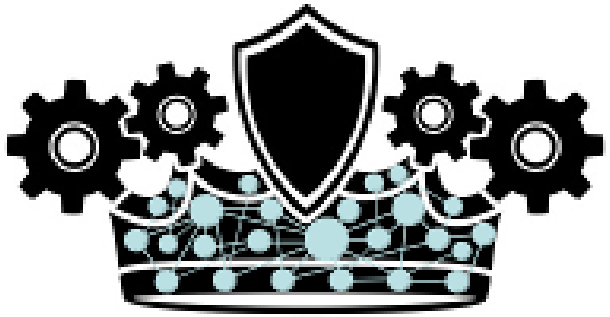
KEEP CALM
Kernel Engines Enable to Prevent Cyber Attacks with Learning Machines



Prevedere gli attacchi cibernetici, prima ancora che essi si verifichino osservando l'attività della rete e/o partendo da una serie storica che registra tali attività malevoli

LUISS Università Guido Carli · GRUPPOISC Information Sharing Company · sogei · CYBER 4.0 CYBERSECURITY COMPETENCE CENTER

Una corposa bibliografia

Abdlhamed M., Kifayat K., Shi Q., Hurst W. (2017) Intrusion Prediction Systems. In: Alsmadi I., Karabatis G., Aleroud A. (eds) Information Fusion for Cyber-Security Analytics. Studies in Computational Intelligence, vol 691. Springer, Cham

Adenso-Díaz B., Laguna M. (2006). Fine-tuning of algorithms using fractional experimental designs and local search. Operations Research, vol. 54, no. 1, pp. 99-114.

Ballarin, A., Gervasi, S., Bacchetti, S., Capponi, U., Costi, S., Gervasi Vidal, K. A., Moore, P. B., Nardone, C., Passali, G., Sagone, F., Signori, M., Vollera, F. (2010). On the Forecasting Abilities of a Time Varying Auto-Adapting Algorithm, Neural Parallel and Scientific Computation, 18, December 2010.

Barreno, M., Nelson, B., Joseph, A., and Tygar, J. (2010) "The security of machine learning," Mach. Learn., vol. 81, pp. 121–148.

Biggio B., Corona I., Nelson B., Rubinstein B. I. P., Maiorca D., Fumera G., Giacinto G., Roli F. (2014). Security Evaluation of Support Vector Machines in Adversarial Environments. In Ma and Guo (Eds.): Support Vector Machines Applications, Springer International Publishing, p. 105–153.

Biggio, B., Fumera, G., Roli, F. (2014), Security evaluation of pattern classifiers under attack. IEEE Transactions on Knowledge and Data Engineering, 26(4), pagg 984-996.

Breiman, L. (2001). "Random Forests". Machine Learning. vol 45, pp. 5–32. doi:10.1023/A:1010933404324.

Buczak A. L., Guven E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. Communications Surveys & Tutorials, IEEE, vol. 18, no. 2, pp. 1153–1176.

Chandola, V., Banerjee, A., and Kumar, V.. (2009). Anomaly detection: A survey. ACM Comput. Surv. 41, 3, Article 15 (July 2009), 58 pages. DOI=http://dx.doi.org/10.1145/1541880.1541882.

Corona I., Giacinto G., Mazzariello C., Roli F., Sansone C. (2009). Information fusion for computer security: State of the art and open issues. Information Fusion, Vol. 10, p. 274-284.

Corona I, Giacinto G, Roli F. (2013) Adversarial attacks against intrusion detection systems: Taxonomy, solutions and open issues. Information Sciences, vol. 239, p. 201-225.

Dash M., Liu H. (1997). Feature Selection for Classification. Intelligent Data Analysis, vol. 1, pp. 131- 156.

Demontis A., Melis M., Biggio B., Maiorca D., Arp D., Rieck K., Corona I., Giacinto G., Roli F. (2017). Yes, Machine Learning Can Be More Secure! A Case Study on Android Malware Detection. IEEE Trans. on Dependable and Secure Computing, (Early Access, published 2 May 2017).

European Foresight Cyber Security Meeting 2016: Public, private academic recommendations to the European Commissions about Internet of Things and Harmonization of duties of care.

Fang, X., Xu, M., Xu, S., Zhao, P.: A deep learning framework for predicting cyber attacks rates; EURASIP Journal on Information Security volume 5, 2019.

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. (1996). From Data Mining to Knowledge Discovery: An Overview. In Advances in Knowledge Discovery & Data Mining, Fayyad, U.M.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R., Eds. AAAI/MIT Press, Cambridge, Massachusetts.

Franceschetti, G. Homeland Security: Threats, Countermeasures, and Privacy Issues. Norwood, MA: Artech House, 2011.

Fumera, G. & Roli, F. (2005). A theoretical and experimental analysis of linear combiners for multiple classifier systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 27, pp. 942-956.

Giacinto G., Perdisci R., Del Rio M., Roli F. (2008). Intrusion detection in computer networks by a modular ensemble of one-class classifiers. Information Fusion, Vol. 9, p. 69-82.

Giacinto G., Roli F., Didaci L. (2003). Fusion of multiple classifiers for intrusion detection in computer networks. Pattern recognition letters, vol. 24, p. 1795-1803.

Goyal, P., Tozammel, H. KSM, Deb, A., Tavabi, N., Bartley, N., Abeliuk, A., Ferrara, E. Lerman, K.: Discovering Signals from Web Sources to Predict Cyber Attack. arXiv preprint arXiv:1806.03342.

Han, J.; Kamber, M. (2001): Data Mining Concepts and Techniques, Morgan Kaufmann Publishers.

Una corposa bibliografia

Hodo, E., Bellekens, X., Hamilton, A., Tachtatzis, C., & Atkinson, R. (2017). Shallow and Deep Networks Intrusion Detection System: A Taxonomy and Survey. arXiv preprint arXiv:1701.02145.

Jones M., Kotsalis G., Shamma J. S. (2013). Cyber-attack forecast modeling and complexity reduction using a game-theoretic framework. In Control of Cyber-Physical Systems (pp. 65-84). Springer International Publishing.

Joseph A.D., Laskov P., Roli F., Tygar J.D., Nelson B. (Eds.) (2013). Machine Learning Methods for Computer Security. Dagstuhl Perspectives Workshop 12371, 2(9), pp. 109—130.

Lee, Y. S., & Tong, L. I. (2011). Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. Knowledge-Based Systems, 24(1), 66-72.

Liu, Y., Sarabi, A., Zhang, J., Naghizadeh, P., Karir, M., Bailey, M., & Liu, M. (2015, August). Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents. In USENIX Security Symposium (pp. 1009-1024).

Mahmood, T., & Afzal, U. (2013, December). Security Analytics: Big Data Analytics for cybersecurity: A review of trends, techniques and tools. In Information assurance (ncia), 2013 2nd national conference on (pp. 129-134). IEEE.

Marchesi, M., Mannaro, K., Uras, S. & Locci, M. (2007). Distributed Scrum in research project management. In Agile Processes in Software Engineering and Extreme Programming, pp. 240-244, Lecture Notes in Computer Science vol 4535, Springer.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. Harvard business review, 90(10), 60-68.

Okutan, A., Yang, S. J., McConky, K.: Forecasting Cyber Attacks with Imbalanced Data Sets and Different Time Granularities. arXiv preprint arXiv:1803.09560.

Perdisci R, Corona I, Giacinto G (2012). Early Detection of Malicious Flux Networks via Large-Scale Passive DNS Traffic Analysis. IEEE Transactions on Dependable and Secure Computing, vol. 9, p. 714-726.

Pluribus One Attack prophecy.. https://www.pluribus-one.it/what/products/attack-prophecy

E. Pontes, A. E. Guelfi, S. T. Kofuji, A. A. A. Silva and A. E. Guelfi, (2011) "Applying multi-correlation for improving forecasting in cyber security," 2011 Sixth International Conference on Digital Information Management, Melbourn, QLD,, pp. 179-186. doi: 10.1109/ICDIM.2011.6093323

Soska, K., & Christin, N. (2014, August). Automatically Detecting Vulnerable Websites Before They Turn Malicious. In USENIX Security Symposium (pp. 625-640).

Sutherland, J. (2001). Agile Can Scale: Inventing and Reinventing SCRUM in Five Companies, Cutter IT Journal, vol 14, pp. 5-11.

Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. Data Classification: Algorithms and Applications, 37.

The Forrester Wave: Security Analytics Platform, Q4 2020, December 1, 2020.

The Global Risk Report 2017 - 12th edition; Global Economic Forum, Geneva, 2017.

Tin Kam Ho, (1998) The Random Subspace Method for Constructing Decision Forests , in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, nº 8, pp. 832–844

Wang, G., Hao, J., Ma, J., & Huang, L. (2010). A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. Expert systems with applications, 37(9), 6225-6232.

Wei-Chao Lin, Shih-Wen Ke, Chih-Fong Tsai, CANN: An intrusion detection system based on combining cluster centers and nearest neighbors, Knowledge-Based Systems, Volume 78, 2015, Pages 13-21, ISSN 0950-7051, http://dx.doi.org/10.1016/j.knosys.2015.01.009.

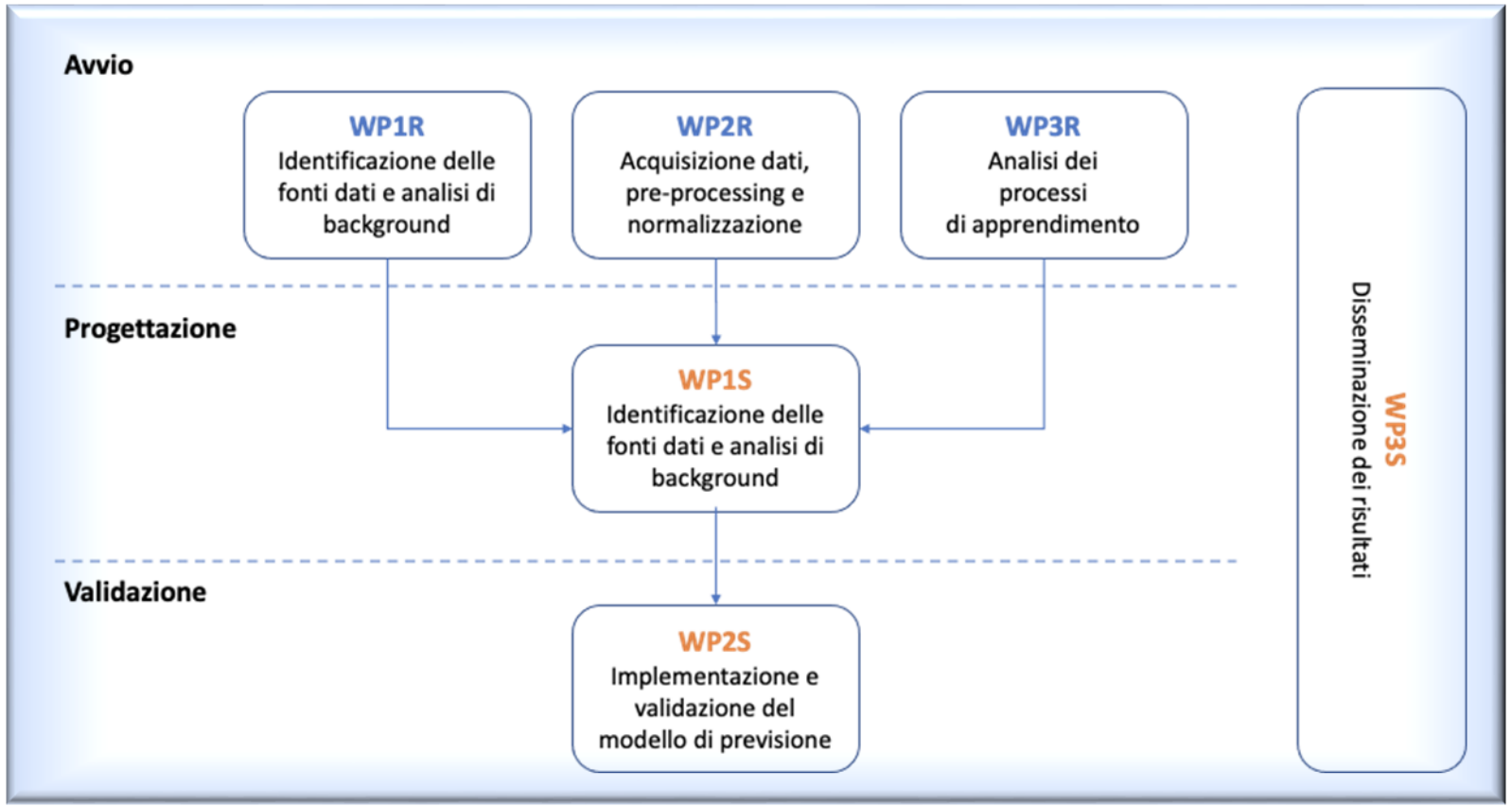
Woolery, L. K., & Grzymala-Busse, J. (1994). Machine learning for an expert system to predict preterm birth risk. Journal of the American Medical Informatics Association, 1(6), 439-446.

Xiaofeng, M., & Xiang, C. (2013). Big data management: concepts, techniques and challenges [J]. Journal of computer research and development, 1(98), 146-169.

Zhang, F., Chan, P. P. K., Biggio, B., Yeung, D. S., Roli, F. (2016). Adversarial Feature Selection Against Evasion Attacks. IEEE Transactions on Cybernetics, 46(3), pagg 766-777.

Zhang, J., Durumeric, Z., Bailey, M., Liu, M., & Karir, M. (2014, February). On the Mismanagement and Maliciousness of Networks. In NDSS.

Keep Calm si articola secondo sei fasi progettuali



Minacce analizzate

Il dati raccolti da fonti pubbliche riguardano una serie di attacchi:

- Analysis,
- Backdoor,
- Denial of Service (DoS),
- Exploits,
- Fuzzers,
- Phishing,
- Reconnaissance,
- ShellCode,
- Worms.

Analysis. L'hacker tenta di accedere alla stessa rete in cui si accede per ascoltare (e acquisire) tutto il traffico della rete stessa. L'hacker può analizzare quel traffico per acquisire informazioni dell'Organizzazione a cui la rete appartiene. È un tipo di attacco passivo.

Backdoor. è un tipo di attacco che sfrutta le vulnerabilità nei sistemi di sicurezza informatica. Queste vulnerabilità possono essere intenzionali o non intenzionali e possono essere causate da errori di progettazione, errori di codifica o malware. Le minacce backdoor vengono spesso utilizzate per ottenere l'accesso non autorizzato a sistemi o dati o per installare malware sui sistemi. La pericolosità di questo tipo di attacco è legata al fatto che la backdoor fa da ponte per malware pericolosi come trojan, ransomware, spyware ecc.

Denial of Service. è un tipo di attacco informatico in cui un attore malintenzionato mira a rendere un computer o un altro dispositivo non disponibile agli utenti interrompendo il normale funzionamento del dispositivo. Gli attacchi DoS sovraccaricano una macchina con un flusso di richieste, fino a quando il traffico normale non è in grado di essere elaborato

Exploits. Gli exploit rappresentano errori nel processo di sviluppo di un software che lasciano delle falle nel sistema di protezione integrato nel software, utilizzabili dai cybercriminali per accedere al software e, partendo da esso, all'intero computer.

Fuzzers. L'attacco di questo tipo tenta di mandare in crash un sistema o di innescare errori fornendo un grande volume di input generati in maniera casuali. Se viene rilevata una vulnerabilità, l'hacker può procedere, di solito, con attacchi DoS, cross-site scripting, injection, ecc.

Phishing. È una particolare tecnica di cracking utilizzata per ottenere l'accesso ad informazioni personali e riservate con la finalità del furto di identità mediante l'utilizzo di messaggi di posta elettronica (o messaggi su social), opportunamente creati per apparire autentici.

Reconnaissance. In questo tipo di attacco, un hacker raccogliere informazioni utilizzando i social network. Gli utenti condividono informazioni personali e gli hacker utilizzano specifiche trappole per attirare il soggetto da attaccare acquisendo informazioni fornite spontaneamente circa il suo sistema.

ShellCode. L'attacco prevede l'utilizzo di una piccola porzione di codice iniettata dentro un sistema al fine di avviare un'istruzione attraverso la quale l'hacker ottiene il controllo del dispositivo. La stragrande maggioranza dei codici shell remoti utilizza connessioni socket TCP/IP.

Worms. È un tipo di programma malevolo la cui caratteristica è quello di infiltrarsi in maniera latente sulle macchine per poi propagarsi, infettando altri sistemi sfruttando le capacità di comunicazione della macchina stessa. È un tipo di virus con capacità autoreplicante.

Algoritmi di machine learning in parallelo

Per ogni attacco, di cui si è acquisito un database, sono stati posti in parallelo diversi algoritmi per valutarne la capacità classificatoria, ma anche per testare le singole performance in termini di velocità nel dare risposte agli input forniti.
Gli algoritmi che hanno dimostrato maggiore capacità discriminatoria sono i seguenti:

- Ada Boost M1
- Hoeffding Tree
- Naïve Bayes
- Radom Forest
- Random Committee
- Random Tree
- Real Ada Boost
- Stochastic Gradient Descent

Gli algoritmi qui elencati vengono utilizzati in ensemble, ovvero simultaneamente, al fine di aumentare la capacità di intercettazione della minaccia.

Algoritmi ad oggi reperibili in letteratura

Algoritmi instance-based
K-nearest neighbors algorithm (KNN),
Learning vector quantization (LVQ),
Self-organizing map (SOM).

Analisi di regressione (Regression analysis)
Regressione logistica (Logistic regression),
Regressione ordinaria ai minimi quadrati (Ordinary least squares regression - OLSR),
Regressione lineare,
Stepwise regression,
Multivariate adaptive regression splines (MARS),
Regularization algorithm:
Ridge regression.
Least Absolute Shrinkage and Selection Operator (LASSO),
Elastic net,
Least-angle regression (LARS),
Classificatori:
Classificatori probabilistici (Probabilistic classifier):
Naive Bayes classifier,
Classificatori binari (Binary classifier),
Classificatori lineari (Linear classifier),
Classificatori gerarchici Hierarchical classifier.

Riduzione di dimensionalità (Dimensionality reduction)
Canonical correlation analysis (CCA)
Analisi fattoriale (Factor analysis),
Estrazione delle caratteristiche (Feature extraction),
Selezione delle caratteristiche (Feature selection),
Analisi delle componenti indipendenti (Independent component analysis - ICA),
Analisi discriminante lineare (Linear discriminant analysis - LDA),
Multidimensional scaling (MDS),
Non-negative matrix factorization (NMF),
Partial least squares regression (PLSR),
Analisi della componente principale (Principal component analysis - PCA),
Regressione della componente principale (Principal component regression - PCR),
Projection pursuit,
Sammon mapping,
t-distributed stochastic neighbor embedding (t-SNE).

Ensemble learning
AdaBoost,
Boosting,
Bootstrap aggregating (Bagging),
Ensemble,
Gradient boosted decision tree (GBDT),
Gradient boosting machine (GBM),
Random Forest,
Stacked Generalization (blending).

Meta learning
Inductive bias,
Metadata.

Apprendimento rinforzato (Reinforcement learning)
Q-learning,
State–action–reward–state–action (SARSA),
Temporal difference learning (TD),
Learning Automata.

Apprendimento supervisionato (Supervised learning)
AODE,
Reti neurali artificiali (Artificial neural network),
Association rule learning algorithms:
Apriori algorithm,
Eclat algorithm,
Case-based reasoning,
Gaussian process regression,
Gene expression programming,
Group method of data handling (GMDH),
Programmazione a logica induttiva (Inductive logic programming),
Instance-based learning,
Lazy learning,
Learning Automata,
Learning Vector Quantization,
Logistic Model Tree,
Minimum message length (decision trees, decision graphs):
Nearest Neighbor Algorithm,
Analogical modeling,
Probably approximately correct learning (PAC) learning,
Ripple down rules,

Bayesiani (Bayesian)
Bayesian knowledge base,
Naive Bayes,
Gaussian Naive Bayes,
Multinomial Naive Bayes,
Averaged One-Dependence Estimators (AODE),
Bayesian Belief Network (BBN),
Bayesian Network (BN).

Apprendimento simbolico (Symbolic machine learning algorithms),
Support vector machines,
Random Forests,
Ensembles of classifiers:
Bootstrap aggregating (bagging),
Boosting (meta-algorithm),
Ordinal classification,
Information fuzzy networks (IFN),
Conditional Random Field,
ANOVA,
Quadratic classifiers,
Boosting:
SPRINT,
Reri bayesiane (Bayesian networks):
Naive Bayes,
Hidden Markov models:
Hierarchical hidden Markov model.

Alberi decisionali (Decision tree algorithms)
Decision tree,
Classification & regression tree (CART),
Iterative Dichotomiser 3 (ID3),
C4.5 algorithm,
C5.0 algorithm,
Chi-squared Automatic Interaction Detection (CHAID),
Decision stump,
Conditional decision tree,
ID3 algorithm,
Random forest,
SLIQ.

Classificatori lineari (Linear classifier)
Fisher's linear discriminant,
Regressione lineare (Linear regression),
Regressione logistica (Logistic regression),
Regressione logistica multinomiale (Multinomial logistic regression),
Naive Bayes classifier,
Percettrone (Perceptron),
Support vector machine.

Apprendimento non supervisionato (Unsupervised learning)
Expectation-maximization algorithm
Vector Quantization
Generative topographic map
Information bottleneck method

Reti neurali artificiali (Artificial neural networks)
Feedforward neural network:
Extreme learning machine,
Convolutional neural network,
Reti neurali ricorrenti (Recurrent neural network):
Long short-term memory (LSTM).
Logic learning machine,
Self-organizing map,
Feedforward neural network:
Extreme learning machine,
Convolutional neural network.

Apprendimento a regole associative (Association rule learning)
Apriori algorithm,
Eclat algorithm,
FP-growth algorithm.

Clustering gerarchico (Hierarchical clustering)
Single-linkage clustering,
Conceptual clustering.

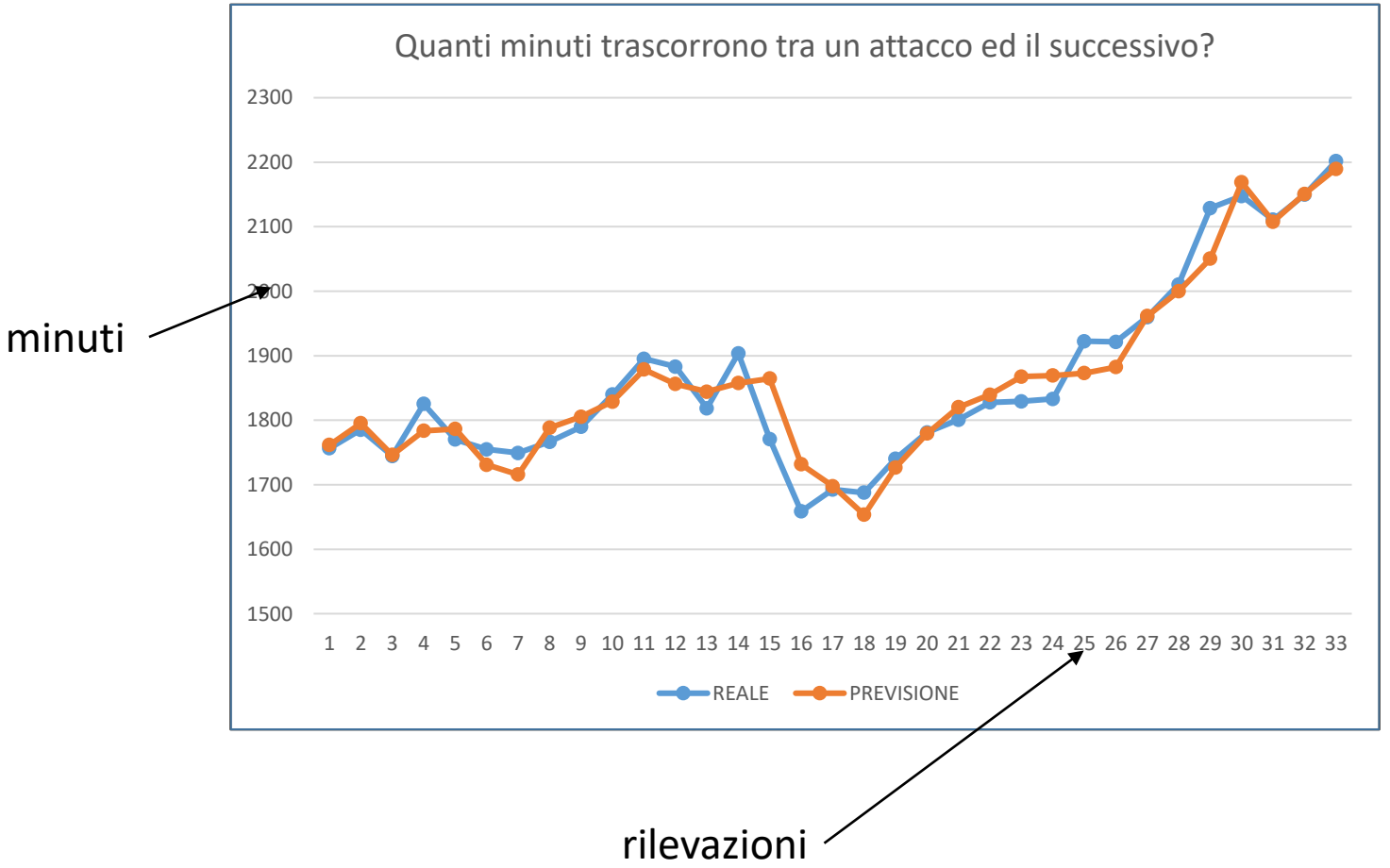
Cluster analysis
BIRCH,
DBSCAN,
Expectation-maximization (EM),
Fuzzy clustering,
Hierarchical Clustering,
K-means clustering,
K-medians,
Mean-shift,
OPTICS algorithm.

Anomaly detection
k-nearest neighbors classification (k-NN),
Local outlier factor.

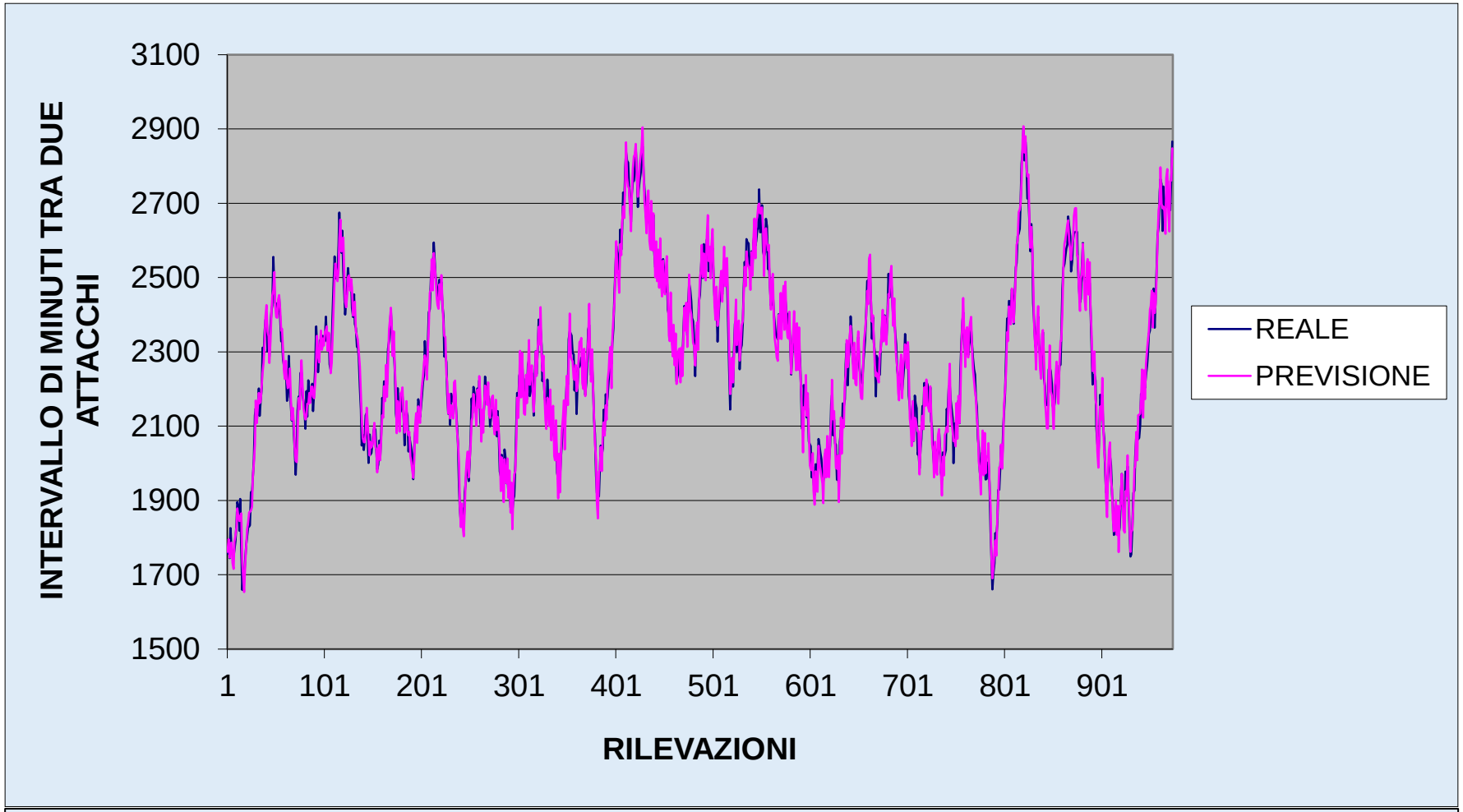
Apprendimento semi-supervisionato (Semi-supervised learning)
Active learning,
Generative models,
Low-density separation,
Graph-based methods,
Co-training,
Transduction.

Deep learning
Deep belief networks,
Deep Boltzmann machines (DBM),
Deep Convolutional neural networks,
Deep Recurrent neural networks,
Hierarchical temporal memory,
Generative Adversarial Networks,
Stacked Auto-Encoders.

Quanti minuti passano tra un attacco ed il successivo?



Prevedere il tempo che passa tra due attacchi



Correttezza della previsione della direzione = 81.67%

LUISS Università Guido Carli · GRUPPOISC Information Sharing Company · sogei · CYBER 4.0 CYBERSECURITY COMPETENCE CENTER