

# Predicción del comportamiento de actividades terroristas basado en Redes Neuronales Profundas

## Proyecto de Tesis I

Ingrid F. Ipanaqué C.<sup>1</sup>

<sup>1</sup>Facultad de Ciencias  
Universidad Nacional de Ingeniería

*Asesor: Juan Carlos Espejo Delzo*

March 3, 2021



# Tabla de Contenido

- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro



# Introducción

## Motivación

- Estudio de las actividades terroristas.
- Interés por el aprendizaje profundo.
- Interés por el el manejo de macrodatos.



# Introducción

## Objetivos

- Entender el funcionamiento de las arquitecturas de red para aprendizaje profundo.
- Implementar un modelo de red neuronal profunda.
- Utilizar la técnica sintética de sobremuestreo de minorías para el preprocesamiento de datos.
- Evaluar el rendimiento del modelo implementado.



# Tabla de Contenido

- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro



# Estado del Arte

## Conceptos previos

- Redes Neuronales Artificiales
- Redes Neuronales Profundas
- Funciones de Activación
  - Rectified Linear Unit (ReLu)
  - Sigmoidea
  - Softmax
- Algoritmos de Optimización
  - Estimación del momento adaptativo (Adam)
- Métricas de evaluación



# Estado del Arte

## Trabajos relacionados

- Comparación de enfoques de aprendizaje automático en la predicción de ataques terroristas. *Palak Agarwal, et al. (2019)*.
- Arquitectura de red neuronal artificial con N capas ocultas: ejemplo de aplicación geofísica. *Jide Nosakare Ogunboa, et al. (2020)*.
- Una red neuronal profunda mejorada para predecir el absentismo laboral. *Syed Atif Ali Shah, et al. (2020)*.



# Tabla de Contenido

- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas**
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro





# Recursos y Herramientas

## Lenguaje y librerías

Para el desarrollo de este proyecto se utilizó el lenguaje de programación Python.

Librerías principales usadas:

- Keras
- Scikit-learn (sklearn)



# Recursos y Herramientas

## Global Terrorism Database (GTD)

Se hizo uso de la fuente desarrollada por el Consorcio Nacional para el Estudio del Terrorismo y las Respuestas al Terrorismo (START). GTD es la base de datos más completa de ataques terroristas del mundo. Es de código abierto y proporciona información sobre ataques terroristas nacionales e internacionales desde 1970 y a la fecha incluye más de 190,000 eventos.



# Tabla de Contenido

- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro



## Preprocesamiento de datos

Orden seguido para el preprocesamiento de los datos.

**1 Selección de atributos:**

181,857 instancias y 34 atributos.

**2 Codificación de etiquetas:**

Uso de la clase LabelEncoder de la librería sklearn.

**3 Manejo de datos faltantes:**

Uso de la clase SimpleImputer de la librería sklearn.  
Reemplaza los datos con la media.

**4 Normalización:**

Uso de la clase MinMaxScalar de la librería sklearn.

**5 Tratamiento de clases desbalanceadas:**

Técnica sintética de sobremuestreo de minorías (SMOTE).



# Preprocesamiento de datos

## Selección de atributos

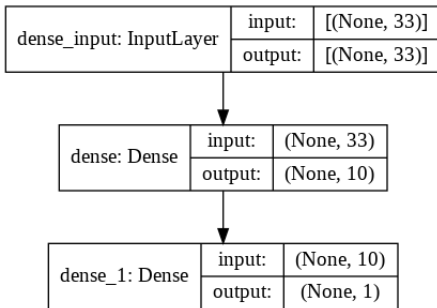
- **Suicidio:** Indica si la actividad terrorista será un suicidio o no.  
Clasificación binaria.
- **Éxito:** Indica si la actividad terrorista tendrá éxito o no.  
Clasificación binaria.
- **Tipo de arma:** Clasifica el tipo de arma utilizado.  
Clasificación multiclase: 13 clases.  
*No existen datos de clase 4 - En blanco.*
- **Región:** Clasifica la región que será blanco de la actividad terrorista.  
Clasificación multiclase: 12 clases.
- **Tipo de ataque:** Clasifica el tipo de ataque realizado.  
Clasificación multiclase: 8 clases.



# Entrenamiento del modelo

## Arquitectura de la Red Neuronal

Figure: Arquitectura de la Red Neuronal.



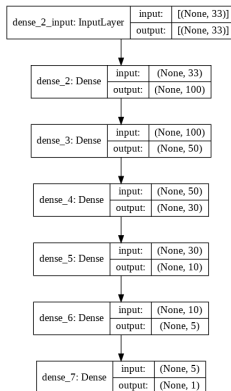
(Fuente: Elaboración propia)



# Entrenamiento del modelo

## Arquitectura de la Red Neuronal Profunda

Figure: Arquitectura de la Red Neuronal Profunda



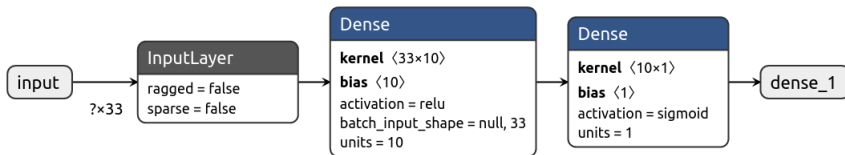
(Fuente: Elaboración propia)



# Entrenamiento del modelo

## Proceso de aprendizaje de la Red Neuronal

Figure: Atributos de la red neuronal de clasificación binaria.



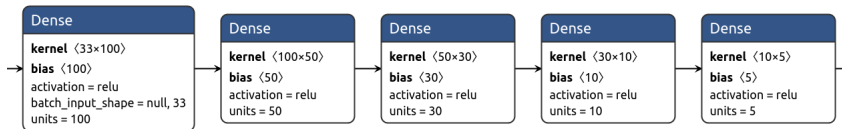
(Fuente: Elaboración propia)



# Entrenamiento del modelo

## Proceso de aprendizaje de la Red Neuronal Profunda

**Figure:** Atributos de la red neuronal profunda de clasificación binaria.



(Fuente: Elaboración propia)

# Tabla de Contenido

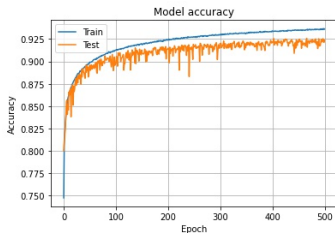
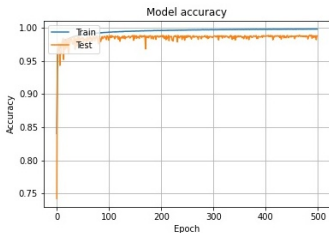
- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados**
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro



# Resultados

## Rendimiento de la Red Neuronal Profunda

**Figure:** Gráfica de *Accuracy* vs. *Épocas* del modelo DNN para el suicidio y éxito.



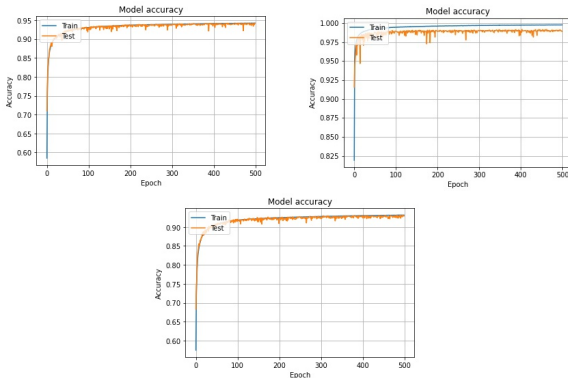
(Fuente: Elaboración propia)



# Resultados

## Rendimiento de la Red Neuronal Profunda

**Figure:** Gráfica de *Accuracy* vs. *Épocas* del modelo DNN para el tipo de arma, región y tipo de ataque.



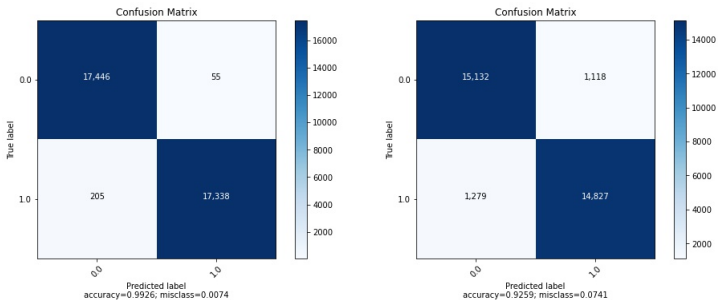
(Fuente: Elaboración propia)



# Resultados

## Matriz de Confusión de la Red Neuronal Profunda

Figure: Matriz de confusión del modelo DNN para el suicidio y éxito.



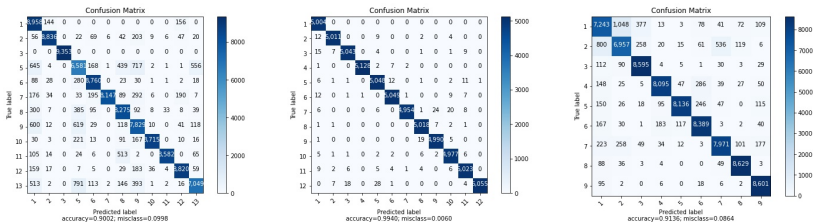
(Fuente: Elaboración propia)



# Resultados

## Matriz de Confusión de la Red Neuronal Profunda

**Figure:** Matriz de confusión del modelo DNN para el tipo de arma, región y tipo de ataque.



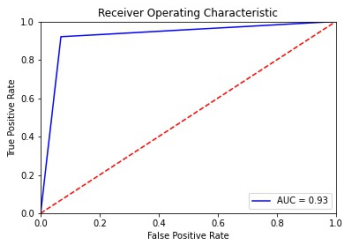
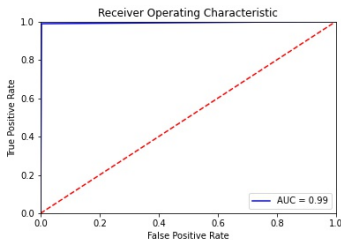
(Fuente: Elaboración propia)



# Resultados

## Curva ROC de la Red Neuronal Profunda

Figure: Curva ROC del modelo DNN para el suicidio y éxito.



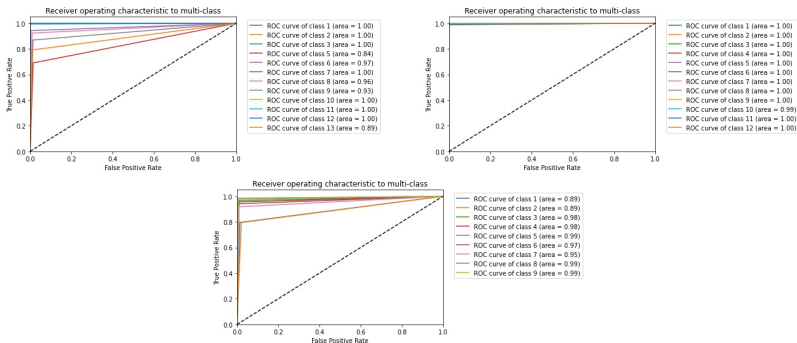
(Fuente: Elaboración propia)



# Resultados

## Curva ROC de la Red Neuronal Profunda

**Figure:** Curva ROC del modelo DNN para el tipo de arma, región y tipo de ataque.



(Fuente: Elaboración propia)

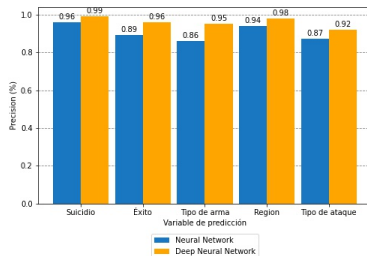
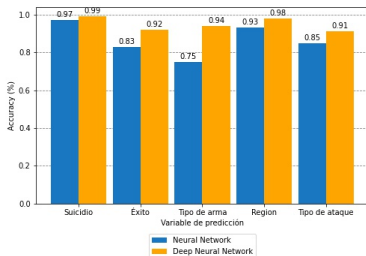




# Resultados

## Comparación en las distintas métricas

**Figure:** Accuracy y Precisión obtenido para la red neuronal y red neuronal profunda.



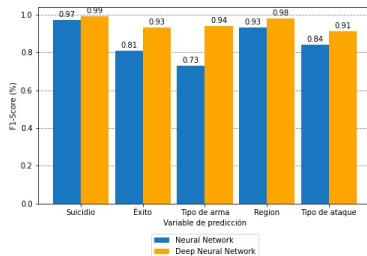
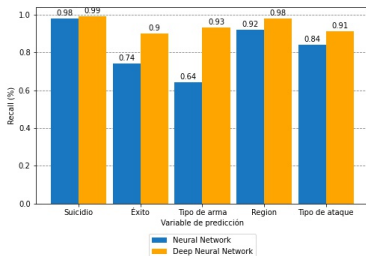
(Fuente: Elaboración propia)



# Resultados

## Comparación en las distintas métricas

**Figure:** Recall y F1-Score obtenido para la red neuronal y red neuronal profunda.



(Fuente: Elaboración propia)



# Resultados

## Comparación con algoritmos tradicionales

**Table:** Comparación del rendimiento de la red neuronal y red neuronal profunda con los algoritmos: Naive Bayes, Regresión Logística y MVS

Algoritmo	Accuracy (%)	Precisión promedio (%)	Recall promedio (%)	F1-Score promedio (%)
Naive Bayes	68.5	67.8	64.8	64.2
Región Logística	77.4	75.0	81.6	77.1
MVS	79.9	75.6	79.8	76.7
ANN	87.0	90.8	82.5	86.0
DNN	95.3	95.0	95.0	95.3



# Tabla de Contenido

- 1 Introducción
- 2 Estado del Arte
- 3 Recursos y Herramientas
- 4 Metodología de desarrollo
  - Preprocesamiento de datos
  - Entrenamiento del modelo
- 5 Resultados
- 6 Conclusiones y Trabajo Futuro
  - Conclusiones
  - Trabajo a Futuro



## Conclusiones

- El uso de Aprendizaje profundo es conveniente para desarrollar un modelo de DNN para la predicción de cinco factores de las actividades terroristas.
- El GTD contiene clases desbalanceadas por lo que es favorable y mejora el rendimiento de la DNN utilizar la técnica de sobremuestreo, como parte del preprocesamiento de datos.
- En comparación con los métodos tradicionales de aprendizaje automático y el modelo de ANN, el modelo desarrollado de DNN resulta ser el más adecuado para el GTD.
- El GTD es un ejemplo para macrodatos, por esta razón se demuestra que el uso de DNN es adecuado, y su rendimiento mejora para procesar macrodatos.



## Trabajo Futuro

- Optimizar el modelo de predicción, mediante la evaluación y análisis de algoritmos de selección de características.
- Mejorar el aprendizaje de la red neuronal profunda mediante distintas técnicas para lidiar con clases desbalanceadas en el conjunto de datos de entrada.
- Posible aplicación en otros campos.



**Gracias**

