



# **Documentazione Caso di Studio**

## **Ingegneria della Conoscenza**

### **SafeFishing**

#### **Gruppo di lavoro**

[Francesco Gassi] - [764485] - [f.gassi2@studenti.uniba.it]

[Alessandro Carli] - [766403] - [a.carli3@studenti.uniba.it]

#### **Repository GitHub:**

[<https://github.com/FrancescoGassi/ProgettoICON>]

**A.A. 2024-25**

# SOMMARIO

1	Introduzione.....	3
1.1	Obiettivi.....	3
1.2	Strumenti.....	3
1.3	Librerie.....	3
2	Architettura del Sistema.....	4
2.1	Struttura Modulare.....	4
2.2	Requisiti Funzionali (RF).....	5
2.3	Requisiti Non Funzionali (RNF).....	5
2.4	Flusso di esecuzione.....	6
3	Rappresentazione della Conoscenza.....	6
4	Classificatore e valutazione.....	8
4.1	Validazione del Modello.....	8
4.2	Funzionamento del Modello Predittivo.....	8
5	Algoritmo del grafo.....	9
5.1	Struttura del Grafo.....	9
5.2	Gestione dell’Inquinamento e Restrizioni di Navigazione.....	9
5.3	Euristica.....	11
6	Interfaccia Utente.....	12
6.1	Funzionalità avanzate.....	13
7	Conclusioni.....	14
7.1	Punti di Forza.....	14
7.2	Aree di Miglioramento.....	14
7.3	Prospettive Future.....	14

---

# 1 Introduzione

Safe-Fishing nasce con l'obiettivo di applicare diversi metodi di apprendimento studiati nel corso di Ingegneria della Conoscenza per monitorare e regolamentare l'attività di pesca nel Mar Mediterraneo, tenendo conto dell'inquinamento ambientale.

Il sistema suddivide il Mar Mediterraneo e il Mar Nero in 10 regioni, seguendo le direttive della Food and Agriculture Organization (FAO) - Area 37. Ogni zona è classificata secondo un livello di inquinamento su una scala da 1 a 5 (molto basso, basso, moderato, alto, molto alto), determinato in base ai fattori inquinanti presenti.

## 1.1 Obiettivi

- **Previsione della presenza di specie ittiche** – Il software stima la probabilità di pesca per diverse specie, aiutando l'utente a individuare le migliori aree in base al livello di inquinamento.
- **Navigazione ottimizzata** – Il sistema suggerisce il percorso ideale per raggiungere una determinata zona, evitando le aree più inquinate.
- **Protezione ambientale** – Minimizza l'esposizione a zone ad alto inquinamento per preservare la qualità dell'acqua e garantire la salute della fauna marina.
- **Algoritmo di ricerca avanzato** – Utilizza un algoritmo ottimizzato per individuare, quando possibile, il tragitto meno inquinato.

Per garantire dinamicità e adattabilità a scenari reali, i livelli di inquinamento vengono assegnati in modo variabile, riflettendo la continua evoluzione delle condizioni ambientali. Nonostante questa variabilità, il sistema mantiene elevate performance e affidabilità, offrendo uno strumento efficace per una pesca più sostenibile.

## 1.2 Strumenti

- **Linguaggio di programmazione:** Python
- **IDE:** PyCharm
- **Hosting:** GitHub
- **Dataset:** CSV per la gestione delle informazioni strutturate

### 1.3 Librerie

NOME	DESCRIZIONE
Pandas	Pandas è una libreria software scritta per il linguaggio di programmazione Python per la manipolazione e l'analisi dei dati. In particolare, offre strutture dati e operazioni per manipolare tabelle numeriche e serie temporali.
Scikit-learn	Scikit-learn (ex scikits learn) è una libreria open source di apprendimento automatico per il linguaggio di programmazione Python. Contiene algoritmi di classificazione, regressione e clustering (raggruppamento) e macchine a vettori di supporto, regressione logistica, classificatore bayesiano, k-mean e DBSCAN, ed è progettato per operare con le librerie NumPy.
Pillow (PIL)	Pillow è una libreria Python per la manipolazione delle immagini che consente di aprire, modificare, elaborare e salvare immagini in vari formati, offrendo funzionalità come ridimensionamento, ritaglio, conversione di colori, applicazione di filtri ed editing grafico. È utile per la visualizzazione della mappa delle zone FAO.
NumPy	NumPy (Numerical Python) è una libreria Python per il calcolo scientifico che fornisce supporto per array multidimensionali e matrici, insieme a una vasta gamma di funzioni matematiche ad alte prestazioni per l'elaborazione numerica e l'algebra lineare.

---

## 2 Architettura del Sistema

### 2.1 Struttura Modulare

MODULO	FUNZIONE
main.py	Interfaccia utente e coordinamento
BaseConoscenza.py	Gestione KB (confini, inquinamento)

Classificatore.py	Previsione probabilità di pesca
Grafo.py	Navigazione con A*
analisi_risultati.py	è uno script che estrae e organizza i dati principali in tabelle chiare per report e analisi rapide.

## 2.2 Requisiti Funzionali (RF)

ID	DESCRIZIONE	PRIORITA'	DOVE SI TROVA
RF 1	Prevedere la probabilità di pesca (%) per una specie ittica in una zona FAO	Alta	<i>Classificatore.py:</i> <i>prevedi_probabilita(),</i> <i>_addestra_modello()</i>
RF 2	Calcolare il percorso ottimale tra due zone, evitando aree con inquinamento "molto alto"	Alta	<i>Grafo.py:</i> <i>ricerca_astar(),</i> <i>genera_grafo()</i>
RF 3	Verificare se una zona ha un specifico livello di inquinamento	Media	<i>BaseConoscenza.py:</i> <i>domanda_inquinament</i> <i>o_zona()</i>
RF 4	Controllare se il passaggio tra due zone è consentito (confini e inquinamento)	Media	<i>BaseConoscenza.py:</i> <i>domanda_passaggio()</i>
RF 5	Mostrare l'elenco delle zone per livello di inquinamento	Bassa	<i>BaseConoscenza.py:</i> <i>trova_zone_per_inquin</i> <i>amento()</i>
RF 6	Visualizzare la mappa statica delle zone FAO	Bassa	<i>main.py:</i> <i>mostra_mappa()</i> <i>(con Pillow)</i>

## 2.3 Requisiti Non Funzionali (RNF)

ID	DESCRIZIONE	METRICA	STATO NEL CODICE
----	-------------	---------	------------------

RNF 1	Tempo di risposta per previsioni e percorsi	< 1 secondo	Ottimizzato (caching modello e grafo)
RNF 2	Validazione input utente (zone e specie valide)	100% input controllati	<code>_chiedi_zona()</code> , <code>_chiedi_pesce()</code>
RNF 3	Portabilità cross-platform	Supporto Windows/Linux/macOS	Uso di <code>os.path</code> e librerie compatibili
RNF 4	Usabilità interfaccia a riga di comando	Messaggi di errore chiari	Prompt guidati e feedback immediato
RNF 5	Scalabilità (aggiunta nuove zone/pesci)	Modifica senza riprogettazione	Dizionari estendibili ( <code>PESCI_VALIDI</code> , <code>ZONE_VALIDI</code> )
RNF 6	Affidabilità del modello di ML	Accuratezza $\geq 85\%$ ( $R^2$ )	Cross-validation in <code>valuta_modello()</code>

## 2.4 Flusso di esecuzione

1. Avvio dell'applicazione e visualizzazione header
2. Presentazione menu principale
3. Selezione di un'opzione da parte dell'utente
4. Esecuzione della funzionalità richiesta
5. Visualizzazione risultati
6. Possibilità di richiedere dettagli ("how")
7. Ritorno al menu principale

Il progetto combina quindi elementi di intelligenza artificiale (classificatore), gestione della conoscenza (knowledge base), algoritmi di grafi (navigazione) e un'interfaccia utente intuitiva.

---

## 3 Rappresentazione della Conoscenza

La knowledge base è stata implementata utilizzando la logica del primo ordine. In questo contesto, una base di conoscenza è costituita da un insieme di proposizioni, chiamate assiomi, che vengono assunte come vere senza necessità di dimostrazione. Questo approccio consente di rappresentare,

all'interno di un sistema informatico, la conoscenza relativa a un determinato dominio.

Grazie a questa struttura, l'utente può:

- **Verificare il livello di inquinamento di una zona** – È possibile interrogare il sistema per sapere se una specifica area presenta un determinato livello di inquinamento. Inoltre, tramite il comando "how", l'utente può richiedere una spiegazione dettagliata sul processo di valutazione.
- **Controllare la possibilità di navigazione tra due zone** – Il software permette di verificare se è possibile spostarsi da una zona all'altra, tenendo conto dei livelli di inquinamento. Anche in questo caso, l'utente può chiedere al sistema di motivare la decisione utilizzando il comando appropriato.

Grazie a questa implementazione basata sulla logica del primo ordine, il sistema garantisce trasparenza nelle decisioni e fornisce una spiegazione chiara dei suoi processi di ragionamento.

```
Inserisci il numero dell'opzione desiderata (0-12) → 12

-----
VERIFICA INQUINAMENTO ZONA
-----
Zona da verificare (es. 1.1): 2.1
Livello da verificare (moltoBasso/basso/moderato/alto/moltoAlto): alto

=====
VERIFICA INQUINAMENTO: 2.1
=====
• Livello attuale: basso
• Livello verificato: alto
-----
RISULTATO: ✗ NO
```

Livello di inquinamento

```
Inserisci il numero dell'opzione desiderata (0-12) → 11

=====
VERIFICA PASSAGGIO TRA ZONE
=====

Zona partenza (es. 1.1) o 'indietro': 1.1
Zona arrivo (es. 1.2): 1.2

=====
VERIFICA PASSAGGIO: 1.1 → 1.2
=====
• Zone confinanti: ☒ (collegate direttamente)
• Inquinamento 1.1: alto ☒
• Inquinamento 1.2: basso ☒
=====
RISULTATO: ☒ YES
```

Passaggio tra zone

---

## 4 Classificatore e valutazione

Il classificatore adottato è un albero decisionale (Decision Tree), implementato in modo da interagire direttamente con l'utente durante l'esecuzione del programma. L'utente dovrà fornire due input principali:

1. La zona in cui desidera pescare.
2. La specie di pesce di interesse.

Sulla base di queste informazioni, il sistema calcolerà e restituirà la probabilità di pesca per quella specifica specie nella zona selezionata.



```
#####
#          LIVELLI DI INQUINAMENTO PER ZONA          #
#####
```

Livello	Zone	Count
alto	3.2, 4.2	2
basso	1.2	1
moderato	2.1, 4.1	2
moltoAlto	3.1, 4.3	2
moltoBasso	1.1, 1.3, 2.2	3

**Commento:**

- La maggior parte delle zone ricade nelle categorie "moltoBasso" (3 zone) e "alto" (2 zone), mentre solo 1 zona ("1.2") è classificata come "basso". Le categorie "moderato" e "moltoAlto" presentano entrambe 2 zone ciascuna. Questo suggerisce che i livelli di inquinamento siano distribuiti in modo irregolare, con una concentrazione maggiore di zone nei livelli più bassi ("moltoBasso") e in quelli più alti ("alto").
- Implicazioni: Le zone classificate come "moltoAlto" (es. 3.1, 4.3) saranno critiche per la navigazione (bloccate nel grafo), mentre le zone "moltoBasso" (es. 1.1, 1.3, 2.2) sono ideali per la pesca sostenibile..
- Dato rilevante: solo 2 zone presentano un livello di inquinamento "moderato" (2.1, 4.1), suggerendo una polarizzazione verso estremi bassi ("moltoBasso") o alti ("moltoAlto" e "alto").

```
#####
#          PRESTAZIONI CLASSIFICATORE          #
#####
```

Metrica	Valore
Accuratezza Media...	0.910
Deviazione Standard	0.086

### Commento:

- Media accuratezza ( $R^2 = 0.91$ ): Il modello predice bene le probabilità di pesca, con un margine di errore contenuto.
- Deviazione standard bassa (0.086): Le prestazioni sono consistenti tra diverse valutazioni (cross-validation).

Conclusione: Il classificatore è affidabile per consigliare zone di pesca, ma potrebbe essere migliorato con più dati (es. condizioni meteo stagionali).

```
#####
#                                     #
#          CONNESSIONI TRA ZONE      #
#####
Da      A
-----
1.1 → 1.2
-----
1.1 → 1.3
-----
1.2 → 1.1
-----
1.2 → 1.3
-----
1.3 → 1.1
-----
1.3 → 1.2
-----
1.3 → 2.2
-----
2.1 → 2.2
-----
2.2 → 1.3
-----
2.2 → 2.1
-----
2.2 → 3.2
-----
3.2 → 2.2
-----
4.1 → 4.2
-----
4.2 → 4.1
-----
```

### Commento:

- Retaggio non orientato: Ogni connessione è bidirezionale (es.  $1.1 \leftrightarrow 1.2$ ), coerente con orientato=False nel grafo.
- Zone isolate: Se una zona non appare (es. 4.3), è perché ha inquinamento "moltoAlto" ed è esclusa automaticamente.
- Percorsi ottimali: La struttura mostra che per passare da 1.1 a 3.1 serve un percorso multistep (es.  $1.1 \rightarrow 1.3 \rightarrow 2.2 \rightarrow 3.1$ ).

MIGLIORI ZONE DI PESCA PER SPECIE		
Specie	Zona	Probabilità
Acciuga	2	92%
Acciuga	4	85%
Acciuga	5	80%
Merluzzo	1	92%
Merluzzo	3	85%
Merluzzo	5	80%
Pesce Spada	2	92%
Pesce Spada	4	85%
Pesce Spada	6	80%

#### Commento:

- Probabilità più alte: Le specie hanno picchi di versi (es. Acciuga: 92% in zona 2, Merluzzo: 92% in zona 1).
- Utilità pratica: I pescatori possono selezionare la zona in base alla specie target (es. Zona 2 per Acciughe).

## 4.1 Validazione del Modello

L'affidabilità del classificatore è garantita dall'uso della k-fold cross-validation, una tecnica che suddivide il dataset in più sottoinsiemi per migliorare la capacità predittiva del modello. In particolare, è stato scelto un valore di  $k = 5$ , un parametro standard nelle librerie di machine learning, che assicura un buon compromesso tra accuratezza e prestazioni computazionali.

Il sistema utilizza come metrica di valutazione il coefficiente di determinazione ( $R^2$ ), che misura la capacità del modello di spiegare la variabilità dei dati. Un valore elevato di  $R^2$  indica una maggiore affidabilità delle previsioni, garantendo così un sistema più preciso ed efficiente per supportare l'utente nella pianificazione dell'attività di pesca.

## 4.2 Funzionamento del Modello Predittivo

Le previsioni del sistema sono basate su un principio chiave: la probabilità di pesca tende a diminuire spostandosi verso le zone più a est. Ad esempio, un tonno sarà più facilmente reperibile nella zona 1.1 rispetto alla zona 4.3. Tuttavia, non tutte le aree con alta probabilità di pesca saranno necessariamente accessibili: se il livello di inquinamento è troppo elevato, la zona potrebbe risultare impraticabile.

L'utente può quindi valutare la convenienza di pescare in una determinata area sulla base della probabilità fornita dal sistema, ma per un'analisi più completa potrà sfruttare anche le altre funzionalità disponibili, come la verifica dell'accessibilità della zona e del livello di inquinamento. Queste verifiche non sono obbligatorie, ma altamente consigliate per ottimizzare l'attività di pesca.

- **Dataset:** probpesca.csv (zone, specie, probabilità %)
- **Addestramento:** 5-fold cross-validation ( $R^2 = 0.85 \pm 0.05$ )

```
Inserisci il numero dell'opzione desiderata (0-12) → 1
-----
CLASSIFICAZIONE ZONE DI PESCA
-----
Inserisci il pesce che vorresti pescare (es: tonno): tonno
Inserisci la zona FAO dove vuoi pescare (es: 1.1): 2.1
```

Input utente

```
Probabilità di pescare tonno in zona 2.1: 85%
```

Output

---

## 5 Algoritmo del grafo

Il grafo implementato ha lo scopo di individuare il percorso ottimale tra un nodo di partenza e un nodo di destinazione (goal). Per questa ricerca è stato scelto l'algoritmo A\*, noto per la sua efficienza nel trovare il cammino più breve considerando anche un costo associato agli spostamenti.

## 5.1 Struttura del Grafo

- Nodi → Corrispondono alle regioni FAO.
- Archi → Rappresentano i collegamenti diretti tra zone confinanti.
- Pesi degli archi → Determinati dal livello di inquinamento delle zone attraversate.

L'utente può cercare un percorso con il minor impatto ambientale, riducendo il rischio di peggiorare le condizioni delle zone già inquinate e garantendo una pesca più sicura, evitando pesce contaminato anche nelle aree di transito.

## 5.2 Gestione dell'Inquinamento e Restrizioni di Navigazione

I livelli di inquinamento sono classificati su una scala di cinque valori:

- Molto basso
- Basso
- Moderato
- Alto
- Molto alto

```
Inserisci il numero dell'opzione desiderata (0-12) → 9
```

```
Livelli di inquinamento:
```

- moltoBasso
- basso
- moderato
- alto
- moltoAlto

```
Premi Invio per continuare...█
```

Il sistema stabilisce le regole di navigazione in base a questi livelli:

**✗ Le zone con inquinamento “molto alto” sono inaccessibili** – Non è possibile attraversarle né raggiungerle, per garantire una pesca sicura e ridurre l'impatto ambientale.

```

Inserisci il numero dell'opzione desiderata (0-12) → 11

=====
VERIFICA PASSAGGIO TRA ZONE
=====

Zona partenza (es. 1.1) o 'indietro': 3.2
Zona arrivo (es. 1.2): 2.2

=====
VERIFICA PASSAGGIO: 3.2 → 2.2
=====
• Zone confinanti: ✓ (collegate direttamente)
• Inquinamento 3.2: moltoAlto ✗
• Inquinamento 2.2: alto ✓

=====
RISULTATO: ✗ NO

```

✓ **Se il livello di inquinamento è inferiore a "molto alto"**, il sistema consente la navigazione e il transito attraverso la zona FAO selezionata. Grazie a questa implementazione, l'algoritmo A\* permette di calcolare il percorso più sicuro, bilanciando la necessità di spostamento con il rispetto delle condizioni ambientali.

```

Inserisci il numero dell'opzione desiderata (0-12) → 11

=====
VERIFICA PASSAGGIO TRA ZONE
=====

Zona partenza (es. 1.1) o 'indietro': 1.1
Zona arrivo (es. 1.2): 1.2

=====
VERIFICA PASSAGGIO: 1.1 → 1.2
=====
• Zone confinanti: ✓ (collegate direttamente)
• Inquinamento 1.1: basso ✓
• Inquinamento 1.2: alto ✓

=====
RISULTATO: ✓ YES

```

### 5.3 Euristica

Il costo del percorso dipende da:

- **Distanza:** Differenza di inquinamento tra zone.
- **Vincoli:** Blocco delle zone con *inquinamento* = *moltoAlto*

**Formula euristica:**

$$h(n) = \frac{|inquinamento(n) - inquinamento(goal)|}{100}$$

```
Inserisci il numero dell'opzione desiderata (0-12) → 10  
  
-----  
CALCOLO PERCORSO OTTIMALE  
-----  
Zona di partenza (es. 1.1): 1.1  
Zona di destinazione (es. 1.2): 4.3  
Percorso ottimale trovato:  
1.1 → 1.3 → 2.2 → 3.1 → 4.1 → 4.2 → 4.3  
  
Premi Invio per continuare... █
```

Esempio di percorso

---

## 6 Interfaccia Utente

SAFEFISHING		
Creato da: Francesco Gassi & Alessandro Carli		
NUM	OPZIONE	DESCRIZIONE
1	Classificazione	Classifica le zone di pesca
2	Valutazione Classificatore	Valuta le prestazioni del classificatore
3	Mostra Zone FAO	Visualizza la mappa delle zone FAO
4	Zone molto basso inquinamento	Elenco zone con inquinamento minimo
5	Zone basso inquinamento	Elenco zone con inquinamento basso
6	Zone moderato inquinamento	Elenco zone con inquinamento moderato
7	Zone alto inquinamento	Elenco zone con inquinamento alto
8	Zone molto alto inquinamento	Elenco zone con inquinamento massimo
9	Lista inquinamenti	Visualizza tutti i livelli di inquinamento
10	Trova percorso	Calcola il percorso tra due zone
11	Verifica passaggio	Controlla se è possibile passare tra due zone
12	Verifica inquinamento	Controlla il livello di inquinamento
0	Esci	Chiudi l'applicazione

Inserisci il numero dell'opzione desiderata (0-12) →

Menù interattivo con 12 funzionalità

### 6.1 Funzionalità avanzate

```
Digita 'how' per maggiori dettagli o INVIO per continuare...
how

=====
SPIEGAZIONE DETTAGLIATA
=====

CONNESSIONE TRA LE ZONE:
- Confinanti: No
  Non esiste un collegamento diretto tra queste zone

LIVELLI DI INQUINAMENTO:
- 1.1: moltoAlto
- 3.2: basso

Nota: Il passaggio è bloccato se una zona ha inquinamento 'moltoAlto'

=====
CONNESSIONI DISPONIBILI PER QUESTE ZONE:
- 1.1
- 1.2
- 1.3
- 2.2
- 3.1
- 3.2
```

Comando *how*: Spiegazioni dettagliate (es: "Perché il passaggio è bloccato?")



Inserisci il numero dell'opzione desiderata (0-12) → 3

APERTURA MAPPA FAO...

La mappa rimarrà aperta. Chiudila manualmente quando hai finito.  
Premi Invio per tornare al menu principale...



Visualizzazione mappa: File *FAO37\_sottozone.png*

---

## 7 Conclusioni

SafeFishing rappresenta una soluzione innovativa che integra tecniche di Intelligenza Artificiale (IA) e gestione della conoscenza per ottimizzare le attività di pesca, garantendo maggiore efficienza e sostenibilità ambientale.

### 7.1 Punti di Forza

- **Accuratezza dell'85%** nelle previsioni di pesca.
- **Pianificazione ottimizzata** dei percorsi, con riduzione dell'impatto ambientale.
- **Architettura modulare e scalabile**, adattabile a diversi contesti operativi.

### 7.2 Aree di Miglioramento

- Integrazione con **fonti dati reali** per una maggiore affidabilità.
- Ampliamento dei **parametri ambientali** considerati nel modello.
- Sviluppo di un'**interfaccia grafica** per migliorare l'esperienza utente.

### 7.3 Prospettive Future

- Adozione in **scenari operativi reali**, con applicazioni nel settore della pesca professionale.
- Estensione del sistema ad **altre aree FAO**, per una copertura geografica più ampia.
- Implementazione su **piattaforme mobili**, rendendo il sistema accessibile in ogni situazione.

SafeFishing costituisce una base solida per lo sviluppo di strumenti avanzati a supporto della pesca sostenibile, con potenziali benefici sia economici che ambientali.