Studio, progettazione e implementazione di un algoritmo genetico per l'individuazione di problemi di privacy in sistemi IoT

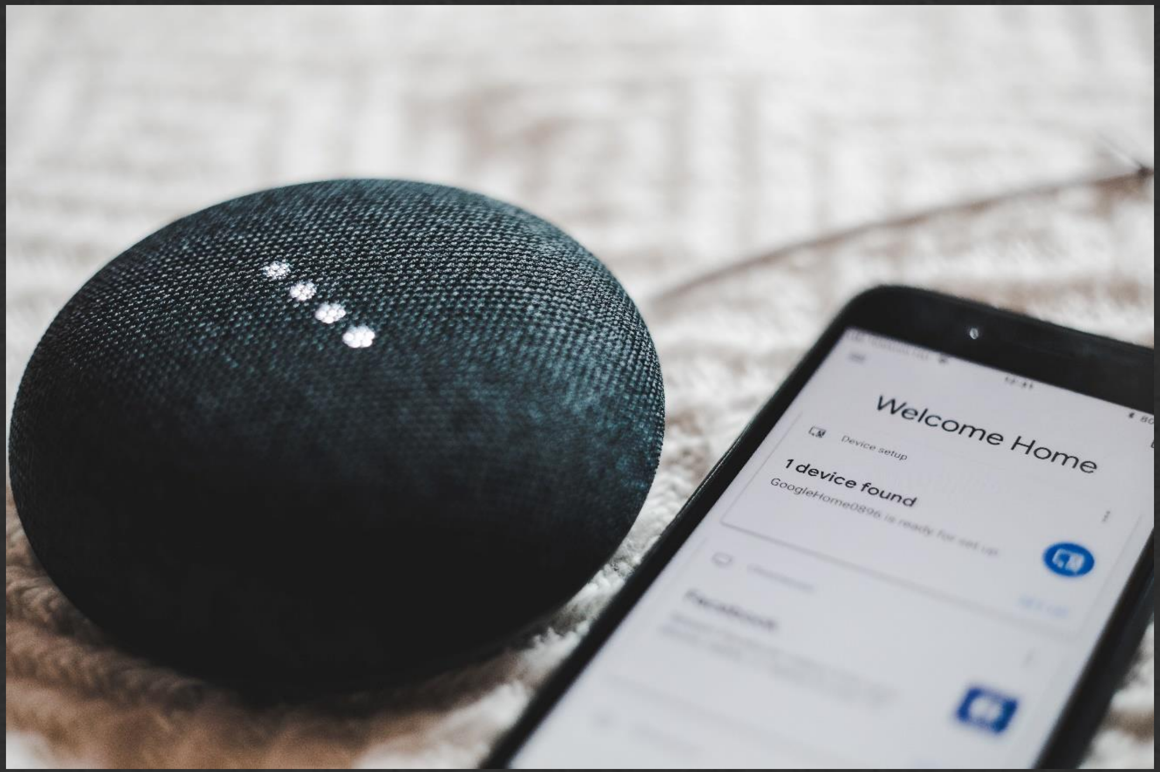
Prof. : Fabio Palomba

Candidato: Davide La Gamba
Mat: 0512106292

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

PDF TESI

Introduzione & Background



d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Introduzione & Background

10 miliardi di dispositivi IoT nel 2021

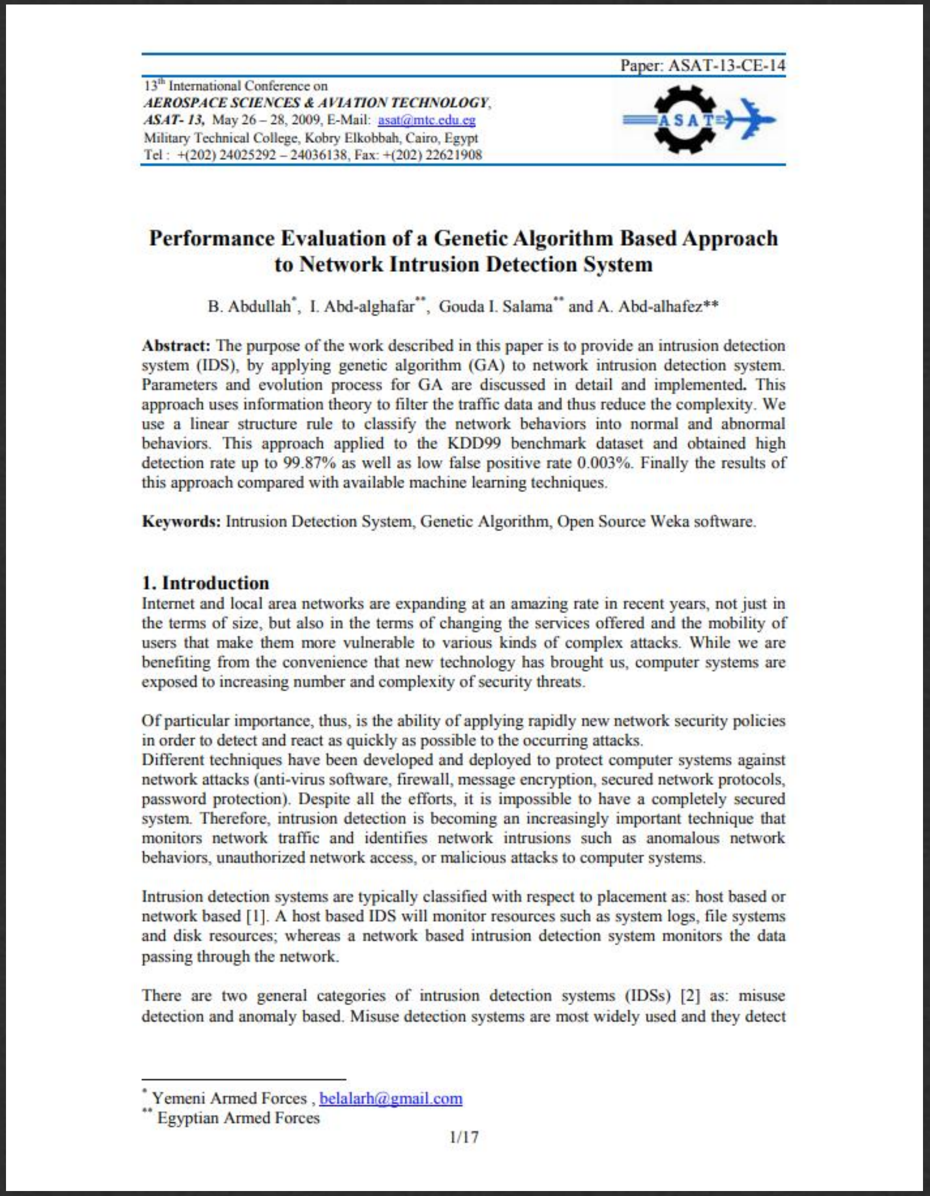
25,4 miliardi di dispositivi nel 2030

1.1 trilioni di dollari spesi nel 2023 per l'IoT

73.1 ZB (1 Zettabyte = circa 1.000.000.000 Terabyte) di dati generati nel 2025

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Introduzione & Background


Paper: ASAT-13-CE-14

13th International Conference on
AEROSPACE SCIENCES & AVIATION TECHNOLOGY,
ASAT- 13, May 26 – 28, 2009, E-Mail: asat@mtc.edu.eg
Military Technical College, Kobry Elkobbah, Cairo, Egypt
Tel : +(202) 24025292 – 24036138, Fax: +(202) 22621908

Performance Evaluation of a Genetic Algorithm Based Approach to Network Intrusion Detection System

B. Abdullah*, I. Abd-alghafar**, Gouda I. Salama** and A. Abd-alhafez**

Abstract: The purpose of the work described in this paper is to provide an intrusion detection system (IDS), by applying genetic algorithm (GA) to network intrusion detection system. Parameters and evolution process for GA are discussed in detail and implemented. This approach uses information theory to filter the traffic data and thus reduce the complexity. We use a linear structure rule to classify the network behaviors into normal and abnormal behaviors. This approach applied to the KDD99 benchmark dataset and obtained high detection rate up to 99.87% as well as low false positive rate 0.003%. Finally the results of this approach compared with available machine learning techniques.

Keywords: Intrusion Detection System, Genetic Algorithm, Open Source Weka software.

1. Introduction
Internet and local area networks are expanding at an amazing rate in recent years, not just in the terms of size, but also in the terms of changing the services offered and the mobility of users that make them more vulnerable to various kinds of complex attacks. While we are benefiting from the convenience that new technology has brought us, computer systems are exposed to increasing number and complexity of security threats.

Of particular importance, thus, is the ability of applying rapidly new network security policies in order to detect and react as quickly as possible to the occurring attacks.
Different techniques have been developed and deployed to protect computer systems against network attacks (anti-virus software, firewall, message encryption, secured network protocols, password protection). Despite all the efforts, it is impossible to have a completely secured system. Therefore, intrusion detection is becoming an increasingly important technique that monitors network traffic and identifies network intrusions such as anomalous network behaviors, unauthorized network access, or malicious attacks to computer systems.

Intrusion detection systems are typically classified with respect to placement as: host based or network based [1]. A host based IDS will monitor resources such as system logs, file systems and disk resources; whereas a network based intrusion detection system monitors the data passing through the network.

There are two general categories of intrusion detection systems (IDSs) [2] as: misuse detection and anomaly based. Misuse detection systems are most widely used and they detect

* Yemeni Armed Forces , belalarh@gmail.com
** Egyptian Armed Forces

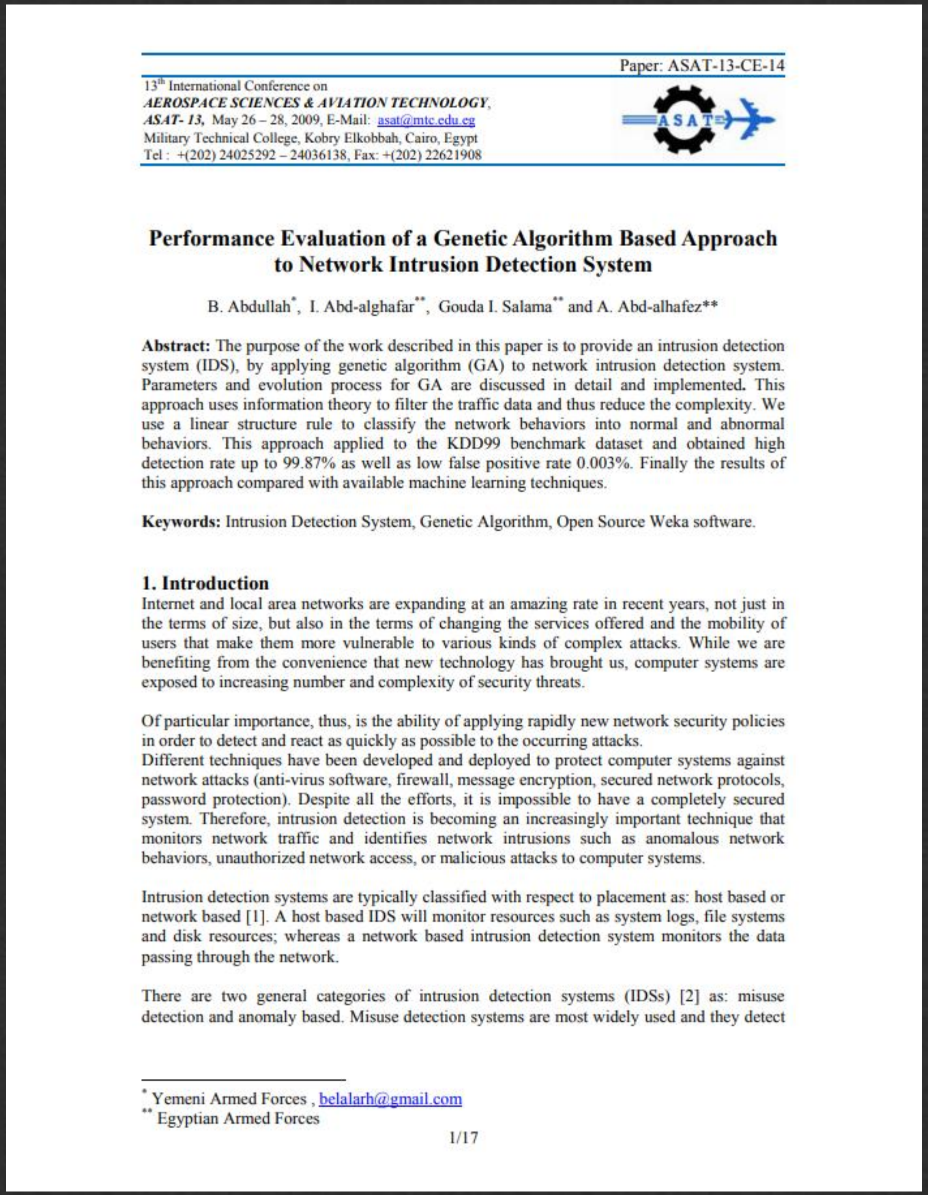
1/17

Lavoro esistente di Al-fuhaidi et al. con algoritmi genetici impiegati per problemi di sicurezza nell’IoT

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Introduzione & Background

sesa lab
SOFTWARE ENGINEERING SALERNO



Paper: ASAT-13-CE-14

13th International Conference on
AEROSPACE SCIENCES & AVIATION TECHNOLOGY,
ASAT- 13, May 26 – 28, 2009, E-Mail: asat@mtc.edu.eg
Military Technical College, Kobry Elkobbah, Cairo, Egypt
Tel : +(202) 24025292 – 24036138, Fax: +(202) 22621908

Performance Evaluation of a Genetic Algorithm Based Approach to Network Intrusion Detection System

B. Abdullah*, I. Abd-alghafar**, Gouda I. Salama** and A. Abd-alhafez**

Abstract: The purpose of the work described in this paper is to provide an intrusion detection system (IDS), by applying genetic algorithm (GA) to network intrusion detection system. Parameters and evolution process for GA are discussed in detail and implemented. This approach uses information theory to filter the traffic data and thus reduce the complexity. We use a linear structure rule to classify the network behaviors into normal and abnormal behaviors. This approach applied to the KDD99 benchmark dataset and obtained high detection rate up to 99.87% as well as low false positive rate 0.003%. Finally the results of this approach compared with available machine learning techniques.

Keywords: Intrusion Detection System, Genetic Algorithm, Open Source Weka software.

1. Introduction

Internet and local area networks are expanding at an amazing rate in recent years, not just in the terms of size, but also in the terms of changing the services offered and the mobility of users that make them more vulnerable to various kinds of complex attacks. While we are benefiting from the convenience that new technology has brought us, computer systems are exposed to increasing number and complexity of security threats.

Of particular importance, thus, is the ability of applying rapidly new network security policies in order to detect and react as quickly as possible to the occurring attacks.
Different techniques have been developed and deployed to protect computer systems against network attacks (anti-virus software, firewall, message encryption, secured network protocols, password protection). Despite all the efforts, it is impossible to have a completely secured system. Therefore, intrusion detection is becoming an increasingly important technique that monitors network traffic and identifies network intrusions such as anomalous network behaviors, unauthorized network access, or malicious attacks to computer systems.

Intrusion detection systems are typically classified with respect to placement as: host based or network based [1]. A host based IDS will monitor resources such as system logs, file systems and disk resources; whereas a network based intrusion detection system monitors the data passing through the network.

There are two general categories of intrusion detection systems (IDSs) [2] as: misuse detection and anomaly based. Misuse detection systems are most widely used and they detect

* Yemeni Armed Forces , belalarh@gmail.com
** Egyptian Armed Forces

Lavoro esistente di Al-fuhaidi et al. con algoritmi genetici impiegati per problemi di sicurezza nell’IoT

Obiettivo: estendere il lavoro in esame, aumentando la capacità di individuazione di problemi di privacy

Risultati: regole di classificazione con Detection Rate migliorato e maggior varietà di attacchi identificati

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato

Individuazione dataset (KDDCUP99)
e lavoro di riferimento (Al-fuhaidi et al.)

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell'approccio utilizzato

Individuazione dataset (KDDCUP99)
e lavoro di riferimento (Al-fuhaidi et al.)

Progettazione algoritmo genetico e obiettivi

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato

Individuazione dataset (KDDCUP99)
e lavoro di riferimento (Al-fuhaidi et al.)

Progettazione algoritmo genetico e obiettivi

Scelta delle codifiche

Descrizione dell'approccio utilizzato

Individuazione dataset (KDDCUP99)
e lavoro di riferimento (Al-fuhaidi et al.)

Progettazione algoritmo genetico e obiettivi

Scelta delle codifiche

Scelta degli altri parametri e degli operatori

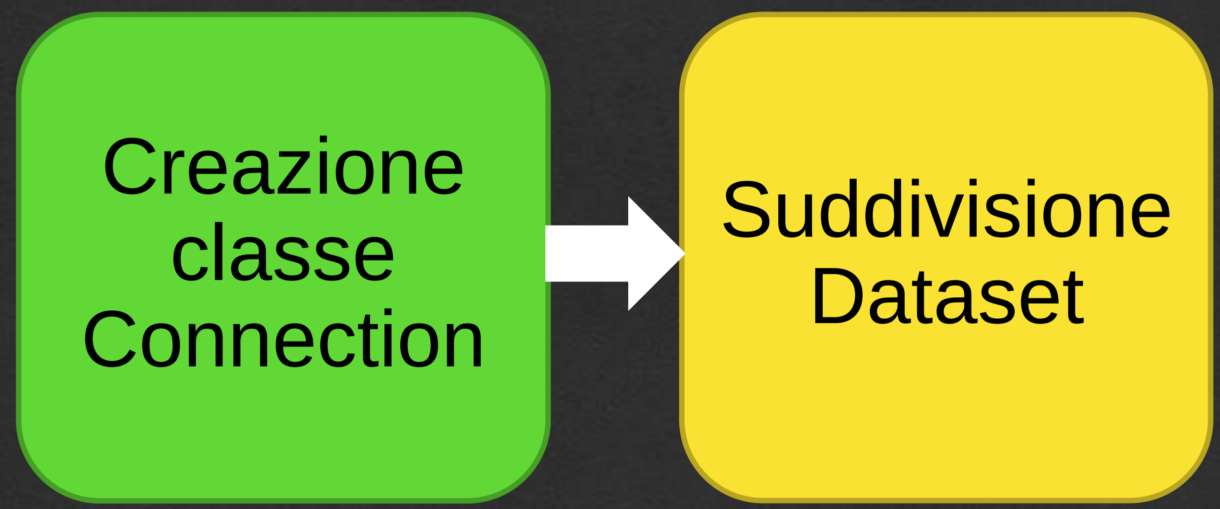
d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato

Creazione classe Connection

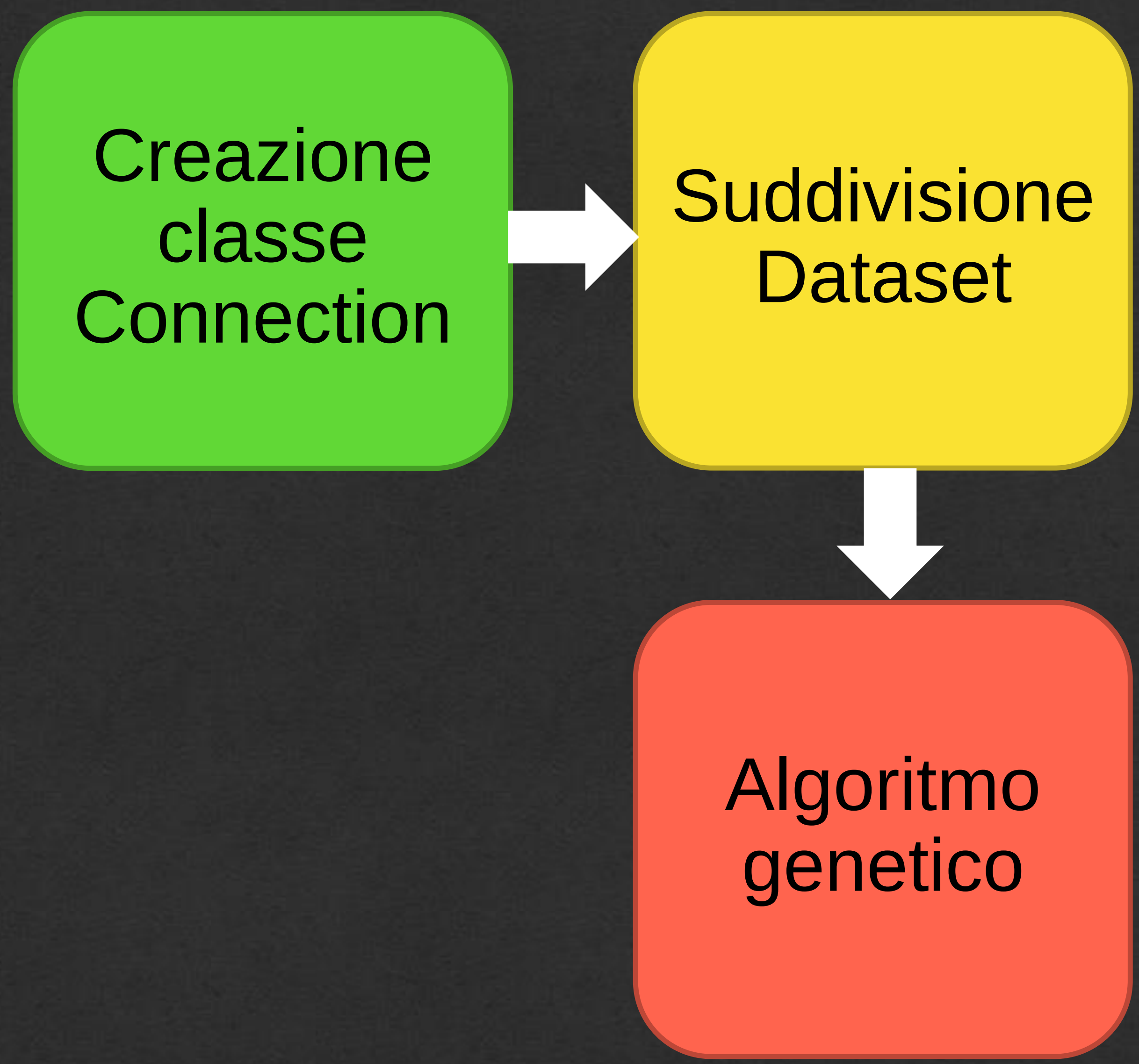
d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato



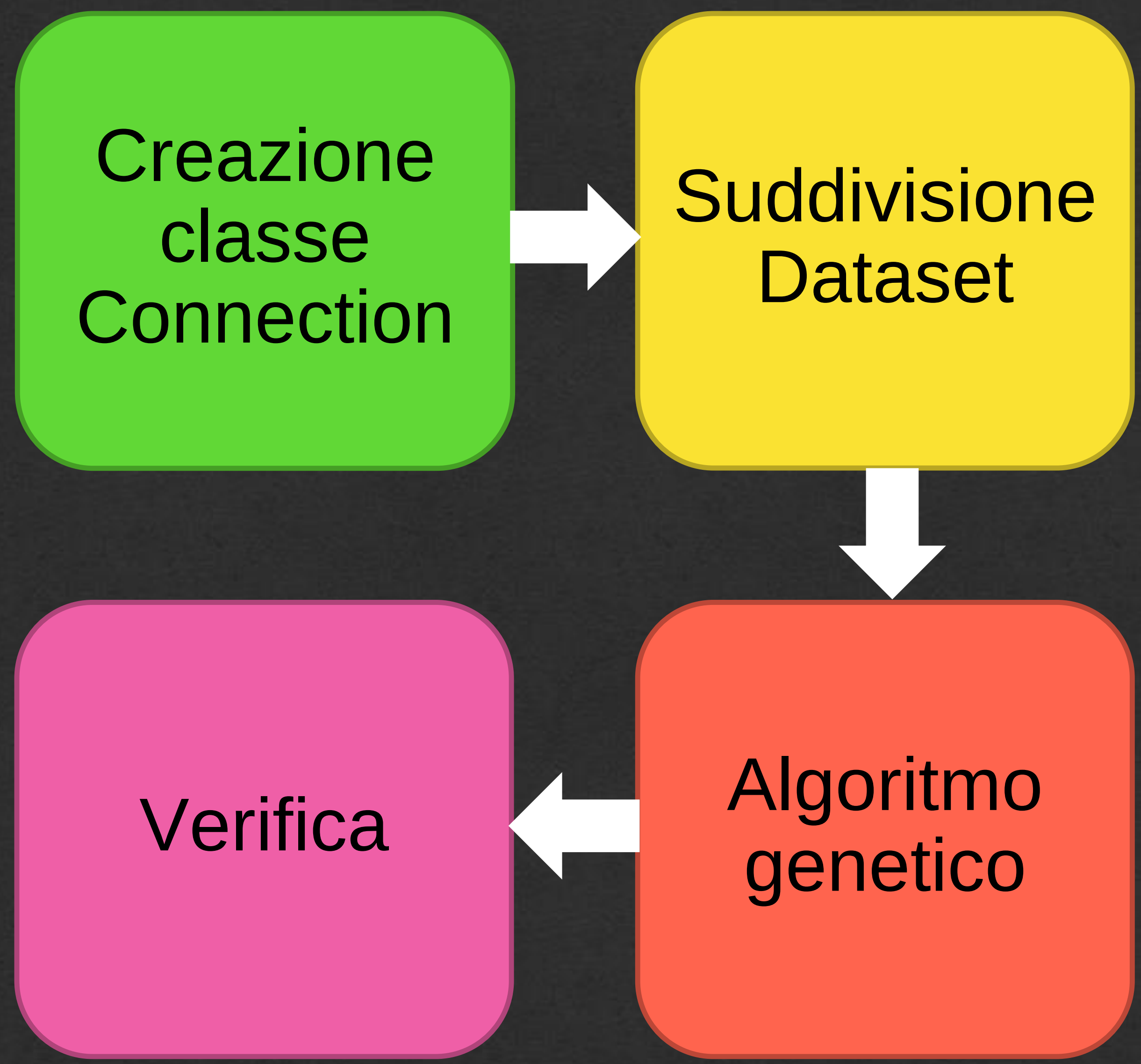
d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato



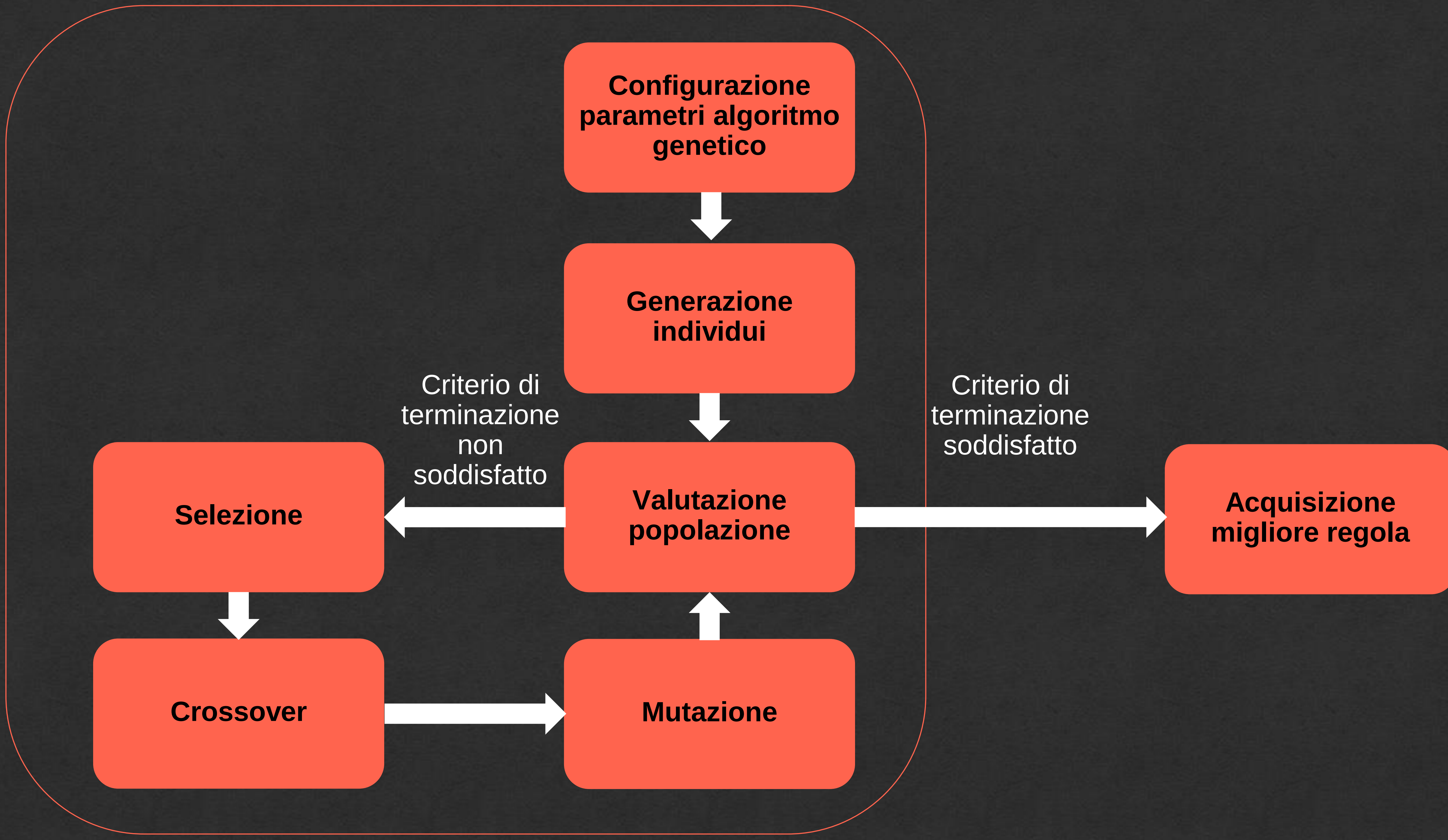
d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato



d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato



d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato

Individuo = serie di geni (valori numerici) raccolti in cromosomi con valori di min e max

Rappresentano sia i valori dei parametri da analizzare per ogni connessione sia i versi delle disuguaglianze

if (duration <= 5855 AND protocolType == icmp AND service == ecr_i AND flag == SF AND … AND sameSrvRate <= 1.00 AND diffSrvRate <= 0.90 AND srvDiffHostRate <=1.00)
{attacco trovato}

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell'approccio utilizzato

Funzione di fitness = $\frac{ab}{A} - \frac{a}{B}$

$ab = attacchi\ correttamente\ individuati$

$a = connessioni\ erroneamente\ individuate$

$A = totale\ degli\ attacchi$

$B = totale\ delle\ connessioni\ normali$

Algoritmo di selezione: Elite Selector

Algoritmo di crossover: Single Point Crossover (probabilità 90%)

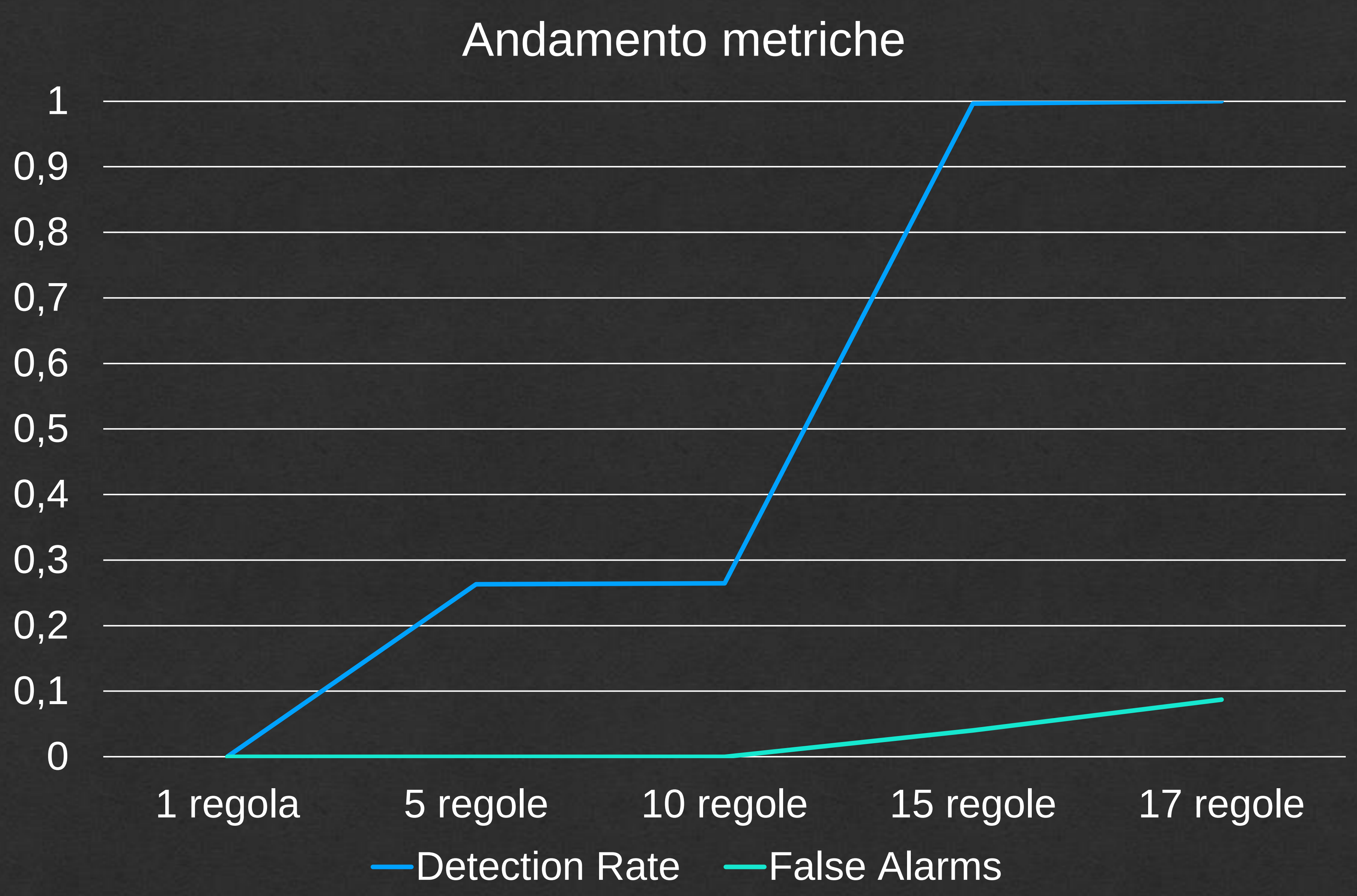
Probabilità di mutazione: 10%

1000 (o 5000) individui

1000 generazioni

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Sperimentazione e Risultati



Numero di regole	17
Detection Rate	0.99989
F-Measure	0.98925
Precision	0.97884
Accuracy	0.98259
Specificity	0.91288
MCC	0.94500
False Alarms	0.08711

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Sperimentazione e Risultati

Metrica	Algoritmo Genetico Soluzione 100%	Naive Bayes	Random Forest
Detection Rate	**0.999**	0,992	1,000

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Sperimentazione e Risultati

Metrica	Algoritmo Genetico Soluzione 100%	Algoritmo Genetico Soluzione 10%	Algoritmo Genetico di Al-fuhaidi et al.
Detection Rate	**99.989%**	**99.938%**	99.87%
False Alarms	**08.711%**	**02.572%**	0.003%

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Pro:

- Flessibilità nella codifica degli individui
- Alta comprensibilità

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Pro:

- Flessibilità nella codifica degli individui
- Alta comprensibilità

Contro:

- Delicata fase di scelta dei parametri
- Prestazioni leggermente inferiori

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Implementazione generica da utilizzare in un contesto reale

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Implementazione generica da utilizzare in un contesto reale

Classificazione specifica degli attacchi

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Implementazione generica da utilizzare in un contesto reale

Classificazione specifica degli attacchi

Utilizzare funzione multi-obiettivo

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

Vi ringrazio per l’attenzione

QR Code Repository GitHub

QR Code PDF Tesi

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Conclusioni e Sviluppi futuri

https://www.flaticon.com/free-icons/code Code icons created by Freepik - Flaticon

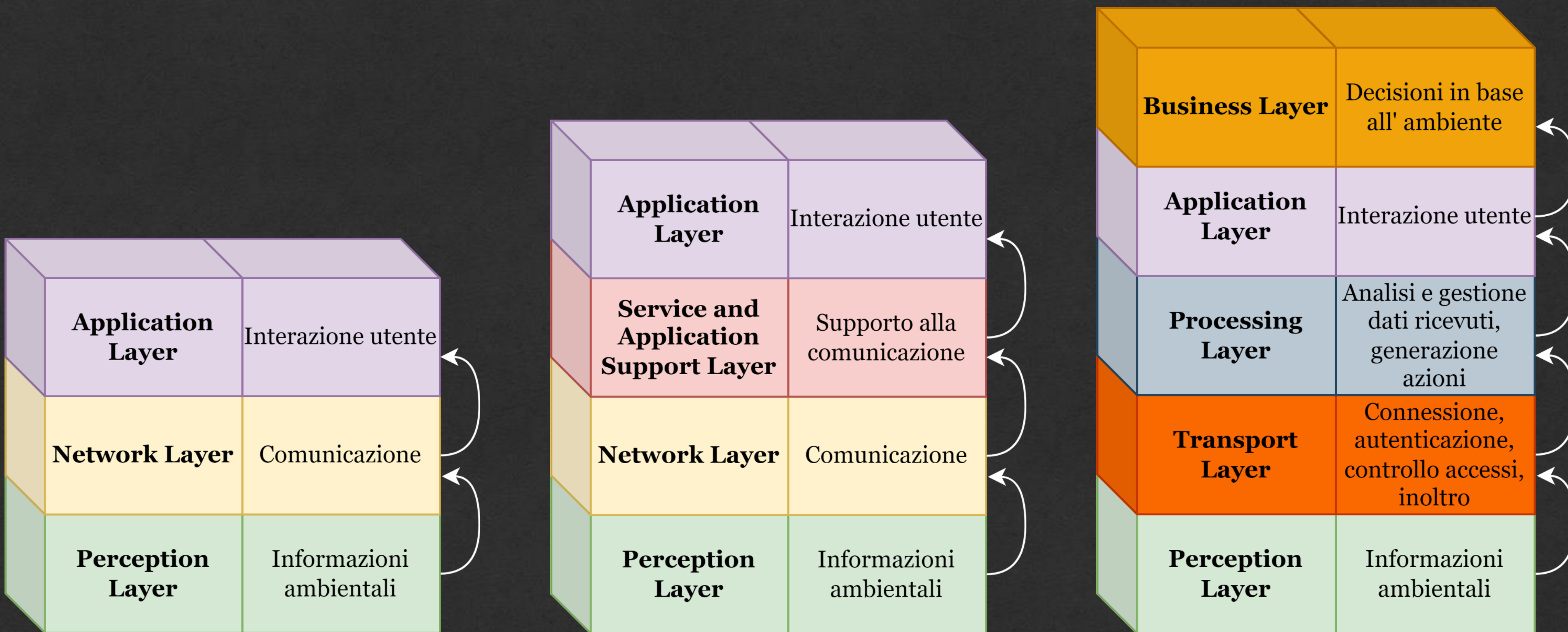
https://www.flaticon.com/free-icons/malware Malware icons created by Smashicons - Flaticon

https://www.flaticon.com/free-icons/world World icons created by Hilmy Abiyyu A. - Flaticon

https://www.flaticon.com/free-icons/gear Gear icons created by Freepik - Flaticon

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Introduzione & Background



Tre livelli | Quattro livelli | Cinque livelli

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

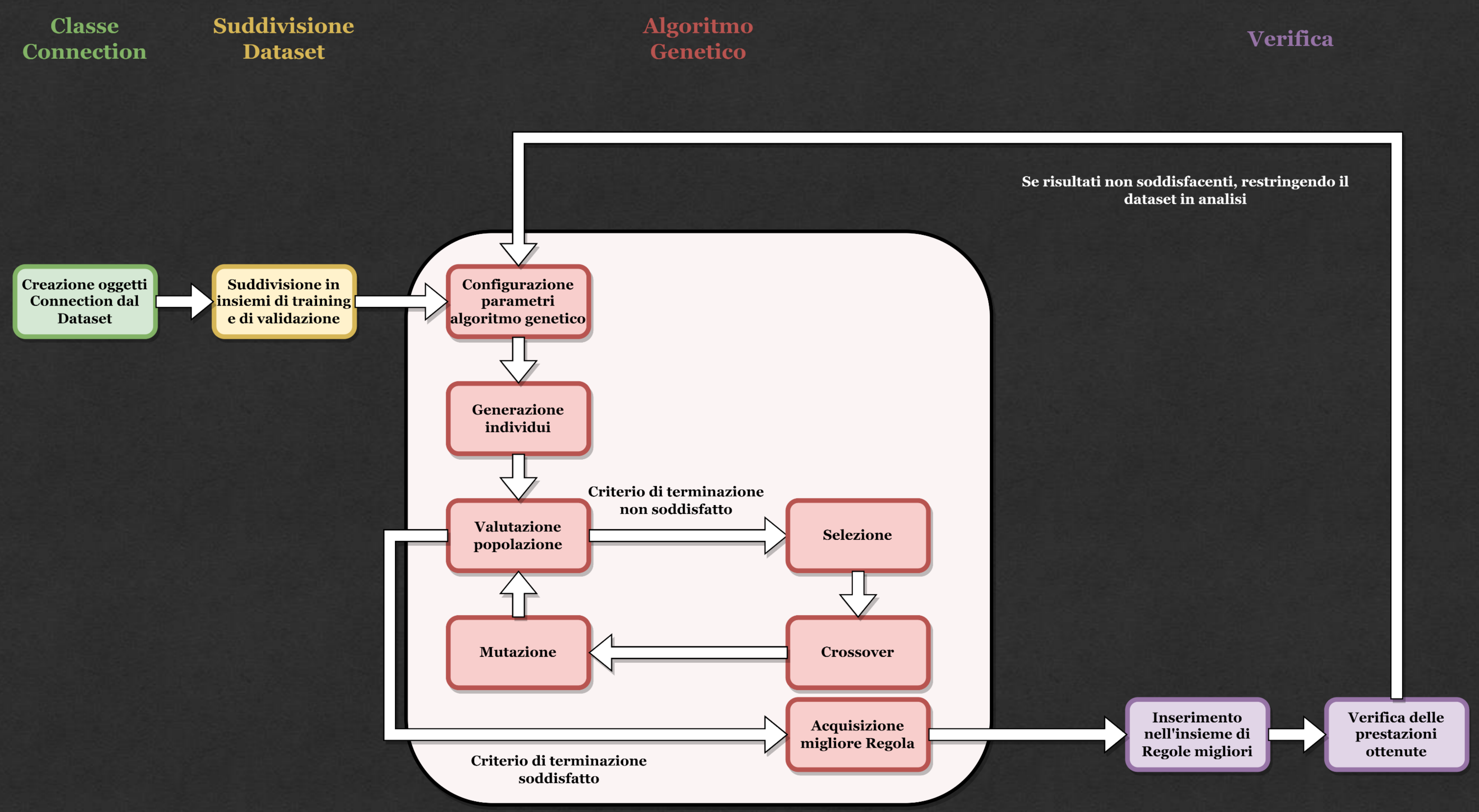
Introduzione & Background

SESA lab SOFTWARE ENGINEERING SALERNO

- Individuazione di vulnerabilità
- Evoluzione di malware
- Feature Selection
- Applicazioni in sistemi blockchain
- Individuazione di intrusioni

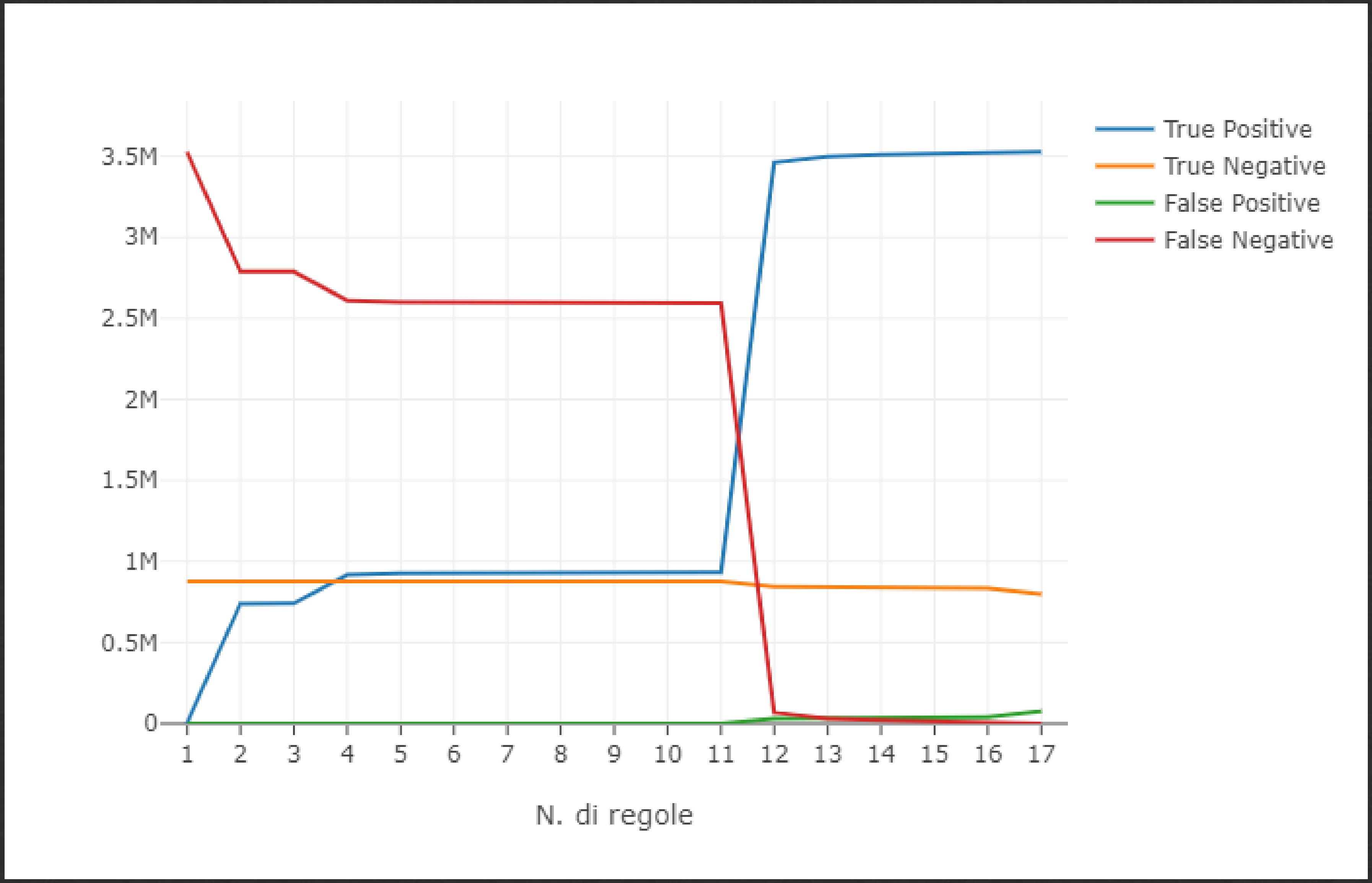
d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Descrizione dell’approccio utilizzato



d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba

Sperimentazione e Risultati



Numero di regole	17
True Positive	3528525
True Negative	799237
False Positive	76267
False Negative	382

d.lagamba@studenti.unisa.it
https://www.linkedin.com/in/davide-la-gamba-b93448244/
https://github.com/davide-lagamba