

# Ejemplo de uso de Redes Neuronales Artificiales (RNAs)

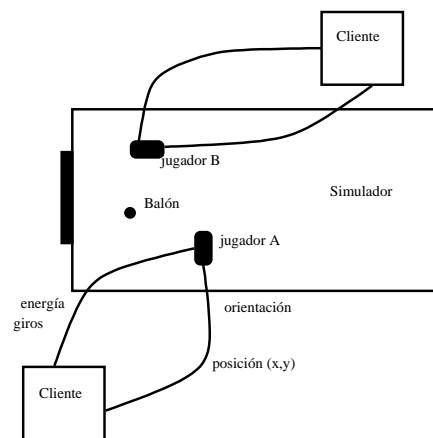
1. Enseñando a un softbot a chutar un balón
2. Escenario típico
3. La tarea  $T$
4. Experimento 1 (velocidad balón fija)
5. Experimento 2 (velocidad balón variable)
6. Experimento 3 (trayectoria balón variable)

# El simulador

## ● Propiedades

1. Realista (fricción, inercia, ruido)
2. Animado
3. Basado en cliente-servidor
4. Clientes usan planificación de rutas (seguimiento balón)

## ● Arquitectura



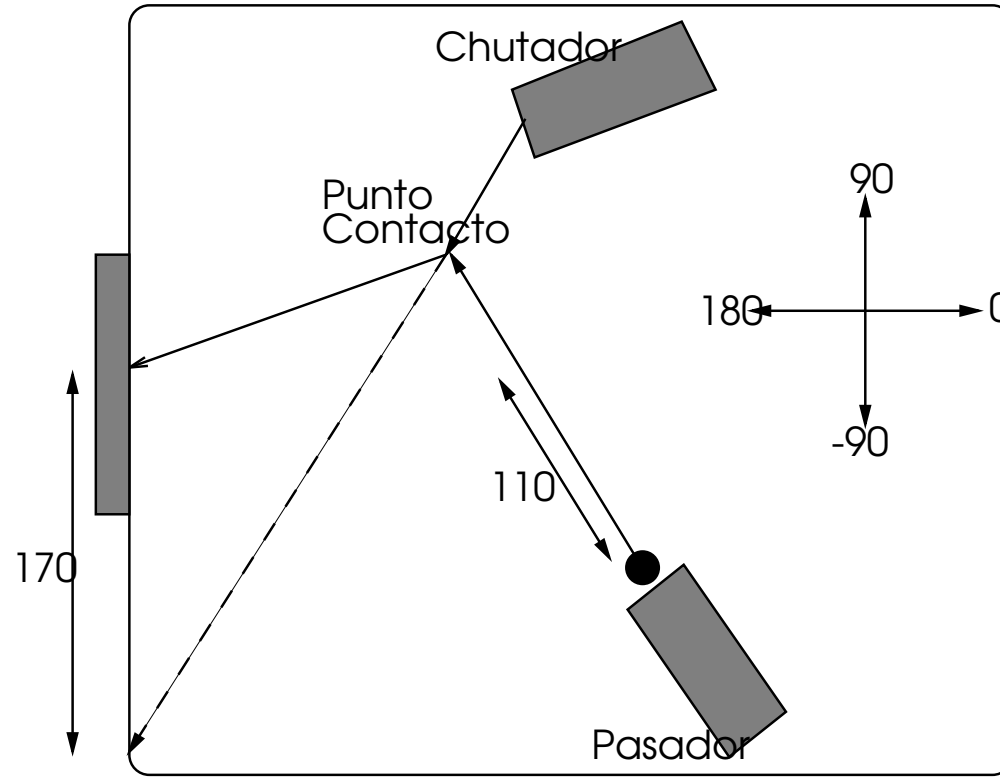
# Enseñando a un softbot a chutar un balón

- Usaremos un simulador de fútbol, construido bajo el paradigma de los sistemas multi-agente, para enseñar a un jugador a chutar.
- Vamos a hacer uso de una red neuronal artificial
- Dos tipos de agentes,
  - un agente que pasa el balón, el **pasador**, y
  - otro que chuta a partir del pase, el **chutador**.
- La velocidad inicial de la pelota en el pase depende de la distancia desde la que el pasador empieza a moverse
- El chutador debe mantener una aceleración constante, una vez que empieza a moverse.
- La tarea  $T$ , a aprender por el chutador es **decidir cuando comenzar a moverse**.

# Escenario típico

- La tarea  $T$  depende de las coordenadas del balón y del chutador  $\rightarrow (x, y, \theta)$ .
- El método mediante el cual el chutador decide sus movimientos se denomina **política de chut**
- El método mediante el que elegir la línea de dirección se denomina política de dirección (*steering policy*)
- Los parámetros  $(x, y, \theta)$  varían de forma aleatoria para generar casos diferentes
  - El rango de variación para la orientación es de  $70^\circ$ .
  - $x$  e  $y$  varían de forma independiente en 40 unidades.
- El ruido es de 2 unidades para  $x$  e  $y$  y 2 grados para  $\theta$

# Escenario típico (II)



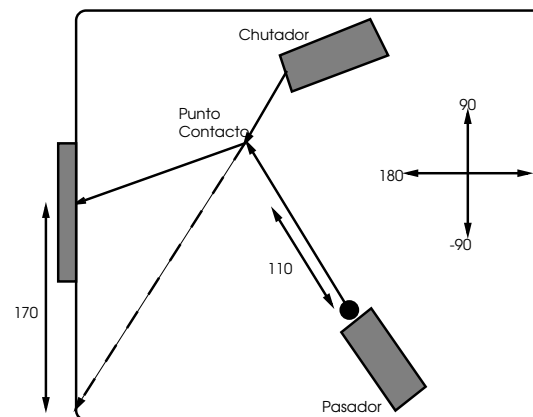
# La tarea $T$

- La dificultad en el aprendizaje de  $T$  depende de
  - Los experimentos se pueden realizar con la misma velocidad para la pelota, o con velocidad variable.
  - La bola puede moverse siempre en la misma trayectoria, o en diferentes trayectorias.
  - La portería puede estar siempre en el mismo sitio o puede moverse arriba ó abajo.
  - Tanto el entrenamiento como la evaluación se pueden realizar en el mismo sitio, o la evaluación puede moverse a otro cuadrante.
- Es posible que no sea necesario aprendizaje inductivo. Si planteamos la política de chut

*cuando el chutador se encuentra en la orientación adecuada, debería comenzar a moverse cuando la distancia de la pelota al punto de contacto es igual o menor que 110 unidades*
- Resultado  $\longrightarrow$  una tasa de aciertos del 60.8%.

# Experimento 1: Velocidad del balón fija

- En los pases el balón se mueve siempre **en la misma trayectoria** y con la **misma velocidad**, para todos los ejemplos de entrenamiento y evaluación.
- Condiciones base para marcar un tanto
  - Trayectoria de la pelota de 140 grados
  - velocidad de la pelota 135 units/sec,
  - intersección de línea de dirección del chutador y banda a 170 unidades de portería



# Experimento 1: diseñando la red

- La salida debe indicar si el chutador debe, o no, comenzar a acelerar en cada *time slice*
- Posible diseño

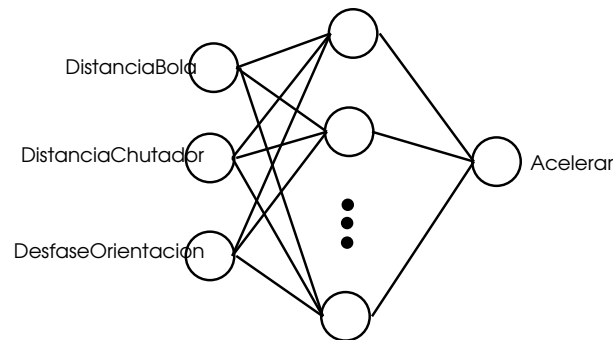
$$f(x_{pelota}, y_{pelota}, \theta_{pelota}, x_{chutador}, y_{chutador}, \theta_{chutador})$$

- Red de 6 entradas → un modelo complejo.
- Mala generalización
- Pero...



# Experimento 1: diseñando la red (II)

- Podemos predecir el punto de contacto de la pelota y el chutador.
- Con lo que podemos calcular
  - Distancia a la bola (`DistanciaBola`): distancia entre la bola y el punto de contacto,
  - Distancia del agente (`DistanciaChutador`): distancia entre el chutador y el punto de contacto
  - Desfase de orientación (`DesfaseOrientacion`): diferencia, en ángulos, entre la orientación del agente y la que considera como óptima.



# Experimento 1: conjunto de entrenamiento

- Objetivo: generar ejemplos, con alta diversidad
- Usaremos una política de chut aleatoria.

*“En cada oportunidad de chut, el chutador se moverá con una probabilidad  $\frac{1}{x}$ ”*

- Conjunto de ejemplos debe ser balanceado
- Determinamos que hay 25 ciclos de simulador que van a corresponder a oportunidades en las cuales el agente puede decidir moverse antes de que la pelota esté a menos de 110 unidades de el punto de contacto (i.e. el chut no sería exitoso entonces).
- Para obtener el valor de  $x$

$$\left(\frac{x-1}{x}\right)^{25} = 0.5$$

## Experimento 1: conjunto de entrenamiento (II)

- Tenemos que  $x$  es más o menos 37.
- Se generan 2990 ejemplos.
- Porcentaje de acierto del 19.7%.
- Solo 590 ejemplos fueron instancias positivas.

# Experimento 1: Entrenamiento y Evaluación

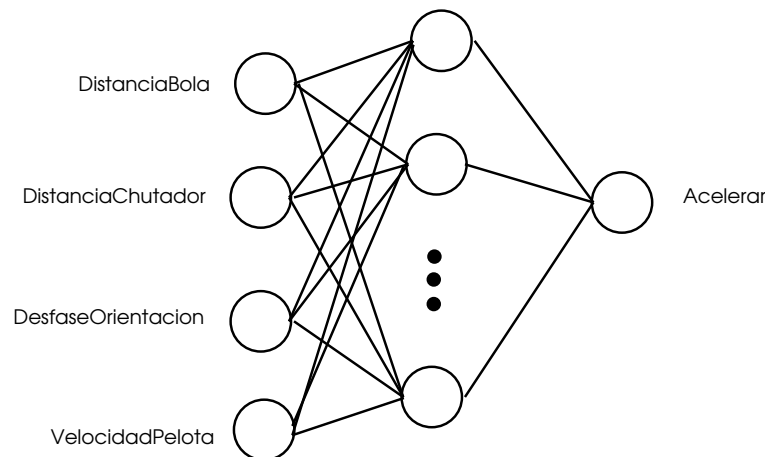
- Se probaron distintas configuraciones de redes neuronales sigmoidales.
- Se varió  $\eta$  y el número de nodos ocultos.
- Los ejemplos de entrenamiento se etiquetaron con un 0.9 y 0.1
- Los pesos  $w_{ij}$  con un valor aleatorio en  $[-0.5, 0.5]$ .
- La red neuronal óptima se obtuvo tras 3000 epochs, con un error RMSE de 0.0386.
- Mejora
  - Conforme la pelota se va acercando a una determinada distancia, la señal de la red se va incrementando hasta llegar a un pico. Una vez pasado ese pico, la señal comienza de nuevo a disminuir hasta llegar a un mínimo.

Acelear cuando  $Output \geq 0.6 \wedge Output \leq Previous - 0.1$

Posición inicial chutador	Política de chut	Éxito
Constante	Simple	100
Cambiante	Simple	60.8
Cambiante	Aleatoria	19.7
Cambiante	Regla de red de 3 nodos	96.5

## Experimento 2: Velocidad del balón variable

- Objetivo: chutador debe aprender a acelerar ante una pelota con velocidad variable.
- Variamos distancia de arranque del pasador en  $[32, 38]$  unidades
- La velocidad de la pelota variaría en  $[110, 180]$  unidades de velocidad.
- Incluimos un parámetro nuevo a la entrada de la red, `VelocidadPelota`.



## Experimento 2: Velocidad del balón variable

- La velocidad se calcula con la diferencia en posición para las 10 últimas unidades de tiempo
- Ejemplos 5737.
- El ratio de aprendizaje ha sido esta vez de 0.001 y
- el número de epochs 4000.

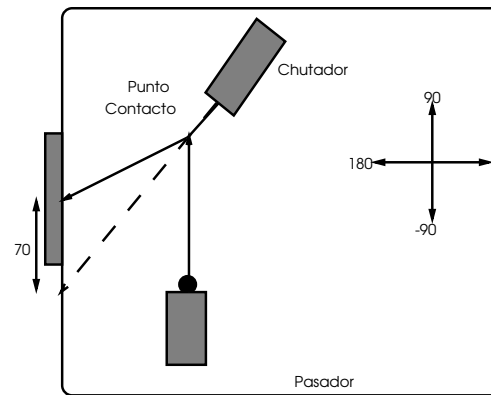
Política de chut	Aciertos
Red de 3 nodos de entrada	49.1
Aleatorio	16.8
Red de 4 nodos de entrada	91.5

# Experimento 3: Trayectoria del balón variable

## ● Hipotesis inicial

*Dado que las entradas a la red de 4 nodos, usada en el experimento 2 son relativas a PuntoContacto, la trayectoria con la que la bola se mueve hacia el chutador no va a afectar negativamente al rendimiento de la red*

## ● Test de la hipótesis → lanzar la pelota con una trayectoria de 90 grados.

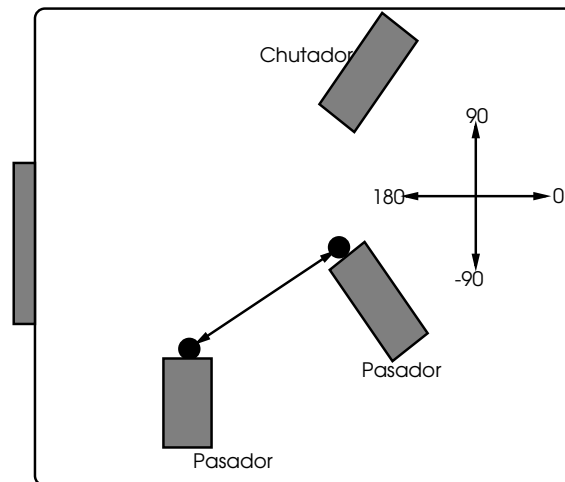


## ● Resultado → efectividad del 96.3%.

*La red de 4 entradas no solo generalizaba en diferentes áreas del campo, sino que también generalizaba a diferentes trayectorias del balón.*

## Experimento 3: Trayectoria del balón variable

- ¿Qué pasa con la variación en la dirección del chutador?
- Problema: decidir la trayectoria de chut.
- La trayectoria del balón variará en  $[82^\circ, 145^\circ]$ .
- La nueva tarea de aprendizaje será la de decidir en qué dirección debería apuntar el chutador.
- Añadiremos una nueva red neuronal a añadir a la anterior





## Experimento 3: Trayectoria del balón variable

- Objetivo → elegir los parámetros para la nueva red.
  1. AnguloPelotaChutador
  2. Salida será AmplitudAngulo
  3. Éxito en 92.8% de las ocasiones
- Si se añade VelocidadBola el rendimiento es de 95.4%.

