Proyecto de Cloud Computing

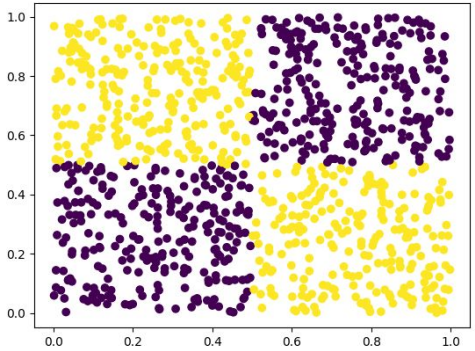
Estudiantes:
- Jorge Rebosio
- Giordano Alvitez
- Cesar Salcedo

Introducción

Proyecto ligero

Objetivo

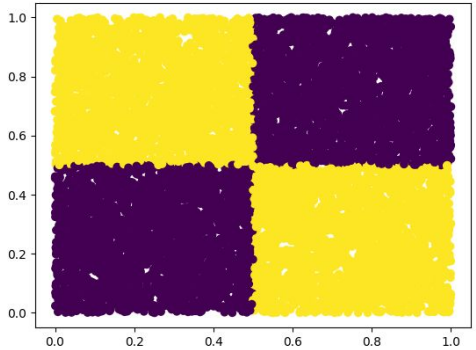
- Tarea: problema XOR de clasificación
- Mostrar cómo es que la capacidad de una red neuronal afecta en su desempeño de aprendizaje



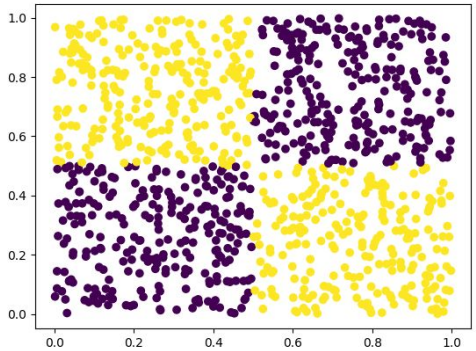
coordenadas (X, Y) → ML Model → ¿resulta en 0 o 1?

Dataset

- Generador de puntos con auto asignación de labels.
- A partir del generador se generan un training set y un test set.

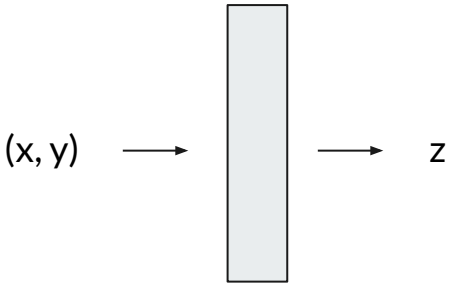


training set

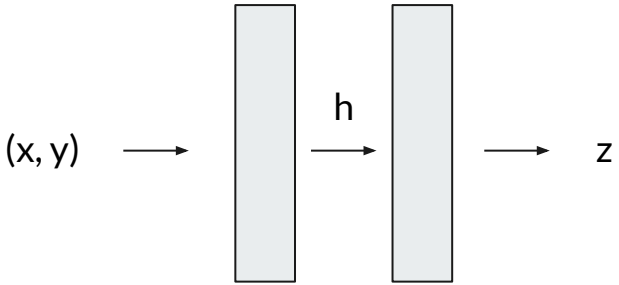


test set

Dos arquitecturas de aprendizaje

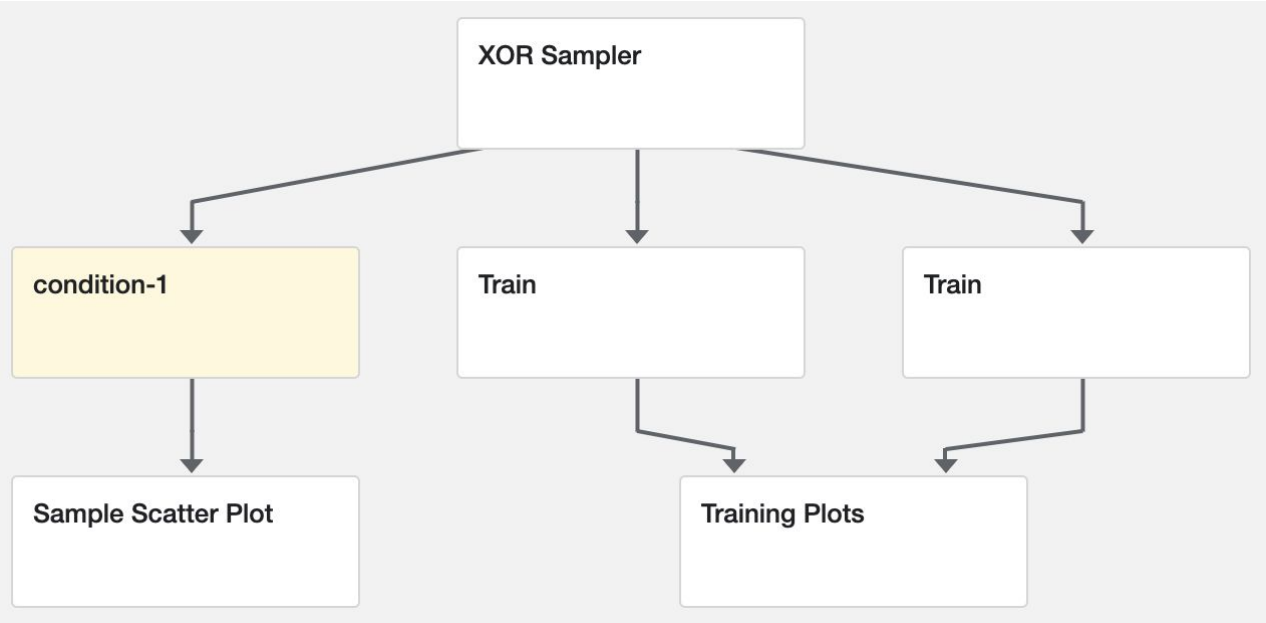


Modelo lineal
(sin hidden layers)

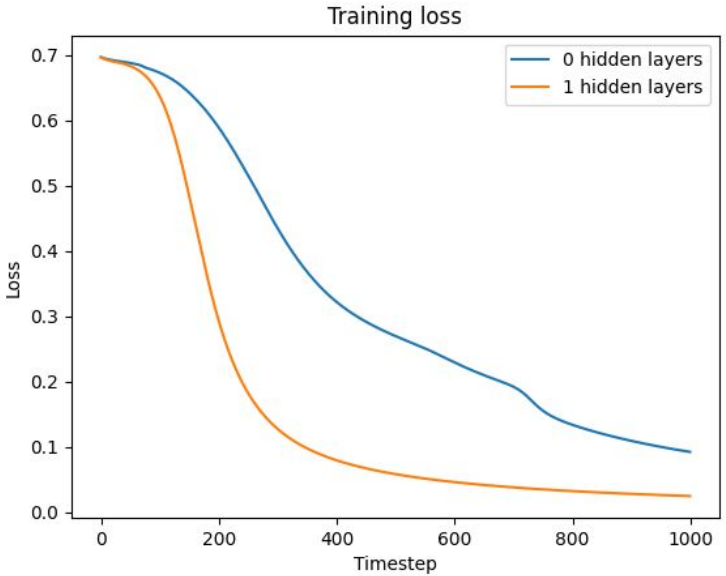


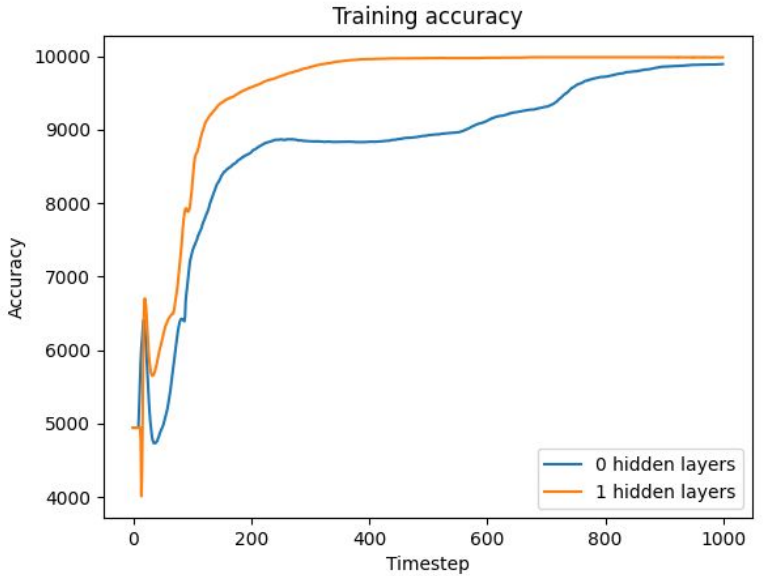
Modelo no lineal
(con un hidden layer)

Pipeline en Kubeflow

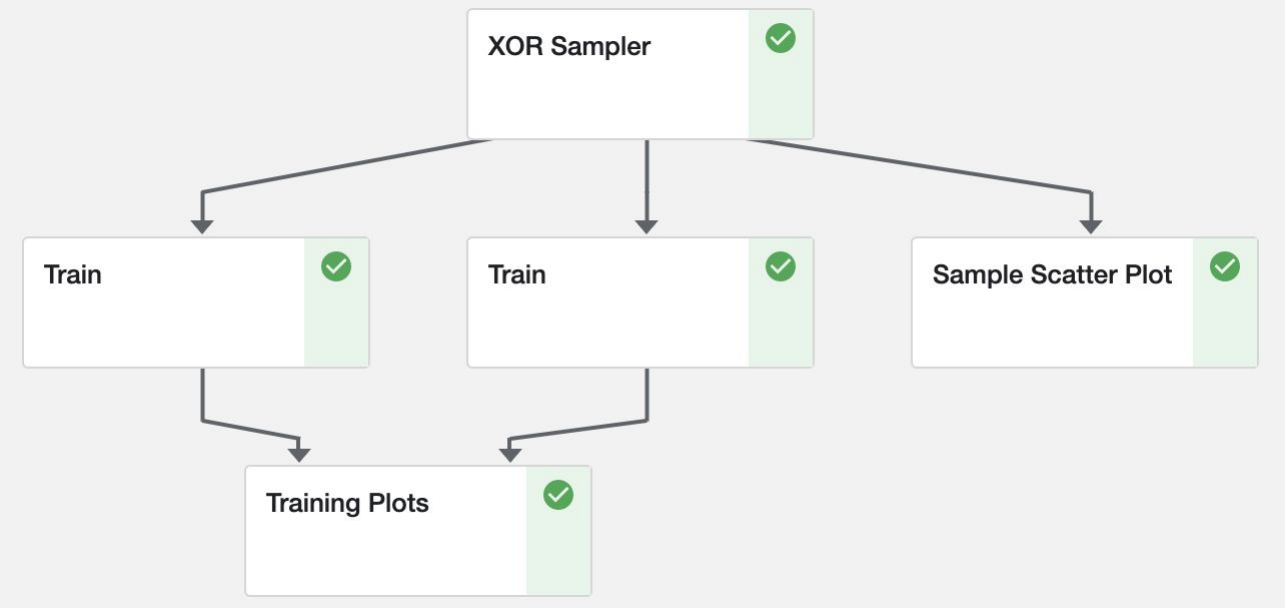


Resultados del entrenamiento





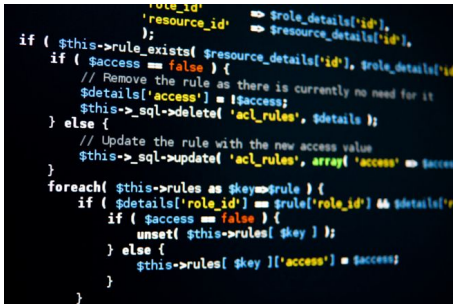
Demo


XOR Sampler
Train
Train
Sample Scatter Plot
Training Plots

Proyecto pesado

Objetivo

Identificar los tipos de vulnerabilidades presentes en un código en C o C++ usando machine learning.



ML Model

- Vulnerability 1 - Yes
- Vulnerability 2 - No
- Vulnerability 3 - Yes
- Vulnerability 4 - Yes

Datos

	SATE IV	Debian	GitHub
Total	121,353	2,806,469	9,532,081
Passing curation	12,001	380,381	955,683
Not vulnerable	6,559 (55%)	364,306 (96%)	907,186 (95%)
Vulnerable	5,442 (45%)	16,075 (4%)	48,497 (5%)

SATE IV :

A collection of test cases in the C/C++ language. It contains examples organized under 118 different CWEs.

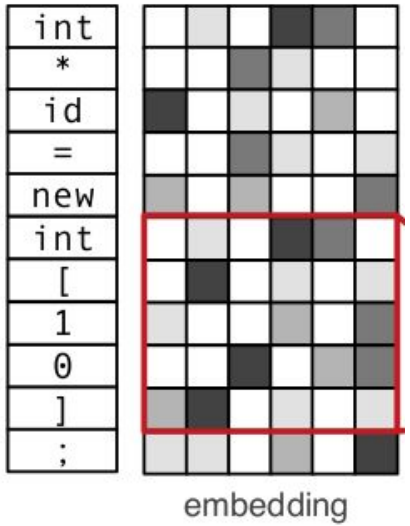
Passing Curation :

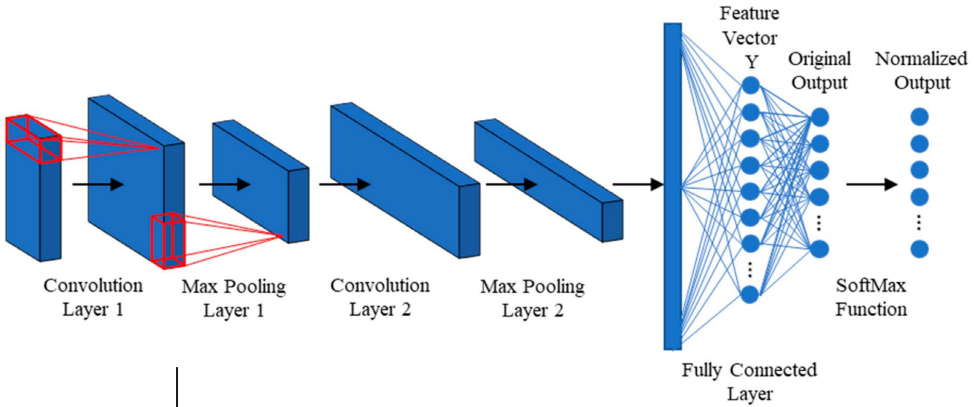
Funciones duplicadas podrían generar overfitting

Cómo se aplicaría ML para clasificarlas debilidades

source code

```
int * data =
new int[10];
```





Word2Vec
Embedding NN

CNN 1D
RNN (GRU)

Métricas

	Predicted Positives	Predicted Negatives
Positives	True Positives	False Negatives
Negatives	False Positives	True Negatives

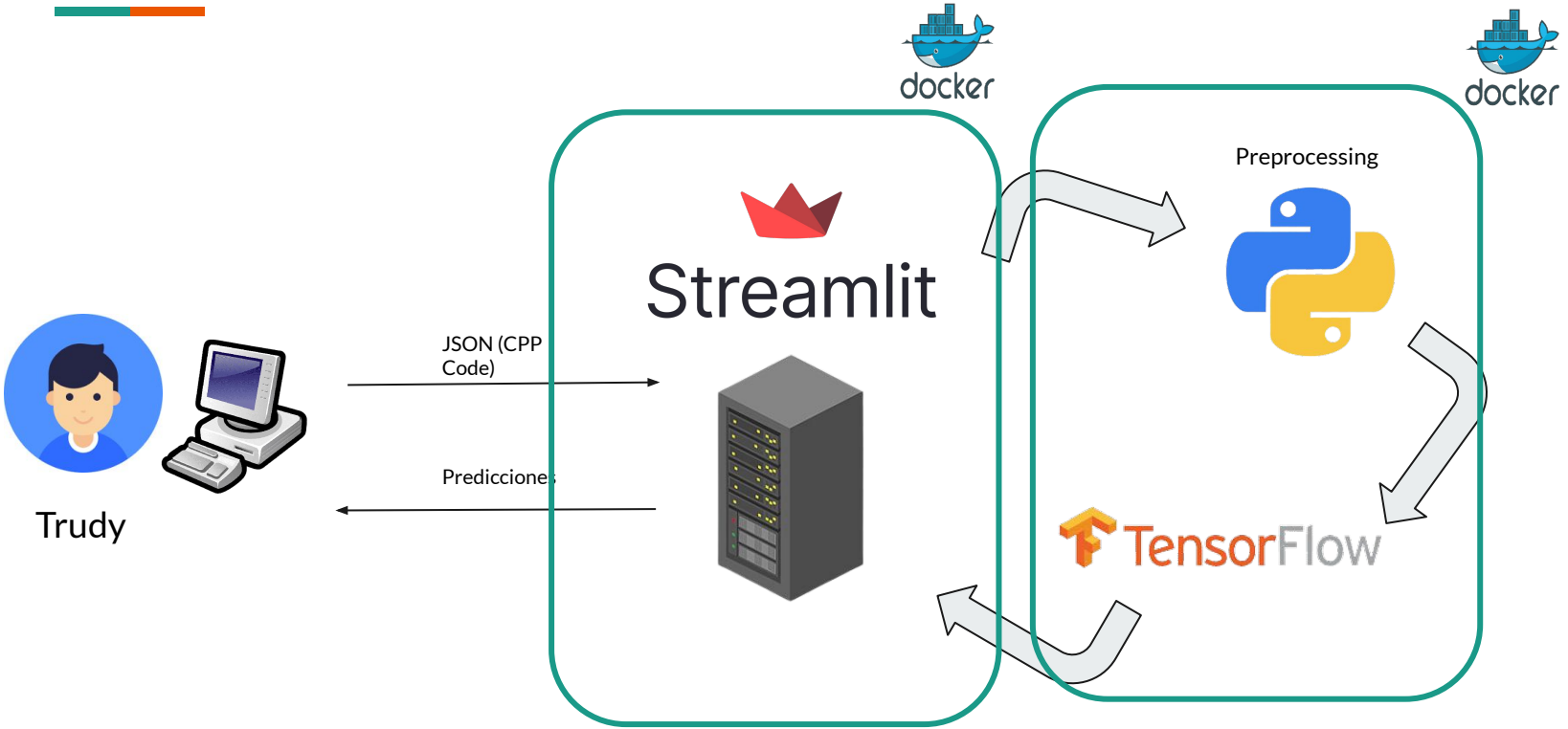
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

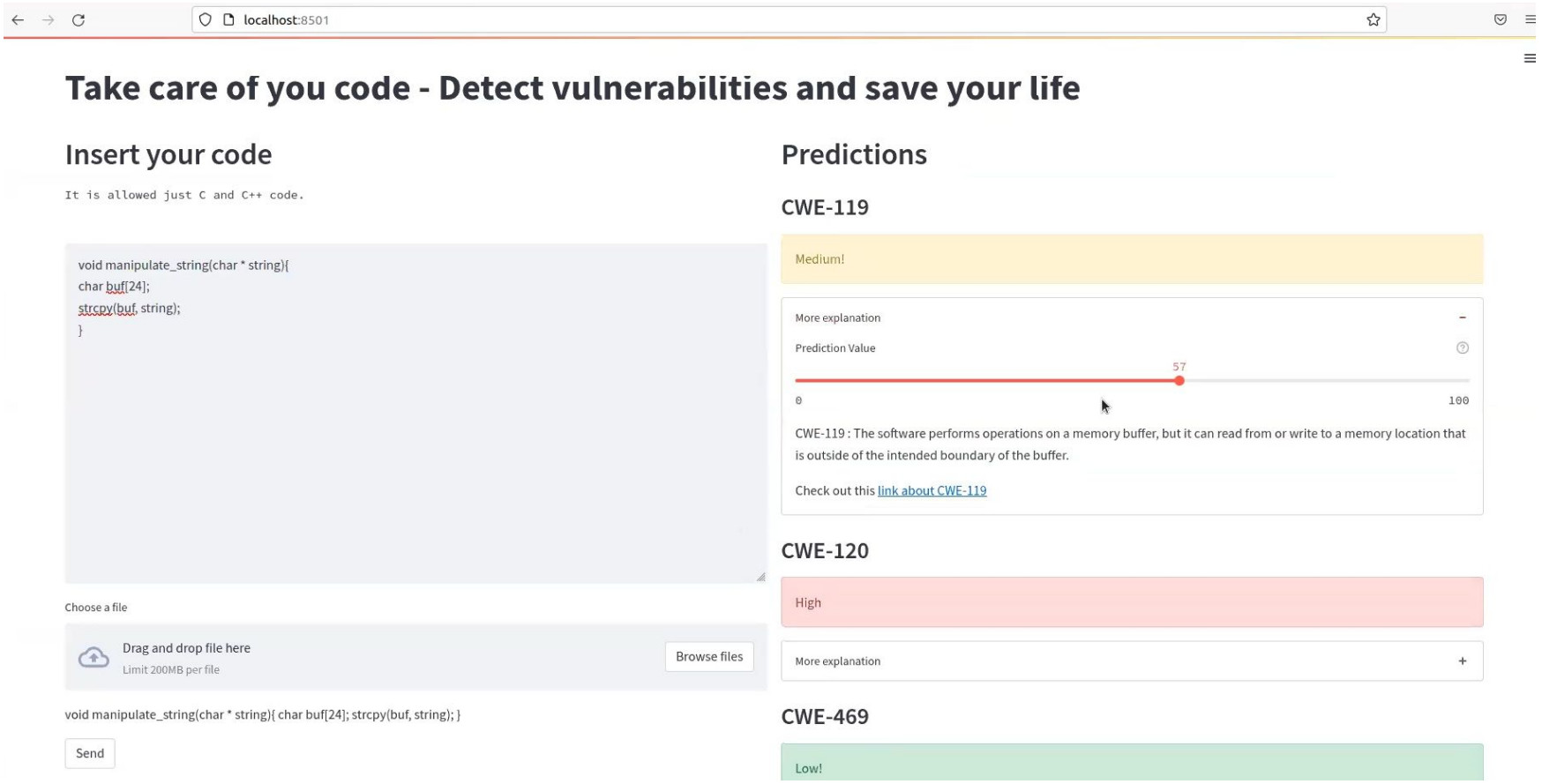
$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Arquitectura



Demo


localhost:8501
Take care of you code - Detect vulnerabilities and save your life
Insert your code
It is allowed just C and C++ code.
void manipulate_string(char * string){
char buf[24];
strcpy(buf, string);
}
Choose a file
Drag and drop file here
Limit 200MB per file
Browse files
void manipulate_string(char * string){ char buf[24]; strcpy(buf, string); }
Send
Predictions
CWE-119
Medium!
More explanation
Prediction Value
57
0
100
CWE-119 : The software performs operations on a memory buffer, but it can read from or write to a memory location that is outside of the intended boundary of the buffer.
Check out this link about CWE-119
CWE-120
High
More explanation
CWE-469
Low!

CONCLUSIONES

¡Gracias por su atención!