

Python Deep Learning

Introducción práctica con
Keras y TensorFlow 2

Jordi Torres



PYTHON DEEP LEARNING

Introducción práctica con Keras y TensorFlow 2

Jordi Torres

Acceda a www.marcombo.info
para descargar gratis
contenido adicional
complemento imprescindible de este libro

Código:

PYTHON4

PYTHON DEEP LEARNING

Introducción práctica con Keras y TensorFlow 2

Jordi Torres



ALPHA EDITORIAL

ALFAOMEGA COLOMBIANA S.A.

Calle 62 No.20-46 esquina, Bogotá

Teléfono (57-1) 746 0102 Fax: (57-1) 210 0122

cliente@alfaomegacolombiana.com

www.alfaomega.com.co

www.alpha-editorial.com

Primera edición: Madrid, 2020
Bogotá, 2020

© Jordi Torres i Viñals

© Alpha editorial

© Alfaomega Colombiana S.A.

© Marcombo S.L.,

Todos los derechos son reservados. Esta publicación no puede ser reproducida total ni parcialmente. No puede ser registrada por un sistema de recuperación de información, en ninguna forma ni por ningún medio, sea mecánico, fotoquímico, electrónico, magnético, electroóptico, fotocopia o cualquier otro, sin el permiso previo y por escrito de la editorial.

Revisor técnico: Ferran Fàbregas

Corrección: Anna Alberola

Directora de producción: M.ª Rosa Castillo

Diseño de la cubierta: Enedénú Diseño Gráfico

ISBN: 978-958-778-640-8 (Colombia)

ISBN: 978-84-267-2828-9 (España)

Hecho en Colombia

Printed and made in Colombia

A Domingo.

Contenido del libro

Prefacio	17
Acerca de este libro	19
PARTE 1: INTRODUCCIÓN	23
CAPÍTULO 1. ¿Qué es el Deep Learning?	25
CAPÍTULO 2. Entorno de trabajo	47
CAPÍTULO 3. Python y sus librerías	57
PARTE 2: FUNDAMENTOS DEL DEEP LEARNING	79
CAPÍTULO 4. Redes neuronales densamente conectadas	81
CAPÍTULO 5. Redes neuronales en Keras	99
CAPÍTULO 6. Cómo se entrena una red neuronal	125
CAPÍTULO 7. Parámetros e hiperparámetros en redes neuronales	139
CAPÍTULO 8. Redes neuronales convolucionales	157
PARTE 3: TÉCNICAS DEL DEEP LEARNING	183
CAPÍTULO 9. Etapas de un proyecto Deep Learning	185
CAPÍTULO 10. Datos para entrenar redes neuronales	205
CAPÍTULO 11. <i>Data Augmentation</i> y <i>Transfer Learning</i>	231
CAPÍTULO 12. Arquitecturas avanzadas de redes neuronales	255
PARTE 4: DEEP LEARNING GENERATIVO	275
CAPÍTULO 13. Redes neuronales recurrentes	277
CAPÍTULO 14. <i>Generative Adversarial Networks</i>	307
Clausura	337
Apéndices	339

ÍNDICE ANALÍTICO

Prefacio	17
Acerca de este libro	19
PARTE 1: INTRODUCCIÓN	23
CAPÍTULO 1. ¿Qué es el Deep Learning?	25
1.1. Inteligencia artificial	26
1.1.1. La inteligencia artificial está cambiando nuestras vidas	26
1.1.2. Clases de inteligencia artificial	28
1.2. Machine Learning	29
1.3. Redes neuronales y Deep Learning	31
1.3.1. Redes neuronales artificiales	31
1.3.2. Las <i>Deep Networks</i> básicas	33
1.4. ¿Por qué ahora?	34
1.4.1. La supercomputación corazón del Deep Learning	34
1.4.2. Los datos, el combustible para la inteligencia artificial	42
1.4.3. Democratización de la computación	43
1.4.4. Una comunidad de investigación muy colaborativa	44
CAPÍTULO 2. Entorno de trabajo	47
2.1. Entorno de trabajo	47
2.2. TensorFlow y Keras	52
2.2.1. TensorFlow	52
2.2.2. Keras	54
CAPÍTULO 3. Python y sus librerías	57
3.1. Conceptos básicos de Python	57
3.1.1. Primeros pasos	57
3.1.2. Sangrado en Python	59
3.1.3. Variables, operadores y tipos de datos	60
3.1.4. Tipos de estructuras de datos	62
3.1.5. Sentencias de control de flujo	65
3.1.6. Funciones	67
3.1.7. Clases	68
3.1.8. Decoradores	70
3.2. Librería NumPy	72
3.2.1. Tensor	72
3.2.2. Manipulación de los tensores	75
3.2.3. Valor máximo en un tensor	77

PARTE 2: FUNDAMENTOS DEL DEEP LEARNING	79
CAPÍTULO 4. Redes neuronales densamente conectadas	81
4.1. Caso de estudio: reconocimiento de dígitos	81
4.2. Una neurona artificial	84
4.2.1. Introducción a la terminología y notación básica	84
4.2.2. Algoritmos de regresión	85
4.2.3. Una neurona artificial simple	86
4.3. Redes neuronales	89
4.3.1. Perceptrón	89
4.3.2. Perceptrón multicapa	90
4.3.3. Perceptrón multicapa para clasificación	93
4.4. Función de activación <i>softmax</i>	94
CAPÍTULO 5. Redes neuronales en Keras	99
5.1. Precarga de los datos en Keras	99
5.2. Preprocesado de datos de entrada en una red neuronal	102
5.3. Definición del modelo	105
5.4. Configuración del proceso de aprendizaje	108
5.5. Entrenamiento del modelo	108
5.6. Evaluación del modelo	110
5.7. Generación de predicciones	112
5.8. Todos los pasos de una tirada	114
CAPÍTULO 6. Cómo se entrena una red neuronal	125
6.1. Proceso de aprendizaje de una red neuronal	125
6.1.1. Visión global	125
6.1.2. Proceso iterativo de aprendizaje de una red neuronal	128
6.1.3. Piezas clave del proceso de <i>backpropagation</i>	130
6.2. Descenso del gradiente	130
6.2.1. Algoritmo básico de descenso del gradiente	131
6.2.2. Tipos de descenso del gradiente	133
6.3. Función de pérdida	135
6.4. Optimizadores	136
CAPÍTULO 7. Parámetros e hiperparámetros en redes neuronales	139
7.1. Parametrización de los modelos	139
7.1.1. Motivación por los hiperparámetros	139
7.1.2. Parámetros e hiperparámetros	140
7.1.3. Grupos de hiperparámetros	141

7.2. Hiperparámetros relacionados con el algoritmo de aprendizaje.....	141
7.2.1. Número de <i>epochs</i>	141
7.2.2. <i>Batch size</i>	142
7.2.3. <i>Learning rate</i> y <i>learning rate decay</i>	142
7.2.4. <i>Momentum</i>	143
7.2.5. Inicialización de los pesos de los parámetros.....	145
7.3. Funciones de activación.....	145
7.4. Practicando con una clasificación binaria.....	148
7.4.1. TensorFlow Playground.....	148
7.4.2. Clasificación con una sola neurona.....	151
7.4.3. Clasificación con más de una neurona.....	152
7.4.4. Clasificación con varias capas.....	154
CAPÍTULO 8. Redes neuronales convolucionales.....	157
8.1. Introducción a las redes neuronales convolucionales.....	157
8.2. Componentes básicos de una red neuronal convolucional.....	159
8.2.1. Operación de convolución.....	159
8.2.2. Operación de <i>pooling</i>	162
8.3. Implementación de un modelo básico en Keras.....	165
8.3.1. Arquitectura básica de una red neuronal convolucional.....	165
8.3.2. Un modelo simple.....	166
8.3.3. Configuración, entrenamiento y evaluación del modelo.....	169
8.4. Hiperparámetros de la capa convolucional.....	170
8.4.1. Tamaño y número de filtros.....	170
8.4.2. <i>Padding</i>	170
8.4.3. <i>Stride</i>	172
8.5. Conjunto de datos Fashion-MNIST.....	173
8.5.1. Modelo básico.....	173
8.5.2. Capas y optimizadores.....	174
8.5.3. Capas de <i>Dropout</i> y <i>BatchNormalization</i>	177
8.5.4. Decaimiento del ratio de aprendizaje.....	179
PARTE 3: TÉCNICAS DEL DEEP LEARNING.....	183
CAPÍTULO 9. Etapas de un proyecto Deep Learning.....	185
9.1. Definición del problema.....	186
9.2. Preparar los datos.....	187
9.2.1. Obtener los datos.....	187
9.2.2. Separar los datos para entrenar y evaluar el modelo.....	190

9.3. Desarrollar el modelo	194
9.3.1. Definir el modelo	194
9.3.2. Configuración del modelo	195
9.3.3. Entrenamiento del modelo	197
9.4. Evaluación del modelo	199
9.4.1. Visualización del proceso de entrenamiento	199
9.4.2. <i>Overfitting</i>	200
9.4.3. <i>Early stopping</i>	201
9.4.4. Evaluación del modelo con los datos de prueba	203
9.4.5. Entrenamiento con MAE	203
CAPÍTULO 10. Datos para entrenar redes neuronales.....	205
10.1. ¿Dónde encontrar datos para entrenar redes neuronales?	205
10.1.1. Conjuntos de datos públicos	206
10.1.2. Conjuntos de datos precargados	207
10.1.3. Conjuntos de datos de Kaggle	207
10.2. ¿Cómo descargar y usar datos reales?.....	208
10.2.1. Caso de estudio: «Dogs vs. Cats»	208
10.2.2. Datos para entrenar, validar y probar	210
10.2.3. Modelo de reconocimiento de imágenes reales	216
10.2.4. Preprocesado de datos reales con ImageDataGenerator	219
10.3. Solucionar problemas de sobreentrenamiento	220
10.3.1. Modelos a partir de conjuntos de datos pequeños	220
10.3.2. Visualización del comportamiento del entrenamiento.....	223
10.3.3. Técnicas de prevención del sobreentrenamiento	227
CAPÍTULO 11. <i>Data Augmentation</i> y <i>Transfer Learning</i>	231
11.1. <i>Data Augmentation</i>	231
11.1.1. Transformaciones de imágenes	231
11.1.2. Configuración de ImageGenerator	232
11.1.3. Código del caso de estudio	234
11.2. <i>Transfer Learning</i>	238
11.2.1. Concepto de <i>Transfer Learning</i>	238
11.2.2. <i>Feature Extraction</i>	239
11.2.3. <i>Fine-Tuning</i>	246
CAPÍTULO 12. Arquitecturas avanzadas de redes neuronales	255
12.1. API funcional de Keras	255
12.1.1. Modelo secuencial.....	255

12.1.2. Modelos complejos.....	258
12.2. Redes neuronales preentrenadas.....	262
12.2.1. Redes neuronales con nombre propio	262
12.2.2. Acceso a redes preentrenadas con la API Keras	263
12.3. Uso de redes preentrenadas con Keras	268
12.3.1. Conjunto de datos CIFAR-10	268
12.3.2. Red neuronal ResNet50	270
12.3.3. Red neuronal VGG19.....	272
PARTE 4: DEEP LEARNING GENERATIVO	275
CAPÍTULO 13. Redes neuronales recurrentes.....	277
13.1. Conceptos básicos de las redes neuronales recurrentes.....	278
13.1.1. Neurona recurrente	278
13.1.2. <i>Memory cell</i>	279
13.1.3. <i>Backpropagation</i> a través del tiempo.....	280
13.1.4. <i>Exploding Gradients</i> y <i>Vanishing Gradients</i>	281
13.1.5. <i>Long-Short Term Memory</i>	282
13.2. Vectorización de texto.....	282
13.2.1. <i>One-hot encoding</i>	283
13.2.2. <i>Word embedding</i>	284
13.2.3. <i>Embedding layer</i> de Keras.....	286
13.2.4. Usando <i>embedding</i> preentrenados.....	286
13.3. Programando una RNN: generación de texto.....	287
13.3.1. <i>Character-Level Language Models</i>	288
13.3.2. Descarga y preprocesado de los datos.....	289
13.3.3. Preparación de los datos para ser usados por la RNN.....	291
13.3.4. Construcción del modelo RNN.....	294
13.3.5. Entrenamiento del modelo RNN	297
13.3.6. Generación de texto usando el modelo RNN.....	299
13.3.7. Generando texto falso de Shakespeare.....	303
CAPÍTULO 14. <i>Generative Adversarial Networks</i>	307
14.1. <i>Generative Adversarial Networks</i>	307
14.1.1. Motivación por las GAN.....	308
14.1.2. Arquitectura de las GAN.....	309
14.1.3. Proceso de entrenamiento	310
14.2. Programando una GAN	311
14.2.1. Preparación del entorno y descarga de datos	314

14.2.2. Creación de los modelos.....	315
14.2.3. Funciones de pérdida y optimizadores	321
14.3. Entrenamiento con la API de bajo nivel de TensorFlow	323
14.3.1. API de bajo nivel de TensorFlow.....	323
14.3.2. Algoritmo de aprendizaje a bajo nivel	324
14.3.3. Entrenamiento de las redes GAN	325
14.3.4. Mejora del rendimiento computacional con decoradores de funciones...	329
14.3.5. Evaluación de los resultados	330
Clausura	337
Apéndices	339
Apéndice A: Traducción de los principales términos	341
Apéndice B: Tutorial de Google Colaboratory	345
Apéndice C: Breve tutorial de TensorFlow Playground.....	359
Apéndice D: Arquitectura de ResNet50	369
Agradecimientos	377
Índice alfabético	379

Prefacio

La educación es el arma más poderosa que puedes usar para cambiar el mundo.

Nelson Mandela

Siempre me ha interesado la tecnología de próxima generación y su impacto, y por ello desde hace un tiempo ha captado mi interés la inteligencia artificial y su relación con tecnologías como Cloud Computing, Big Data o la supercomputación, áreas en las que llevo investigando e impartiendo docencia desde hace más 30 años.

Sin duda, los avances tecnológicos en inteligencia artificial, junto con el resto de tecnologías mencionadas, ya están aquí; eso nos permite construir una sociedad que mejora la vida de las personas, aunque también es cierto que la perspectiva del futuro cercano de estas tecnologías presenta alguna que otra incertidumbre.

Sin embargo, estoy convencido de que conseguiremos, a medida que nos vayamos encontrando con nuevos problemas debido a estas nuevas tecnologías, encontrar como sociedad sus soluciones. Para ello, es clave que todos y cada uno de nosotros consigamos una mejor comprensión de estos nuevos temas que están revolucionando la informática y podamos darle el uso correcto, además de saber explicarlos.

Y este es el propósito de este libro que, gracias al ingente esfuerzo de la editorial Marcombo, presenta de forma elegante y ordenada una revisión y ampliación del contenido del libro abierto *Deep Learning - Introducción práctica con Keras*¹, junto con diferentes entradas de mi blog² que perseguían abrir el conocimiento a todos aquellos que quieren usarlo para mejorar. Porque para mí el acceso al conocimiento es un derecho humano.

¹ *Deep Learning - Introducción práctica con Keras* (parte 1 y parte 2), WATCH THIS SPACE Book Series. Acceso abierto: <https://torres.ai/deeplearning>.

² Enlace a mi blog personal: <https://torres.ai/blog>.

PYTHON DEEP LEARNING

Nos centramos en Deep Learning, una de las áreas más activas actualmente en el ámbito de la inteligencia artificial, y elemento tecnológico central en gran parte de las innovaciones actuales que están cambiando profundamente el funcionamiento de nuestro mundo.

Como reza el título, este libro presenta una introducción práctica a Deep Learning, no un tratado exhaustivo sobre el tema, pues está dirigido a un lector que dispone de conocimientos en programación pero que aún no ha tenido la oportunidad de iniciarse en estos nuevos conceptos claves en el futuro de la informática, y que quiere descubrir cómo funciona esta tecnología.

Ya les avanzo que el libro será una invitación a usar el teclado de su ordenador mientras va aprendiendo: nosotros lo llamamos *learn by doing*, y mi experiencia como profesor en la UPC me indica que es una aproximación que funciona muy bien para alguien que trata de iniciarse en un nuevo tema. Por esta razón, el libro reducirá todo lo posible la parte teórico-matemática, aunque es estrictamente necesario recurrir a ciertos detalles teóricos para ofrecer un conocimiento sólido al lector con explicaciones claras que complementen los ejemplos de código.

Déjenme acabar diciéndoles: ¡gracias por estar leyendo este libro! El simple hecho me reconforta y justifica mi esfuerzo en escribirlo. Aquellos que me conocen saben que la formación de las personas es una de mis pasiones, y que me mantiene con vigor y energía.

Jordi Torres i Viñals, a 25 de enero de 2020.

Acerca de este libro

A quién va dirigido el libro

En este libro el lector o lectora encontrará una guía para adentrarse de manera práctica a Deep Learning con la ayuda de la librería TensorFlow, la cual aprenderá a usar con el objetivo de desarrollar y evaluar modelos Deep Learning. Aunque Deep Learning se sustenta en fascinantes matemáticas, estas no son estrictamente necesarias para poder iniciarse, ni siquiera para crear proyectos que generen valor a la empresa, gracias a librerías Python como TensorFlow.

Por ello, este libro se centrará en temas prácticos y concretos para descubrir al lector o lectora el apasionante mundo que se abre con el uso de Deep Learning. Debemos tener siempre en mente que solo podremos examinar una pequeña parte, pues es imposible mostrar su alcance total en un único libro; tan solo mantenerse al día de las últimas investigaciones, tecnologías o herramientas que van apareciendo es casi misión imposible o, como diría un amigo inglés, *like drinking from a fire hose*, como beber de una manguera contra incendios.

En vez de centrarse en conceptos teóricos del Deep Learning que pueden resultar abrumadores para un lector o lectora que no dispone de conocimientos previos mínimos en Machine Learning, el enfoque de usar ejemplos de código práctico introducidos de forma lineal para explicar los conceptos fundamentales permite iniciarse en el Deep Learning a un mayor número de personas. Para ello, se propone involucrar al lector invitándole a que esté con el teclado delante y a que vaya probando lo que va leyendo en el libro.

Qué contiene este libro

El libro se organiza en 14 capítulos agrupados en 4 partes, que se aconseja que se lean en orden la primera vez, ya que van guiando al lector o lectora y lo introducen gradualmente a los conocimientos necesarios para seguir los ejemplos prácticos; en todos los casos se intenta ser lo más conciso posible. Al ser un libro introductorio, consideramos que es mucho más útil este enfoque que no uno más formal. A pesar de todo, hemos intentado que el índice del libro exprese un mapa razonablemente ordenado de los conceptos del área, complementado con un exhaustivo índice alfabético al final del libro para que el lector o lectora pueda encontrar rápidamente aquellos conceptos más concretos una vez ya familiarizado con el tema.

La primera parte del libro nos prepara con los conocimientos básicos para comprender sin dificultad el contenido de las siguientes partes. En el capítulo 1 se explica qué es el Deep Learning a nivel general y se describe el contexto y los

PYTHON DEEP LEARNING

ingredientes que han permitido este creciente interés en esta área. En el capítulo 2 presentamos el entorno de trabajo que se usará para poder probar los códigos que a lo largo del libro se van presentando. Recordemos que se trata de un libro con un eminente componente práctico y esto requiere poder probar por uno mismo los códigos. Finalmente, en el capítulo 3 hacemos un breve repaso para todos aquellos lectores y lectoras que no tengan una buena base de Python; se presentan aquellos aspectos importantes de conocer, de este lenguaje de programación, para poder programar las redes neuronales.

En la segunda parte del libro se agrupan las explicaciones de los aspectos más fundamentales del Deep Learning. En el capítulo 4, aprovechando que se presenta uno de los tipos de redes neuronales básicos en Deep Learning —las densamente conectadas—, se presentan todos aquellos conceptos fundamentales que es necesario conocer para poder seguir aprendiendo conceptos más avanzados. En el capítulo 5 se sugiere a los lectores y lectoras que se pongan delante del teclado, y se explica cómo programar los conceptos explicados en el anterior capítulo usando las API Keras de la librería TensorFlow. En los dos siguientes capítulos, el 6 y 7, pasamos a describir los conceptos teóricos que hay detrás de las redes neuronales, para ofrecer al lector y lectora una base sólida del tema. Finalmente, en el capítulo 8 se describe y programa una de las redes neuronales más populares en el mundo del Deep Learning dado su aplicabilidad en gran número de aplicaciones actuales de visión por computador.

En la tercera parte del libro se repasan diferentes aspectos técnicos muy importantes en la programación de redes neuronales. En el capítulo 9 se revisan las etapas habituales en un proyecto de Deep Learning de principio a fin. En el siguiente capítulo, el 10, descubrimos dónde podemos encontrar datos para entrenar redes neuronales y se trata desde un punto de vista práctico la problemática que presenta el sobreajuste de modelos. En el capítulo 11 tratamos técnicas para mitigar este sobreajuste, como son el *Data Augmentation* o *Transfer Learning*. Finalmente, en el capítulo 12 nos familiarizaremos con arquitecturas avanzadas de redes neuronales e introduciremos la API de Keras que permite su creación.

En la cuarta parte del libro entraremos a explicar conceptos más avanzados sobre Deep Learning alrededor de modelos generativos, que generan ejemplos de datos parecidos a los que se han usado para entrenarse pero que son totalmente creados por el modelo. En el capítulo 13 mostraremos cómo programar un generador de texto nuevo y aprovecharemos para explicar las redes neuronales recurrentes, uno de los tipos de redes neuronales de gran popularidad en estos momentos. Finalmente, en el capítulo 14 introduciremos las redes conocidas como *Generative Adversarial Networks*, que están generando un gran interés puesto que son las responsables en gran medida de las *fake news*. Aprovecharemos este capítulo para mostrar algunos conceptos avanzados sobre el modelo de ejecución en TensorFlow.

El libro acaba con una clausura donde se tratan algunos temas importantes para que el lector o lectora reflexione sobre las implicaciones del Deep Learning.

Al final se incluyen varios apéndices que complementan o amplían algunos aspectos presentados en los capítulos del libro para aquellos lectores o lectoras que lo requieran.

Requisitos para seguir el libro

Como ya hemos mencionado, esta obra pretende ser una introducción; por ello, no es necesario que el lector o lectora sea un experto en Python, solo ha de tener, evidentemente, conocimientos de programación e interés en aprender por su cuenta detalles del lenguaje cuando no los entienda.

Tampoco se necesita ser un experto en Machine Learning, pero está claro que puede ser muy útil conocer unos conceptos básicos sobre el tema. Solo se presuponen los conocimientos básicos en matemáticas de cualquier estudiante de bachillerato. A partir de ellos, a lo largo del libro se introducen y repasan muy brevemente los conceptos más importantes de Machine Learning que se puedan requerir como conocimientos previos.

Pero lo más importante, en cuanto a prerrequisitos, ¡es tener interés por aprender!

Acerca de los anglicismos usados en este libro

Un comentario recurrente que he recibido sobre el contenido de mis publicaciones previas es acerca del uso de muchos anglicismos en la terminología técnica en detrimento de su versión traducida. Me gustaría hacer notar al lector o lectora que en este campo informático que avanza tan rápidamente a menudo no existe en español una traducción consensuada o una palabra equivalente con el mismo significado; además, algunos de los términos en inglés no siempre se traducen de la misma manera en los diferentes países de habla hispana, por lo que el uso de un término u otro podría dar lugar a confusiones.

Intentaré, en la medida de lo posible, usar la versión traducida de todos los términos siempre que esto no pueda producir dudas al lector o lectora. En todo caso, presentaré también la versión en inglés para ayudar a los lectores y lectoras a familiarizarse con esos términos, puesto que creo que les será muy útil para poder avanzar por su cuenta en este tema tan dinámico en el que los avances más recientes se encuentran mayormente en literatura escrita solo en lengua inglesa. En esta línea, también mantendré los códigos Python usados en este libro con nombres de variables en inglés para familiarizar al lector o lectora con algunos nombres de variables que disfrutaban de un cierto consenso en la comunidad Deep Learning y facilitarles, de esta manera, el poder seguir otros códigos en Python referenciados en este libro.

En el apéndice A del libro hemos hecho un esfuerzo de revisión de aquellos términos que disponen de una traducción más o menos estable; los hemos enumerado para así ayudar al lector o lectora a familiarizarse con ellos, ya que — insisto— es fundamental conocer el término inglés para poder seguir los avances en este campo en el que todo cambia tan rápido.

Relacionado con este tema, quiero hacer notar al lector o lectora que los números se representan en notación anglosajona, con el punto reservado para los decimales. De esta manera se mantiene coherencia con los códigos Python usados en este libro.

PYTHON DEEP LEARNING

Descargar el código de ejemplos y figuras en color

En la parte inferior de la primera página del libro, el lector o lectora encontrará el código de acceso que le permitirá descargar localmente los códigos de ejemplo de este libro, los conjuntos de datos requeridos en los ejemplos de este libro, la lista de referencias para poder acceder a los links, así como las imágenes en color. Alternativamente, los códigos pueden descargarse también del repositorio GitHub del libro:

<https://github.com/JordiTorresBCN/python-deep-learning>

Página web del libro

En la página web <https://JordiTorres.ai/python-deep-learning>, se irá incluyendo material suplementario de nuevos temas que vayan apareciendo y que creamos conveniente poner a disposición del lector o lectora, así como una lista con las fe de erratas que puedan aparecer en el libro con su correspondiente corrección.

PARTE 1:

INTRODUCCIÓN

CAPÍTULO 1.

¿Qué es el Deep Learning?

Se está considerando la inteligencia artificial como la nueva revolución industrial, corazón de lo que algunos llaman industria 4.0. Pues bien, Deep Learning es el motor de este proceso y, a continuación, centraremos el tema y veremos que la inteligencia artificial (*Artificial Intelligence* en inglés) ya está aquí y por qué ha venido para quedarse.

Los sistemas informáticos actuales ya traducen textos en cualquier idioma, responden automáticamente correos electrónicos o crean *fake news* que nos están volviendo locos a todos. Esto se debe, en gran medida, a una parte de la inteligencia artificial que se denomina Deep Learning (traducido a veces como «aprendizaje profundo»). El término Deep Learning agrupa una parte de técnicas de aprendizaje automático (*Machine Learning* en inglés) que se basan en modelos de redes neuronales y cuya gran aplicabilidad se ha mostrado recientemente en multitud de usos por parte la industria.

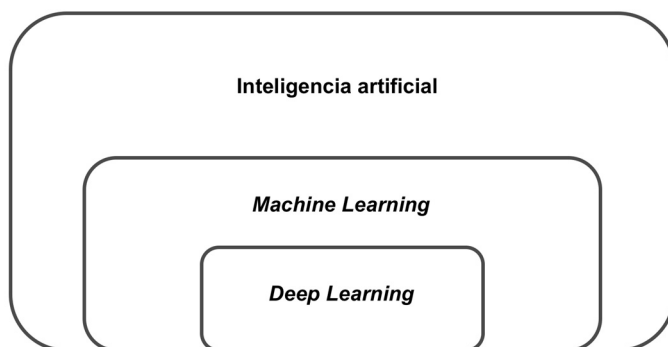


Figura 1.1 Deep Learning es un subconjunto de Machine Learning, que es solo una parte de la inteligencia artificial, aunque en estos momentos quizás es la más dinámica y la que está haciendo que la inteligencia artificial esté nuevamente en pleno auge.

El objetivo de este capítulo es ayudar al lector o lectora a enmarcar el Deep Learning, que ha surgido de la investigación en inteligencia artificial y Machine Learning. Para ello, empezaremos explicando de forma general qué se entiende por Machine Learning e inteligencia artificial, marco donde se engloba el Deep Learning, tal como se muestra en la Figura 1.1. Todo ello sin entrar en detalles de definiciones y categorizaciones académicas, quedándonos en una visión general suficiente para poder entrar a la parte práctica de esta apasionante disciplina.

1.1. Inteligencia artificial

1.1.1. La inteligencia artificial está cambiando nuestras vidas

Nos encontramos ante vertiginosos avances en la calidad y prestaciones de una amplia gama de tecnologías cotidianas: en el caso del reconocimiento de voz automática (*Automated Speech Recognition* en inglés, ASR), la transcripción de voz a texto ha experimentado avances increíbles, y ya está disponible en diferentes dispositivos de uso doméstico. Estamos interactuando cada vez más con nuestros ordenadores (y todo tipo de dispositivo) simplemente hablando con ellos.

También ha habido avances espectaculares en el procesado de lenguaje natural (*Natural Language Processing* en inglés, NLP). Por ejemplo, simplemente haciendo clic en el símbolo de micro de Google Translate, el sistema transcribirá a otro idioma lo que está dictando. Google Translate ya permite convertir oraciones de una lengua a otra en un gran número de pares de idiomas, y ofrece traducción de texto para más de un centenar.

Incluso más relevante en el ámbito del NLP es lo que ocurre con el texto predictivo y la redacción automática en proyectos como el GPT-2³ de la fundación OpenAI, en el que los propios creadores decidieron abortar su publicación completa en abierto a inicios del 2019, por considerar que era una herramienta tan poderosa que podía por ejemplo ser usada para fabricar potentes *fake news*. Finalmente, a finales de año liberaron la versión completa⁴.

A su vez, los avances en la visión por computador (*Computer Vision* en inglés, CV) también son enormes: ahora nuestros ordenadores, por ejemplo, pueden reconocer imágenes y generar descripciones textuales de su contenido en segundos. O la perfección que están alcanzando los generadores de rostros artificiales, que permite que se mezclen personajes reales y ficticios con total realismo.

Estas tres áreas (ASR, NLP y CV) son cruciales para dar rienda suelta a las mejoras en robótica, drones o automóviles sin conductor. La inteligencia artificial está en el corazón de toda esta innovación tecnológica, que últimamente avanza tan rápidamente gracias a Deep Learning.

³ Better Language Modells and Their Implications. Open AI, febrero 2019. <https://openai.com/blog/better-language-models/> [Consultado: 18/11/2019].

⁴ GPT-2: 1.5B Release. [online] Disponible en: <https://openai.com/blog/gpt-2-1-5b-release/> [Consultado: 26/11/2019].

Y todo ello a pesar de que la inteligencia artificial todavía no se ha desplegado ampliamente; es difícil hacerse una idea del gran impacto que tendrá, al igual que en 1995 lo era el imaginarse el impacto futuro de Internet. En aquel entonces, la mayoría de la gente no veía que Internet fuera a ser relevante para ellos ni cómo iba a cambiar sus vidas.

Personas como Sundar Pichai, director ejecutivo de Google, dicen que el impacto de la inteligencia artificial en la historia de la humanidad es comparable con el de la electricidad y el fuego⁵. Para él, la inteligencia artificial es una de las cosas más importantes en las que la humanidad está trabajando y opina que, al igual que la gente aprendió a utilizar el fuego para los beneficios de la humanidad, también necesitó superar sus desventajas.

Quiero creer que Pichai es muy optimista respecto a la inteligencia artificial y que está convencido de que podría usarse para ayudar a resolver algunos de los retos que tiene la humanidad encima de la mesa. Quizás esta comparativa sea una exageración, eso solo lo sabremos con el tiempo; pero yo de ustedes le tendría puesto el ojo a la inteligencia artificial, porque algo está cambiando, y a todos nos conviene estar atentos a lo que se avecina.

Pero, ¿a qué nos referimos cuando hablamos de inteligencia artificial? Una extensa y precisa definición (y descripción de sus ámbitos) se encuentra en el libro de Stuart Russell⁶ y Peter Norvig⁷ titulado *Artificial Intelligence, a modern approach*⁸, el texto sobre inteligencia artificial más completo y, sin duda para mí, el mejor punto de partida para tener una visión global del tema. Pero intentando hacer una aproximación más generalista (propósito de este libro), podríamos aceptar una definición simple en la que por inteligencia artificial entendamos aquella inteligencia que muestran las máquinas, en contraste con la inteligencia natural de los humanos. En este sentido, una posible definición concisa y general de inteligencia artificial podría ser «el esfuerzo para automatizar tareas intelectuales normalmente realizadas por humanos».

Como tal, el área de inteligencia artificial es un campo muy amplio que abarca muchas áreas del conocimiento relacionadas con el aprendizaje automático; incluso se incluyen muchos más enfoques no siempre catalogados como aprendizaje automático. Además, a lo largo del tiempo, a medida que los computadores han sido cada vez más capaces de «hacer cosas», han ido cambiando las tareas o tecnologías consideradas «inteligentes».

⁵ El CEO de Google defiende que la inteligencia artificial tendrá más impacto que la electricidad o el fuego. El país 23/01/2018. [online]. Disponible en: https://elpais.com/tecnologia/2018/01/21/actualidad/1516570888_812262.html [Consultado: 23/11/2019].

⁶ Stuart J. Russell. Wikipedia. [online]. Disponible en: https://en.wikipedia.org/wiki/Stuart_J._Russell [Consultado: 16/04/2018].

⁷ Peter Norvig Wikipedia. [online]. Disponible en: https://en.wikipedia.org/wiki/Peter_Norvig [Consultado: 16/04/2018].

⁸ Artificial Intelligence: A Modern Approach (AIMA) ·3rd edition, Stuart J Russell and Peter Norvig, Prentice hall, 2009. ISBN 0-13-604259-7.

Esto explica por qué desde los años 50 la inteligencia artificial ha experimentado varias oleadas de optimismo, seguidas por la decepción y la pérdida de financiación e interés (épocas conocidas como *AI winter*⁹), seguidas luego de nuevos enfoques, éxito y financiación. Además, durante la mayor parte de su historia, la investigación en inteligencia artificial se ha dividido en subcampos basados en consideraciones técnicas o herramientas matemáticas concretas y con comunidades de investigación que no se comunicaban suficientemente entre sí.

Pero sin duda estamos ante una nueva época de expansión de la inteligencia artificial con resultados muy llamativos. Por ejemplo, un grupo de investigación de DeepMind ya en el 2016 consiguió que las máquinas aprendieran solas (sin intervención humana) a vencer a los humanos jugando a complejos juegos de mesa como el Go¹⁰. Otro ejemplo es la página web de acceso público *ThisPersonDoesNotExist.com* que, aplicando los resultados de un artículo de investigación¹¹, muestra lo fácil que es para la inteligencia artificial generar caras falsas increíblemente realistas para cualquier humano. Les propongo que lo prueben.

1.1.2. Clases de inteligencia artificial

Creo que queda justificado el entusiasmo que genera la inteligencia artificial, pero también es cierto que se escriben muchas exageraciones acerca de la misma en los medios de comunicación. Una de las razones es que con «inteligencia artificial» se expresan muchas cosas, pero podríamos destacar dos de ellas, dos ideas separadas que se refieren a cosas muy diferentes.

Por un lado, casi todo el progreso que estamos viendo en la inteligencia artificial —que se encuentra detrás de los avances antes mencionados— en el mundo académico se agrupa en lo que se denomina inteligencia artificial débil (*Artificial Narrow Intelligence* en inglés). Pero con inteligencia artificial también nos referimos a un segundo concepto, etiquetado en el mundo académico como inteligencia artificial fuerte (*Artificial General Intelligence* en inglés). Este tipo de inteligencia es aquella que considera que las máquinas pueden hacer cualquier cosa que un humano pueda hacer, o ser superinteligentes y hacer incluso más cosas.

Si bien hay muchos progresos en el área de la inteligencia artificial débil, no hay casi ninguno en lo que se refiere a la inteligencia artificial fuerte. Pero el rápido progreso en la inteligencia artificial débil, que es increíblemente valioso, ha hecho que los medios de comunicación a veces concluyan que hay mucho progreso también en la fuerte, lo cual no es cierto en estos momentos. La inteligencia artificial fuerte es un ámbito en el que los investigadores e investigadoras pueden trabajar, pero en el que se está aún

⁹ AI winter. Wikipedia. [online]. Disponible en: https://en.wikipedia.org/wiki/AI_winter [Consultado: 16/04/2018].

¹⁰ Reinforcement Learning: 10 Breakthrough Technologies 2017. MIT Technology Review. [online]. Disponible en: <https://www.technologyreview.com/s/603501/10-breakthrough-technologies-2017-reinforcement-learning/> [Consultado: 16/11/2019].

¹¹ A Style-Based Generator Architecture for Generative Adversarial Networks. NVIDIA. [online]. Disponible en: <https://arxiv.org/pdf/1812.04948.pdf> [Consultado: 16/11/2019].

muy lejos de conseguir un gran conocimiento; pueden pasar décadas o cientos de años, quién sabe.

Humildemente, creo que todavía estamos muy lejos de una máquina que sea tan capaz como los humanos de aprender a dominar muchos aspectos de nuestro mundo. Tomemos un ejemplo: incluso una niña de tres años puede aprender cosas de una manera en la que las computadoras no pueden hacerlo por ahora; una niña de tres años en realidad ¡domina la física intuitivamente! Por ejemplo, sabe perfectamente que cuando tira una bola al aire esta caerá. O cuando derrama algunos líquidos espera el desastre resultante. Sus padres no necesitan enseñarle las leyes de Newton, o hablarle de las ecuaciones diferenciales que definen la trayectoria de los objetos; esta niña de tres años descubre todas estas cosas sola, sin supervisión.

Ahora bien, hay autores que consideran que incluso solo con la inteligencia artificial débil nos dirigimos rápidamente hacia una situación en la que los sistemas informáticos tomarán decisiones por nosotros, y piden que nos preguntemos qué sucederá cuando esos sistemas dejen de lado la estrategia humana en favor de algo totalmente desconocido para nosotros. Creo que todos y cada uno de nosotros debemos pensar en ello.

1.2. Machine Learning

Machine Learning, traducido al castellano como «aprendizaje automático» (aunque yo voy a mantener su nombre en inglés en este libro para concretar más cuando hablo de la disciplina), es en sí mismo un gran campo de investigación y desarrollo. En concreto, Machine Learning se podría definir como «el subcampo de la inteligencia artificial que proporciona a los ordenadores la capacidad de aprender sin ser explícitamente programados, es decir, sin que estos necesiten que el programador indique las reglas que deben seguir para lograr su tarea, sino que la hace automáticamente».

Generalizando, podemos decir que Machine Learning consiste en desarrollar para cada problema un «algoritmo» de predicción para un caso de uso particular. Estos algoritmos aprenden de los datos con el fin de encontrar patrones o tendencias para comprender qué nos dicen estos datos y, de esta manera, construir un modelo para predecir o clasificar los elementos.

Dada la madurez del área de investigación en Machine Learning, existen muchos enfoques bien establecidos para el aprendizaje automático por parte de máquinas. Cada uno de ellos utiliza una estructura algorítmica diferente para optimizar las predicciones basadas en los datos recibidos. En resumen, Machine Learning es un amplio campo con una compleja taxonomía de algoritmos que se agrupan, en general, en tres grandes categorías: aprendizaje supervisado (*supervised learning*), aprendizaje no supervisado (*unsupervised learning*) y aprendizaje por refuerzo (*reinforcement learning*).

Nos referimos a aprendizaje supervisado cuando los datos que usamos para el entrenamiento incluyen la solución deseada, llamada etiqueta (*label*). En este caso el aprendizaje radica en aprender un modelo (o función) que mapea una entrada a una salida. Un escenario óptimo permitirá que el modelo, una vez entrenado, determine correctamente las etiquetas para datos de entrada no vistos anteriormente. Este

PYTHON DEEP LEARNING

modelo se construye con un algoritmo que iterativamente va afinando el modelo mediante la generación de predicciones sobre los datos de entrenamiento y va comparando estas predicciones con la respuesta correcta que el algoritmo conoce. De aquí que se denomine aprendizaje supervisado. En cambio, nos referimos a aprendizaje no supervisado cuando los datos de entrenamiento no incluyen las etiquetas, y es el algoritmo el que intentará clasificar la información por sí mismo.

Hablamos de aprendizaje por refuerzo (*reinforcement learning* en inglés) cuando el modelo se implementa en forma de un agente que deberá explorar un espacio desconocido y determinar las acciones a llevar a cabo mediante prueba y error; aprenderá por sí mismo gracias a las recompensas y penalizaciones que obtiene de sus acciones. El agente debe crear la mejor estrategia posible para obtener la mayor recompensa en tiempo y forma. Este aprendizaje permite ser combinado con Deep Learning y está ahora mismo muy presente en el mundo de la investigación, como lo demuestran los últimos progresos en áreas tan diversas como el reconocimiento de imágenes, los coches que conducen solos o los juegos complejos como Go o Starcraft.

Para resaltar la diferencia de los tipos de aprendizaje volvamos al ejemplo de la niña y a cómo su abuela la puede ayudar a aprender. La abuela de la niña podría sentarse con ella y enseñarle pacientemente ejemplos de gatos, de perros, etc. (actuando como en el aprendizaje supervisado), o recompensarla con un aplauso por resolver un rompecabezas de bloques de madera (como en el aprendizaje por refuerzo). Pero en realidad, durante una gran parte de su tiempo, la niña se dedica a explorar el mundo ingenuamente, dándole sentido a su entorno a través de la curiosidad, el juego y la observación. Es decir, aprende sin supervisión a partir de los datos que observa sin tener un propósito en particular.

Es importante notar que ambos paradigmas de aprendizaje —no supervisado y por refuerzo— requieren que el entrenamiento esté diseñado por un humano. En el caso del aprendizaje no supervisado, se definen los «objetivos» (por ejemplo, obtener la etiqueta correcta para una imagen); en el caso del aprendizaje por refuerzo, son las «recompensas» por un comportamiento exitoso (como obtener una puntuación alta en un juego). Por lo tanto, los entrenadores humanos, aunque no intervienen directamente en el bucle del aprendizaje, sí definen los límites del aprendizaje en ambos casos.

Pero, como decíamos, la taxonomía es muy amplia y podríamos incluso mencionar otros, como el caso del aprendizaje autosupervisado (*self-supervised learning*), que podríamos considerar una instancia específica del aprendizaje supervisado. Es como un *supervised learning* pero sin las etiquetas anotadas por humanos, sino por sistemas automáticos a partir de otros datos o heurísticas. En todo caso, podríamos catalogar en este grupo las técnicas que no requieren a humanos en el bucle.

Para acabar esta sección, quisiera recalcar al lector o lectora que es importante tener en cuenta que la distinción entre los diferentes tipos de aprendizaje puede ser borrosa a veces; estas categorías son un continuo sin fronteras sólidas. En capítulos siguientes continuaremos con más detalle la explicación sobre Machine Learning. Pero invito al lector o lectora que quiera saber más del tema a continuar con el libro *Python Machine Learning*¹² de esta misma colección.

¹² Sebastian Raschka y Vahid Mirjalili (2019). *Python Machine Learning*. Editorial Marcombo.

1.3. Redes neuronales y Deep Learning

Como decíamos en el anterior apartado, avances como el reconocimiento de voz, el procesado de lenguaje natural o la visión por computador son cruciales en muchas áreas que están cambiando el futuro próximo. Estos avances y la nueva popularidad de la inteligencia artificial se deben en gran parte a los avances del Deep Learning durante este decenio. En este apartado presentamos una visión global del tema para acotar a qué nos referimos cuando hablamos de Deep Learning; a lo largo del libro ya entraremos en más detalle.

1.3.1. Redes neuronales artificiales

Como ya hemos avanzado, un caso especial de algoritmos de Machine Learning son las redes neuronales artificiales, que en cierta manera son un intento de imitar la actividad en capas de neuronas en la neocorteza, que es la parte del cerebro humano donde ocurre el pensamiento. Estas redes neuronales aprenden estructuras jerárquicas y niveles de representación y abstracción para comprender los patrones de datos que provienen de varios tipos de fuentes, como imágenes, vídeos, sonido o texto.

En estos modelos de redes neuronales, como veremos, las abstracciones de nivel superior se definen como la composición de la abstracción de nivel inferior. Una de las mayores ventajas del Deep Learning es su capacidad de aprender automáticamente la representación de características en múltiples niveles de abstracción. Esto permite que un sistema aprenda funciones complejas asignadas desde el espacio de entrada al espacio de salida sin muchas dependencias de las funciones creadas por humanos. Aunque, precisamente, esto implica que a menudo la explicabilidad del modelo —o más bien de los resultados de esto— no sea fácilmente explicable, lo cual se convierte en un problema en ciertas áreas de aplicación, como veremos en el apartado de clausura.

La idea que hay detrás de una neurona artificial, conceptualmente hablando, es bastante simple. Tiene una o más entradas y una salida. Dependiendo del valor de esas entradas, la neurona puede «dispararse». De manera simplista, «dispararse» significa que la salida pasa de estar apagada a encendida (se puede pensar que es un interruptor binario que va de 0 a 1). En el caso concreto de Deep Learning, las estructuras algorítmicas antes mencionadas permiten modelos que están compuestos de múltiples capas de procesamiento (construidas con neuronas artificiales) para aprender representaciones de datos, con múltiples niveles de abstracción que realizan una serie de transformaciones lineales y no lineales que, a partir de los datos de entrada, generen una salida próxima a la esperada. El aprendizaje —supervisado en este caso— consiste en obtener los parámetros de esas transformaciones y conseguir que esas transformaciones sean óptimas, es decir, que la salida producida y la esperada difieran lo mínimo posible.

Una aproximación gráfica simple a una red neuronal Deep Learning es la que se muestra en la Figura 1.2.

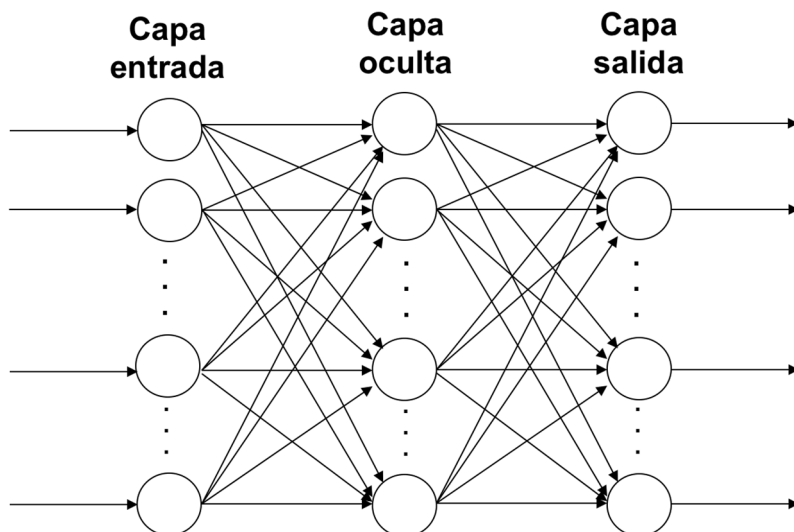


Figura 1.2 Una red neuronal se considera Deep Learning cuando tiene una o más capas ocultas.

En concreto, en esta figura se representa una red neuronal artificial con 3 capas: una de entrada (*input layer*) que recibe los datos de entrada, una de salida (*output layer*) que devuelve la predicción realizada, y las capas que tenemos en medio. Estas últimas se llaman capas ocultas (*hidden layers*) y podemos tener muchas, cada una con distinta cantidad de neuronas. Veremos más adelante que las neuronas, representadas por los círculos, estarán interconectadas unas con otras de diferente manera entre las neuronas de las distintas capas.

En general, hoy en día estamos manejando redes neuronales artificiales con muchísimas capas que, literalmente, están apiladas una encima de la otra; de aquí el concepto de *deep* (profundidad de la red), donde cada una de ellas está compuesta, a su vez, por muchísimas neuronas, cada una con sus parámetros que, a su vez, realizan una transformación simple de los datos que reciben de neuronas de la capa anterior para pasarlos a las de la capa posterior. La unión de todas permite descubrir patrones complejos en los datos de entrada.

Como veremos en detalle más adelante, los avances en Deep Learning han mejorado drásticamente el estado de la técnica en reconocimiento de voz, reconocimiento de objetos visuales, detección de objetos y muchos otros dominios, y han puesto la inteligencia artificial en el foco de interés de las empresas; de aquí el gran interés que ahora mismo suscitan.

Antes de acabar, me gustaría comentar la magnitud del problema que conlleva programar en estos momentos los algoritmos de Deep Learning: diferentes capas sirven para diferentes propósitos, y cada parámetro e hiperparámetro importa mucho en el resultado final. Esto lo hace extremadamente complicado a la hora de intentar afinar la programación de un modelo de red neuronal, lo cual puede parecer más un arte que una ciencia a quienes se adentran por primera vez en el área. Pero esto no

implica que sea algo misterioso, si bien es cierto que queda mucho por investigar, sino que simplemente hacen falta muchas horas de aprendizaje y práctica.

1.3.2. Las *Deep Networks* básicas

Ya hemos dicho que el Deep Learning es una técnica donde la información se procesa en capas jerárquicas para comprender representaciones y características de datos en niveles crecientes de complejidad. En la práctica, todos los algoritmos de Deep Learning son redes neuronales que comparten algunas propiedades básicas comunes, como que todas consisten en neuronas interconectadas que se organizan en capas.

En lo que difieren es en la arquitectura de la red (la forma en que las neuronas están organizadas en la red) y, a veces, en la forma en que se entrenan. Con eso en mente, enumeramos a continuación las principales clases de redes neuronales que presentaremos a lo largo del libro. Aunque no es una lista exhaustiva, representan la mayor parte de los algoritmos en uso hoy en día:

- Perceptrón multicapa (MLP, del inglés *Multi-layer perceptron*): un tipo de red neuronal con capas densamente conectadas que veremos en el capítulo 4.
- Redes neuronales convolucionales (CNN del inglés *Convolutional Neural Networks*): una CNN es una red neuronal con varios tipos de capas especiales, como veremos en el capítulo 8. Hoy en día este tipo de red está siendo muy usada por la industria en diferentes tipos de tarea, especialmente de visión por computador.
- Redes neuronales recurrentes (RNN del inglés *Recurrent Neural Networks*): este tipo de red tiene un estado interno (o memoria) que se crea con los datos de entrada ya vistos por la red. La salida de una RNN es una combinación de su estado interno y los datos de entrada. Al mismo tiempo, el estado interno cambia para incorporar datos recién entrados. Debido a estas propiedades, las redes neuronales recurrentes son buenas candidatas para tareas que funcionan en datos secuenciales, como texto o datos de series de tiempo. Presentaremos las RNN en el capítulo 13.

Y, para finalizar, para quien requiera una explicación más formal del tema —de la misma manera que antes les mencionaba la obra de Stuart Rusell y Peter Novig como libro base de inteligencia artificial—, para Deep Learning nos encontramos con un excelente libro, titulado *Deep Learning*¹³, realizado por Ian Goodfellow, Yoshua Bengio y Aaron Corville, que es el «campamento base» en estos momentos para el aprendizaje del tema en más profundidad.

¹³ Deep Learning. I Goodfellow, Y. Bengio and A Corville. MIT Press 2016. Disponible también en acceso libre en <http://www.deeplearningbook.org> [Consulta: 20/01/2018].

1.4. ¿Por qué ahora?

John McCarthy acuñó el término inteligencia artificial en la década de los 50 y fue uno de los padres fundadores de la inteligencia artificial junto con Marvin Minsky. También en 1958 Frank Rosenblatt construyó un prototipo de red neuronal, que llamó el *Perceptron*. Además, las ideas clave de las redes neuronales Deep Learning para la visión por computador ya se conocían a finales de los 80 del siglo pasado; también los algoritmos fundamentales de Deep Learning para series temporales como LSTM (que trataremos más adelante) ya fueron desarrollados en 1997, por poner algunos ejemplos. Entonces, ¿por qué este *boom* de la inteligencia artificial?

Sin duda, la computación disponible ha sido el principal desencadenante, y por ello dedicamos una sección a explicar algunos de los últimos avances. Pero, además, otros factores han contribuido a desencadenar el potencial de la inteligencia artificial y las tecnologías relacionadas. A continuación, vamos a presentar estos factores de manera que nos permita comprender el porqué de este auge de la inteligencia artificial que solo ha hecho que empezar.

1.4.1. La supercomputación corazón del Deep Learning

Mi carrera profesional se ha desarrollado siempre en el marco de la supercomputación, y en 2006 empecé también a investigar cómo la supercomputación podía contribuir a mejorar los métodos de Machine Learning.

Pero fue en septiembre de 2013, momento en el que ya disponía de una base sobre Machine Learning y en el que empecé a centrar mi interés en Deep Learning, cuando cayó en mis manos el artículo *Building High-level Features Using Large Scale Unsupervised Learning*¹⁴, escrito por investigadores de Google. En este artículo presentado en el congreso International Conference in Machine Learning del año anterior, los autores explicaban cómo entrenaron un modelo Deep Learning en un clúster de 1000 máquinas con 16 000 *cores*. Me interesó muchísimo e impresionó ver cómo la supercomputación permitía acelerar este tipo de aplicaciones y aportaba tanto valor a este campo. En ese momento añadimos este foco en el *roadmap* de nuestro grupo de investigación.

En el año 2012, cuando estos investigadores de Google escribieron este artículo, disponíamos de supercomputadores que permitían resolver problemas que hubieran sido intratables unos pocos años antes debido a la capacidad de computación que se había incrementado siguiendo la ley de Moore¹⁵. Por ejemplo, el computador al que yo tenía acceso en el año 1982, donde ejecuté mi primer programa con tarjetas perforadas, era un Fujitsu que permitía ejecutar algo más de

¹⁴ Quoc Le and Marc'Aurelio Ranzato and Rajat Monga and Matthieu Devin and Kai Chen and Greg Corrado and Jeff Dean and Andrew Ng, *Building High-level Features Using Large Scale Unsupervised Learning*. International Conference in Machine Learning, ICML 2012 [online]. Disponible en: <https://arxiv.org/abs/1112.6209> [Consultado: 12/02/2018].

¹⁵ Ley de Moore. Wikipedia. [online]. Disponible en: https://es.wikipedia.org/wiki/Ley_de_Moore [Consultado: 12/03/2018].

un millón de operaciones por segundo. Treinta años después, en el 2012, el supercomputador MareNostrum que teníamos por aquel entonces en el Barcelona Supercomputer Center - Centro Nacional de Supercomputación¹⁶ (BSC) era solo 1 000 000 000 de veces más rápido que el ordenador en el que yo empecé.

Con la actualización de aquel año, el supercomputador MareNostrum presentaba un rendimiento máximo teórico de 1.1 Petaflops (1 100 000 000 000 000 operaciones de coma flotante por segundo¹⁷). Lo conseguía con 3056 servidores con un total de 48 896 *cores* y 115 000 *Gibabytes* de memoria principal total albergados en 36 *racks*. Por aquel entonces, el supercomputador MareNostrum estaba considerado como uno de los más rápidos del mundo, concretamente ocupaba la trigésimosexta posición en la lista TOP500¹⁸, que se actualiza cada medio año y ordena los 500 supercomputadores más potentes del mundo. En la fotografía de la Figura 1.3 se pueden observar los *racks* de computación del MareNostrum que se albergaban en la capilla de Torres Girona del campus nord de la UPC en Barcelona¹⁹.

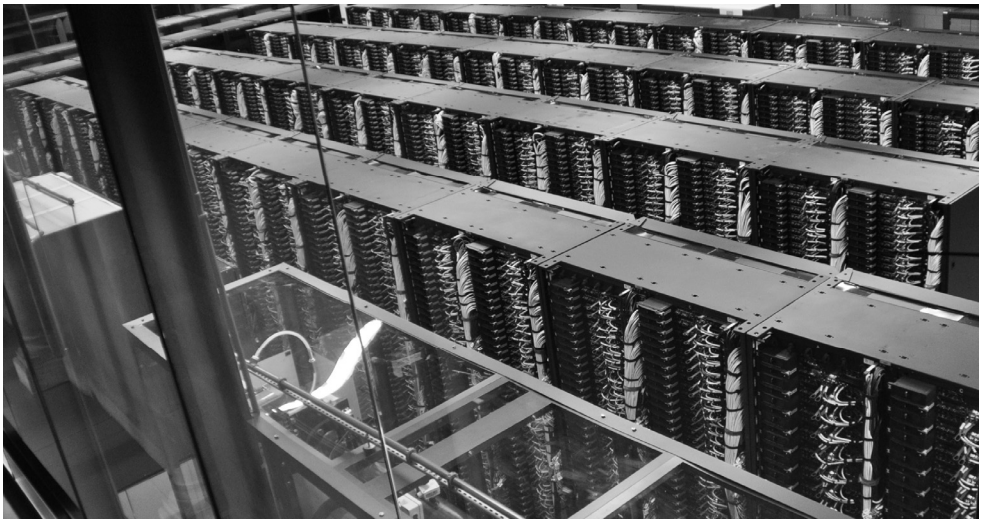


Figura 1.3 Fotografía del MareNostrum 3, que se albergaba en la capilla de Torres Girona del campus nord de la UPC en Barcelona.

¹⁶ Página web del Barcelona Supercomputing Center. <http://www.bsc.es> [Consultado: 12/12/2019].

¹⁷ Operaciones de coma flotante por segundo. Wikipedia. [online]. Disponible en: https://es.wikipedia.org/wiki/Operaciones_de_coma_flotante_por_segundo [Consultado: 12/03/2018].

¹⁸ Top 500 List – November 2012. [online] Disponible en: https://www.top500.org/list/2012/11/?_ga=2.211333845.1311987907.1527961584-375587796.1527961584 [Consultado: 12/12/2019]

¹⁹ MareNostrum 3. Barcelona Supercomputing Center. [online]. Disponible en: <https://www.bsc.es/marenostrum/marenostrum/mn3> [Consultado: 12/03/2018].