International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

SNAKE CATCHER - SNAKE RESCUE AND AWARENESS APP

Laximi Deulkar[1], Deepesh Madgaonkar[2], Virisha Kalangutkar[3], Nikhil Polekar[4], Shripad Alvi[5], Shreedatta Sawant[6], Rajesh Gauns[7]

[1,2,3,4,5] *Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa*
[6, 7] *Assistant Professor, Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa*

Abstract - *Effective data storage management is more important than ever. Storage management ensures data availability to users. It enhances performance and safeguards against data loss. It also makes sure that data is secure from outside danger, unintentional errors from humans, and system crashes. This project focuses on building an application with a database to help utilize and sort the data according to the user. All the data being uploaded will get saved on the application, plus the verified snake catchers will have their respective accounts on which their personal details will be displayed so the locals can contact them. There will be an option provided for the users(snake catcher) to add, delete, update and search for the data. Information about poisonous snakes and their medicinal treatment will be displayed on the application. The application will help to predict the snake species established on the characteristics like length and shape of its body, colour, and pattern, scale texture, etc. The symptoms produced after the snake bite are also helpful to predict the snake species.*

These predictions can be done using classification algorithms like J48.

***Key Words*: Classifications, decision tree, data management, identification.**

1. INTRODUCTION

There exists a community of people who rescues snakes. Every time after the rescue they need to update the data like the location, date, time, image, etc. on their existing WhatsApp group and one person has to note down all these details in a book which is a very tedious task. So the data they are updating every day is not properly getting managed and stored and is also insecure. So effective data storage management is more important. Storage management assures data is obtainable to users when they need it. It enhances performance and guards against data loss.

It also makes certain that data is secure from external threats, human mishandles and system crashes. This project focuses on building an application with a database to help utilize and sort the data according to the user. All the data being uploaded will get saved on the application, plus the verified snake rescuers will have their respective accounts on which their personal details will be displayed so the locals can contact them. Information about poisonous snakes and their medicinal treatment will be displayed on the application. The application will help to predict the snake species based on the image being uploaded or the attributes like type, head shape, colour and pattern, scale texture, etc. These predictions can be done using classification algorithms like J48/C4.5.

1.1 Problem Definition

When a rescuer catches a snake it is very difficult to maintain the data or keep track of the data of the user who calls them for snake rescue. The contact details may get lost in a hurry and can create a state of confusion.

The problem here is that there is no connection between the rescuer and the people. When there is an emergency like a snake bite or a snake invading our homes there is no contact or a rescuer to reach on spot on time. Many deaths have occurred due to this. There is no proper medication on time or delay of the patient in reaching the health care centres.Snake killing is the most common activity that reduces the snake population and affects the natural animal life system.

1.2 Problem Specification

Here we try to solve the problem of rescuers by managing the data and transactions between the user and the rescuer. This application will help users to connect with rescuers by finding nearby rescuers quickly and avoiding any further emergencies. This will also help rescuers to keep a record of the user and their contact details.

It will also have other features like first aid, and hospital details which can help users to give proper first aid on time and can save a person's life. The rescued snakes can be left in the uninhabited area which will ensure their safety too. Getting to know more about snakes and their species will help people to handle any emergencies.

2. LITERATURE SURVEY

2.1 Past Research

Epidemiology and Clinical Profile of Snakebites in Goa and Surrounding Areas[1]

Snakebites are a very common medical emergency in Goa which can occur in any season. There is no specific person or occupation with a risk of snakebite. Snakebite primarily occurs in homemakers, students, or aged people. There are many symptoms that can be seen after a bite like vomiting, breathing, breathlessness, chest pain, etc. The majority receive anti-venom within 6 hours of intoxication. Snakebite is a mistreated general health hazard globally as well as in India. Many people die or are disabled due to snakebites. Anti-snake venom (ASV) 5 is the only snake bite treatment available in hospitals. Due to poor access to health facilities, scarcity of anti-venom, lack of adequate knowledge and training of doctors, late reporting to health centres leads to delays in administration.

Snake species identification by using natural language processing[2]

This paper helps us to understand that people describing a snake using natural language and the medical administrator understanding which type of snake it is can go wrong and due to that some wrong anti-venom given will cause more severe morbidity and mortality.

So to predict what species of the snake using natural language in which the person describes the snake in words and based on their visual view this research paper used different classification algorithms. First, a lot of preliminary processing, character extraction, and classification was done on the data. Then four machine learning algorithms which are naïve Bayes, K-nearest Neighbor, support vector Machine, and Decision trees J48 was used for training and classification of data.

Results show that the J48 algorithm gained the highest classification precision of 71.6% accurate prediction for the Natural Language Processing-Snake data set with high accuracy and recollection.

2.2 Present Research

J48 Classification (C4.5 Algorithm) in a Nutshell[3]

This algorithm is based on concept learning systems and is a supervised learning algorithm. J in J48 stands for java. J48 is a non-propriety Java application of C4.5. J48 permits classification via either decision trees or rules developed from them. This is a classification algorithm that yields decision trees based on Information theory. This adopts a non-backtracking strategy in which decision trees are produced in a top-down recursive divide-and-conquer technique. It starts with a training collection of tuples and their associated category labels. The training set is recursively partitioned into shorter subsets as the tree is being assembled.

This implementation has many more features including interpreting the unknown values, decision tree pruning, constant attribute value spans, derivation of rules, etc.

Performance tuning of J48 Algorithm for prediction of soil fertility [4]

Soil fertility is a major issue in India. Modern research and data mining techniques should be used for predicting the soil fertility of Indian soil. A large dataset which are samples of soil was used to find out different types of fertility aspects and measures. Using the data analyses were done using three different classification algorithms which are NBTree, simple CART, and J48. NBTree is a classification based on Naive Bayes classifiers. Simple CART is a multivariate decision tree learning method that generates either classification or regression trees, depending on whether the subjected variable is categorical or numeric. And J48 is a non-propriety Java application of the C4.5 algorithm in the Weka data mining tool. C4.5 is a program that creates a decision tree based on a collection of labelled input data. The accuracy of the J48 algorithm for estimating soil fertility was leading, therefore it was utilised as a weak learner.

So now with that accuracy, they can use other meta strategies like feature selection and enabling in the weka tool.

Attribute selection reduces dataset size by removing irrelevant/redundant attributes and Boosting is a machine learning meta-algorithm *for performing supervised learning.*

3. Prerequisites

3.1 Decision Tree

A decision tree is a tree-like structure, where each sub-node (non-leaf tuple) denotes a check on a feature, each branch depicts an answer to the test, and an individual leaf node (or terminal node) has a category label. The highest single node in the tree is the root node.

Provided a tuple, X, for which the connected class label is unspecified, the feature values of the tuple are checked against the decision tree. From the root to the leaf node the path is tracked, which retains the class prediction for that tuple. For classification decision tree algorithms have been used in many application areas such as medicine, manufacturing and production, financial analysis, astronomy, and molecular biology.

During tree structure, attribute selecting techniques are used to choose the attributes that best partition the tuples into discrete classes.

When decision trees are constructed, many of the branches may predict noise or outliers in the training data. Tree pruning attempts to identify and eradicate such branches, with the main purpose of enhancing classification precision on hidden data.

3.2 C4.5 / J48 Algorithm

The C4.5 algorithm is a classification algorithm that produces decision trees based on information theory. It is a continuation of Ross Quinlan's earlier ID3 algorithm also known in Weka as J48, J standing for Java. The decision trees developed by C4.5 are used for classification, and for this reason, C4.5 is usually referred to as a statistical classifier..C4.5 adopts a greedy process in which decision trees are constructed in a top-down recursive divide-and-conquer approach.

I. Calculate the appearance frequency of individual attributes in the survey data.

2. Calculate the Entropy value of each attribute.

3. Compute the Information Gain value using the known Entropy value calculated beforehand.

4. Calculate the Split Info value of each attribute.

5. Calculate Gain Ratio value using Information Gain value and Split Info value.

6. Choose the biggest Gain Ratio and make it the root node.

7. Eliminate the attributes that have been chosen before, and repeat the calculations of Entropy value, Information Gain value, Split Info value, and Gain Ratio value by choosing the biggest Gain Ratio and making it the tree internal node.

8. Repeat all the calculation processes until all the attributes are categorized into classes.

9. After all the trees are categorized into the classes, show the initial decision tree and generate the initial decision rule.

Algorithm

Input: training samples, training labels, attributes

- Building root nodes for the tree
- If all the samples are positive, stop after a tree with a root node, and label it with (+)
- If all the samples are negative, stop after a tree with a root node, label it with (-)
- If the attributes are missing, stop after a tree with a root node, label it with the frequently appearing-value in training label
- For the others:

A ← attribute that classifies samples with the best result

(based on gain ratio)

Decision attribute for a root node ← A

For each V_i value that is possible for A

- Additional branch under the root related with A = V_i
- Determine sample Svi as a subset from the samples having V_i value for attribute A
- If sample SV_i is missing
    - Under the branch, add a leaf node and the label = the frequently appearing-value in training label
    - For others, add a new branch under the current branch

Pseudo-Code for C4.5

1. First, notice the base
2. For per attribute X, find the normalized data gain ratio by dividing between X.
3. Assume that X is an attribute with the highest normalized data gain.
4. Create a decision node that separates on attribute X.
5. Repeat it on the sublists acquired by splitting the attribute X, and count these nodes as children of the node.

Mathematical Model for C4.5 Algorithm

$$\textbf{SplitInfo}_A(D) = -\sum_{i=1}^{v} \frac{|D_i|}{|D|} \times \textbf{Log}_2\left(\frac{|D_i|}{|D|}\right)$$

$$\textbf{GainRatio(A)} = \frac{\text{Gain}(A)}{\text{SplitInfo}(A)}$$

3.3 Sample Dataset

type	*color*	*head shape*	*eye shape*	*Snake name*
venomous	*brown*	*triangular*	*round*	*Cobra*
venomous	*black*	*triangular*	*round*	*Cobra*
venomous	*brown*	*triangular*	*vertical*	*Saw-scaled viper*
venomous	*reddish*	*triangular*	*pear-shaped*	*Saw-scaled viper*
non-venomous	*yellow*	*broader*	*round*	*Rat snake*
non-venomous	*brown*	*broader*	*round*	*Rat snake*
mild-venomous	*brown*	*triangular*	*vertical*	*Cat snake*
non-venomous	*brown*	*triangular*	*round*	*Wolf snake*
non-venomous	*brown*	*broader*	*vertical*	*Wolf snake*
venomous	*reddish*	*triangular*	*vertical*	*Saw-scaled viper*

`4. Design

4.1 Admin side log-in

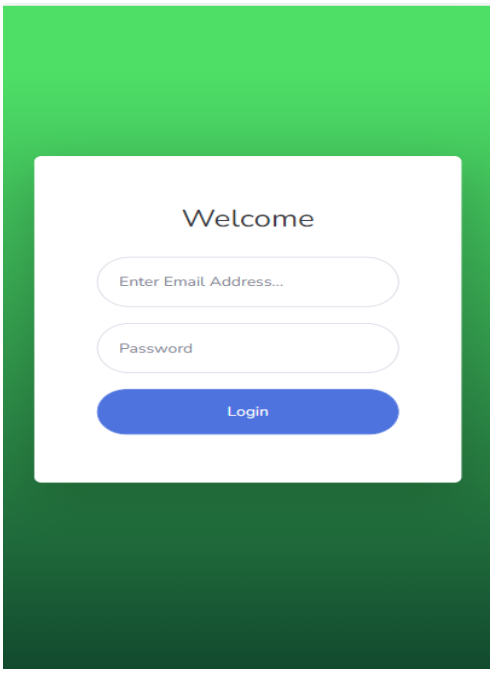

fig - 4.1: Admin side log-in

4.3 Registration (User/Rescuer)

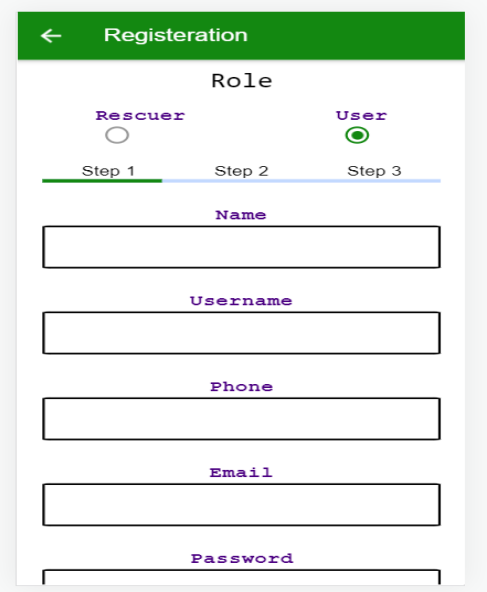

Fig - 4.2: Registration (User/Rescuer)

4.3 Log-in (User/Rescuer)

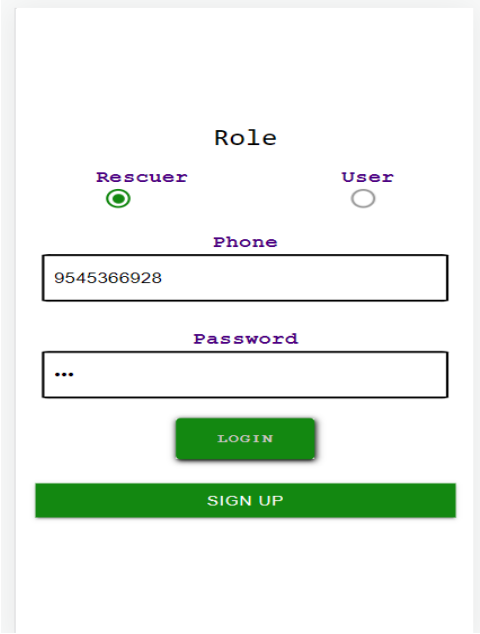

Fig - 4.3: Log-in (User/Rescuer)

4.4 Home page

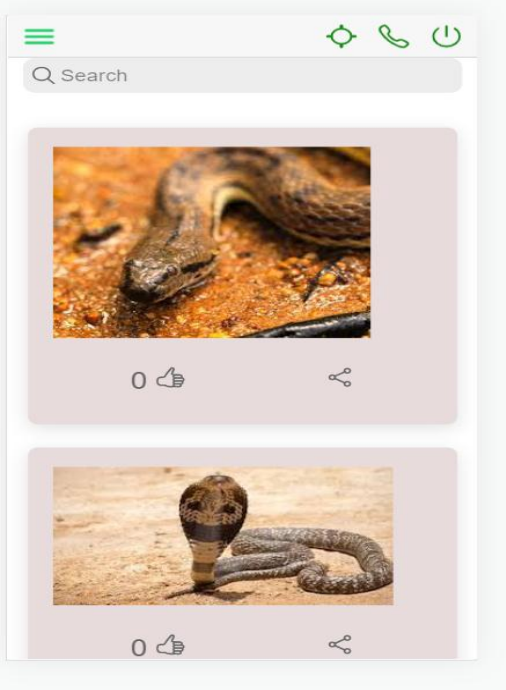

Fig - 4.4: Home page

4.5 Search rescuer page

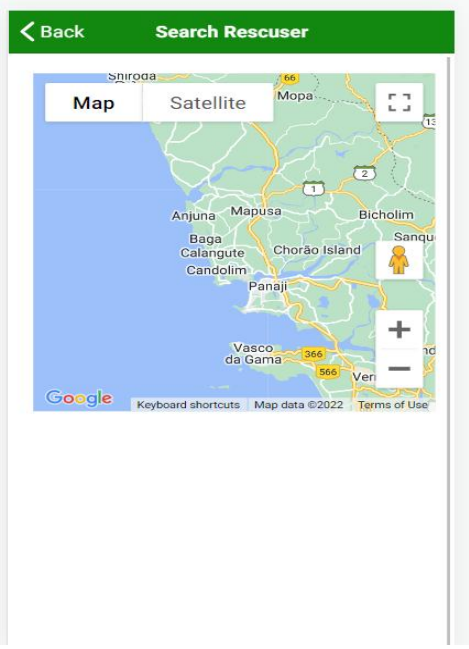

Fig - 4.5: Search rescuer page

4.6 Identification

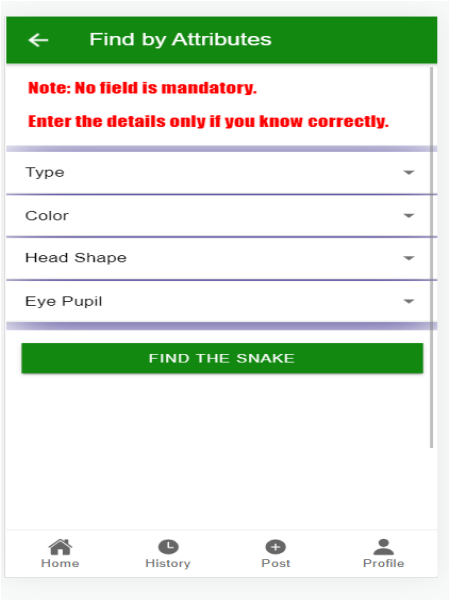

Fig - 4.6: Identification page

5. Conclusion

On the completion of this project there will be an app which will help snake rescuers to maintain their data properly and will have a platform to show their achievements. This will also help users to get in contact with the rescuers very fast. This will also help to save the lives of people bitten by snakes. This app will provide a lot of knowledge about snakes to users. This project will showcase the dynamism of rescuers and appreciate what they do for society and protect the reptiles. This app will help to keep data secure and in data loss problem-solving. The killing of snakes will reduce due to the faster availability of rescuers.

REFERENCES

[1] Narvencar, K., Favas, T. T., & Dias, A. (2020). **Epidemiology and Clinical Profile of Snakebites in Goa and Surrounding Areas.** The Journal of the Association of Physicians of India, 68(3), 28-32.

[2] Rusli, N. L. I., Amir, A., Zahri, N. A. H., & Ahmad, R. B. (2019). **Snake species identification by using natural language processing**. Indones. J. Electr. Eng. Comp. Sci, 13, 999-1006.

[3] Gholap, J. (2012). **Performance tuning of J48 Algorithm for prediction of soil fertility**. arXiv preprint arXiv:1208.3943.

[4] Nilima Khanna, Professor Amrinder Arora, **J48 Classification (C4.5 Algorithm) in a Nutshell**, George Washington University.