# Detección de intrusiones en redes de datos mediante redes neuronales

Gianfranco Fagioli
Nicolás Arato

Tutor: Leandro Vignolo

# Introducción

Sistemas de detección de intrusos (IDS) -> NIDS

Función de los IDS (trabaja en conjunto con otras aplicaciones)

Base de datos KDD cup 99 - Tercer Concurso Internacional de Descubrimiento de Conocimientos y Herramientas de Minería de Datos. (Subconjunto mejorado NSL-KDD)

Diferentes enfoques: Algoritmos genéticos- Deep Learning- Sistemas híbridos

# Data set NSL-KDD - Preprocesamiento

- **Modificación/codificación**
- Normalización
- Desbalance de clases

| Tipo | Codificación |
|---|---|
| Normal | 000 |
| Neptune | 001 |
| Smurf | 010 |
| warezclient | 011 |
| ipsweep | 100 |
| nmap | 101 |
| portsweep | 110 |
| satan | 111 |

# Data set NSL-KDD - Preprocesamiento

- Modificación/codificación
- **Normalización**
- Desbalance de clases

$$x_i = \frac{v_i - min(v_i)}{max(v_i) - min(v_i)}$$

# Data set NSL-KDD - Preprocesamiento

- Modificación/codificación
- Normalización
- **Desbalance de clases**

| Clase | Cantidad de paquetes |
|---|---|
| normal | 21264 |
| neptune | 13207 |
| smurf | 811 |
| warezclient | 308 |
| ipsweep | 1143 |
| nmap | 473 |
| portsweep | 947 |
| satan | 1147 |

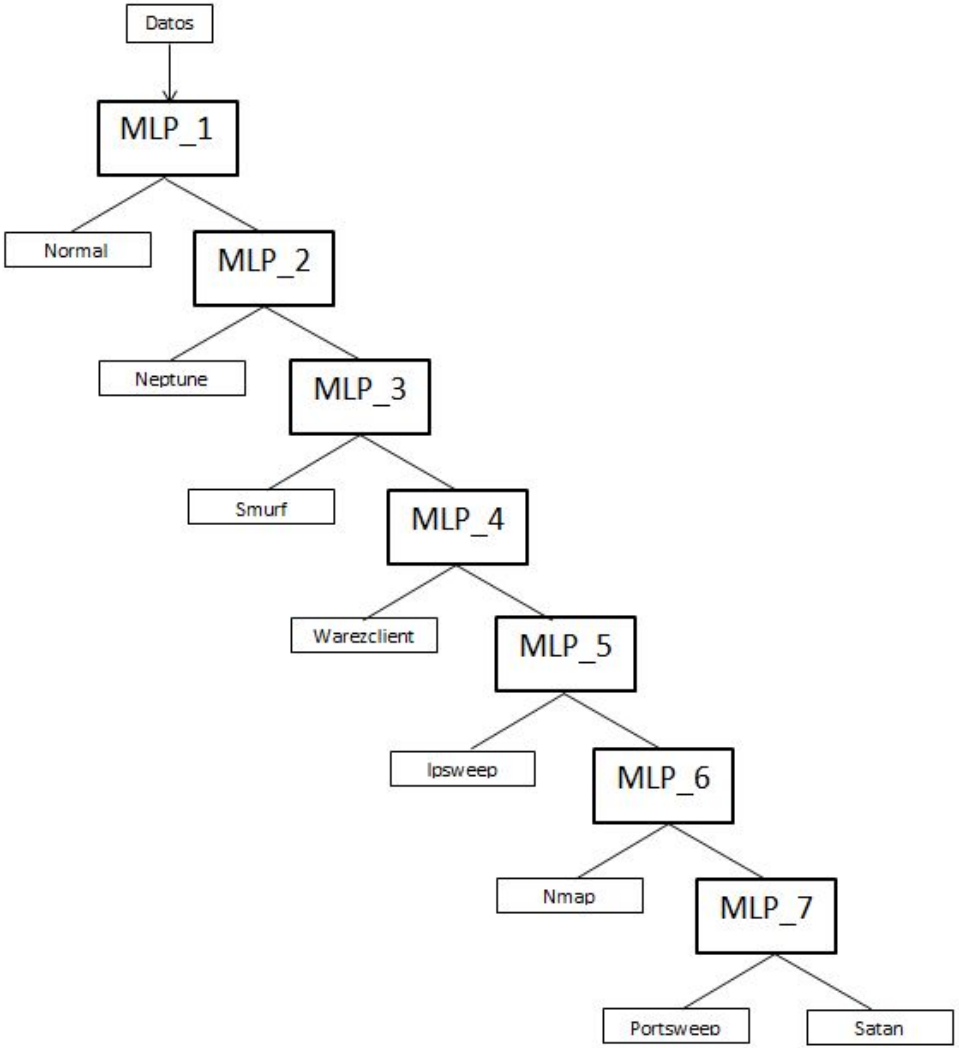
# Propuesta

- Clasificador Jerárquico para detección de intrusiones de red.

- Comparación con MLP.

# MLP Jerárquico

-Validación cruzada (Hold out)



# Resultados: Matriz de confusión promedio

| | nor | nep | smu | war | ips | nmap | por | sat |
|---|---|---|---|---|---|---|---|---|
| nor | 4238.2 | 1.4 | 3.2 | 6.2 | 9.6 | 2.6 | 3.4 | 4.6 |
| nep | 2.6 | 2624 | 0 | 0 | 0 | 0.4 | 0.8 | 0 |
| smu | 2.4 | 0 | 155.2 | 0 | 0 | 0.2 | 0 | 0 |
| war | 0 | 0.2 | 0 | 65.2 | 0 | 0 | 0 | 0 |
| ips | 3.2 | 0 | 0 | 0 | 219.2 | 5 | 0 | 0.2 |
| nmap | 0.2 | 0 | 0 | 0 | 1.2 | 90.2 | 0 | 2 |
| por | 1.8 | 0 | 0 | 0 | 0.4 | 0 | 187.4 | 1.2 |
| sat | 1.2 | 0 | 0 | 0 | 0 | 1.6 | 0.8 | 224.2 |

# Medidas de desempeño

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

$$Missrate = \frac{fn}{fn + tp} = 1 - Recall$$

# Resultados

| | recall | precision | miss_rate | accuracy |
|---|---|---|---|---|
| **MLP1** | 0.9979 | 0.9910 | 0.0021 | 0.9949 |
| **MLP2** | 0.9974 | 0.9890 | 0.0026 | 0.9937 |
| **MLP3** | 0.9971 | 0.9914 | 0.0029 | 0.9947 |
| **MLP-J1** | 0.9955 | 0.9912 | 0.0045 | 0.9939 |
| **MLP-J2** | 0.9967 | 0.9904 | 0.0033 | 0.9942 |
| **MLP-J3** | 0.9968 | 0.9914 | 0.0032 | 0.9946 |

# Conclusión

- MLP vs MLP Jerárquico
- Resultados similares
- Pruebas
- Variación de parámetros libres

# Gracias