

The title is surrounded by a circular arrangement of decorative elements. These include several white diagonal lines of varying lengths and pairs of small triangles in teal and orange colors, some pointing inwards and some outwards.

AI Unplugged

Desconectamos la
Inteligencia Artificial

Actividades y material didáctico
sobre inteligencia artificial

Annabel Lindner
Stefan Seegerer



Prólogo

La IA (Inteligencia Artificial) se está convirtiendo en un tema de creciente importancia social. Reacciones políticas como la publicación de la estrategia de IA del Gobierno alemán a finales de 2018 son un indicador de esto. Pero lo más importante es que ya estamos interactuando con sistemas de IA como si fuera lo más natural del mundo, por ejemplo, al utilizar asistentes lingüísticos como Siri o Alexa. Sin embargo, según las encuestas, más del 50% de los alemanes no saben qué es la inteligencia artificial.

Para abordar esta cuestión, hemos redactado una colección de Actividades Unplugged relacionadas con el tema de la IA. Las Actividades Unplugged ofrecen enfoques que ayudan a alumnos de todas las edades a familiarizarse con las ideas y los conceptos de la informática de forma activa y sin necesidad de utilizar un ordenador.

Este folleto contiene cinco actividades que se pueden utilizar para enseñar ideas y conceptos de inteligencia artificial a alumnos de todas las edades.

Hoy en día, la IA se realiza principalmente a través del aprendizaje automático, pero la inteligencia artificial es mucho más que esto: La IA no sólo se refiere a aspectos técnicos, sino que también plantea cuestiones de relevancia social. Este folleto muestra las posibilidades de tratar estos temas con niños y adultos.

Si tiene alguna pregunta, comentario u observación sobre el material, no dude en ponerse en contacto con nosotros en aiunplugged@dig4all.de.

Algunas actividades necesitan material adicional. Las plantillas para imprimir se pueden encontrar aquí:

<https://aiunplugged.org>





Índice

Clasificación con Árboles de Decisión	4
#aprendizajeprofundo	8
Reinforcement Learning	12
Back to the Roots	16
El Test de Turing	20
Más Ideas	23

Clasificación con Árboles de Decisión

El juego del mono bueno y el mono malo

Grupo de destinatarios

Nivel de la escuela primaria, nivel de la escuela secundaria

De eso se trata

¿Cómo es que un ordenador puede tomar decisiones de forma autónoma? ¿Cómo decide un ordenador si una persona es deportista, si debe adquirir un préstamo, etc.? Estos procesos de clasificación son una aplicación frecuente de la IA. En esta actividad, los alumnos tienen la oportunidad de crear su propio modelo de clasificación utilizando un árbol de decisión. Al final, se selecciona el mejor de los modelos de los estudiantes para realizar otras tareas de clasificación.

Estas ideas están detrás

- La IA clasifica los datos basándose en patrones.
- La IA utiliza el modelo de clasificación que mejor se ajusta a los datos dados.
- Los modelos de clasificación no son perfectos.
- Ciertas combinaciones de características indican una determinada categoría.

Lo que se necesita

- Tarjetas de monos (preparación: recortar la plantilla, alternatively utilice la versión digital)
- Pizarra con imanes o tablón de notas



Así es como funciona

Los alumnos examinan la pertenencia de una serie de ejemplos (datos de entrenamiento) a una categoría. Para ello, desarrollan criterios que pueden utilizarse para clasificar nuevos elementos. Posteriormente, los modelos resultantes se ponen a prueba con nuevos ejemplos (datos de prueba) y se determina la precisión de la predicción.

Contexto

Somos cuidadores de animales en un zoo y nos encargamos de alimentar a los monos. Los monos son muy adorables, pero tenemos que tener cuidado porque algunos monos muerden. Ya sabemos si los monos del zoo muerden. Sin embargo, pronto se incorporarán nuevos monos al grupo y tenemos que pensar en cómo averiguar cuáles son los nuevos monos que muerden y cuáles no, preferiblemente sin acercarnos demasiado a sus dientes.

Descripción de la actividad

Según el grupo de destinatarios, se elige la versión de juego elemental con 20 tarjetas de dibujo (azul) o la versión avanzada con 40 tarjetas de dibujo (azul y verde). Estos 20 o 40 monos son todos los animales del zoo, es decir, ya sabemos si van a morder. Se dividen en datos de entrenamiento y de prueba. A partir de los datos de entrenamiento, pensamos en criterios que determinen si los monos muerden y comprobamos su fiabilidad a partir de los datos de prueba. Los datos de entrenamiento están divididos en las dos categorías morder y no morder - fijados en la pizarra. Los datos de prueba no se revelan al principio. Puedes pensar en reglas para distinguir a los monos tú mismo o utilizar una de las propuestas que aparecen a continuación (también es posible utilizar subconjuntos reducidos). Las reglas que se aplican en los ejemplos se ilustran con árboles de

Versión 1 (azul)



Versión 1 datos de entrenamiento

muerden: 6, 7, 8, 15

no muerden: 1, 2, 4, 9, 12, 14, 17, 18

Versión 1 datos de prueba

muerden: 3, 5, 11, 19

no muerden: 10, 13, 16, 20

decisión. En primer lugar, haga que sus alumnos se den cuenta de los detalles en los que podrían fijarse ilustrando el procedimiento con un ejemplo. Por ejemplo, compare las tarjetas de mono 01 a 04 y 05 a 08. En este ejemplo, la forma de la boca es un indicio para los monos que muerden, pero no los ojos (Fig. 1). Como alternativa, con los alumnos mayores, también puede utilizar la versión sencilla del juego (Versión 1) para demostrar las reglas y los procedimientos necesarios.

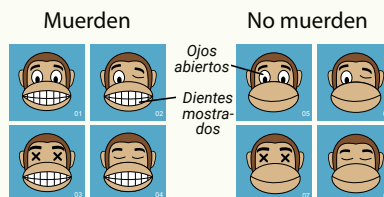
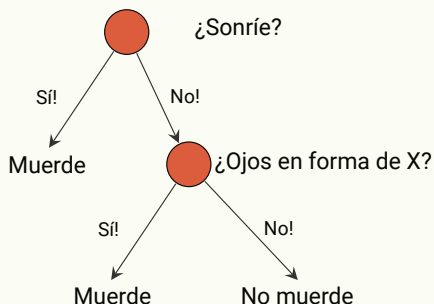
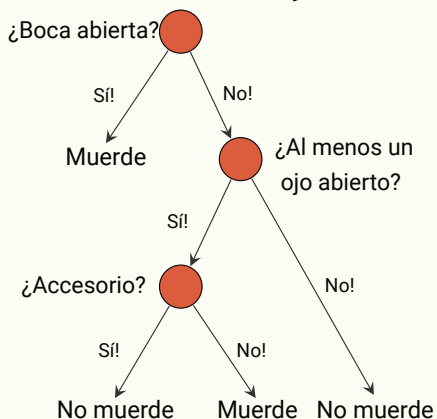


Fig. 1: En este ejemplo sencillo, todos los monos con dientes mostrados están mordiendo.

Los alumnos forman equipos de dos y utilizan los datos de entrenamiento para desarrollar criterios para distinguir los monos que muerden de los que no lo hacen. Estos criterios deben anotarse claramente para que otro equipo pueda aplicarlos posteriormente a nuevos ejemplos. Una posibilidad de registrar los criterios es un árbol de decisión. El objetivo debe ser que la existencia o ausencia de una determinada característica permita una clara asignación a uno de los grupos. El uso



Versión 2 (azul & verde)



de árboles de decisión es opcional; alternatively, también es posible anotar explícitamente las reglas de decisión. Al final de la fase de entrenamiento, los criterios formulados se intercambian con otro equipo. Ahora, se les enseña a los alumnos las imágenes de los monos restantes (datos de prueba) una tras otra. Para cada imagen, los equipos deciden si el mono morde o no utilizando el esquema de reglas desarrollado por sus compañeros. Cada equipo anota sus decisiones. Después de mostrar todos los monos, se evalúa qué equipo ha valorado mejor el comportamiento de morde de los monos. Los alumnos se dan cuenta de que muchos modelos de clasificación clasifican correctamente a la mayoría de los monos, pero que es difícil clasificar correctamente a todos los animales. Por lo tanto, para nosotros, como cuidadores de animales, es inteligente utilizar el modelo más acertado a la hora de alimentar a los nuevos monos, aunque eso no garantice que nunca nos vayan a morde.

En la versión avanzada, la imagen nº 21 (véase la Fig. 2) puede utilizarse para

Versión 2 datos de entrenamiento

muerden: 1, 2, 5, 9, 10, 14, 15, 16, 17, 28, 33, 35, 36
no muerden: 4, 7, 12, 19, 22, 23, 24, 25,
30, 32, 37, 38, 39, 40

Versión 2 datos de prueba

muerden: 6, 13, 18, 34
no muerden: 3, 8, 11, 20, 21, 26, 27, 29, 31

ilustrar los problemas de un sistema de IA cuando el valor característico de un elemento difiere significativamente de los datos de entrenamiento. No tenemos experiencia con las características de la imagen nº 21, porque este mono tiene una forma de boca nueva y desconocida. Por lo tanto, no es posible realizar una atribución adecuada del mono. En la práctica, el comportamiento de un sistema de IA es muy difícil de predecir en este caso. En lugar de la imagen nº 21 también se puede utilizar la imagen de un animal diferente para destacar las diferentes características del nuevo elemento. Posteriormente, estos ejemplos pueden aplicarse a la realidad: imprevisto, un banco no concede un crédito a un determinado cliente, el coche autopropulsado injustamente reconoce las hojas en la carretera como una situación peligrosa y frena. En estas situaciones, el sistema de IA también puede ser peligroso, si no es comprensible cómo se tomaron estas decisiones.

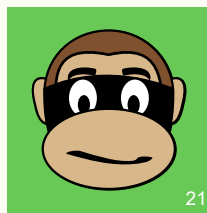


Fig. 2: Para el mono nº 21, no se puede derivar ningún criterio explícito de los datos.

Información de trasfondo

La formación de categorías es posible gracias al reconocimiento de patrones repetitivos en elementos individuales. Pero, ¿cómo se relacionan estos aspectos con la inteligencia artificial?

En el aprendizaje supervisado, el sistema de IA observa una serie de pares de entrada y salida (datos de entrenamiento) y aprende cómo se relacionan entre sí, así como qué patrones son típicos de cada categoría. Este conocimiento se utiliza entonces para clasificar nuevos elementos en las categorías. Los datos de prueba, cuyas categorías conocemos, pero el sistema de IA no, se utilizan para determinar la calidad del modelo de clasificación aprendido.

El mismo principio se utiliza para las redes neuronales y otras aplicaciones de IA. Este procedimiento puede dar lugar a diversos problemas, porque ningún

modelo es perfecto. Dependiente de los datos de entrenamiento, el modelo de clasificación puede sobreponderar o descuidar ciertas características de los datos, de modo que no es posible hacer afirmaciones generales y una clasificación correcta de los elementos desconocidos. Una gran cantidad de datos de entrenamiento puede ayudar a reducir estos efectos, pero no siempre conduce a resultados más precisos, ya que un exceso de datos de entrenamiento también puede dar lugar a un sobreajuste. En este caso, el sistema de IA aprende los datos de entrenamiento "de memoria" y ya no es capaz de generalizar a nuevos datos.

Tiene sentido abordar estos aspectos del aprendizaje automático como parte de la actividad. Cuando apliquen sus reglas en la fase de prueba, deje que los alumnos expliquen qué características han utilizado para clasificar a los monos. Esto ilustrará que los alumnos han creado diferentes conjuntos de reglas.



Fig. 3: Los datos de entrenamiento se dividen en dos categorías.

#aprendizajeprofundo

Reconocimiento de imágenes con redes neuronales

Grupo de destinatarios

Nivel de la escuela secundaria

De eso se trata

¿Cómo es posible que un ordenador "reconozca" cosas? ¿Cómo decide un ordenador si una foto muestra un gato? ¿Cómo puede distinguir los edificios de las personas? Reconocer objetos por su forma o su aspecto es muy fácil para el ser humano. Para el ordenador, que, por ejemplo, puede utilizarse en un coche autopulsado para reconocer los objetos de su entorno, esto representa una tarea compleja. En esta actividad, los alumnos tienen la oportunidad de descubrir cómo los ordenadores reconocen el contenido de las imágenes.

Estas ideas están detrás

- Las redes neuronales asignan inputs a outputs específicos: Datos, como las imágenes, se clasifican, por ejemplo, asignando términos a los objetos de la imagen.
- Las redes neuronales constan de diferentes capas de abstracción que pueden identificar características cada vez más complejas.
- Las clases de objetos a las que se quiere reconocer deben ser conocidas por el sistema de IA.

Lo que se necesita

Tarjetas con fotos de casas, gatos y coches para cada grupo.



Así es como funciona

Los alumnos reconstruyen el proceso de reconocimiento de imágenes de una red neuronal (simplificada). Asumen los papeles de las diferentes capas de dicha red. Extraen características de una fotografía y clasifican la imagen. Reconocen los límites del sistema y consideran qué modificaciones de la red son necesarias para conseguir mejores resultados con su red.

Contexto

Como humanos, nos fijamos mucho de lo que vemos. Si vemos un gato, sabemos inmediatamente que es un gato, si vemos un perro, lo reconocemos enseguida. Un ordenador, en cambio, no lo detecta tan fácilmente, pero puede aprenderlo, igual que nosotros en la infancia. Al ordenador se le muestran muchas fotos de perros, pero también de otros animales. Con esta información, aprende qué patrones de una imagen pueden servir para distinguir un perro de un gato. Si se le entrena adecuadamente, el ordenador no sólo puede etiquetar automáticamente las imágenes, sino también detectar el cáncer de piel o -integrado en un coche- reaccionar ante los obstáculos en la carretera

Descripción de la actividad

Comience por discutir cómo un ordenador podría reconocer el contenido de una imagen. Las respuestas suelen referirse a reglas definidas o a una comparación con una base de datos de imágenes, pero hoy en día los ordenadores lo hacen de otra manera. Divida a los alumnos en grupos de tres, cada grupo recibe una pila de tarjetas fotográficas. En cada grupo hay tres rollos, cada uno de los cuales representa una capa de una red neuronal (véase la Fig. 4). Las tareas de los roles son las siguientes:

A escoge una imagen de la pila de tarjetas fotográficas (¡B y C no deben ver

la imagen!), crea dos bocetos diferentes de la misma (30 segundos cada uno) y se los pasa a B. Es importante que C no vea los bocetos.

B recibe los bocetos de A y comprueba si se incluyen formas cuadradas, triangulares o redondas. A continuación, B transmite la información recogida a C.

C evalúa la información recibida utilizando la siguiente tabla y anuncia si la foto original es una casa, un coche o un gato.

	Forma cuadrada?	Forma triangular?	Forma redonda?
Casa	Sí	Sí	No
Coche	Sí	No	Sí
Gato	No	Sí	Sí

Finalmente, A determina si la solución es correcta.

Deje que los alumnos prueben el juego en diferentes roles. También es posible que un papel sea desempeñado por dos alumnos o que se asigne explícitamente cada papel a tres alumnos, que representen cada uno una neurona (es decir, un nudo de la capa) y ya no una capa entera.

Después de un breve ensayo, reparta entre los alumnos algunas imágenes que no correspondan a las categorías que la red puede reconocer o que tengan características que no permitan una clasificación clara. Por ejemplo, una imagen de un perro no es reconocida correctamente por la red, simplemente

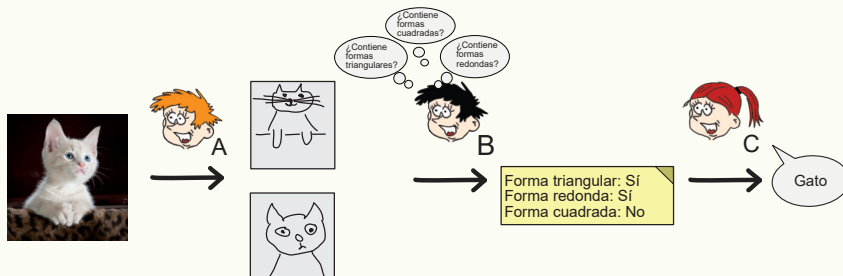


Fig. 4: Los papeles de los alumnos

porque la red no conoce la categoría "perro". A partir de este hallazgo, los estudiantes están estudiando ahora cómo se puede cambiar y ampliar la red para que reconozca perros u otros objetos en el futuro. En primer lugar, hay que introducir una nueva categoría de output. Pero al mismo tiempo, el número de características identificadas por la red ya no es suficiente. En consecuencia, se añaden más características que permitan distinguir las categorías, o se combinan varias características para formar un patrón más complejo. Esta combinación implica, en última instancia, la adición de más capas en la red neuronal. La Fig. 7, por ejemplo, puede utilizarse para ayudar a los alumnos a entender cómo combinar características simples para formar patrones más complejos.

Información de trasfondo

Cuando es difícil traducir un problema concreto en reglas lógicas, las redes neuronales artificiales se utilizan a menudo en la resolución de problemas. Ejemplos típicos son la comprensión de textos o el reconocimiento de objetos en imágenes. La idea de diseño de las redes neuronales artificiales tiene su origen en la neurobiología y se basa en la estructura del cerebro humano. De forma análoga a una célula nerviosa humana, que procesa varios estímulos y transmite un impulso, una neurona artificial también trata varias entradas y puede transmitir una señal. A las aristas de input se les asigna un peso, es decir, tienen una influencia variable en el output de la neurona. Una neurona artificial es, pues, algo parecido a una neurona humana, pero funciona más bien como una simple calculadora: multiplica los pesos de los bordes y los valores de entrada, los suma y transmite un resultado.

Del mismo modo que en el sistema nervioso humano, muchas neuronas artificiales están conectadas y forman una red de esta manera. Las neuronas se organizan en capas. Dependiendo de la complejidad de un problema, una red puede constar de dos o más capas. En la situación inicial de esta actividad, hay

tres capas. Si una red tiene más capas entre la de input y la de output, esto se llama aprendizaje profundo.

En la práctica, el reconocimiento y la clasificación de imágenes suelen funcionar mediante las llamadas redes neuronales convolucionales, que están especializadas en el reconocimiento de patrones y, por tanto, son muy adecuadas para clasificar imágenes. Este tipo de redes neuronales se caracterizan por utilizar las llamadas convoluciones para extraer características y patrones de los datos de input. Hoy en día, estas redes pueden clasificar imágenes más rápido que los humanos.

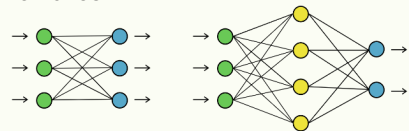


Fig. 5: Ejemplo de redes neuronales artificiales - izquierda: red simple, derecha: red "más profunda".

¿Cómo funciona exactamente? Las fotos digitales se componen de pequeños elementos de color -píxeles- dispuestos en una cuadrícula. Cada píxel tiene un valor de color determinado. A diferencia de los humanos, para el ordenador las fotos son sólo valores numéricos. Inicialmente, las redes de reconocimiento de imágenes tratan de reconocer características simples. Para ello, se colocan filtros sobre la imagen. Esto es similar a lo que hacemos en los programas de edición de fotos, por ejemplo, cuando utilizamos un filtro de paso alto (véase la Fig. 6). En definitiva, se trata de un cálculo matemático, que captura varios píxeles y calcula un nuevo píxel.



Fig. 6: Aplicación de un filtro



Fig. 7: Las redes neuronales identifican características cada vez más complejas

Dependiendo del filtro, se puede detectar, por ejemplo, se pueden juntar los píxeles con valores de brillo similares para formar bordes. En un nivel posterior, se extraen características como líneas horizontales y verticales, círculos o esquinas. Los programas habituales de tratamiento de imágenes, como Gimp, permiten introducir estos filtros como una matriz, lo cual es una buena posibilidad para explorar estos efectos.

En el juego, los bocetos realizados por los alumnos sirven también de filtro, ya que extraen los elementos centrales del objeto representado en la foto. A continuación, los bocetos se utilizan para identificar las formas geométricas en las imágenes. Sin embargo, utilizar sólo tres características es una gran simplificación en comparación con una red neuronal real, que tiene varios millones de neuronas en una multitud de capas.

En los niveles iniciales de una red neuronal, hay una multitud de filtros simples y bastante geométricos. A continuación, estos patrones se combinan (también mediante la aplicación de filtros) en patrones más complejos. En los "niveles más profundos" no sólo se reconocen las esquinas y los bordes, sino también partes de los objetos, como los ojos o el pelo, y finalmente incluso objetos completos, como perros o gatos.

Mientras se procesa, se descarta repetidamente información (superflua), ya que, por ejemplo, la posición exacta de una línea diagonal en una imagen tiene poco interés para el reconocimiento de un objeto en muchos casos. Al final, un valor de probabilidad indica la probabilidad de que una imagen pueda asignarse a una categoría específica.

Sin embargo, una red neuronal no puede reconocer fácilmente el contenido de cualquier imagen. Más bien, el marco de aplicación es muy limitado: La red neuronal debe ser "entrenada" primero con un número muy grande de imágenes (varios miles). Aprende qué características son decisivas para las imágenes que pertenecen a una determinada categoría. Así, la red neuronal sólo puede clasificar correctamente las imágenes cuya categoría ya conoce. Por ejemplo, una red que debe distinguir entre perros y gatos no puede reconocer otros animales, sino que los clasifica en una de las dos categorías conocidas. Sin embargo, una red neuronal entrenada puede cumplir su tarea mucho más rápido que un ser humano. Por eso, los métodos de reconocimiento de imágenes ya se utilizan, por ejemplo, en los coches autopropulsados para detectar diversos objetos en el tráfico rodado (tráfico en sentido contrario, peatones, etc.) o en la detección del cáncer de piel.

Reinforcement Learning

Vence al cocodrilo

Grupo de destinatarios

Nivel de la escuela primaria, nivel de la escuela secundaria

De eso se trata

Ya conocemos los ordenadores que pueden jugar al ajedrez y vencer a los jugadores humanos con una facilidad superior. El juego de tablero chino Go, en cambio, se ha considerado tan complejo que solo un humano puede dominarlo... hasta que Google utilizó AlphaGo para amedrentar a los jugadores humanos profesionales. En esta actividad, vemos cómo los ordenadores pueden aprender estrategias para los juegos, aunque solo conozcan las reglas del juego.

Estas ideas están detrás

- Los ordenadores pueden aprender por "recompensa" y "castigo".
- Los ordenadores evalúan los beneficios de las acciones aleatorias basándose en la recompensa y el castigo.
- Los ordenadores aprenden estrategias o secuencias de acciones buscando la máxima recompensa.

Lo que se necesita

- Por cada pareja de alumnos: 1 tablero de "mini ajedrez", 3 tarjetas de mono y 3 de cocodrilo, 1 resumen de posibles jugadas
- Caramelos de distintos colores (por ejemplo, fichas de chocolate) o fichas de papel para evaluar las jugadas en 4 colores diferentes (amarillo, rojo, naranja, azul; aprox. 20 por color)



Así es como funciona

Dos estudiantes juegan una partida de "mini ajedrez" entre ellos. Uno de los estudiantes asume el papel de un ordenador "de papel". Al principio, el ordenador selecciona sus jugadas al azar, pero poco a poco aprende con un sistema de fichas de caramelo qué jugadas le ayudan a ganar y cuáles acaban en derrota. Utilizando la estrategia que se desarrolla de este modo, el ordenador va mejorando con el tiempo.

Esta actividad se basa en una idea de CS4Fun. (<http://www.cs4fn.org/machinelearning/sweetlearningcomputer.php>)

Contexto

¿Cómo aprendemos a jugar a los juegos de mesa o a los videojuegos? Quizá observamos a otros jugando o probamos cómo ciertas acciones o movimientos influyen en la partida. Como ganamos con más frecuencia, mejoramos en el juego. Desarrollamos estrategias para determinar qué jugadas tienen más éxito en determinadas situaciones del juego. Del mismo modo, un ordenador aprende a jugar.

Descripción de la actividad

El juego sigue unas sencillas reglas de ajedrez: Cada pieza se mueve como un peón, es decir, sólo puede avanzar en línea recta y sólo puede vencer a las piezas contrarias en diagonal. Un alumno se hace cargo de los monos y actúa como jugador humano. Otro alumno asume el papel del ordenador en forma de cocodrilos.

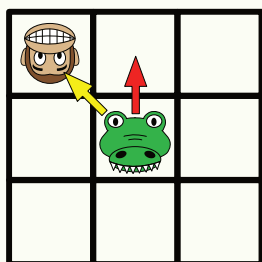


Fig. 8: Movimientos posibles de una pieza

Un bando ha ganado si consigue...

- llevar una pieza al otro lado del tablero de juego.
- vencer a todas las piezas contrarias.
- que el adversario no pueda hacer más movimientos en la siguiente jugada.

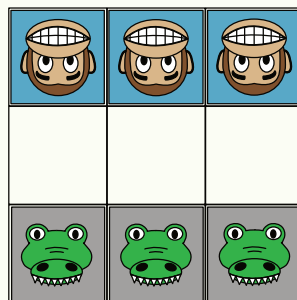


Fig. 9: Tablero antes de que comience el juego

Como preparación, los impresiones de las opciones de movimiento del ordenador se pone delante del jugador que se hace cargo de los cocodrilos. A continuación, se distribuyen fichas de chocolate (o de otro material) sobre estos movimientos. Para cada flecha de color, se coloca una ficha del color correspondiente en la zona de la derecha de cada situación de juego (ver Fig. 11).

El jugador humano comienza. Puede moverse libremente según las reglas del juego. Luego es el turno de los cocodrilos. El jugador compara el tablero actual con los movimientos posibles y selecciona la situación de juego adecuada entre las alternativas dadas. Para una orientación más rápida, se indica el turno al que pertenece cada situación. En la primera ronda sólo hay que considerar las dos posibilidades del turno 1, en la segunda ronda las 10 jugadas del turno 2 y en la tercera ronda las 7 jugadas del turno 3. Las situaciones de juego simétricas no se indican dos veces. A continuación, el jugador del cocodrilo cierra los ojos y elige al azar una de las fichas colocadas junto a la situación de juego correspondiente y la desplaza al tablero representado. El color de la ficha determina qué movimiento se realiza y el jugador mueve la ficha según la flecha del mismo color. Si, por ejemplo, se saca

una ficha roja, el cocodrilo se mueve siguiendo la flecha roja.

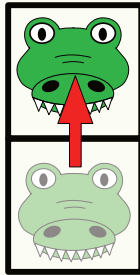


Fig. 10: El cocodrilo se mueve a lo largo de la flecha roja.

Este procedimiento se repite hasta determinar el ganador de la ronda. Antes de jugar una nueva ronda, el ordenador ajusta su estrategia de la siguiente manera:

- Los cocodrilos han ganado: Una ficha adicional del color del último turno ganador se coloca en la casilla de ese turno.
- Los monos han ganado: La ficha de chocolate que determinó el último movimiento del jugador del cocodrilo se retira. El jugador del mono puede comerlo.

Además, todas las fichas se colocan de nuevo a la derecha del campo de juego correspondiente.

Opcional: Para simplificar las reglas, también puedes dejar de añadir fichas adicionales cuando los cocodrilos ganen.

Información de trasfondo

Al principio, el ordenador apenas tendrá posibilidades de ganar porque elige sus acciones al azar (eligiendo una ficha con

los ojos cerrados). Cuantas más partidas termine el ordenador, mejor será: "aprende" qué movimientos ayudan a ganar y qué movimientos deben evitarse porque acabaron en derrota en el pasado. De este modo, la estrategia del ordenador se va perfeccionando poco a poco.

Dado que el ordenador es castigado por perder y recompensado por ganar, también se habla de aprendizaje por refuerzo, es decir, de aprendizaje por recompensa y castigo:

- Castigo = Quitar un caramelo en una partida que ha llevado a la derrota.
- Refuerzo = Añadir un caramelo en un turno que llevó a la victoria.

Este procedimiento "clasifica" las jugadas que resultaron en una derrota, de modo que en algún momento sólo quedan las jugadas "buenas". En la práctica, las estrategias que no llevan al éxito no se eliminan inmediatamente, sino que sólo se reduce la probabilidad de que se produzcan. De este modo, el sistema de IA aprende gradualmente qué estrategia es la más adecuada en qué situaciones, pero no excluye inmediatamente las estrategias individuales que no han conducido al éxito en todos los casos. Aunque este procedimiento se simplifica en el juego al eliminar inmediatamente las jugadas que han llevado a la derrota, nunca puede ocurrir que se eliminen todas las jugadas posibles para una situación de juego. Para cada situación hay al menos una acción posible que no conduce a una derrota inmediata.

De este modo, los ordenadores pueden aprender a ganar un juego simplemente conociendo las reglas del juego o sus posibles inputs. Por ejemplo, si un ordenador aprende a jugar al videojuego

Super Mario, al principio sólo pulsará las teclas al azar. Esto podría hacer que el ordenador se detuviera durante minutos o se encontrara con el mismo oponente varias veces. Analiza los objetos o píxeles de la imagen y reacciona con inputs. Su objetivo es maximizar los puntos conseguidos en el juego, que actúan como recompensa. Cuanto más se mueva el ordenador hacia la derecha, mayor será la ganancia. Con el tiempo, aprenderá, por ejemplo, que saltar aumenta su recompensa si un oponente está inmediatamente a su derecha, ya que avanza más en el nivel saltando sobre el oponente. De este modo, el rendimiento de un sistema de IA mejora poco a poco en una partida, porque el sistema siempre intenta maximizar su recompensa (o más exactamente: una determinada función).

Como parte de la descontextualización, haga que los alumnos analicen cómo se desarrolla el comportamiento del ordenador. Debe quedar claro que el ordenador llega a una estrategia de juego eficiente evaluando acciones puramente aleatorias. Después, por ejemplo, se puede utilizar un vídeo sobre el juego Super Mario (ver página web) para mostrar cómo se produce el aprendizaje por refuerzo en una red neuronal. Deje que los alumnos reflexionen sobre los límites de las estrategias aprendidas por el sistema. Esta actividad puede combinarse muy bien con el juego *Volver a las raíces: El ajedrez de los cocodrilos* y la IA clásica para destacar el contraste entre los sistemas de aprendizaje y las aplicaciones tradicionales de la IA, como los sistemas basados en reglas.

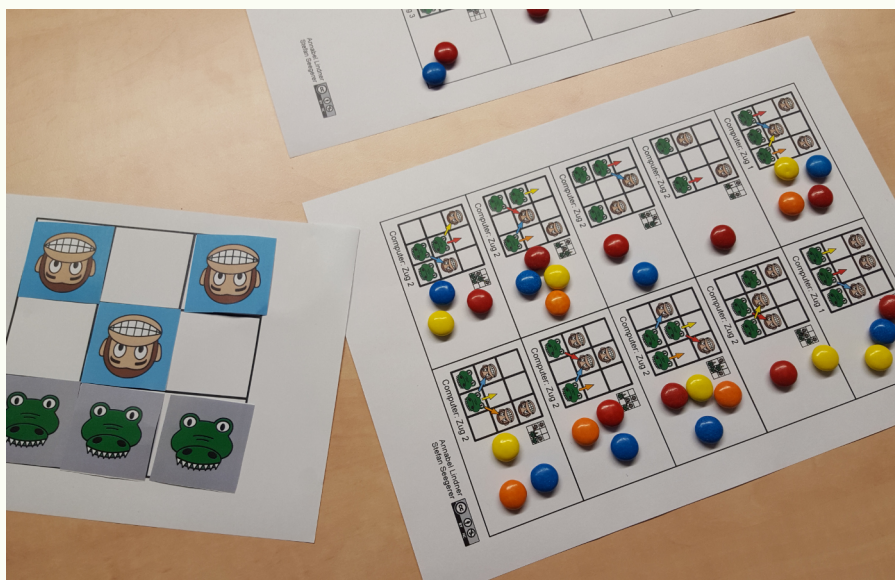


Fig. 11: Configuración del juego: La distribución de las fichas de chocolate muestra las estrategias aprendidas.

Back to the Roots

El ajedrez de los Cocodrilos y la IA clásica

Grupo de destinatarios

Nivel de la escuela primaria, nivel de la escuela secundaria

De eso se trata

Los ejercicios anteriores tratan de sistemas de IA que aprenden. Pero eso no es todo lo que ofrece la IA: Los orígenes de la IA se encuentran en la lógica y en la idea de formalizar el conocimiento mediante una descripción matemática, poniéndolo así a disposición de las máquinas. En esta actividad se muestran las diferencias entre la IA que puede aprender y los enfoques tradicionales, así como los límites de estos sistemas. Para ello, la actividad precedente sobre el Aprendizaje por Refuerzo se realiza con un sistema experto y así se ilustra los enfoques tan diferentes.

Estas ideas están detrás

- El conocimiento debe ser representable de manera formal para poder ser procesado automáticamente.
- Los sistemas expertos pueden combinar reglas y hechos para generar nuevos conocimientos.
- Estos sistemas de IA no toman decisiones independientes, sino que trabajan según las reglas de la lógica.
- Los sistemas de IA disponen de mecanismos de procesamiento para deducir automáticamente la información de los conocimientos existentes.

Lo que se necesita

- Por pareja de alumnos: 1 tablero de "mini ajedrez", 3 cartas de mono y 3 de cocodrilo, 1 resumen de las reglas para la siguiente jugada

Así es como funciona

Como en la actividad de Aprendizaje por Refuerzo, dos estudiantes juegan una partida de "mini ajedrez". Uno de los alumnos asume el papel de ordenador "de papel". Si, como se recomienda, esta actividad se combina con la versión anterior de Aprendizaje por Refuerzo, es una buena idea intercambiar los papeles. Sin embargo, en lugar de elegir sus movimientos al azar, el ordenador trabaja ahora según unas reglas predefinidas, que se ponen a su disposición en forma de copias.

Contexto

¿Cómo se puede programar un ordenador para que juegue a juegos de mesa o de vídeo? Los ordenadores sólo pueden "entender" las reglas de un juego y actuar en consecuencia si éstas se presentan de forma que un ordenador pueda procesarlas. Por lo tanto, el conocimiento debe representarse formalmente para que pueda ser procesado por la máquina (por ejemplo, mediante términos matemáticos). En este caso, el ordenador puede evaluarlo con la ayuda de la lógica y derivar sus acciones de ello. Por tanto, los sistemas de IA no son realmente inteligentes, sino que utilizan ingeniosamente distintas posibilidades para derivar su comportamiento de los conocimientos disponibles.

Descripción de la actividad

El juego sigue unas reglas de ajedrez sencillas y tiene las mismas condiciones básicas que las descritas en la actividad de aprendizaje por refuerzo: Cada pieza se mueve como un peón, es decir, sólo puede moverse en línea recta y sólo puede golpear a las piezas del adversario en diagonal. Un alumno se hace cargo de los monos y actúa como un jugador humano. Otro alumno asume el papel del ordenador en forma de los cocodrilos.

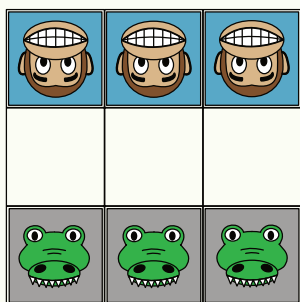


Fig. 12: Tablero antes de que comience el juego

Un bando ha ganado si consigue...

- llevar una pieza al otro lado del tablero de juego.

- vencer a todas las piezas contrarias.
- que el adversario no pueda hacer más movimientos en la siguiente jugada.

En preparación, el jugador que se hace cargo de los cocodrilos recibe una copia de las reglas de los movimientos del ordenador. Éstas sustituyen a las opciones de movimiento y a las fichas de la actividad de aprendizaje por refuerzo. El jugador humano comienza. Puede moverse libremente según las reglas del juego. Luego es el turno de los cocodrilos. El jugador compara la situación actual del juego con la "tabla de reglas" y selecciona el escenario apropiado entre las 10 opciones, las situaciones simétricas no aparecen dos veces. A continuación, realiza el movimiento que exige la regla.

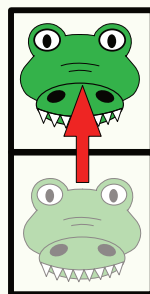


Fig. 13: El cocodrilo se mueve junto a la flecha roja

Este procedimiento se repite hasta que se determina un ganador. Se pueden jugar varias rondas para comprobar si el ordenador es capaz de ganar siempre con la ayuda de sus reglas.

A continuación, presente a los alumnos la versión ampliada del juego de ajedrez del cocodrilo con 4x4 campos (véase el sitio web). Destaque que las reglas de las que dispone el ordenador ya no son suficientes. Deje que los alumnos establezcan una comparación con la variante de ajedrez en miniatura de la actividad de aprendizaje por refuerzo, en la que el ordenador está aprendiendo: incluso aquí el "conocimiento" del ordenador ya no es suficiente. Aquí es

donde ambos sistemas alcanzan sus límites.

Ahora los alumnos consideran qué procedimientos son necesarios para adaptar los "conocimientos" del ordenador a la versión 4x4 del juego. Se dan cuenta de que el ordenador basado en reglas tiene que ser ajustado manualmente por los humanos añadiendo nuevas reglas sobre las mejores jugadas. En comparación, el sistema de aprendizaje puede aprender el mejor comportamiento para el campo 4x4 de la misma manera que lo hizo en la actividad anterior, es decir, evaluando el comportamiento aleatorio, tan pronto como se hayan añadido todas las nuevas jugadas posibles. Por tanto, el sistema de aprendizaje necesita una nueva fase de entrenamiento en la que se adquieren implícitamente los conocimientos para el juego ampliado, mientras que las nuevas reglas deben añadirse explícitamente al sistema experto basado en reglas. En este caso, el ser humano primeramente tiene la tarea de determinar la mejor jugada para cada situación de juego y luego formalizar todas las reglas de forma completa y exhaustiva. Esto no es necesario para el sistema de aprendizaje. Esta representación formal explícita de las reglas que controlan las acciones del sistema es muy difícil y en el caso de problemas complejos y con múltiples pasos puede ser imposible. En este contexto, los sistemas de IA de aprendizaje -en comparación con los sistemas expertos- ofrecen la gran ventaja de que pueden determinar tales procedimientos de forma "independiente". Además, pueden incluso detectar correlaciones en los datos que no pueden ser identificadas por los humanos, por ejemplo, porque es necesario considerar miles de registros de datos para encontrarlas.

Información de trasfondo

La lógica y el procesamiento del conocimiento desempeñan un papel importante en muchas áreas de la informática y son, además, temas centrales de la inteligencia artificial.

Dado que el lenguaje natural es ambiguo y demasiado diverso para ser un medio apropiado para hacer accesible el conocimiento a las máquinas, la cuestión sobre la mejor representación posible para las máquinas es de importancia decisiva desde los inicios de la IA.

En este contexto, los métodos de esta forma tradicional de IA se basan en la representación simbólica del conocimiento, es decir, en la representación explícita del conocimiento en los sistemas informáticos, por ejemplo con la ayuda de la lógica. Esto permite la representación inequívoca, uniforme y precisa del conocimiento, necesaria para el procesamiento con un ordenador. Este tipo de métodos de representación se utilizan, por ejemplo, en los sistemas expertos basados en reglas, que todavía hoy desempeñan un papel en las aplicaciones comerciales. En estos sistemas, las afirmaciones lógicas que representan hechos y conocimientos sobre reglas se utilizan para sacar conclusiones sobre cómo tiene que actuar el ordenador automáticamente. En esta actividad, los hechos corresponden a la situación actual del juego y el conocimiento de las reglas a las instrucciones de qué movimiento es útil.

La base de hechos representa afirmaciones válidas. Un conjunto de reglas representadas en forma de "si... entonces" constituye la base de reglas. Formalmente, esta forma "si... entonces" puede expresarse, por ejemplo, mediante la lógica proposicional. Un sistema de control (motor de inferencia) selecciona las reglas adecuadas a partir de los hechos, las evalúa y actúa en consecuencia. En los sistemas expertos complejos, las conclusiones extraídas de las reglas pueden servir también como hechos de entrada para otras

reglas y contribuir así a la ampliación de la base fáctica. En esta actividad, la tarea del sistema de control es asumida por el alumno, que desempeña el papel de ordenador. Con la ayuda de las descripciones de las jugadas, que representan la base de reglas, tiene que deducir la siguiente jugada a partir de la situación actual del juego.

En un sistema basado en reglas, este procedimiento se denomina encadenamiento hacia delante o guiado por datos, porque se intenta alcanzar un objetivo aún desconocido a partir de hechos. En cambio, existe el encadenamiento hacia atrás, que intenta demostrar una hipótesis. En el juego se juega "hacia atrás". Partiendo de una situación de juego en la que el ordenador ha ganado, intentamos deducir los movimientos necesarios para llegar a ella.

Incluso el aprendizaje automático, que es el método dominante en la inteligencia artificial actual y que ya ha

sustituido a los sistemas expertos y a otras aplicaciones tradicionales de la IA en muchos ámbitos, no puede prescindir de la representación del conocimiento. En tecnologías como las redes neuronales, sin embargo, esto se hace de forma implícita, por lo que se habla de sistemas subsimbólicos en este caso: Se entrena un comportamiento sistemático y se adquiere así una especie de conocimiento implícito sobre las correlaciones subyacentes. Sin embargo, es difícil conocer los procesos de solución concretos de estas redes, ya que las reglas generales que sustentan los datos sólo están representadas indirectamente en la red neuronal, por ejemplo en los pesos de los bordes y los umbrales de activación de las neuronas. Se puede encontrar más información sobre las redes neuronales en la actividad [#deeplearning](#). CS4FN también ofrece otra actividad que modela un sistema experto para las suertes y cruces de una manera no conectada.

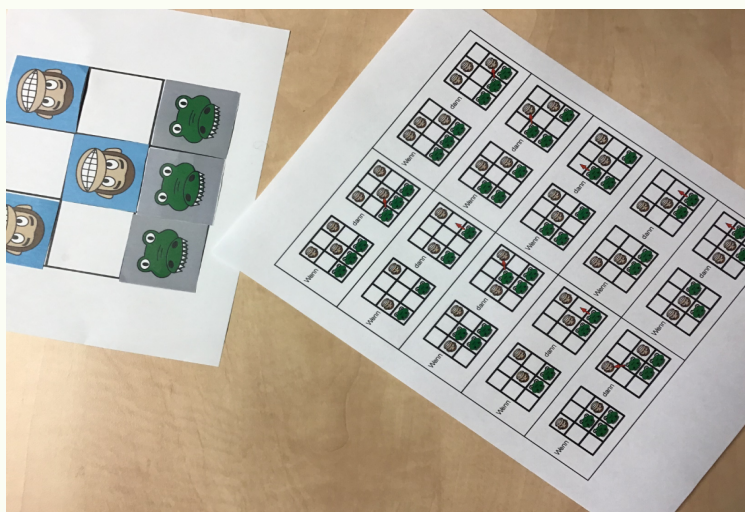


Fig. 14: Configuración del juego: Las reglas para la persona que se hace cargo del ordenador están claramente definidas.

El Test de Turing

"¡Oh, qué bueno es que nadie sepa que soy un ordenador!"

Grupo de destinatarios

Nivel de la escuela secundaria

De eso se trata

¿Cómo debe comportarse una máquina para ser considerada inteligente? ¿Qué significa exactamente la inteligencia artificial? Los investigadores llevan trabajando en estas cuestiones desde los inicios de la Inteligencia artificial. Con el test de Turing, Alan Turing tuvo una idea sobre cómo determinar si una máquina es inteligente en 1950. Esta actividad recrea el test de Turing con los alumnos y quiere estimular el debate sobre si los ordenadores pueden mostrar algo parecido a la inteligencia humana. También revela lo fácil que es dejarse engañar por una máquina mediante ejemplos de "inteligencia" bien elegidos.

Estas ideas están detrás

- Los sistemas inteligentes utilizan determinadas estrategias para imitar el comportamiento humano.
- Se necesitan métodos especiales para evaluar la inteligencia de las máquinas.
- La definición de inteligencia (artificial) no está clara.

Lo que se necesita

- Hojas de trabajo o transparencias con las preguntas del test de Turing para toda la clase
- Una copia de las respuestas a las preguntas del test de Turing
- 4 estudiantes voluntarios en los papeles de ordenador (1x), persona (1x) y corredores (2x)

Así es como funciona

En esta actividad, los alumnos juegan a un juego de preguntas y respuestas en el que intentan distinguir un ordenador de un ser humano haciendo preguntas y analizando las respuestas. Un alumno asume el papel de ordenador y otro simplemente reacciona como una persona. Son interrogados por sus compañeros y la clase tiene que determinar quién representa cada papel basándose en sus respuestas.

Esta actividad tiene su origen en los materiales originales de CS Unplugged. Estos están licenciados bajo Creative Commons CC-BY-SA por Bell, Witten y Fellows. El material original ha sido adaptado en esta descripción, que también tiene licencia CC-BY-SA.

Contexto

Durante siglos, los filósofos han discutido si una máquina es capaz de alcanzar la inteligencia humana o si el cerebro humano es quizás sólo una máquina muy buena. Algunos piensan que la inteligencia artificial es una idea absurda, otros creen que acabaremos desarrollando máquinas tan inteligentes como nosotros. La inteligencia artificial tiene mucho potencial, pero por otro lado, la idea de máquinas inteligentes también alimenta temores.

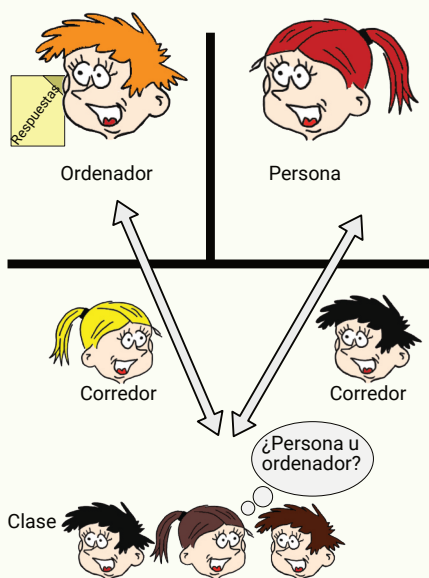


Fig. 15: Diseño del test de Turing

Descripción de la actividad

Antes de empezar el juego, discuta con los alumnos si consideran que los ordenadores son inteligentes o si suponen que los ordenadores serán inteligentes algún día. Pregúnteles cómo se decide si un ordenador es inteligente y presente brevemente el test de Turing, que se simula en la actividad.

Para preparar la actividad, se seleccionan cuatro voluntarios que asumen los papeles de un ordenador y una persona (véase la figura 15). Además, hay dos corredores que aseguran el buen desarrollo del juego y están equipados con un papel y un bolígrafo para anotar las respuestas. El profesor asigna en secreto los papeles de "persona" y "ordenador" antes de que estos dos alumnos salgan del aula y se dirijan a dos salas distintas (también se pueden utilizar tabiques, pero hay que asegurarse de que los alumnos no se vean). El alumno que asume el papel del ordenador recibe una copia de las respuestas a las preguntas del test de Turing. Cada uno de los corredores es responsable de un papel, que también se mantiene en secreto.

Ahora, la clase tiene que averiguar qué alumno ha asumido el papel del ordenador. Para ello, seleccionan una pregunta por ronda de la hoja de trabajo distribuida, que se formulará al ordenador y a la persona. Una vez elegida la pregunta, los alumnos deben explicar por qué consideran que esa pregunta es adecuada para distinguir al ordenador del ser humano. Esta argumentación es el elemento central de la tarea, ya que la clase reflexiona sobre cómo pueden diferir las respuestas de una persona y de un ordenador "inteligente".

A continuación, los corredores plantean la pregunta a sus compañeros en las otras salas y vuelven a la clase con sus respuestas. La persona está obligada a responder a la pregunta de forma breve y honesta, es decir, a dar una respuesta humana. El ordenador, por su parte, selecciona la respuesta adecuada de la hoja de trabajo. Si las instrucciones están escritas en cursiva, el ordenador tiene que elaborar una respuesta por sí mismo (por ejemplo, la hora actual). Al transmitir las respuestas dadas, los corredores deben tener cuidado de no revelar con quién están interactuando.

La clase discute ahora qué repuesta probablemente provenga de un ordenador. Repita el proceso con algunas preguntas más, si es posible, hasta que la clase pueda tomar una decisión clara sobre quién es el ordenador. Si la clase no puede distinguir con seguridad entre un humano y un ordenador, el ordenador ha pasado la prueba de Turing.

Información de trasfondo

Aunque ningún programa informático actual dispone de algo parecido a la inteligencia general, la cuestión de si los ordenadores son básicamente capaces de ella sigue sin respuesta. Esto se debe principalmente al hecho de que la propia definición de inteligencia se discute de forma polémica.

En este contexto, el matemático británico Alan Turing propuso en 1950 un método para determinar la inteligencia de una máquina sin necesidad de una definición exacta de inteligencia. El llamado test de Turing permite al ordenador demostrar su "inteligencia". El escenario de la prueba es similar a la actividad descrita anteriormente: Un interrogador interactúa tanto con una persona como con un ordenador a través de un chat. Si no puede distinguir de forma fiable entre ambos, el ordenador ha pasado la prueba de Turing. Dado que la comunicación tiene lugar a través del chat, el ordenador no puede revelar sus características físicas, como el tono de voz. Un ejemplo muy conocido de este sistema de interacción es el chatbot Eliza. Las respuestas dadas por un estudiante en el papel del ordenador no son muy diferentes a las dadas por un

programa informático "inteligente". Algunas de las respuestas dejarán en evidencia al ordenador muy rápidamente: un humano apenas podrá dar la raíz de 2 a 20 dígitos. Otras preguntas, en las que el ordenador utiliza siempre un determinado patrón de respuesta, lo revelarán sólo después de algún tiempo. Por ejemplo, las respuestas a las preguntas "¿Te gusta XY?" no son llamativas cuando se ven de forma independiente. Sin embargo, si se combinan varias preguntas de este tipo, queda claro que el ordenador trabaja con fórmulas para generar respuestas a partir de las preguntas. Las respuestas también pueden mostrar que el ordenador ha interpretado mal una pregunta, aunque esto también podría ocurrirle a un ser humano. Muchas respuestas son vagas y una investigación más profunda dejaría claro que el ordenador no ha entendido realmente el contenido de la pregunta. Además, a menudo es más seguro que el ordenador responda con un "no lo sé" (por ejemplo, a la pregunta sobre la raíz de 2). Esto finge rasgos humanos, pero también puede llevar a un desenmascaramiento si esta táctica se utiliza con demasiada frecuencia o con preguntas demasiado simples. Las respuestas retardadas y erróneas, por ejemplo a los problemas de aritmética, también pueden confundir al interrogador. Así, los ordenadores son capaces de fingir su capacidad de hablar, por ejemplo mediante respuestas formulistas, reflejando las afirmaciones del interlocutor, reacciones a palabras clave, el uso de modismos y la reanudación de temas, pero esto es sólo una fachada que es fácil de ver.

Más Ideas

Los enlaces y detalles sobre estas actividades pueden encontrarse en nuestra página web.

Reconocimiento facial

Nuestra puerta de entrada puede distinguirnos del cartero, nuestro software de gestión de fotos etiqueta automáticamente a nuestros amigos: el reconocimiento facial es una aplicación habitual de la IA. Para ello, la tecnología debe ser tan flexible como para reconocernos incluso en invierno con gorra y en verano con gafas de sol. Esta actividad transmite este principio a través de personajes de dibujos animados.

Mono, Sherlock Mono

¿Cómo se puede representar el conocimiento de manera que un ordenador pueda "entenderlo" y sacar conclusiones lógicas de él? La lógica y la representación formal del conocimiento son de gran importancia en este caso. Por ello, los sistemas de IA no son realmente "inteligentes", sino que utilizan inteligentemente diferentes posibilidades para representar el conocimiento. Este tipo de representación del conocimiento también se puede representar en rompecabezas lógicos: Los rompecabezas correspondientes requieren la combinación de diferentes hechos según ciertas reglas, para luego encontrar una solución.

Cerebro en una bolsa

En esta actividad, los alumnos simulan ellos mismos el funcionamiento de una red neuronal con cuerdas y rollos de papel higiénico. La red final es entonces capaz de jugar un juego.

Aprendizaje sin Supervisión

Además del Aprendizaje Supervisado y del Aprendizaje por Refuerzo, también existen los llamados procedimientos de Aprendizaje No Supervisado: los ordenadores aprenden sin objetivos previamente conocidos y sin recompensas. A partir de un conjunto de puntos de datos se pueden identificar categorías (por ejemplo, clientes con alto potencial de compra en tiendas web) o anomalías (por ejemplo, actividades sospechosas en servidores web). Utilice tiza para dibujar una cuadrícula de coordenadas (por ejemplo, en el patio del colegio) y pida a sus alumnos que se posicionen adecuadamente en la cuadrícula utilizando los dos ejes. Dependiendo de los ejes seleccionados, se pueden identificar no sólo clusters, sino también valores atípicos o anomalías.



FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Pie de Imprenta

Editor:

Cátedra de Enseñanza de la Informática
Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstraße 3
91058 Erlangen
<https://aiunplugged.org>

Redacción y diseño:

Annabel Lindner, Stefan Seegerer

Traducción:

Annabel Lindner

Todos los textos y gráficos de este folleto (excepto el emblema de la FAU o lo que se indique) tienen licencia CC BY NC 3.0, lo que significa que puede editar, reproducir y distribuir el material en cualquier formato o medio, pero no comercialmente. Lo único que tienes que hacer es nombrar al autor.