
SCIENCE • INTELLIGENCE • INNOVATION
Detect · Predict · Design · Validate · Deploy
AMK
ARGUMENTATIVE AI
TUTOR PLATFORM
Human–AI Cooperative Execution
AI CERTIFICATION STUDENTS
FINAL CAPSTONE PROJECTS

Benchmark Lecture Note 17
Drone-Based Aquaculture Threat Detection, Pond Risk Classification, and ROI Protection

AI Certification Program and AMK Research Lab Program

Revised benchmark case-study edition using ensemble models, hierarchical clustering, farthest-first clustering, and drone computer vision in Python

Core methods	Scenario goal	Outputs
Random Forest, Voting, Agglomerative clustering, farthest-first coverage selection, YOLO/OpenCV	Detect threat agents, identify the most affected pond, and estimate ROI decline	Threat counts, risk class, most affected pond, representative monitoring points, management actions

Instructional note: this is a synthetic teaching scenario for monitoring, decision support, and safe human intervention. It is not an autonomous harmful-response design.

1. Learning outcomes

- Explain why ensemble models can improve robustness over a single estimator in a surveillance-style classification task.
- Use hierarchical clustering to group ponds or pond zones by shared ecological and operational risk patterns.
- Use farthest-first style point selection to choose representative and well-separated drone monitoring points.
- Design a Python workflow that joins drone detections, pond features, risk scoring, and ROI interpretation.
- Frame outputs using responsible AI principles, public safety, human oversight, and documented uncertainty.

2. Benchmark case-study statement

A three-pond fish farm breeding catfish and tilapia faces repeated losses, neighborhood concern, and declining return on investment. Field observations suggest intrusions from crocodiles, monitor lizards, snakes, and occasional birds of prey. Manual monitoring is slow, inconsistent, and risky.

The benchmark problem is to build an AI-assisted drone monitoring pipeline that detects visible threat agents, identifies the most affected pond, groups pond zones by risk profile, and translates the technical outputs into a safe management response and business-impact view.

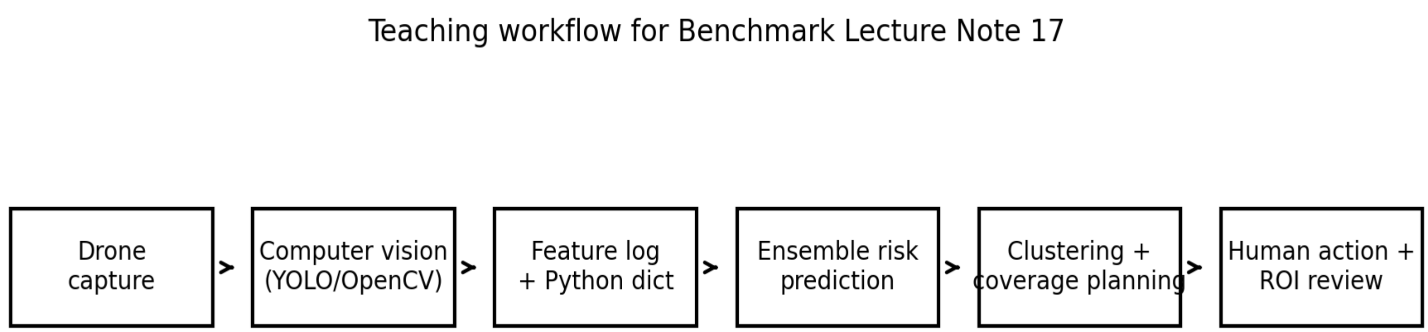

Figure 1. Teaching workflow used in the revised benchmark note: drone capture, computer vision, structured feature logging, ensemble prediction, clustering, and human-led response.

1.3 Scenario Illustration

Figure 17.1 presents an AI-generated conceptual replication of the aquaculture threat-detection scenario used in this benchmark lecture note. The figure helps learners visualize the relationship between drone surveillance, pond-level monitoring, visible threat agents, and fish population stress conditions.

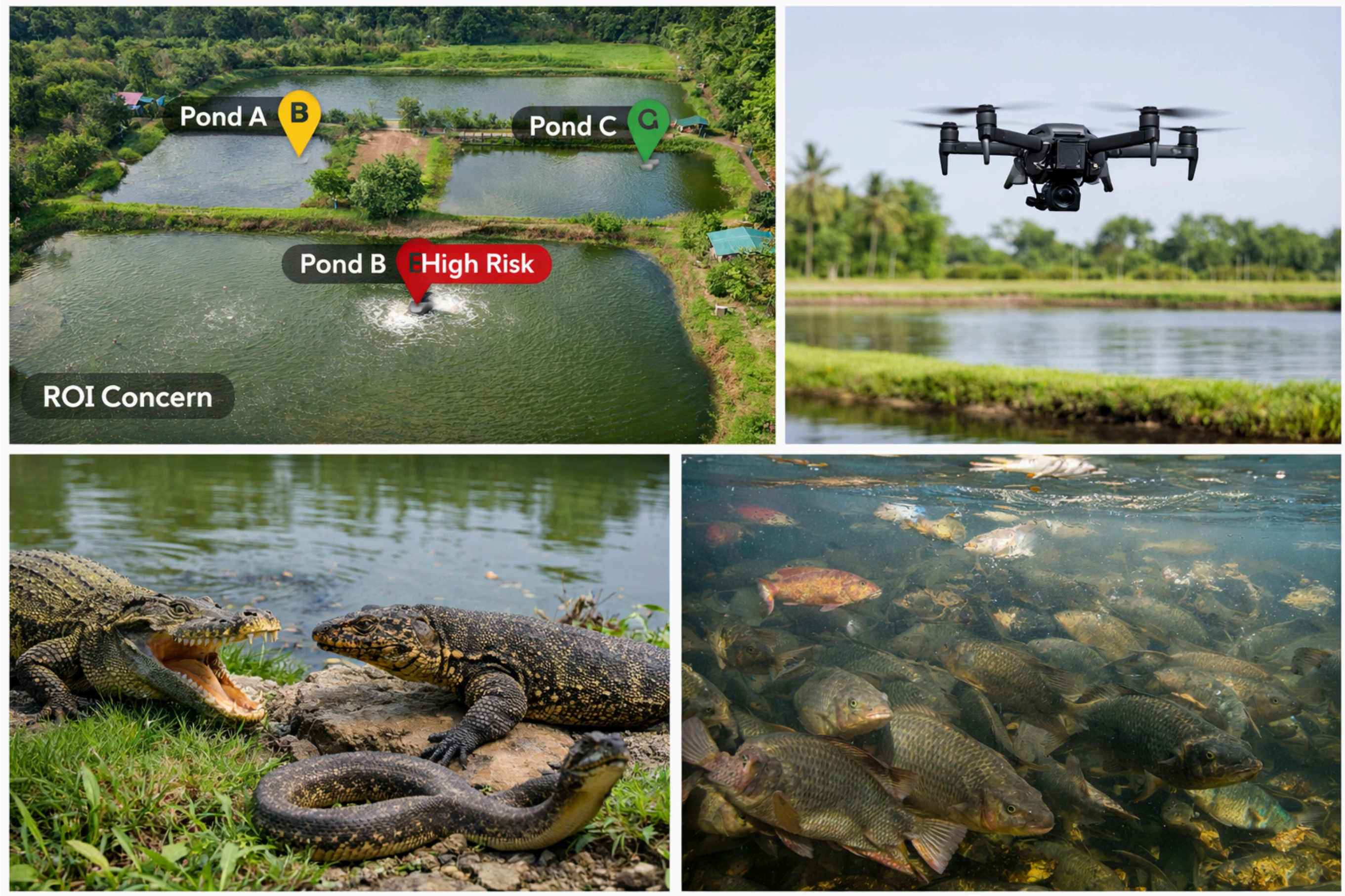

Figure 17.1. *AI-generated conceptual illustration of a drone-based aquaculture monitoring scenario showing three ponds, drone surveillance, predatory wildlife near pond edges, and fish-density conditions for teaching and simulation purposes. Pond B is highlighted as the highest-risk pond based on the sample monitoring scenario.*

This visual supports the case study by showing how an AI-enabled drone monitoring system can observe multiple ponds, detect threat agents near pond boundaries, and link visual evidence to pond-level risk scoring, clustering analysis, and ROI assessment. In the sample scenario, Pond B is treated as the most affected pond because it shows the strongest disturbance pattern and the highest projected operational loss.

Instructional note: This image is a simulated teaching visual created for classroom demonstration, benchmark scenario explanation, and capstone concept development. It is not an actual site photograph or live drone surveillance record.

The figure supports the analytical workflow described in this lecture note, where computer vision detects visible threat agents, clustering methods group pond-risk patterns, and ensemble models estimate the most affected pond and likely ROI decline.

3. AMK governance alignment

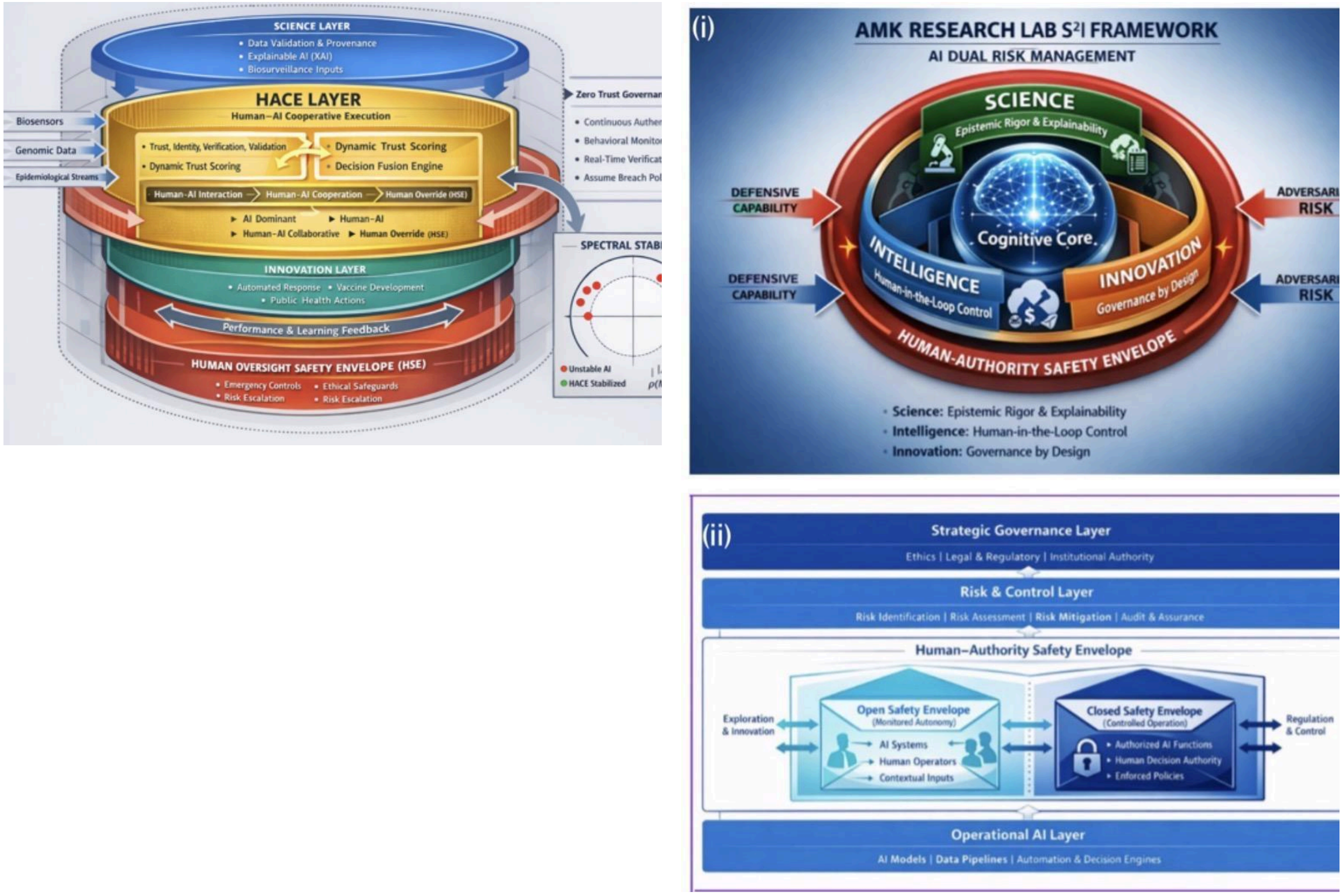

Figure 2. AMK governance visuals used to anchor the scenario in human authority, S2I governance, HACE collaboration, and stability-aware control.

Recommended governance posture: the drone system may detect and rank risk, but confirmation and intervention remain human-led, logged, and proportionate. This aligns with NIST AI RMF functions such as Govern, Map, Measure, and Manage, which frame AI deployment as a lifecycle risk-management activity rather than a one-time model build.

4. Teaching dataset and feature design

The note uses a synthetic teaching dataset so that the class can focus on method selection, traceable outputs, and responsible deployment without claiming that the numbers are real farm measurements.

Feature	Type	Interpretation	Example values
crocodile_count	Count	Visible crocodile detections by pond	2, 5, 1
monitor_lizard_count	Count	Visible monitor-lizard detections by pond	3, 4, 1
snake_count	Count	Visible snake detections by pond	1, 3, 1
bird_prey_count	Count	Bird-of-prey activity near the pond edge	2, 2, 1
fish_distress_score	Ordinal	Surface distress or abnormal movement signal	6, 9, 4
historical_loss_score	Ordinal	Past fish-loss pattern from farm records	5, 9, 3

5. Visual analytics and synthetic outputs

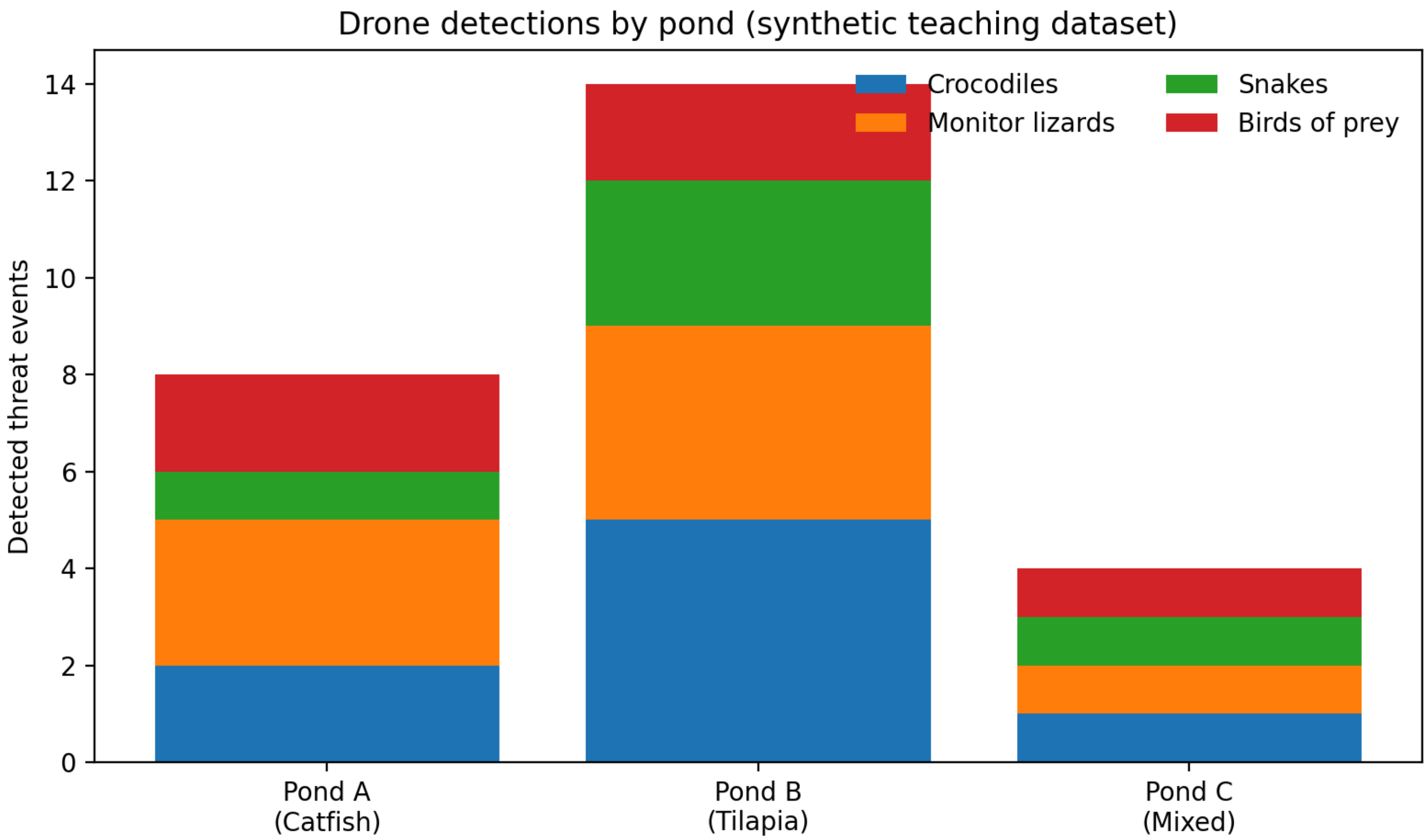

Figure 3. Synthetic stacked detections by pond. Pond B accumulates the largest number of reptile and predator events.

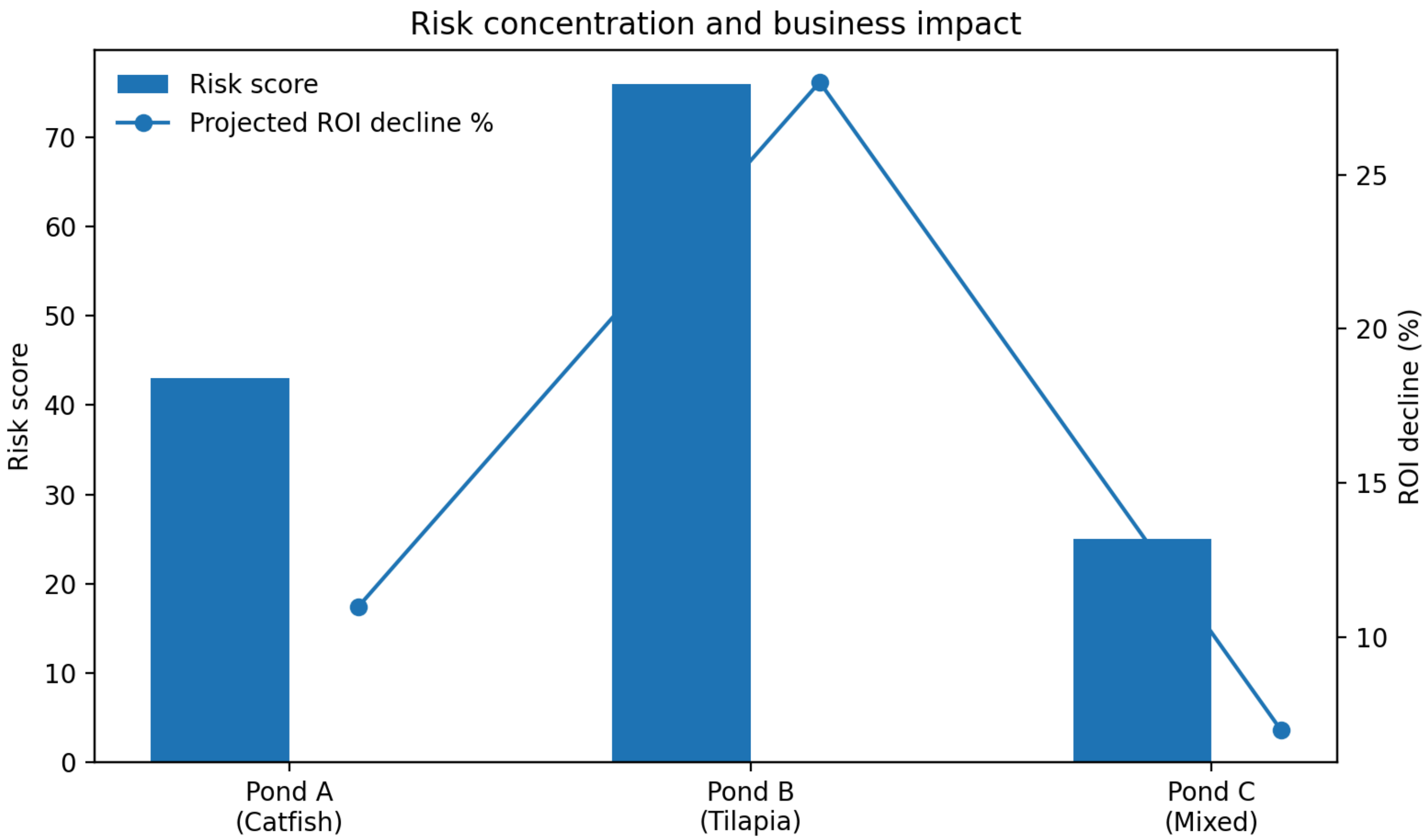

Figure 4. Risk concentration and projected ROI decline. The teaching scenario marks Pond B as the highest-priority operational risk.

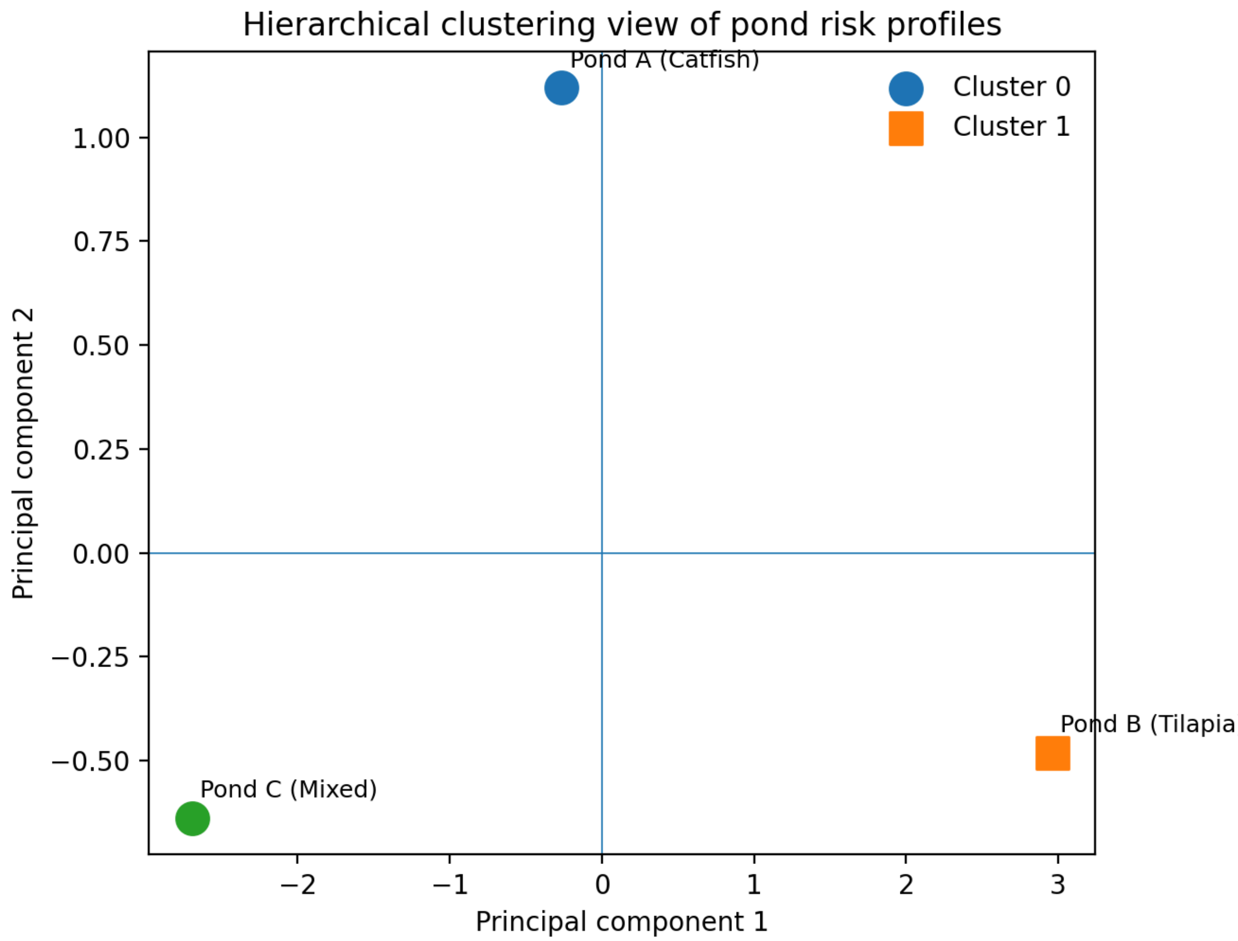

Figure 5. Hierarchical clustering view. A separate cluster around Pond B is a signal that its threat profile differs materially from the other ponds.

6. Worked benchmark output

Output item	Worked result
Likely threat agents	Crocodiles, monitor lizards, snakes, and occasional birds of prey
Most affected pond	Pond B (Tilapia)
Why Pond B ranks highest	Highest predator count, strongest fish-distress signal, highest historical-loss score, and the largest projected ROI decline
Risk classes	Pond A = medium, Pond B = high, Pond C = low to medium
Representative monitoring points	Pond B plus one lower-risk pond chosen for coverage contrast under farthest-first selection
Human action plan	Reinforce the Pond B perimeter, reduce edge vegetation, add evening monitoring, and escalate suspected crocodile activity to qualified human responders

Interpretation: the point of the benchmark is not only to name the most exposed pond, but also to move from detections to explainable managerial action. In an international benchmark context, the strongest answer is a transparent answer that links model output to a bounded human decision.

7. Sample Python workflow

```
import pandas as pd
from sklearn.ensemble import RandomForestClassifier, VotingClassifier
```

```python
from sklearn.cluster import AgglomerativeClustering
from sklearn.preprocessing import StandardScaler

data = pd.DataFrame({
    'pond': ['A', 'B', 'C'],
    'crocodile_count': [2, 5, 1],
    'monitor_lizard_count': [3, 4, 1],
    'snake_count': [1, 3, 1],
    'bird_prey_count': [2, 2, 1],
    'fish_distress_score': [6, 9, 4],
    'historical_loss_score': [5, 9, 3],
    'roi_decline_label': [1, 1, 0]
})

X = data.drop(columns=['pond', 'roi_decline_label'])
y = data['roi_decline_label']

rf = RandomForestClassifier(n_estimators=50, random_state=42)
ensemble = VotingClassifier(estimators=[('rf', rf)], voting='hard')
ensemble.fit(X, y)
pred = ensemble.predict(X)

scaler = StandardScaler()
clusters = AgglomerativeClustering(n_clusters=2).fit_predict(
    scaler.fit_transform(X)
)

print(pd.DataFrame({
    'pond': data['pond'],
    'predicted_high_loss_risk': pred,
    'cluster': clusters
}))
```

Sample output

```
  pond  predicted_high_loss_risk  cluster
0    A                         1        0
1    B                         1        1
2    C                         0        0

Most affected pond: B
Representative monitoring points: B and C
```

Teaching point: students should explain why a model can flag Pond A as medium-high while still ranking Pond B as the most affected pond. That discussion trains them to separate binary classification from richer, multi-factor operational scoring.

8. Responsible AI, public safety, and citation notes

- This benchmark is appropriate when framed as aquaculture threat detection, pond risk classification, and ROI protection - not as an autonomous harmful-response system.
- Outputs should support lawful, proportional, and human-led action. Wildlife handling, public notification, and any physical intervention should be escalated to qualified personnel and aligned with local regulation.
- Because the scenario uses synthetic data, performance values are illustrative. In a real deployment, students should document uncertainty, data-collection limits, model drift risk, and false-alarm costs.

References

1. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).
2. scikit-learn developers. Ensemble methods user guide. Official scikit-learn documentation.
3. scikit-learn developers. AgglomerativeClustering. Official scikit-learn documentation.
4. Python Software Foundation. Data structures - dictionaries. Official Python documentation.
5. NetworkX developers. Shortest paths. Official NetworkX documentation.
6. Ultralytics. Python usage and predict mode. Official YOLO documentation.
7. OpenCV. OpenCV-Python tutorials and object-detection module documentation.

AMK benchmark verdict: Pond B is the most affected pond in the worked example, and the appropriate response is reinforced monitoring, explainable reporting, and human-led intervention.

8. Sample Python Workflow

```
import pandas as pd

import numpy as np

from sklearn.ensemble import RandomForestClassifier, VotingClassifier

from sklearn.cluster import AgglomerativeClustering

from sklearn.metrics import classification_report

from sklearn.preprocessing import StandardScaler

from sklearn.linear_model import LogisticRegression

from sklearn.tree import DecisionTreeClassifier


# -----------------------------

# Sample pond monitoring dataset

# -----------------------------

data = pd.DataFrame({

  "pond": ["A", "B", "C"],

  "crocodile_count": [2, 5, 1],

  "monitor_lizard_count": [3, 4, 1],

  "snake_count": [1, 3, 1],

  "bird_prey_count": [2, 2, 1],
```

```
    "fish_distress_score": [6, 9, 4],
    "turbidity_score": [5, 8, 4],
    "veg_density_score": [7, 9, 5],
    "historical_loss_score": [5, 9, 3],
    "roi_decline_label": [1, 1, 0]  # 1 = high loss risk, 0 = lower risk
})

# ----------------------------
# Supervised learning section
# ----------------------------
X = data.drop(columns=["pond", "roi_decline_label"])
y = data["roi_decline_label"]

rf = RandomForestClassifier(n_estimators=50, random_state=42)
dt = DecisionTreeClassifier(random_state=42)
lr = LogisticRegression()

ensemble = VotingClassifier(
    estimators=[("rf", rf), ("dt", dt), ("lr", lr)],
    voting="hard"
)

ensemble.fit(X, y)
pred = ensemble.predict(X)

print("=== Classification Output ===")
print(pd.DataFrame({
    "pond": data["pond"],
    "predicted_high_loss_risk": pred
```

```
}))

print("\n=== Classification Report ===")
print(classification_report(y, pred, zero_division=0))

# -----------------------------
# Risk score interpretation
# -----------------------------
data["risk_score"] = (
    data["crocodile_count"] * 3 +
    data["monitor_lizard_count"] * 2 +
    data["snake_count"] * 2 +
    data["bird_prey_count"] * 1 +
    data["fish_distress_score"] * 2 +
    data["historical_loss_score"] * 3
)

most_affected = data.loc[data["risk_score"].idxmax(), "pond"]
print(f"\nMost affected pond: Pond {most_affected}")

# -----------------------------
# Hierarchical clustering
# -----------------------------
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

hc = AgglomerativeClustering(n_clusters=2)
clusters = hc.fit_predict(X_scaled)
```

```
data["hierarchical_cluster"] = clusters

print("\n=== Hierarchical Clustering Output ===")

print(data[["pond", "risk_score", "hierarchical_cluster"]])


# ----------------------------

# Simple farthest-first selection

# ----------------------------

def farthest_first(points, k=2):

    selected = [0]

    while len(selected) < k:

        distances = []

        for i in range(len(points)):

            if i in selected:

                distances.append(-1)

                continue

            min_dist = min(np.linalg.norm(points[i] - points[j]) for j in selected)

            distances.append(min_dist)

        selected.append(int(np.argmax(distances)))

    return selected


selected_points = farthest_first(X_scaled, k=2)

print("\nRepresentative farthest-first pond indices:", selected_points)

print("Representative ponds:", data.iloc[selected_points]["pond"].tolist())
```

9. Example Output Interpretation

A likely output from the sample code is:

- Pond B predicted as highest loss-risk pond
- Pond B assigned the highest risk score
- Hierarchical clustering groups Pond B apart from lower-risk ponds
- Farthest-first selection picks Pond B and one lower-risk pond as representative monitoring points

10. Responsible AI and Safety Note

This scenario should remain a **monitoring and decision-support system**. It should not automate harmful wildlife response. Recommended outputs should support:

- safe human inspection
- lawful environmental management
- public safety notification
- farm recovery planning
- transparent and explainable actions

11. Suggested Capstone Title

An AI-Driven Drone Monitoring Framework for Aquaculture Threat Detection, Pond Risk Classification, and ROI Protection

12. Short Results Template for the Lecture Note

Sample Findings:
The AI-assisted drone monitoring framework detected multiple predatory wildlife events across three ponds. Pond B showed the highest concentration of crocodile and reptile activity, elevated fish-distress signals, and the greatest projected ROI decline. Ensemble prediction classified Pond B as the highest-risk pond, while hierarchical clustering separated Pond B from the lower-risk pond group. Farthest-first clustering identified Pond B and Pond C as representative monitoring zones for strategic drone coverage. These results suggest that AI-supported monitoring can improve situational awareness, reduce delayed response, and support more targeted farm management decisions.

1.3 Illustration

Figure 17.1 presents an AI-generated conceptual replication of the aquaculture threat-detection scenario used in this benchmark lecture note. The figure helps learners visualize the relationship between drone surveillance, pond-level monitoring, visible threat agents, and fish-population stress conditions within a three-pond aquaculture environment. It also supports discussion of how computer vision, clustering methods, and ensemble learning can be integrated into a practical monitoring workflow for farm protection and operational decision support.

[Insert Figure 17.1 here — centered, width 6.0 to 6.5 inches]

Figure 17.1. AI-generated conceptual illustration of a drone-based aquaculture monitoring scenario showing three ponds, drone surveillance, predatory wildlife near pond edges, and fish-density conditions for teaching and simulation purposes. Pond B is highlighted as the highest-risk pond based on the sample monitoring scenario.

Instructional note. This image is a simulated teaching visual created for classroom demonstration, benchmark scenario explanation, and capstone concept development. It is not an actual site photograph or live drone surveillance record.

The figure supports the benchmark case study by showing how an AI-enabled drone monitoring framework can observe multiple ponds, detect threat agents near pond boundaries, and connect visible environmental patterns to pond-level risk scoring, clustering analysis, and ROI assessment. In the sample scenario, Pond B is treated as the most affected pond because it exhibits the strongest disturbance pattern, the highest observed concentration of predatory wildlife indicators, and the greatest projected operational loss.

1.4 Sample Analytical Output

Figure 17.2 presents a sample analytical output from the proposed AI framework. The chart summarizes how observed threat activity, fish-distress indicators, and historical loss variables can be transformed into composite pond-risk scores for management interpretation. This type of output is valuable because it converts raw drone observations and environmental monitoring signals into a decision-support view that can be understood by farm managers, safety stakeholders, and learners in applied AI programs.

[Insert Figure 17.2 here — centered, width 5.8 to 6.3 inches]

Figure 17.2. Sample model-output chart showing pond-level risk scores generated from drone-observed threat activity, fish-distress indicators, and historical loss variables. In the benchmark scenario, Pond B records the highest composite risk score and is therefore classified as the most affected pond.

Interpretation note. The chart is an instructional benchmark output designed to show how ensemble models and supporting analytics may summarize pond-level threat exposure. Values are illustrative and intended for teaching, simulation, and capstone development.

The analytical output demonstrates the logic of the proposed framework. First, drone imagery and video observations are processed through a computer-vision layer to detect visible wildlife activity and pond-surface disturbances. Next, extracted features are combined with contextual variables such as historical fish-loss patterns, environmental exposure, and pond conditions. These variables are then used to generate a composite risk score for each pond. In the benchmark scenario, Pond B shows the highest combined score, reflecting greater predator frequency, stronger fish-distress signals, and elevated historical loss indicators. This makes Pond B the priority location for intensified monitoring and management response.

1.5 Sample Pond-Risk Results Table

Table 17.1 summarizes the sample pond-level results produced by the benchmark scenario. The table combines biological, environmental, and operational indicators into a single management-oriented view. This structured summary is important because it allows farm operators and learners to move from visual detection to actionable interpretation. Rather than simply identifying the presence of wildlife, the framework supports a broader assessment of pond exposure, operational disruption, and expected business impact.

Table 17.1
Sample pond-risk classification results for the drone-based aquaculture monitoring scenario

Pond	Main Fish Type	Detected Threat Agents	Composite Risk Score	Risk Class	Estimated ROI Impact	Recommended Action
Pond A	Catfish	Monitor lizards, occasional birds	62	Medium	Moderate decline	Increase routine drone patrols and inspect boundary vegetation.
Pond B	Tilapia	Crocodiles, snakes, monitor lizards, birds	89	High	Significant decline	Prioritize reinforced monitoring, rapid inspection, and protective controls.
Pond C	Mixed/Secondary Stock	Snakes, occasional birds	48	Low-Medium	Limited decline	Maintain monitoring and review seasonal intrusion patterns.

Note. Risk classifications are based on illustrative benchmark variables including predator detections, fish-distress indicators, environmental exposure, and projected economic impact. The table is designed for teaching and scenario demonstration rather than field-certified deployment.

Table 17.1 identifies Pond B as the highest-risk pond because it shows the strongest concentration of threat indicators and the greatest projected operational loss. Pond A is classified as medium risk because it shows recurring but less severe wildlife pressure, while Pond C remains in the low-to-medium range due to more limited intrusion patterns. The table also demonstrates how AI-generated outputs should be connected to practical human decisions. In this scenario, the recommended response is not autonomous action against wildlife, but rather enhanced monitoring, human inspection, reinforced controls, and lawful safety management. This framing keeps the benchmark aligned with responsible AI, public safety, and human-centered oversight.