Aquatic Catch Predictor.

Zixuan Yang

Motivation.

1. **Personal Inspiration:**

As an avid fishing enthusiast, I encountered a frustrating experience where a fishing app predicted excellent fish activity, leading me to invite friends for a fishing trip. Unfortunately, not a single fish was caught that day. This sparked the idea for a more reliable and data-driven solution.

2. **Significance:**

Who cares?

Recreational and competitive anglers aiming to improve their skills and catch rates.

Why does it matter?

Saves time and resources, creating a more enjoyable and effective fishing experience.

Objectives.

Portable Device: Create a compact system capable of predicting fish activity and recommending optimal fishing spots.

Environmental Analysis: Integrate data from temperature, pressure, and weather conditions to assess fish behavior.

Real-time Insights: Use GPS and live weather updates to identify fishing spots with the highest success potential.

Enhanced Experience: Reduce uncertainty and improve catch rates through data-driven decision-making using decision trees for prediction.

Related Work.

Existing Solutions: Current fishing apps, offer predictions based on general weather data and historical trends.

Limitations: These apps often lack real-time data integration, location-specific accuracy, and personalized recommendations for anglers.

Inconsistent Performance: The variability in prediction accuracy underscores the need for a more reliable, data-driven solution tailored to individual fishing scenarios.

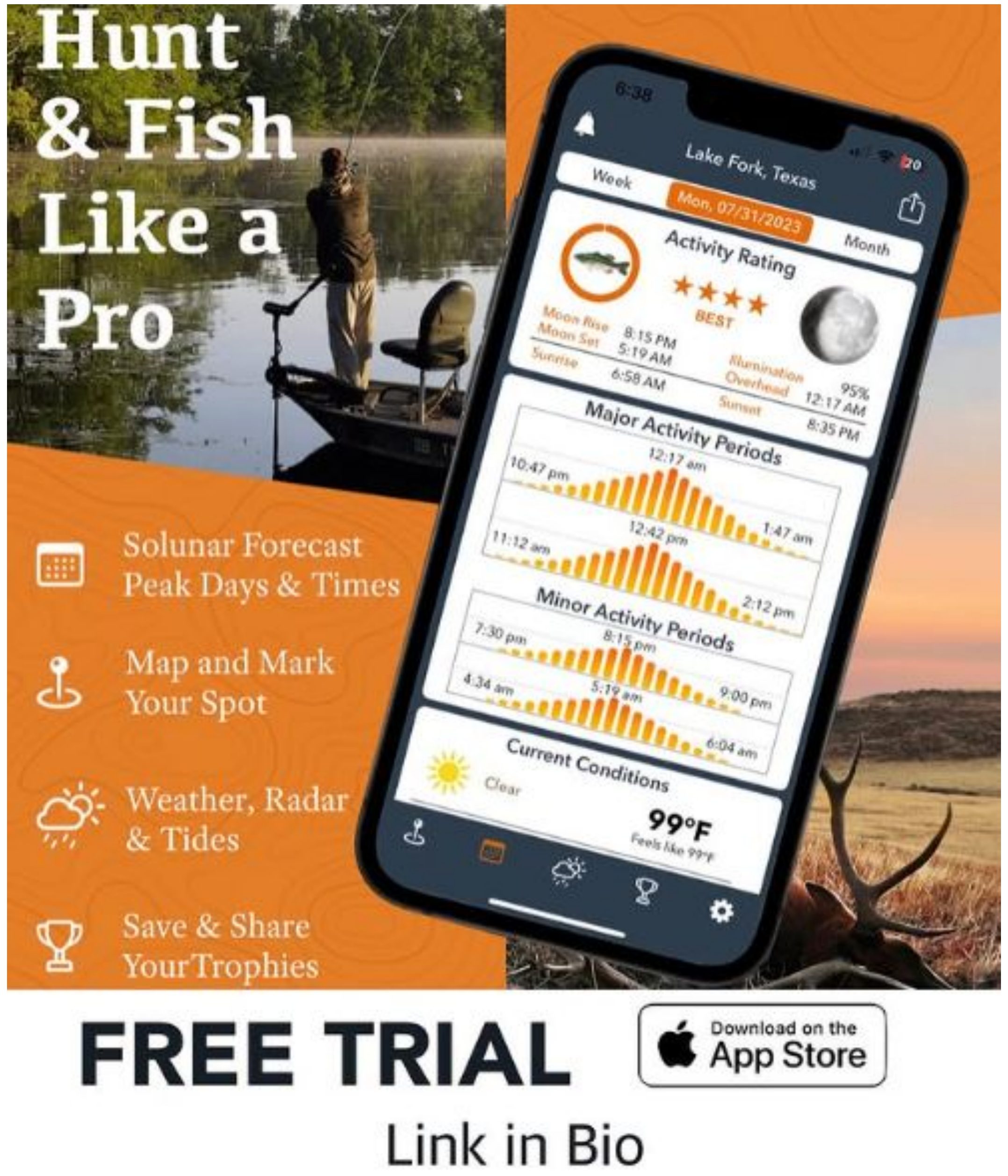


Fig.1 fishing app example

Novelty.

What's New:

Scientifically Collected Data: Predictions are supported by scientifically collected data, ensuring high reliability and accuracy in results.

Offline Functionality: Operates seamlessly without internet access, using onboard sensors and pre-trained models to provide predictions in any environment.

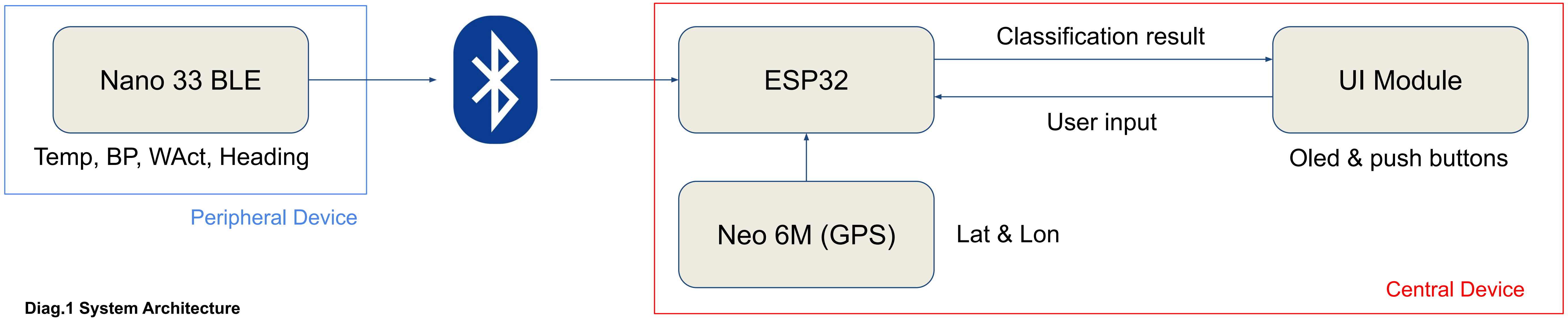
Decision Tree Predictions: Utilizes decision tree algorithms for precise fish activity forecasts, providing a significant improvement over generalized or heuristic methods.

Technical Approach.

The system consists of a **peripheral device** and a **central device**, optimized for specific roles:

Peripheral Device:

- Uses the **Arduino Nano 33 BLE** & **500mah Lipo battery** with onboard sensors to collect:
    - **Water temperature** (temperature sensor).
    - **Atmospheric pressure** (barometric sensor).
    - **Water surface activity** (IMU).
    - **Orientation** (magnetometer).
- Sends data to the central device via **Bluetooth**.



Diag.1 System Architecture

Nano 33 BLE
Clock Speed: 64 MHz
Flash Memory: 1 MB
RAM: 256 KB

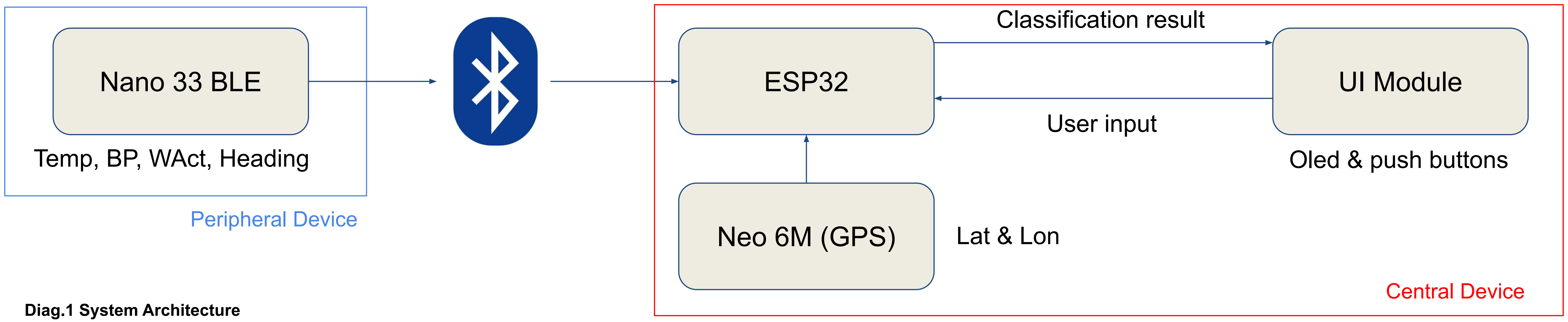
Technical Approach.

Central Device:

- Built with an **ESP32** for its computational power to deploy machine learning models.
- Equipped with:
  - **Neo-6M GPS module** for location tracking.
  - **OLED screen** for displaying predictions.
  - **Push buttons** for user interaction.
  - **500mah Lipo battery**
- Processes data and provides fishing predictions using decision trees.

Communication:

- **Bluetooth link** enables data transfer between the Nano 33 BLE and ESP32.



ESP32
Clock Speed: Up to 240 MHz
Flash Memory: 4 MB
RAM: 520 KB

Diag.1 System Architecture

Methods. data collection

Participants:

Three individuals with varying fishing skill levels: beginner, intermediate, and expert.

Locations:

Data collected simultaneously at three different spots using peripheral devices connected to the nRF App on smartphones for logging.

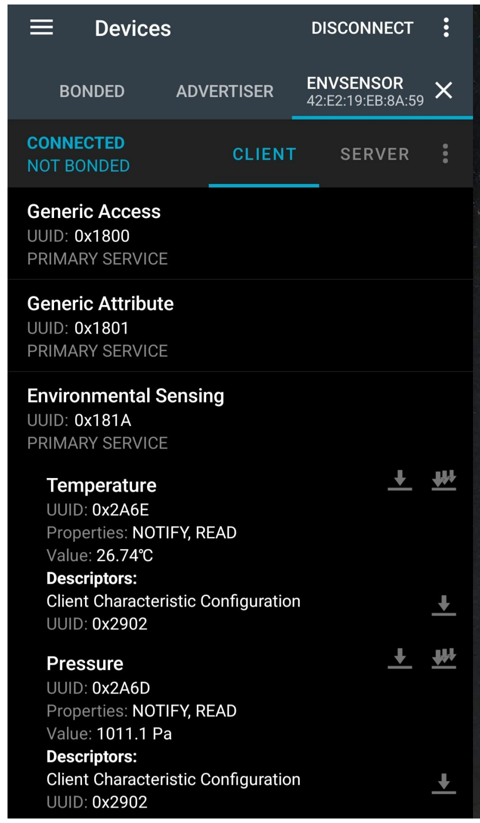


Fig.2 nRF App

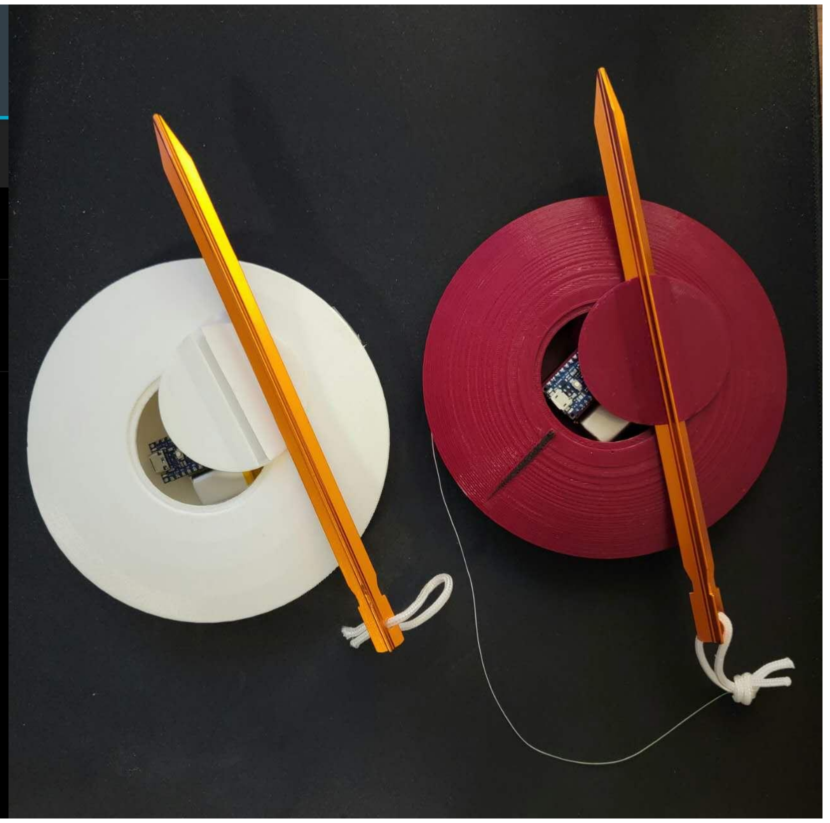


Fig.3 Peripheral Devices (prototype_left)

Methods. data collection

Process:

- Each session lasts 30 minutes, participants rotate between locations every 10 minutes.
- This ensures that each location has data contributed by all three participants, enhancing the objectivity and reliability of the dataset.
- During data collection, efforts were made to cover the entire lake; however, a section in the North-West area was obstructed by reeds, making it inaccessible for data collection.

Duration:

- Data collected over 6 days, with 3 hours of collection each day.
- Each day yields 18 data entries, totaling 108 data points across the experiment.

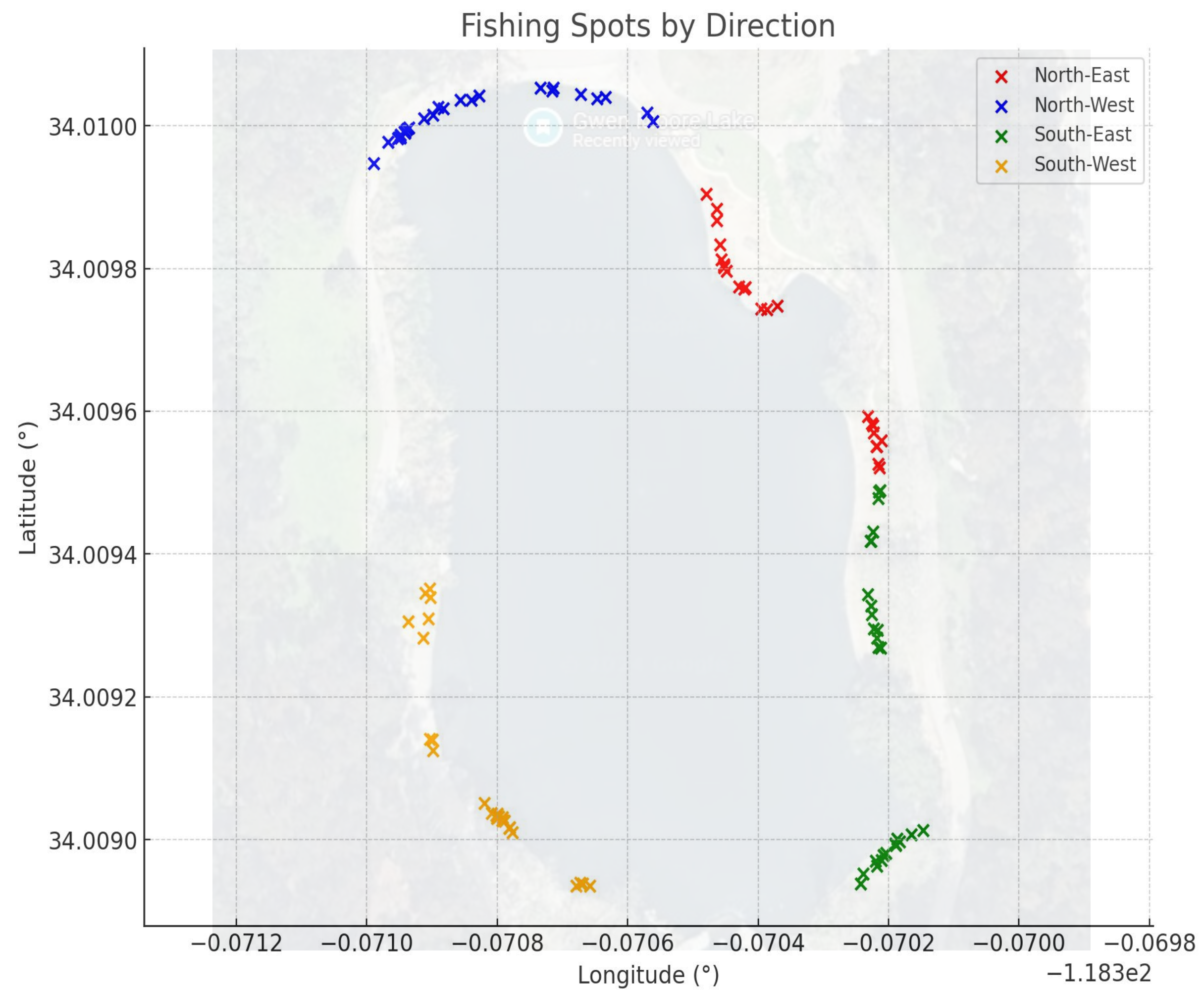


Fig.4 Collected Fishing Spots

Methods. data processing

Data Aggregation:

- Collected data from the nRF App is exported and consolidated into an Excel spreadsheet for analysis.

Bite Rate Calculation:

- Bite rate is determined by summing the catch data from all three participants at each location during a 30-minute session.
- This provides a comprehensive measure of fish activity across skill levels at each spot.

Data Quality:

- No missing values were identified during the data review process.
- GPS data occasionally showed slight shifts, requiring manual adjustments to ensure accuracy.

Date	Latitude	Longtitude	Temperature	Pressure	Water Activity	Bite Rate	Fishing Spot Location
10/23/2024	34.009027	-118.370792	26.43	1014.80	0	4.67	South-West
10/23/2024	34.008993	-118.370189	26.71	1014.70	1	2.67	South-East
10/23/2024	34.009478	-118.370216	26.29	1014.80	0	4.67	South-East
10/23/2024	34.009806	-118.370452	26.68	1014.70	1	4	North-East
10/23/2024	34.010042	-118.370828	26.21	1014.70	1	4.67	North-West
10/23/2024	34.009125	-118.370899	27.59	1013.50	1	3.33	South-West
10/23/2024	34.008938	-118.370243	26.84	1014.40	1	2.67	South-East
10/23/2024	34.009977	-118.370967	27.32	1014.00	1	4.67	North-West
10/23/2024	34.009583	-118.370226	26.89	1014.30	1	4	North-East
10/23/2024	34.009559	-118.370211	27.27	1014.00	1	4.33	North-East
10/23/2024	34.009947	-118.370989	26.92	1014.40	1	4.67	North-West
10/23/2024	34.008963	-118.370218	27.56	1013.90	1	3	South-East
10/23/2024	34.009748	-118.370371	27.67	1013.60	1	4.33	North-East
10/23/2024	34.010026	-118.370891	26.98	1013.20	1	4.67	North-West
10/23/2024	34.009991	-118.370943	26.87	1014.70	0	4.67	North-West
10/23/2024	34.009141	-118.370904	27.04	1013.30	1	3	South-West
10/23/2024	34.009983	-118.370952	27.52	1013.60	0	4.67	North-West
10/23/2024	34.009797	-118.370448	26.93	1013.20	1	4	North-East
10/27/2024	34.010040	-118.370634	25.66	1011.60	1	1	North-West
10/27/2024	34.010036	-118.370840	27.77	1011.70	1	4.67	North-West
10/27/2024	34.009487	-118.370215	27.72	1012.00	1	3	South-East
10/27/2024	34.009418	-118.370228	26.88	1011.80	1	2.67	South-East
10/27/2024	34.008938	-118.370669	26.16	1012.10	1	3	South-West
10/27/2024	34.009031	-118.370792	25.84	1011.80	1	0	South-West
10/27/2024	34.009013	-118.370147	24.58	1011.40	1	0	South-East
10/27/2024	34.010024	-118.370883	27.97	1012.10	1	4.67	North-West
10/27/2024	34.009345	-118.370911	25.44	1011.50	0	1	South-West

Fig.5 dataset

Methods. model training

- Trained a decision tree model in Python to predict fish bite rates.
- Input features: temperature, pressure, water surface activity, and geographical location.
- Validated the model using an 80/20 train-test split.

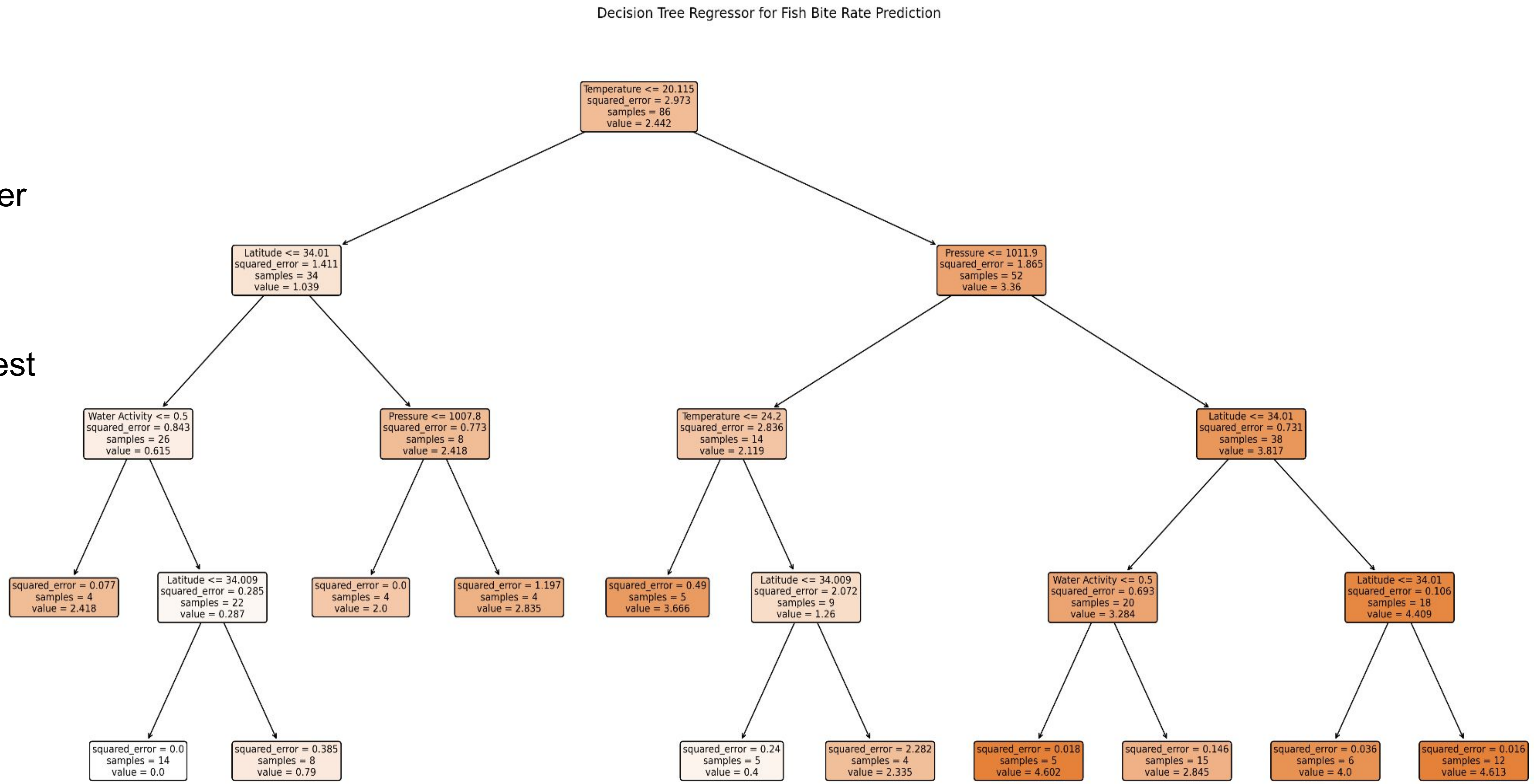


Fig.5 Decision Tree Model

Methods. model tuning

- Directly training the decision tree without constraints results in training MSE being significantly smaller than testing MSE, indicating overfitting.

```
# Initialize and train the decision tree regressor
decision_tree_regressor_latest_upload = DecisionTreeRegressor(random_state=42)
decision_tree_regressor_latest_upload.fit(X_train_latest_upload, y_train_latest_upload)
```

- To address this, max depth and min samples leaf parameters are introduced to limit the tree's complexity and reduce overfitting.
  - Achieved best performance with max_depth=4 and min_samples_leaf=4, balancing accuracy and generalization.

```
# Initialize and train the decision tree regressor
decision_tree_regressor_latest_upload = DecisionTreeRegressor(max_depth=4, min_samples_leaf=4, random_state=42)
decision_tree_regressor_latest_upload.fit(X_train_latest_upload, y_train_latest_upload)
```

Demo.

Vid.1 Prediction Mode

Vid.2: Location Suggestion & Navigation Mode

Experimental Evaluation. Metrics

MSE: Evaluates the average squared error between predicted and actual values to measure accuracy.

RMSE: Square root of MSE, providing an interpretable error in the same units as the target variable.

Relative Error: Measures the difference between testing and training MSE as a percentage of the training MSE.

Prediction Speed: Measure the time taken for the system to compute a prediction.

Power Efficiency: Evaluate the system's energy consumption during operation.

- Average power consumption (mW).
- Battery life (hours) on a given power source.

Experimental Evaluation. Results

Testing MSE: 0.2498 **Training MSE:** 0.2751

Testing RMSE: 0.4998 **Training RMSE:** 0.5245

Relative Error: 10.12% (Training MSE: 0.2751 vs. Testing MSE: 0.2498)

Average Prediction Speed: 15 ms

Central Device Power Efficiency:

- Average power consumption 408.6 mW.
- Battery life 4.77 hours.

Memory use: 90%

Peripheral Device Power Efficiency:

- Average power consumption 70.2mW.
- Battery life 27.76 hours.

Memory use: 23%

```
Training MSE: 0.27509203488372086
Training RMSE: 0.5244921685628117

Testing MSE: 0.24980314772727272
Testing RMSE: 0.49980310896119157

Feature Importances:
       Feature  Importance
2  Temperature    0.557417
0     Latitude    0.193384
3     Pressure    0.133084
4  Water Activity 0.116116
1   Longtitude    0.000000
```

Fig.6 Model Metrics

```
17:57:52.464 -> Perdiction Starts
17:57:56.552 -> Perdiction Ends
17:57:56.588 -> Perdiction Starts
17:57:59.697 -> Perdiction Ends
17:57:59.697 -> Perdiction Starts
17:58:03.746 -> Perdiction Ends
17:58:03.746 -> Perdiction Starts
17:58:07.872 -> Perdiction Ends
17:58:07.872 -> Perdiction Starts
17:58:10.982 -> Perdiction Ends
```

Fig.7 Serial Monitor output of prediction time

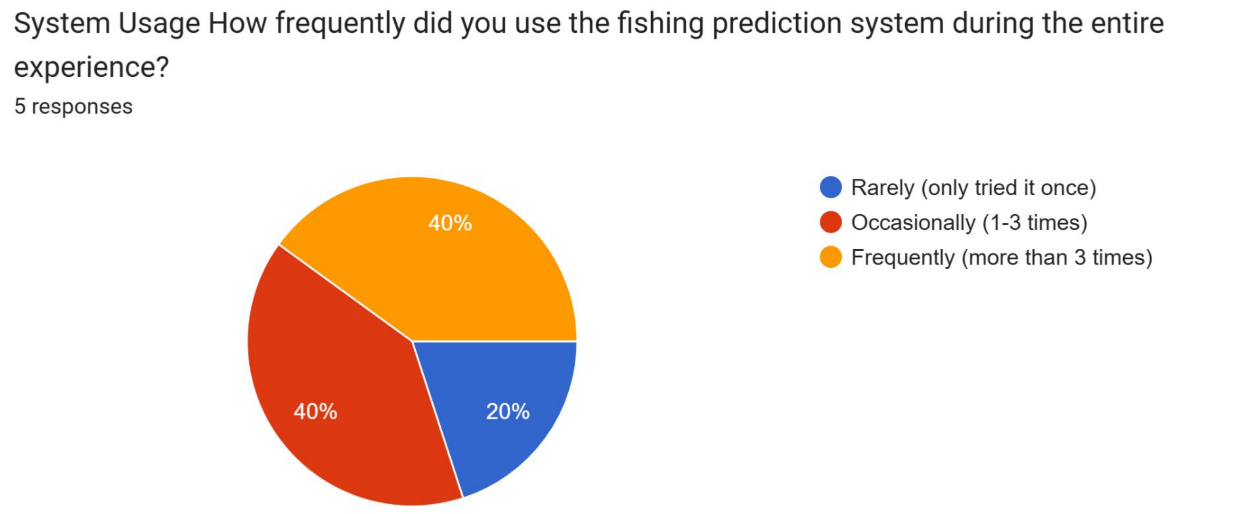
Experimental Evaluation. Finding

Beginners without any experiences:

- Show a higher level of dependence on the system and use it more frequently.
- Their actual catch rates are generally lower than the system's predictions due to lack of experience.
- Rely heavily on fishing location recommendations for guidance.

Intermediate and Advanced Users:

- Use the system less frequently but demonstrate higher alignment between their catch rates and the system's predictions.
- Depend more on fish activity predictions to refine their strategies.



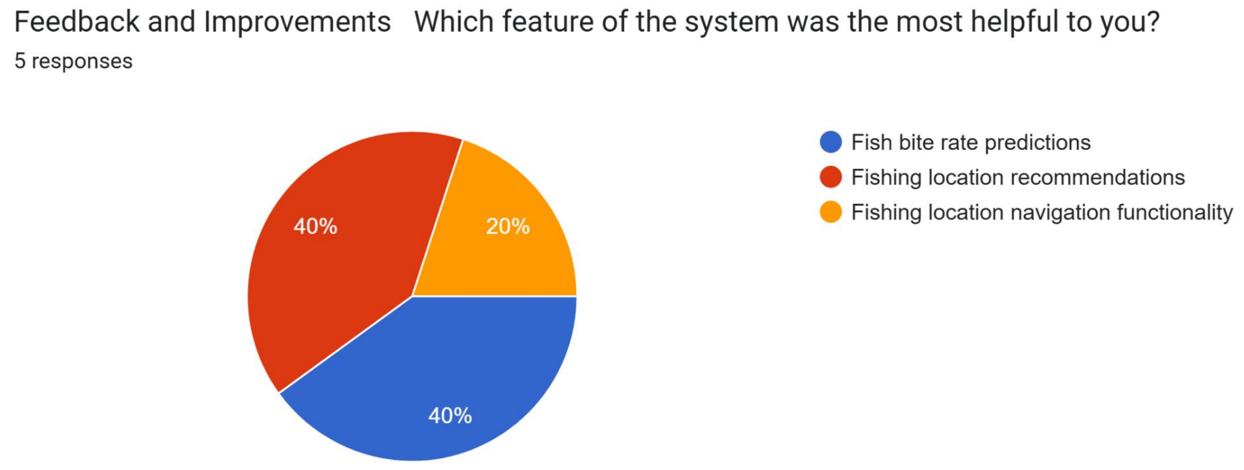


Fig.8 Example of User Study Results

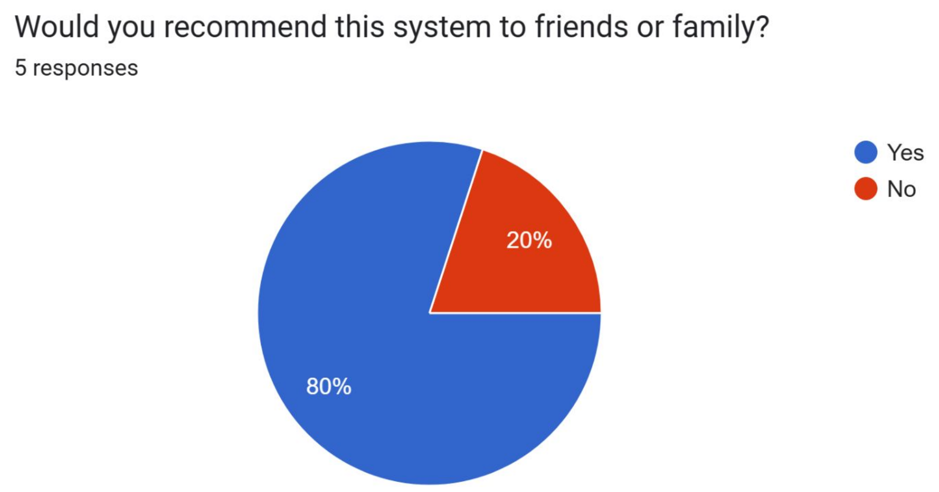
Conclusions.

- The system helps both beginners and experienced users by providing useful fishing predictions.
  - Beginners benefit the most from location recommendations and use the system more frequently.
  - Experienced users use the system less often but rely more on fish activity predictions to improve their results.

- The decision tree model works well, with a low testing error and accurate predictions in different conditions.

- The system’s real-time predictions and low power use make it reliable for long fishing trips.

- User feedback shows the system is easy to use and helpful, with 80% of users willing to recommend it to others.



Discussions.

Hardware Improvements:

- Replace the current GPS module with an Adafruit GPS module for lower power consumption and support for low-power modes.
- Optimize the device enclosure to make it more compact and portable, improving user convenience.

Data Collection:

- Expand data collection to include more sessions and diverse water bodies, improving model accuracy and supporting more fishing conditions.

Future Features:

- Include measurements like water depth or dissolved oxygen to improve prediction accuracy.
- 加上能让用户设定自己的水平，系统可以根据用户的设定scale最终的预测结果

Question & Answer .