





Corso di Laurea in Informatica

Costruzione di Modelli di Predizione di Vulnerabilità con Algoritmi Genetici: un'Indagine Preliminare

Prof. Fabio Palomba

Dott. Emanuele lannone

Alfonso Cannavale

Mat.: 0512108068

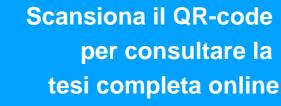


a.cannavale7@studenti.unisa.it



https://github.com/alfcan

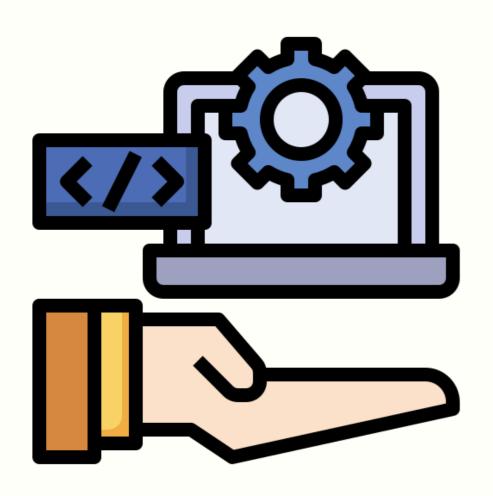




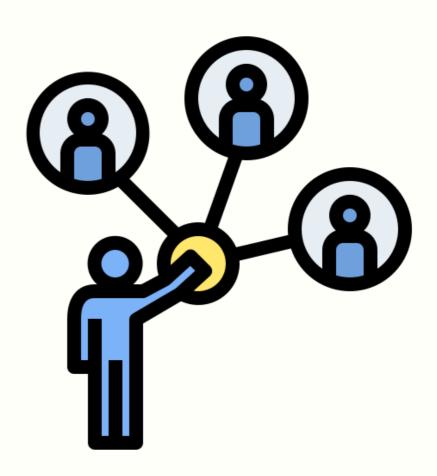


Contesto









Incremento dei servizi digitali

La società si affida sempre di più al software





https://github.com/alfcan



Qual è la problematica?





Presenza di vulnerabilità software

Identificazione e gestione delle vulnerabilità













Identificazione delle vulnerabilità

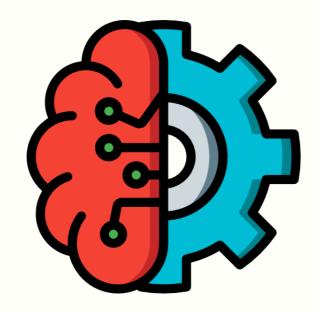




Analisi Statica







Modelli di predizione di vulnerabilità



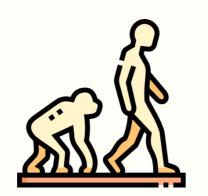






Gli algoritmi genetici



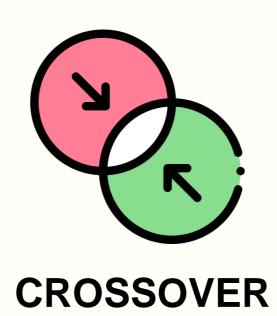


Ispirati alla Teoria dell'Evoluzione



Operatori di Ricerca









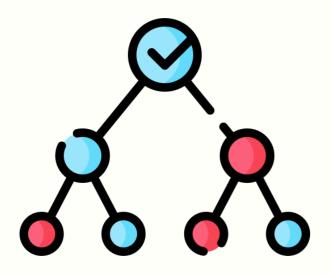


https://github.com/alfcan



Codifica degli individui





WHY?

- Facile da comprendere
- Facilmente codificabile in individui ricombinabili

Composite Pattern





INNER NODE:

- nodi figli
- condizione: '<=' o '>'
- valore soglia



LEAF:

label



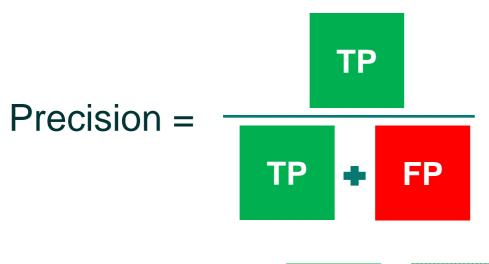


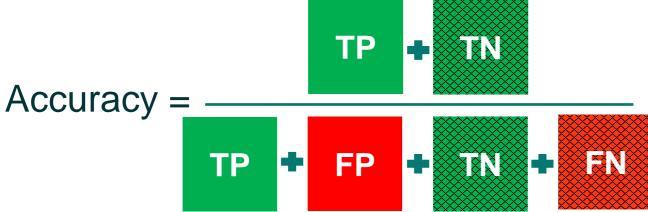


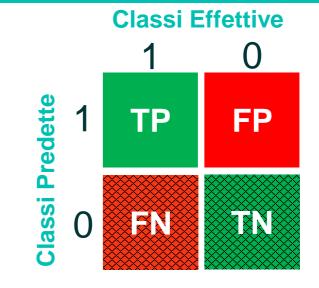
Funzioni Obiettivo

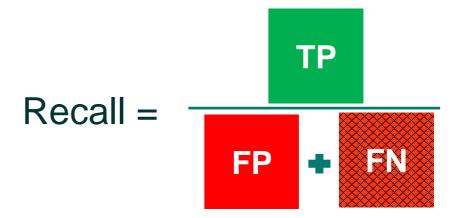


Problema di ottimizzazione a singolo obiettivo









F-measure =
$$2 \times \frac{precision \times recall}{precision + recall}$$







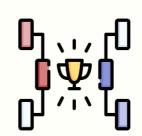
https://github.com/alfcan

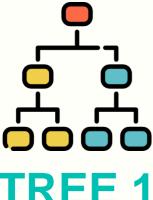


Operatore di Selezione



Tournament Selection



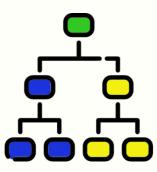


Precision 69%

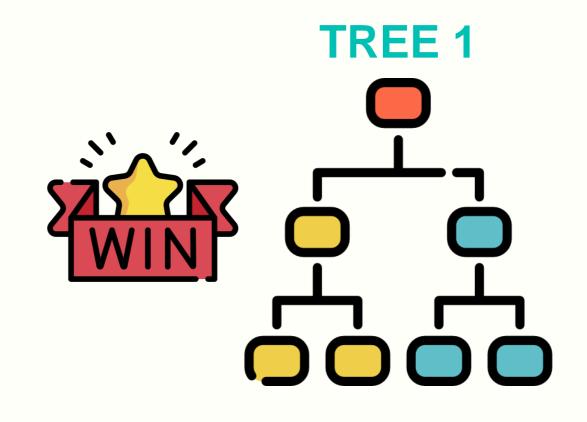
TREE 1

VS

TREE 2



Precision 63%







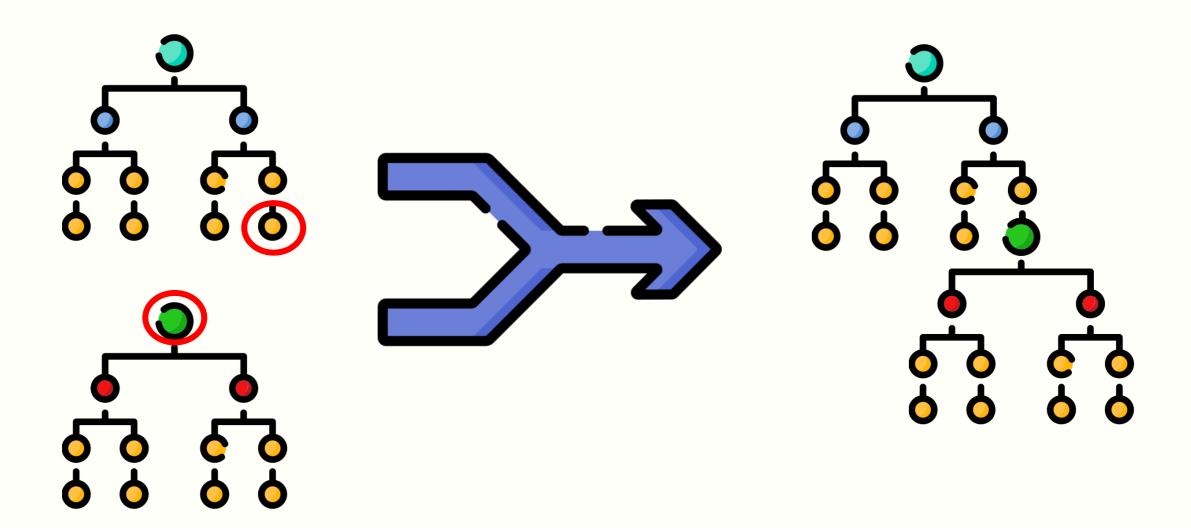




Operatore di Crossover



Tree Crossover









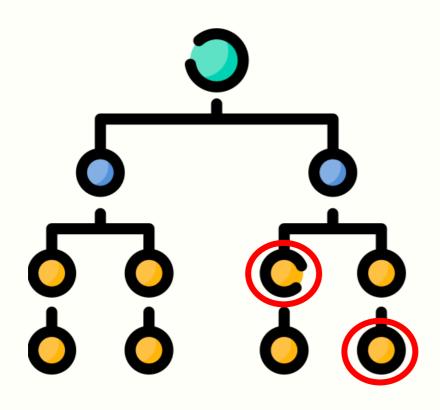




Operatore di Mutazione



Tree Mutation



NODO INTERNO Condizione: '<=' Valore soglia: 2,3 NODO INTERNO Condizione: '>'
Valore soglia: 2,3

NODO INTERNO Condizione: '>=' Valore soglia: 1,3

NODO FOGLIA Etichetta: 1 NODO FOGLIA Etichetta: 0









Qualche esperimento...





Ottimizziamo la precision, la recall, l'accuracy o la f-measure?

Proviamo a cambiare i parametri...





Cosa succede se cambio il criterio di arresto?









Ottimizzazione della F-measure



Decision Tree GA vs Decision Tree CART

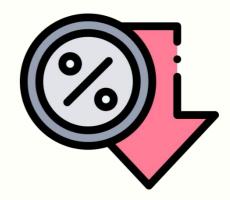






Precision: 23% vs 8,2%

Recall: 12,5% vs 17%





F-measure: 16,6% vs 11%

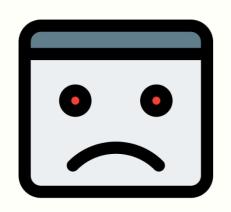






Riflessioni finali





Le prestazioni dei modelli sono molto basse, ma...



L'utilizzo di algoritmi genetici per la definizione di VPMs ha un impatto positivo sulle prestazioni dei modelli











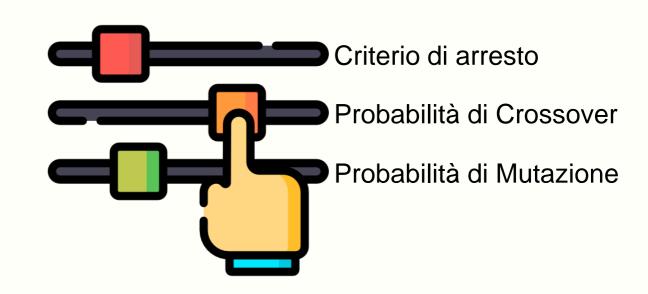


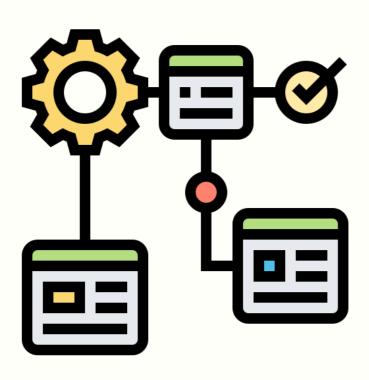
Limitazioni e Sviluppi Futuri



LIMITAZIONI

 Sensibilità della scelta dei parametri





SVILUPPI FUTURI

- Analisi della struttura dei modelli definiti tramite GA
- Definizione di altri modelli predittivi

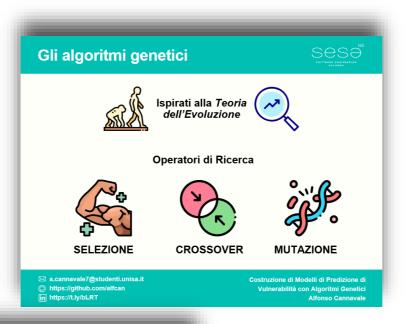
















Questa tesi ha contribuito a piantare un albero in Kenya



Costruzione di Modelli di Predizione di Vulnerabilità con Algoritmi Genetici: un'Indagine Preliminare



Grazie!

Alfonso Cannavale

a.cannavale7@studenti.unisa.it https://github.com/alfcan





