

Predicción de eventos comportamentales utilizando series de tiempo

REUNIÓN ANUAL DE LA UNIÓN MATEMÁTICA ARGENTINA
SECCIÓN ESTADÍSTICA Y SUS APLICACIONES
ROCÍO GUADALUPE FONSECA - ROCIO.FONSECA@MI.UNC.EDU.AR
23 DE SEPTIEMBRE DE 2022

El problema y los pasos para resolverlo

- Identificación del problema
- Adquisición de la señal
- creación de una base de datos
- Elección de una red neuronal
- Entrenamiento de la red neuronal
- Análisis de los resultados

Identificación del problema

Estudiar la dinámica temporal del comportamiento a partir de señales provenientes de videograbaciones y datos provenientes de acelerómetros

Adquisición de la señal

Se trabajo en el laboratorio utilizando codornices japonesas, cámaras y acelerómetros para conseguir las señales que luego derivarán en una base de datos para algoritmos de aprendizaje automático.

Creación de una base de datos

Características de los datos

Gran cantidad de datos por segundo

Se trabaja con codornices japonesas, que son vertebrados pequeños con alta tasa metabólica por lo que se necesita una alta resolución, se está trabajando con 25 datos por segundo. [1], [2], [4]

Series multivariadas

Se tienen datos de aceleración de tres coordenadas x , y , z y si se utiliza una sola serie se perderá información, así que se necesita trabajar con las tres series a la vez.

Eventos de interés poco frecuentes

Se estudian eventos reproductivos, sacudidas y otras conductas. Si bien estos eventos son detectables a partir de la inspección del video, debido a la alta resolución es tedioso y se tiene mayor error del observador.

La red neuronal LSTM

Las redes neuronales Long-Short Time Memory, usualmente llamadas redes LSTM, son un tipo de red neuronal recurrente, cuya característica principal es que la información puede permanecer introduciendo bucles en el diagrama de la red, por lo que, pueden recordar estados previos y utilizar esta información para decidir cuál será el siguiente paso.

Utilización de la red LSTM

Ventajas

- Tiene la capacidad de recordar información de estados anteriores, lo que las hace útiles para series de tiempo comportamentales ya que se sabe que el comportamiento animal posee correlaciones de largo alcance. [1] [2]
- Usan una serie de puertas que controlan la información que entra a la red, la que se almacena y la información que sale.

Problemas de la database

Los eventos de interés representan menos del 5% del total de los datos.

Resultados al clasificar eventos con esta red

- 97,5% de buena clasificación.
- Distingue los tres tipos de eventos
- Es necesaria una gran contribución de los estados anteriores, este caso tiene el 80%
- La tasa de aprendizaje inicial utilizada es del 0.5

Analicemos la matriz de confusión

Predicciones de la red

clase 0
(otros eventos)

263286

2386

3843

clase 1
(reproductivos)

303

47

20

clase 2
(sacudidas)

105

4

4

clase 0

clase 1

clase 2

Clase real (observada)

- 97.5% bien clasificado
- 0.1% de la clase 2 bien clasificada y 1.9% de la clase 1
- Si se clasifica todo como clase 0 se tiene 97.7%

El problema de eventos poco frecuentes

La matriz de confusión nos sugiere buscar otro método para abordar el problema, algunas de las opciones son:

- Investigar otros algoritmos de aprendizaje automático.
- Utilizar otros métodos para detectar eventos raros propuestos en la literatura. [3]
- Transformar convenientemente los datos o agregar características.

Muchas gracias

Referencias

1. Guzman DA, Flesia AG, Aon MA, Pellegrini S, Marin RH, Kembro JM. The fractal organization of ultradian rhythms in avian behavior. *Sci Rep.* 2017;7(1):684.
2. Kembro JM, Flesia AG, Gleiser RM, Perillo MA, Marin RH. Assessment of long-range correlation in animal behavior time series: The temporal pattern of locomotor activity of Japanese quail (*Coturnix coturnix*) and mosquito larva (*Culex quinquefasciatus*). *Phys A.* 2013; 392:6400-13.
3. Lilly JM. 2017 Element analysis: a wavelet-based method for analysing time-localized events in noisy time series. *Proc. R. Soc. A* 473: 20160776. <http://dx.doi.org/10.1098/rspa.2016.0776>
4. Simian C. Etogramas de alta resolución temporal y uso de acelerómetros para la detección automática de comportamientos reproductivos en codornices japonesas (*Coturnix coturnix japonica*) [Tesis de grado]. Biblioteca de la FCEFYN, UNC: Universidad Nacional de Córdoba; 2020.