or invalid input. For example, it can be designed with resilience to low memory conditions.
- **Security:** The software is able to withstand hostile acts and influences.

## 4.2 Architecture Design

Architecture design process is concerned with establishing a basic structural framework that identifies the major components of a system and the communication between there components. Architectural design is a creative process where you try to establish a system organisation that will satisfy the functional and non-functional system requirements. During the architectural design process, system architects have to make a number of fundamental decisions that profoundly affect the system and its development process.

Some key elements in software architectural model are:

- **Rich:** for the viewpoint in question, there should be sufficient information to describe the area in detail
- **Rigorous:** the architect has applied a specific methodology to create this particular model, and the resulting model 'looks' a particular way.
- **Diagram:** in general, a model may refer to *any* abstraction that simplifies something for the sake of addressing a particular viewpoint.
- **Standards:** standards work when everyone knows them and everyone uses them. This allows a level of communication that cannot be achieved when each diagram is substantially different from another.
- **Primary Concern:** it is easy to be too detailed by including many different needs in a single diagram.
- **Illustrate:** the idea behind creating a model is to communicate and seek valuable feedback.
- **Specific set of trade-offs:** the architecture trade-off analysis method (ATAM) methodology describes a process whereby software architecture can be peer-reviewed for appropriateness.

System implementation is the phase in the software development lifecycle where the designed system is translated into a working and functional reality. This process involves coding, testing, integration of components, and deployment of the software or system. It aims to bring the conceptual design into practical use, addressing technical specifications and user requirements. System implementation involves meticulous attention to detail, collaboration among development teams, thorough testing procedures, and often requires adapting to unforeseen challenges. Successful implementation results in a fully operational system that meets the defined objectives, paving the way for subsequent maintenance and optimization efforts.

## V. RESULTS AND DISCUSSION

This section presents the outputs and findings of the project, detailing the methodologies used to process the sequences and derive results. Additionally, a comparative analysis of the accuracy achieved by various machine learning models is conducted to evaluate their performance.

**5.1 Home Page:** In this web page, all the tabs related to kids hobby prediction app are displayed. Its main page of the project is as shown below in Figure 3.

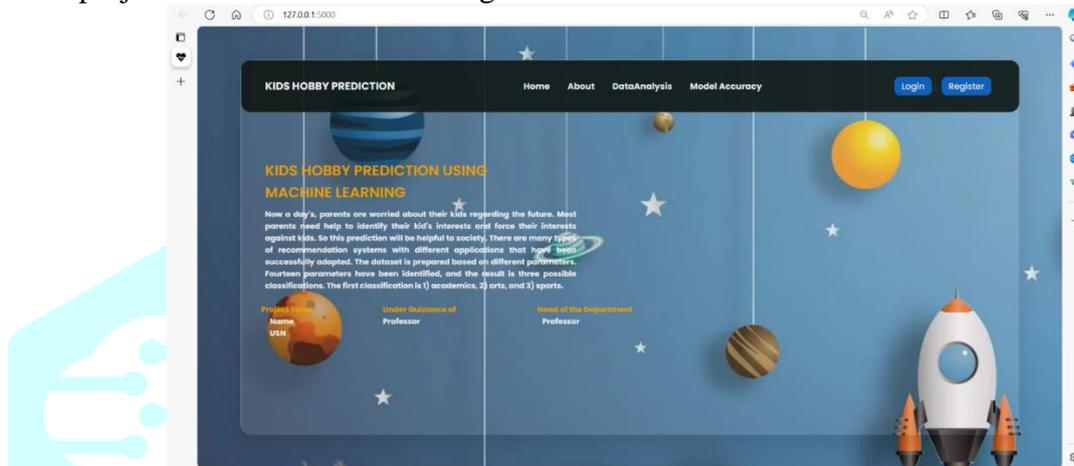


Figure 3. Home Page

**5.2 Signup Page:** The new user can register with this form and entered the proper credentials to make yourself registered as shown in Figure 4.

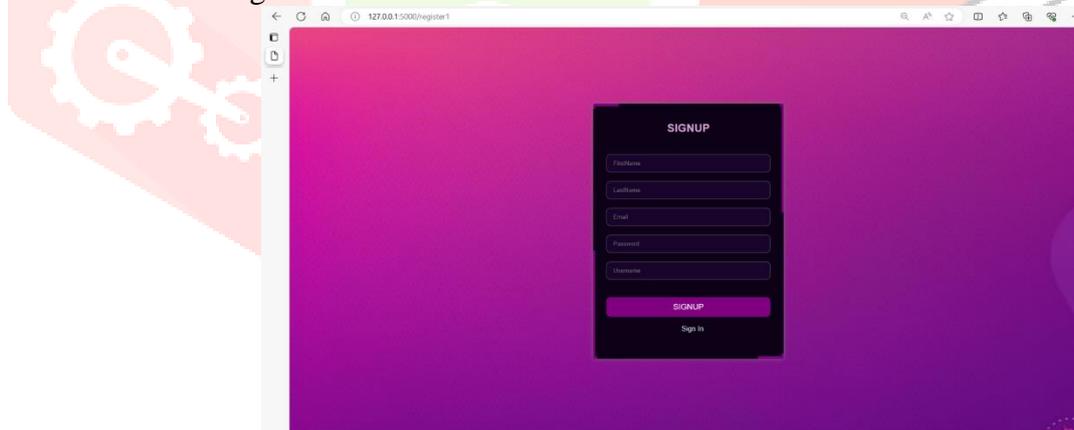


Figure 4. Signup Page

**5.3 Login Page:** Through this form, existing user can login into the web application by entering proper username and password (Figure 5).

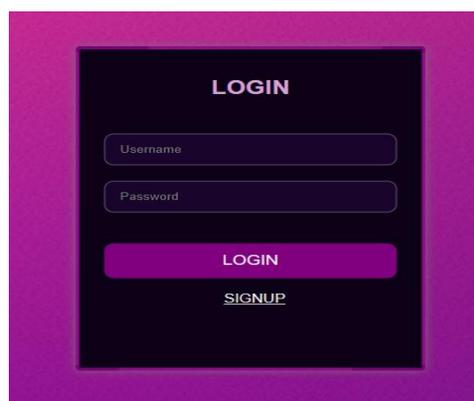


Figure 5. Login Page

**5.4 About Us:** In this page, gives description about the project is as shown below in figure 6.



Figure 6. About Us

**5.5 Data Analysis:** Here, all the plots/graphs that have been done in the backend code are displayed. It's data analysis have to be done with kids hobby dataset (Figure 7).

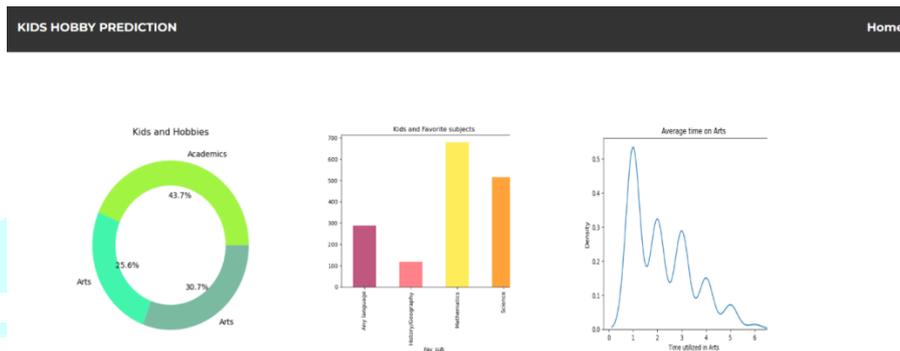


Figure 7. Data Analysis

**5.6 Model Accuracy:** Here, the different machine learning algorithm applied to the kids hobby prediction application is shown (Figure 8). The table shows the comparison model.

Models Name	Accuracy
Logistic Regression	88.16%
ADA Boost Classifier	87.85%
Decision Tree Classifier	89.41%
GaussianNB	90.65%
<b>Random Forest Classifier</b>	<b>91.90%</b>
Support Vector Classifier	89.41%

Figure 8. Model Accuracy

**5.7 Prediction Page:** Here, is a place where we can enter the input feature to the system to find the proper results as shown in figure 9-12.

Figure 9. Input

Want's to pursue his/her career in sports?

Yes No

Regular in his/her sports activities?

Yes No

Love creating fantasy paintings?

Yes No

Won art competitions?

Yes No

Time utilized in Arts? (in hours)

1

SUBMIT

Figure 10. Input

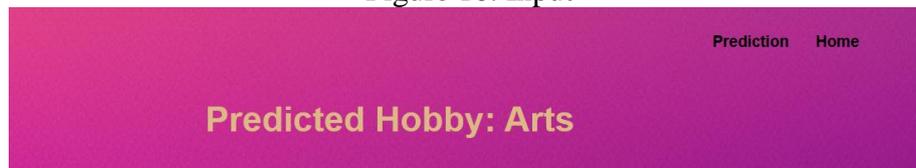


Figure 11. Prediction\_1

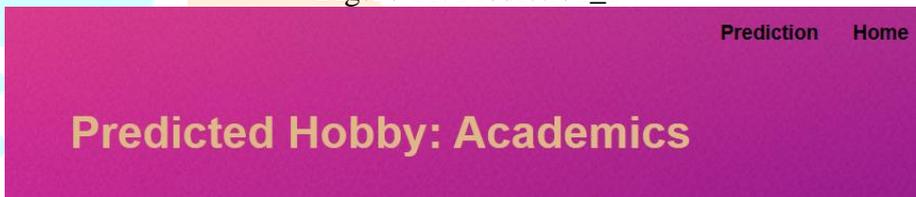


Figure 12. Prediction\_2

## VI. CONCLUSION

Any country's future is in young kids. The kids have to grow with different parameters of multiclass neural parameters and their values. The parents should not be the decider for the kid's future. They have to felicitate the kids' future. So this kind of work is helpful for a wealthy society and a healthy lot. The work concentrated on three types of prediction 1) academics, 2) arts and 3) sports. Countries like India couldn't provide much encouragement for sports, especially for the kids. They have colossal potential. Their only problem is correct prediction and motivation. Now a day's, machine learning-based prediction has entered into all the areas of our day-to-day life. We assure you that this kind of work will bring some change in the parents as well as students. A total of 1600 dataset has been used for training and modelling the different algorithm such as Logistic Regression (88.16%), Adaptive Boosting Classifier (87.85%), Decision Tree Classifier (89.41%), Gaussian Naïve Bayes Classifier (90.65%), Random Forest Classifier (91.90%) and Support Vector Classifier (89.41%). We have developed the python-based flask web application to demonstrate the predictions.

## VII. REFERENCES

- [1] Sugeng Riyadi, Hartono, Wanayumini, Predicting Children's Talent Based On Hobby Using C4.5 Algorithm And Random Forest, Vol. 2 No. 1 (23): Proceeding of International Conference on Information Science and Technology Innovation (ICoSTEC) 2023.
- [2] Prof. Priyanka Shahane, Prutha Rinke, Taniksha Datar, Soham Badjate , Student's Career Interest Prediction using Machine Learning, International Research Journal of Engineering and Technology (IRJET), Volume: 09 Issue: 11 | Nov 2022.
- [3] Koji Bando, Kenji keta, Twitter User's Hobby Estimation Based on Sequential Statements Using Deep Neural Networks, International Journal of Machine Learning and Computing, Vol. 9, No. 2, April 2019.
- [4] Akanksha Pandey, L S Maurya, Prediction of Undergraduate Students' Career Using Various Machine Learning And Ensemble Learning Algorithms, Webology (ISSN: 1735-188X) Volume 18, Number 6, 2021.
- [5] Improved Recommender System for Kid's Hobby Prediction Using Different Machine Learning Techniques, 2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT).