Detecting Trajectory Spoofing Attacks on AIS

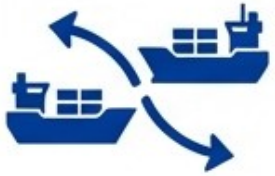
Autori: Andrea La Rocca, Cristina Tomaciello

Il Sistema AIS

Cos'è l'AIS? La spina dorsale della navigazione moderna.

Il Sistema di Identificazione Automatica (AIS) è lo standard globale dell'Organizzazione Marittima Internazionale (IMO) che permette alle navi e alle stazioni a terra di scambiare dati cinematici in tempo reale via radiofrequenze VHF.

Funzioni Fondamentali:



Prevenzione Collisioni

Aumenta la sicurezza della navigazione per evitare incidenti in mare.

Gestione del Traffico

Permette alle autorità di monitorare e gestire il traffico marittimo globale.

Sicurezza e Logistica

Cruciale per operazioni di salvataggio, logistica e monitoraggio di attività illecite.

La Minaccia: Un Protocollo Intrinsecamente Vulnerabile

Guerra Ibrida e Disinformazione

Attività Illegali
(Contrabbando, Pirateria)

Attacchi Cinetici
(Collisioni indotte)

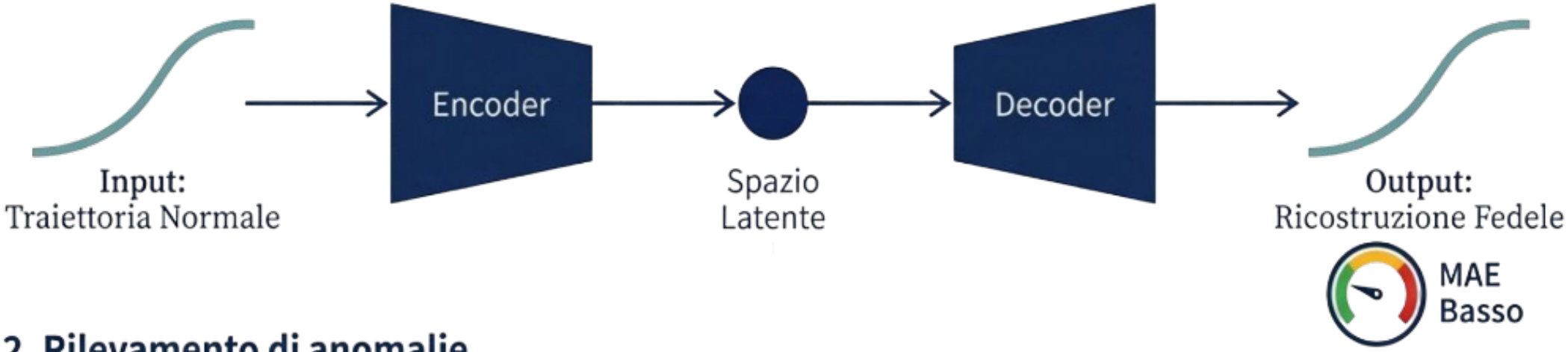
Mancanza di Autenticazione e Cifratura

Tassonomia delle Minacce: 4 Scenari di Attacco

Per validare il sistema in assenza di dataset pubblici di attacchi reali, sono state iniettate **quattro tipi di minacce sintetiche** in traiettorie legittime.

Strategia di Difesa: Anomaly Detection

1. Addestramento su dati normali

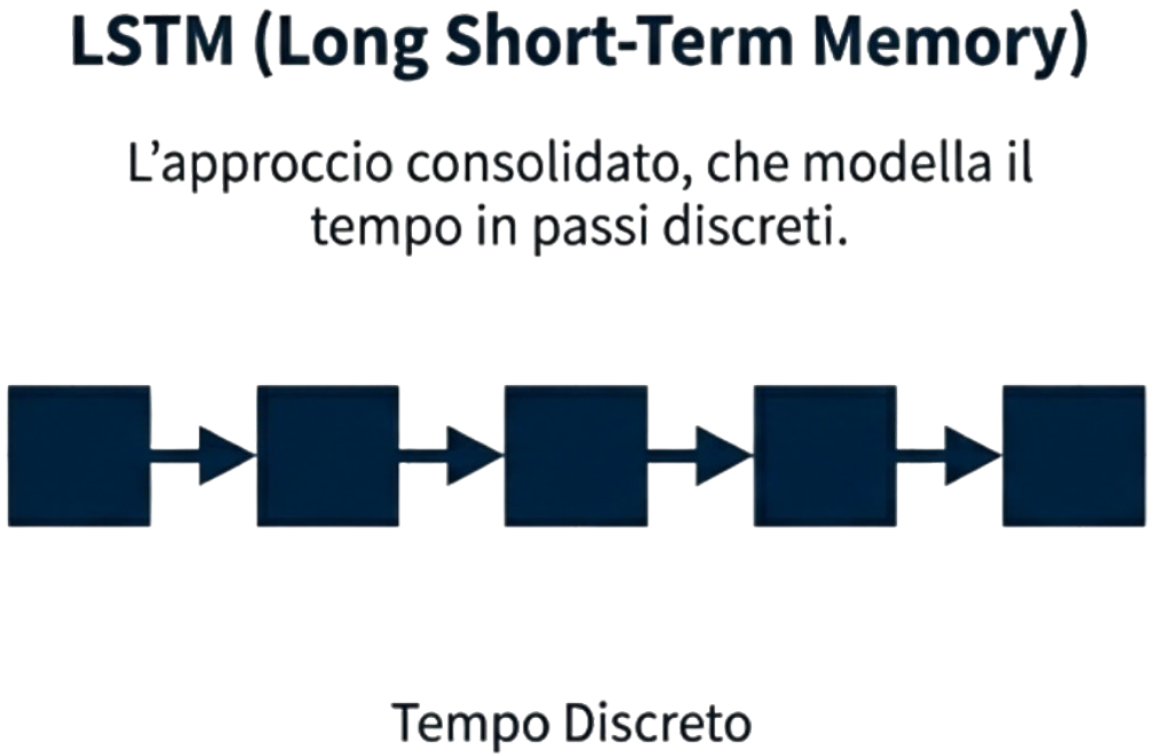


2. Rilevamento di anomalie

Un alto valore di MAE è il segnale che smaschera l'attacco.

L’AutoEncoder LSTM (Long Short-Term Memory)

- **Standard per Serie Temporali:** L'architettura LSTM è la soluzione più consolidata per gestire dati sequenziali, perché risolve il problema della "memoria corta" (*vanishing gradient*) delle reti ricorrenti classiche.
- **Il "Cervello" della Cella:** Gestisce il flusso di informazioni tramite tre *Gate* logici:
  1. **Forget Gate:** Decide cosa dimenticare dal passato.
  2. **Input Gate:** Decide quali nuove informazioni memorizzare.
  3. **Output Gate:** Determina l'output per il passo successivo.
- Configurazione:
  - **128 unità per layer**
  - **Ottimizzatore:** Adam (*LR*=0.001)



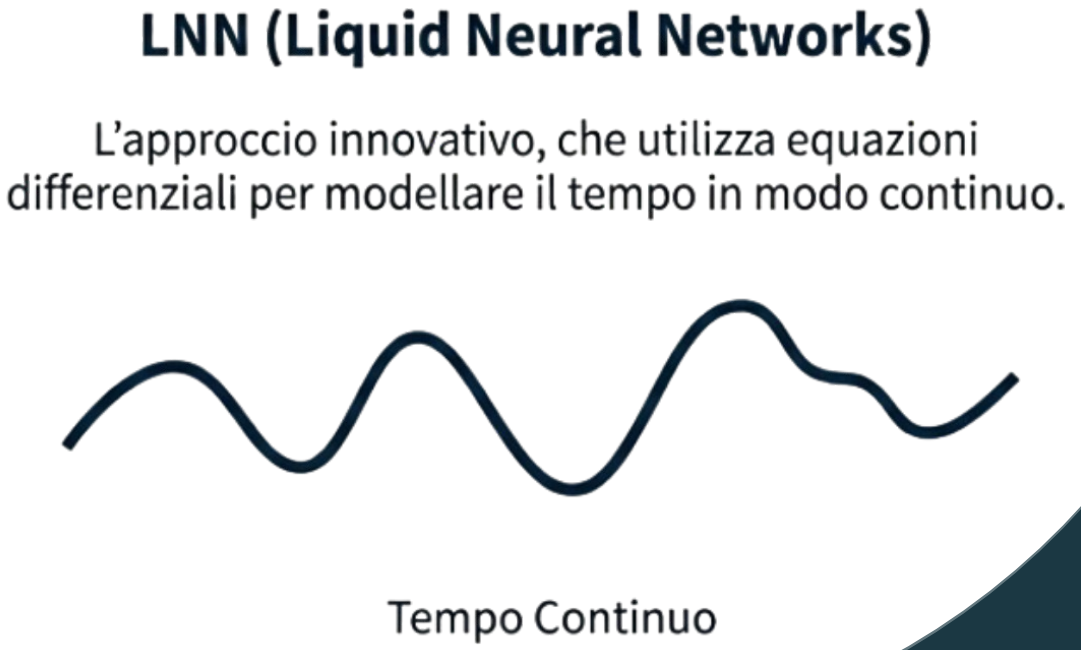
L’ AutoEncoder LNN (Liquid Neural Network)

- A differenza delle reti classiche, le LNN non lavorano a tempo discreto ma utilizzano **Equazioni Differenziali Ordinarie (ODE)** per modellare l'evoluzione del sistema, adattandosi perfettamente alla natura fluida del movimento di una nave.

Caratteristiche essenziali:

- **Celle CfC (Closed-form Continuous-time):** Il cuore della rete. Queste celle hanno il compito di calcolare lo stato del sistema in qualsiasi istante. La loro peculiarità è di non limitarsi a ricordare cosa è successo, ma di riuscire a modellare il sistema fisico nel tempo, anche tra i punti di campionamento discreti.

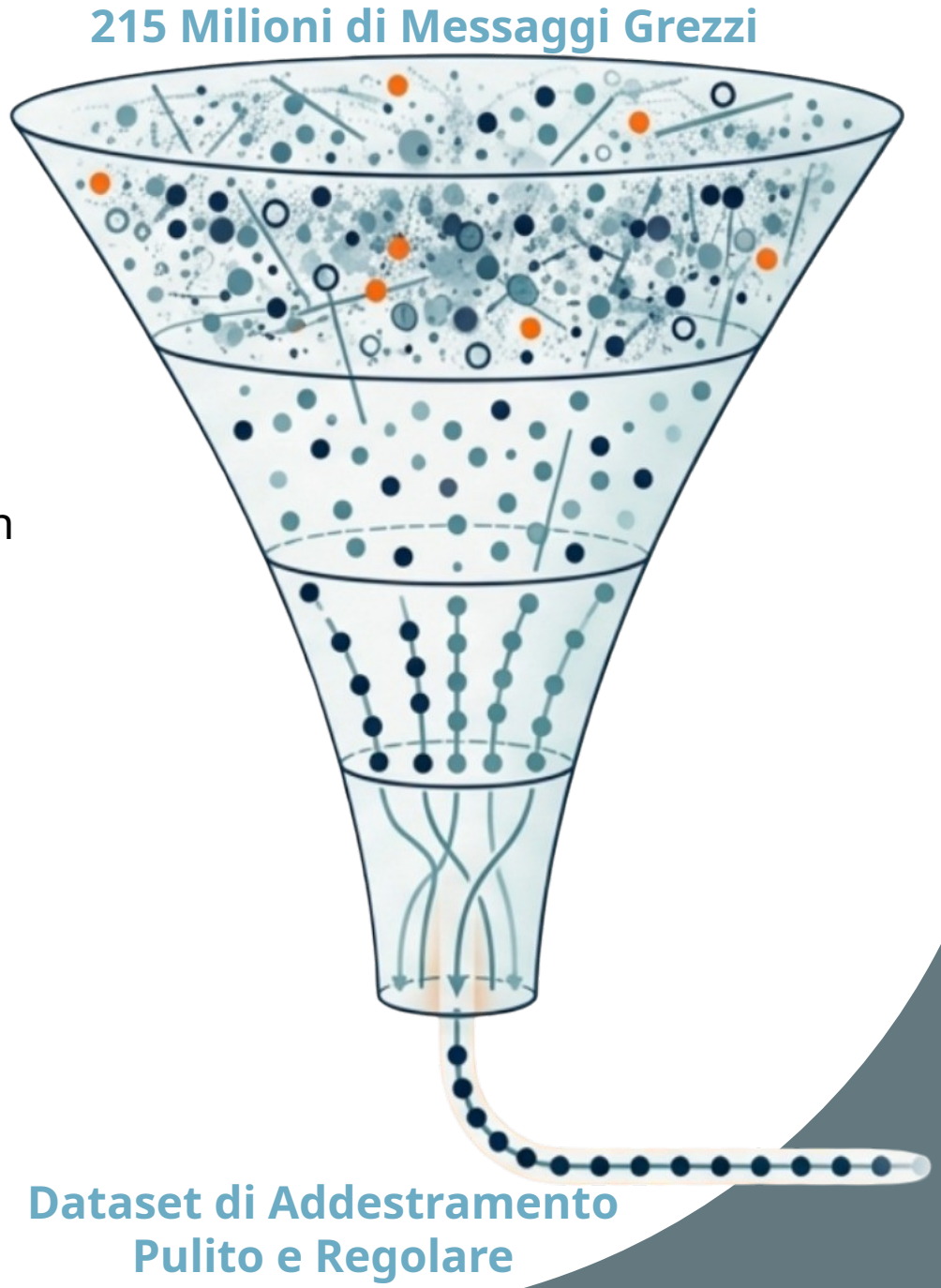
- **Efficienza AutoNCP:** Utilizza un **Wiring Sparso** (connettività non totale), ispirato ai cervelli biologici. Questo comporta l’utilizzo di meno parametri da addestrare conferendo al modello una maggiore robustezza al rumore e una migliore capacità di generalizzazione rispetto alle reti dense.



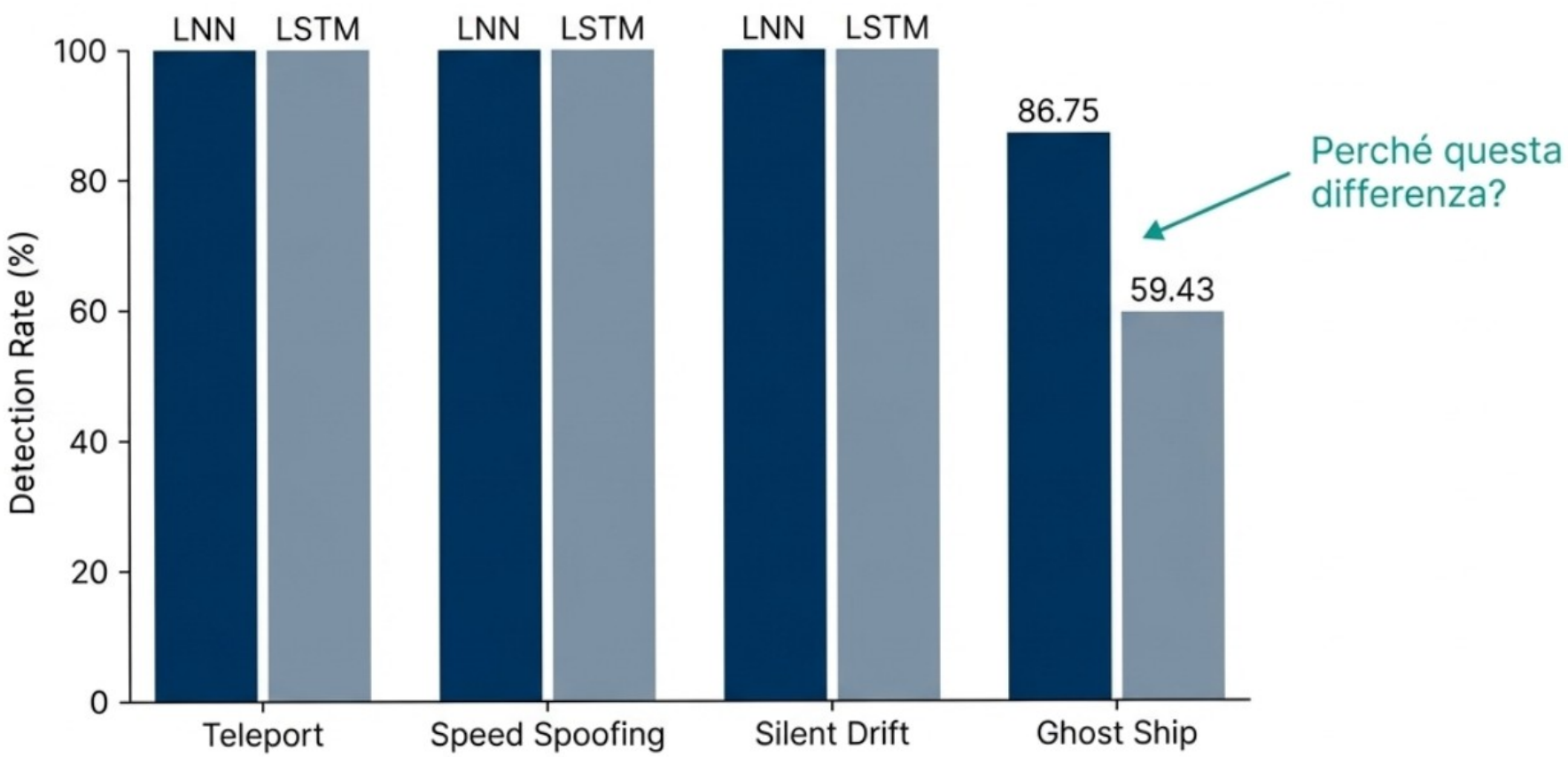
DataSet & Preprocessing

Fonte Dati: Dataset pubblico della U.S Maritime Admistration (MARAD), da Luglio 2024 a Giugno 2025

- **Data Cleaning:** Rimozione di navi stazionarie (SOG<2 nodi) e dati anomali dei sensori.
- **Segmentazione Temporale:** Suddivisione delle traiettorie con interruzioni superiori a 60 minuti.
- **Interpolazione (Snap-to-Grid):** Ricampionamento dei dati su una griglia temporale fissa di 1 minuto.
- **Preparazione Input:**
  - **Normalizzazione:** StandardScaler (media 0, varianza 1) per stabilizzare il training.



Performance di Rilevamento: Recall



Non Sensibilità, ma Stabilità del Modello

Entrambe le reti, LSTM e LNN, hanno raggiunto **un'ottima convergenza** durante il training.

Tuttavia, l'analisi della distribuzione dell'errore ha rivelato due aspetti importanti:

• **Risposta all'Attacco**: Di fronte all'incoerenza cinematica di un attacco Ghost Ship, entrambe le reti reagiscono generando un errore di ricostruzione assoluto simile (≈ 0.026).

• **Comportamento sul Normale**: La LNN presenta una varianza (σ) inferiore sul traffico lecito rispetto alla LSTM.

Considerazione :
La LNN non è più sensibile all'attacco ma è più **stabile** sul traffico normale, consentendo una soglia di allarme più precisa.

Il Punto Cieco dei Modelli

Analisi Apprendimento Effettivo

I modelli non agiscono come sistemi di monitoraggio geografico assoluto.
La loro funzione è quella di validatori di coerenza fisica locale.

Hanno appreso la relazione matematica che lega velocità (SOG) e rotta (COG) a una variazione di coordinate (ΔLat, ΔLon), imparando le leggi dell'inerzia in un vuoto spaziale, senza contesto geografico.

Cosa l'IA Rileva

Incoerenza SOG/COG vs Spostamento

Cosa l'IA Ignora

Traiettorie geograficamente impossibili

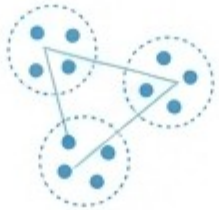
La protezione contro attacchi geograficamente assurdi ma fisicamente coerenti richiede un approccio diverso.

Sviluppi Futuri

1. Ancoraggio Geografico: Superare la 'Cecità Spaziale'

L'evoluzione del modello in un **Conditional Autoencoder** per risolvere il suo limite intrinseco. Vincolando l'output alla posizione iniziale, il sistema impara non solo la coerenza cinematica, ma anche la consapevolezza geografica assoluta.

Obiettivo: Passare da un validatore di moto differenziale a un sistema di monitoraggio geografico completo.



2. Intelligenza Contestuale: Soglie di Allarme Dinamiche

Abbandonare la soglia di allarme statica in favore di un sistema adattivo. Attraverso tecniche di **Clustering**, il sistema identificherà lo stato operativo della nave (es. navigazione, manovra) e applicherà soglie specifiche per ogni contesto.

Obiettivo: Massimizzare la sensibilità agli attacchi più subdoli e ridurre drasticamente i falsi positivi.

3. Diagnosi Azionabile: Dall'Errore alla Causa

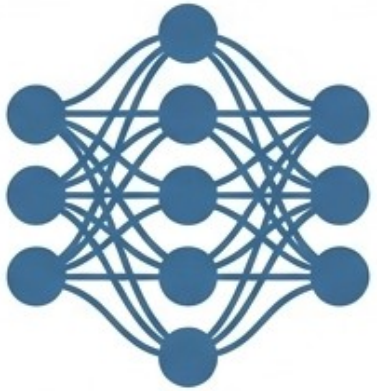
Trasformare il sistema da una 'Black Box' a uno strumento di supporto decisionale. Invece di un errore aggregato, il sistema fornirà un **vettore di errore dettagliato** per ogni feature (posizione, velocità, rotta).

Obiettivo: Offrire una diagnosi immediata e chiara sulla natura dell'anomalia, permettendo una risposta rapida ed efficace.

Conclusioni

1. Efficacia del Deep Learning

E’ stata dimostrata l’eccellente capacità dei modelli AutoEncoder (LSTM & LNN) nel rilevare anomalie che violano la coerenza cinematica, raggiungendo il **100% di detection rate** su attacchi macroscopici (Teleport, Speed Spoofing) e progressivi (Silent Drift).



2. Superiorità delle LNN

Le **Liquid Neural Networks** si sono dimostrate l’architettura superiore. Grazie alla modellazione a tempo continuo, presentano una **minore varianza sul traffico normale.** Questo si traduce in una maggior stabilità e nella possibilità di calibrare soglie di allarme più precise.

3. La Prossima Frontiera

La “cecità spaziale” è la principale limitazione emersa. Il futuro della difesa cyber-marittima basata su AI risiede nell’**ancorare l’analisi cinematica al contesto geografico reale,** evolvendo da validatori di moto a sistemi di monitoraggio **pienamente consapevoli.**

Grazie per l’attenzione

Lavoro svolto da:

Andrea La Rocca - andrea.larocca4@studio.unibo.it
Cristina Tomaciello - cristina.tomaciello@studio.unibo.it