



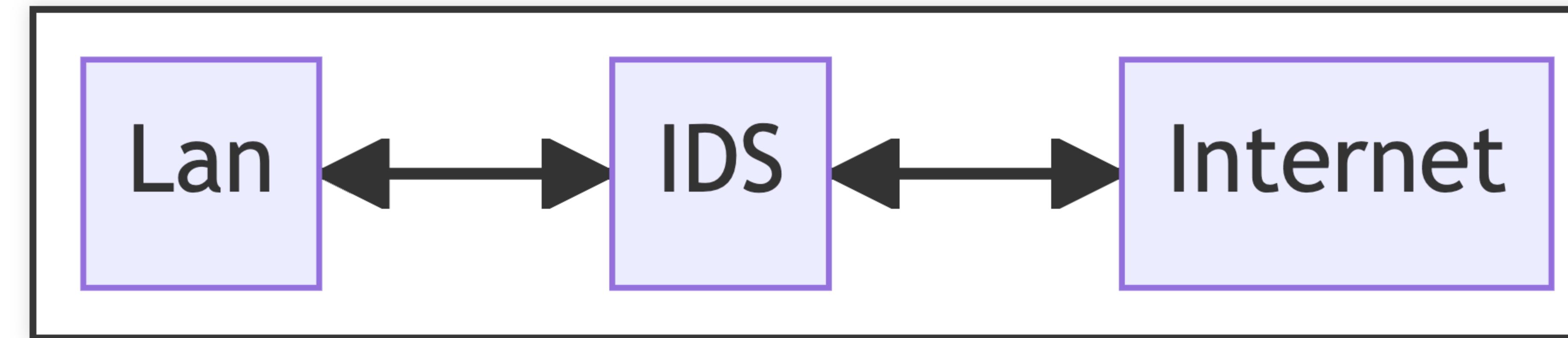
# **Utilizzo di Soft-Brownian-Offset per la generazione di attacchi ai fini dell'addestramento di rilevatori di intrusioni**

**Guglielmo Bartelloni**  
**Relatore: Andrea Ceccarelli**



# INTRUSION DETECTION SYSTEMS (IDS)

Applicazione che monitora continuamente la rete per identificare attività malevole.





Nell'ultimo decennio si è iniziato ad utilizzare algoritmi di  
**Machine Learning** per gli IDS.



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# MACHINE LEARNING

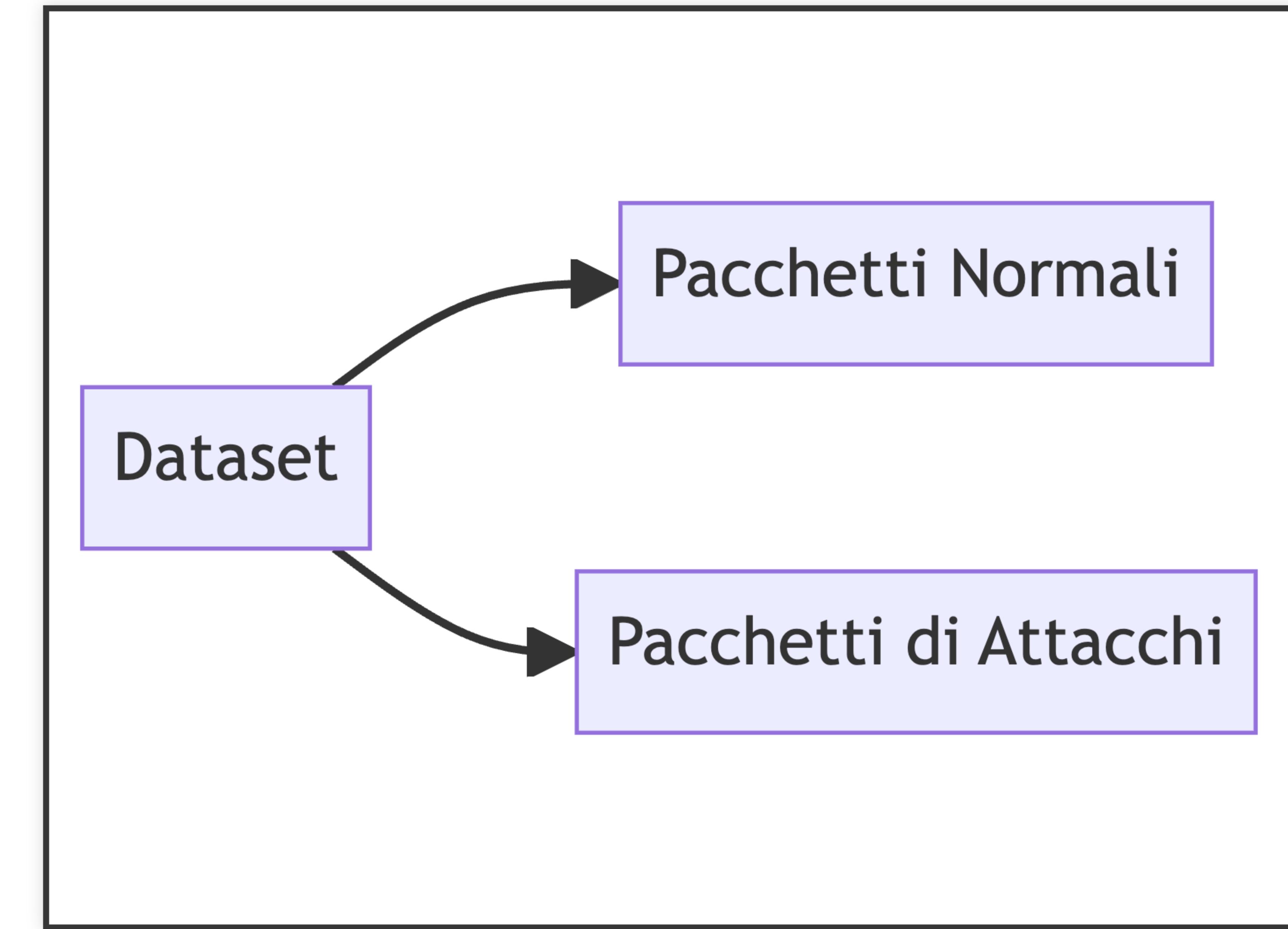
La branca dell'**Intelligenza Artificiale** che sviluppa modelli per permettere alle macchine di imparare dai dati.



- I modelli di Machine Learning sono sensibili ai dati di addestramento.
- I dataset non contengono tutti i possibili attacchi in una rete



# DATASET DI UN IDS





# POSSIBILE METODO MIGLIORAMENTO

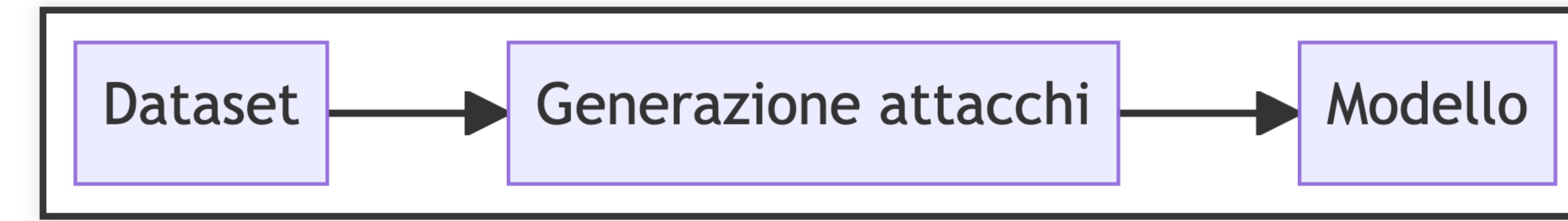
**Generare** nuovi dati a partire da quelli già esistenti per  
migliorare i modelli di Machine Learning

- Soft-Brownian-Offset



# SCOPO DELLA TESI

Cercare di migliorare un IDS utilizzando Soft-Brownian-Offset





UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# **SOFT BROWNIAN OFFSET (2021)**

Algoritmo di generazione di dati creato inizialmente per la  
generazione di eventi anomali.



# DATASET UTILIZZATI

- Adfanet
- CICIDS18



## APPROCCI DI GENERAZIONE

Generazione a partire dalla tipologia di dati:

- Solo pacchetti normali
- Solo pacchetti attacchi
- Dataset completo



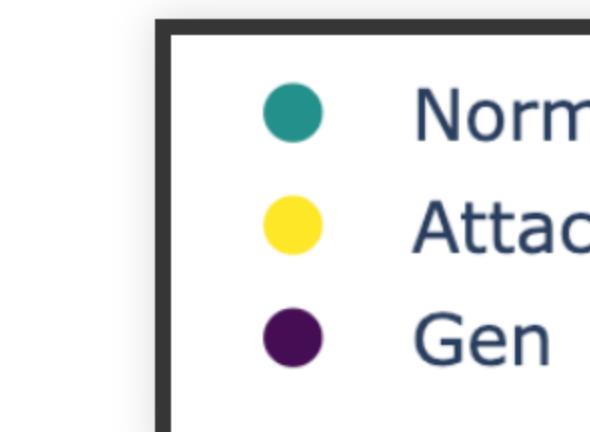
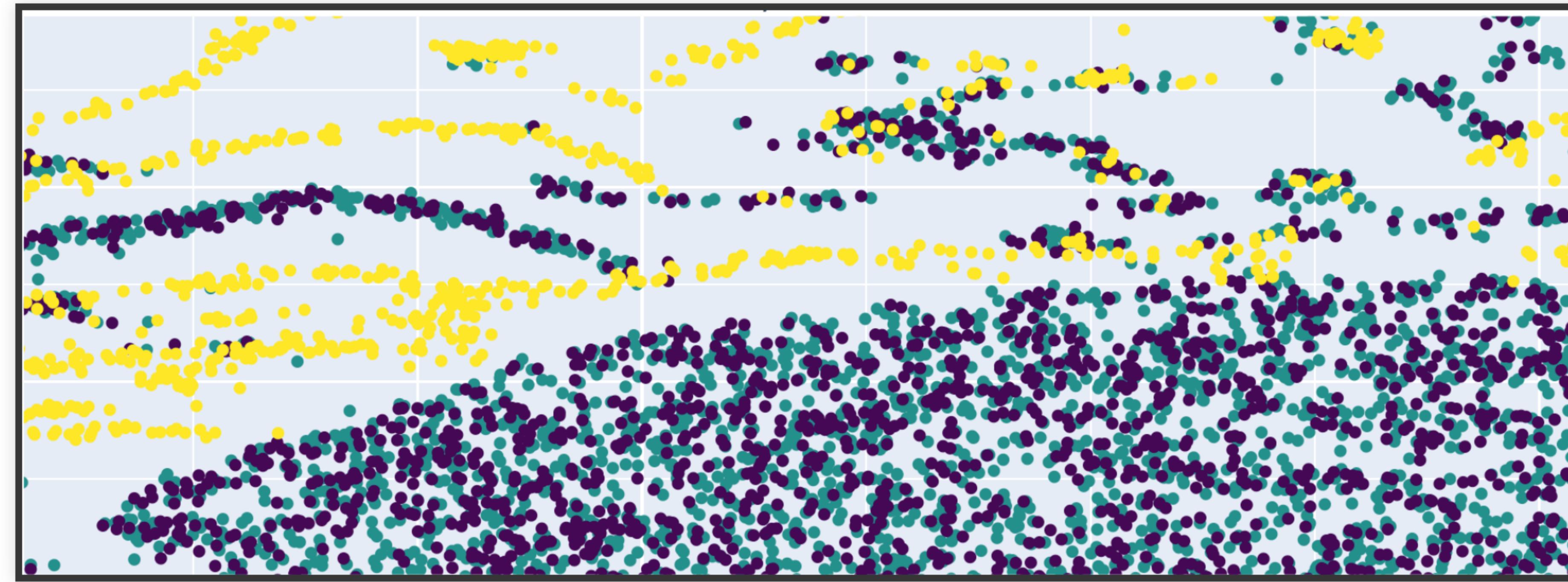
UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# GRAFICI

Per una valutazione qualitativa degli approcci di generazione.

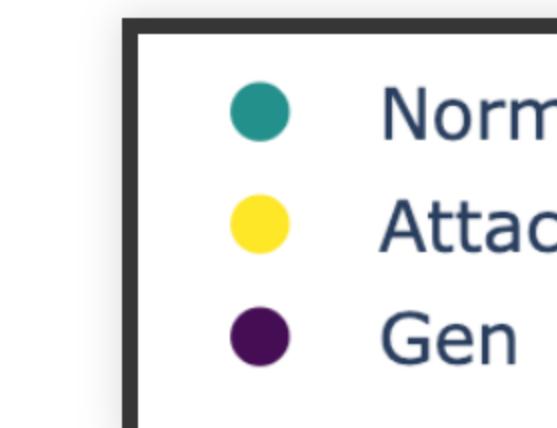
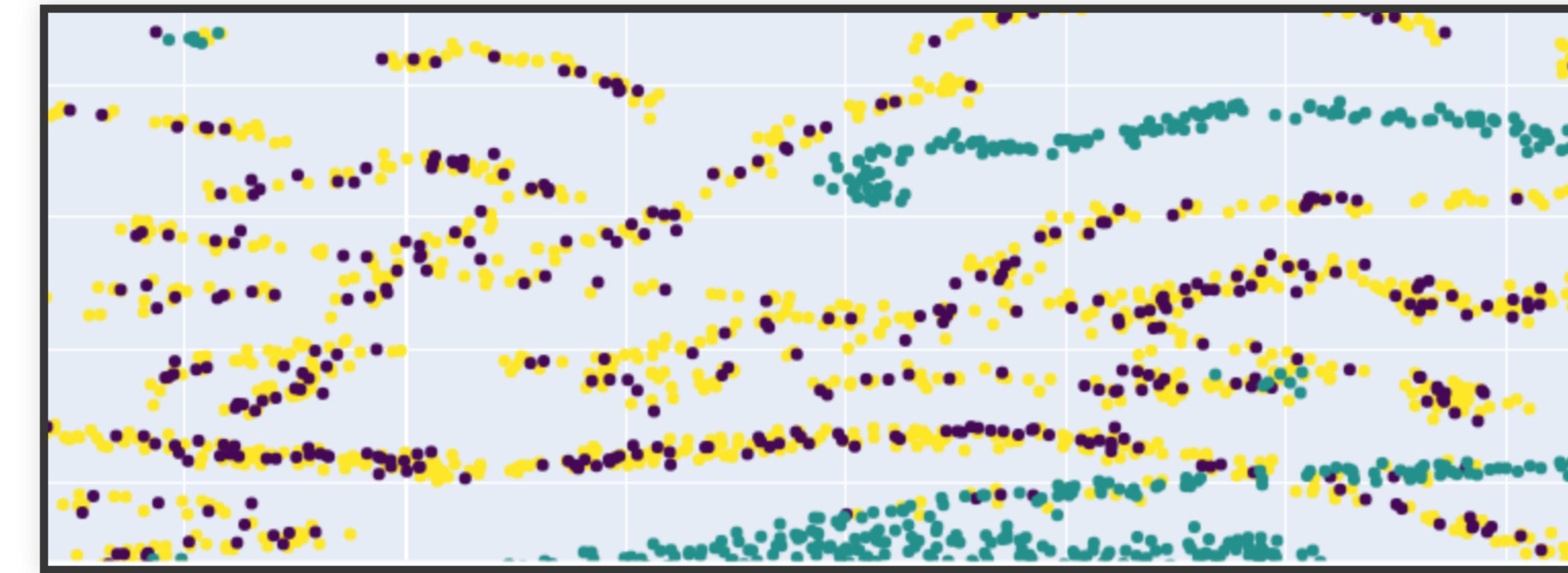


# Generazione a partire da pacchetti normali





# Generazione a partire da pacchetti di attacchi





# APPROCCI DI ADDESTRAMENTO DEL MODELLO

Addestramento di XGBoost usando:

- Dataset completo + Dati sintetici
- Solo pacchetti normali + Dati sintetici
- Solo dataset (senza dati generati)



## ADFANET (MCC)

	<b>Pacchetti normali + Gen</b>	<b>Dataset Completo + Gen</b>	<b>Solo Dataset</b>
Gen Normali	0.3337	0.99839	0.99842
Gen Attacchi	0.4404	0.99865	0.99842
Gen Completo	0.3452	0.99854	0.99842



## ADFANET (MCC)

	Pacchetti normali + Gen	Dataset Completo + Gen	Solo Dataset
Gen Normali	0.3337	0.99839	0.99842
Gen Attacchi	0.4404	<b>0.99865</b>	0.99842
Gen Completo	0.3452	0.99854	0.99842



## CICIDS (MCC)

	<b>Pacchetti normali + Gen</b>	<b>Dataset Completo + Gen</b>	<b>Solo Dataset</b>
Gen Normali	-0.1153	0.92772	0.93596
Gen Attacchi	-0.1366	0.93428	0.93596
Gen Completo	-0.1206	0.92493	0.93596



## CICIDS (MCC)

	Pacchetti normali + Gen	Dataset Completo + Gen	Solo Dataset
Gen Normali	-0.1153	0.92772	<b>0.93596</b>
Gen Attacchi	-0.1366	0.93428	<b>0.93596</b>
Gen Completo	-0.1206	0.92493	<b>0.93596</b>



## CONCLUSIONI

Soft Brownian Offset è efficace nel caso di dataset **semplici**.

In dataset **complessi** invece l'algoritmo non presenta miglioramenti.

In quest'ultimo caso è necessario rivolgersi ad algoritmi differenti.



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**GRAZIE PER L'ATTENZIONE**