

Introducción

En este libro el lector encontrará una guía para adentrarse de manera práctica al *Deep Learning* con la ayuda de la librería Keras, la cual aprenderá a usar con el objetivo de desarrollar y evaluar modelos *Deep Learning*. Aunque el *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 Keras.

Por ello, este libro se centrará en temas prácticos y concretos para descubrir al lector el apasionante mundo que se abre con el uso del *Deep Learning*., teniendo siempre en mente que solo podremos examina 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.

A quién va dirigido el libro

Esta es una obra introductoria, ideada inicialmente para dar soporte a mi docencia en la UPC a alumnos y alumnas de ingeniería informática pero con pocos o ningún conocimiento de *Machine Learning*. Pero a su vez, este libro puede ser también útil a ingenieros e ingenieras que ya dejaron las aulas y se encuentran trabajando pero requieren, o simplemente les apetece aprender sobre este tema.

Por ello el libro está escrito en forma de “distancia corta”, por así decirlo, con el lector, como si estuviéramos en una de mis clases en la UPC. Para facilitarlo, intentaré introducir los conocimientos y conceptos en forma lineal, intentando involucrar al lector al requerirle que esté en todo momento con el teclado delante y probando lo que le voy contando.

No obstante, esta no es una obra para todo el mundo, y es importante que el lector calibre bien sus expectativas; aquí no va a encontrarse explicaciones sobre la teoría fundamental de redes neuronales artificiales, ni se ofrecen detalles de cómo funcionan internamente los algoritmos relacionados con el tema. Por tanto, no es recomendable a quienes ya estén programando redes neuronales en algún otro entorno que no sea Keras, y si fuese esta librería su campo de interés, quizá este es un camino demasiado largo para conocerla.

Organización del libro

El libro se organiza en capítulos que deben ser leídos en orden, ya que van guiando al lector y lo introducen gradualmente a los conocimientos imprescindibles para seguir los ejemplos prácticos, intentando ir al grano y ser lo más conciso posible. Al ser un libro introductorio, considero que es mucho más útil este enfoque que no uno más formal. A pesar de todo, he intentado que el índice del libro exprese un mapa razonablemente ordenado de los principales conceptos del área.

El libro está compuesto por ocho capítulos que mezclan el contenido teórico y práctico, creando un viaje donde se repasan conceptos de las tres familias de arquitecturas de redes neuronales con las que el lector acabará estando familiarizado: redes neuronales densamente conectadas (*densely connected networks*), redes neuronales convolucionales (*convolutional neural networks*) y, finalmente redes neuronales recurrentes (*recurrent neural networks*).

Concretamente, la obra está concebida en 8 capítulos, ni más ni menos, porque son el número de sesiones de clase que dedico a este tema. Por ello, he intentado incluir todos los conceptos que considero básicos en ellos intentando equilibrar al máximo la carga de trabajo que se requiere para poder introducir los conceptos teóricos y prácticos que se proponen. En algún capítulo se dedica más tiempo a “teoría” y en algún otro a “práctica”, pero la idea es que en total queden

compensados.

El primer capítulo contiene una motivación al tema y las instrucciones para preparar el entorno de trabajo para poder seguir los detalles de código que se irán presentando a continuación.

En el segundo capítulo, a partir de un caso de estudio, se explican los conceptos básicos de una red neuronal. Luego introducimos las redes neuronales densamente conectadas y mostramos su implementación en **Keras**.

En el tercer capítulo presentamos cómo se realiza el proceso de aprendizaje de una red neuronal, adentrándonos en algunos de sus componentes más importantes. La segunda parte de capítulo invita al lector a que practique los conocimientos adquiridos con la web *TensorFlow playground*.

En el cuarto capítulo vemos de dónde se obtienen los datos cuando se quiere entrenar una red neuronal y cómo se preprocesan antes de alimentar a la red neuronal con ellas. Se plantean en detalle dos técnicas de vectorización muy utilizadas en los proyectos reales.

En el quinto capítulo analizamos desde un punto de vista absolutamente práctico problemas habituales a los que nos tenemos que enfrentar cuando estamos tratando con modelos *Deep Learning* como puede ser la prevención del sobreajuste de los modelos, el temido *Overfitting*. Se presentarán diferentes técnicas básicas para casos reales.

En el capítulo sexto, ya preparado el lector con una base importante de cómo son las redes neuronales, presentaremos e implementaremos las redes neuronales convolucionales, una de las familias más populares en estos momentos.

El capítulo séptimo lo centraremos en uno de los problemas más habituales en *Deep Learning* en el mundo real, como es el no tener suficientes datos para entrenar nuestras redes. Se explicarán varias técnicas frecuentemente utilizadas en estos momentos en proyectos reales.

Y, finalmente, en el capítulo octavo introducimos las redes neuronales recurrentes, que a pesar de tener requerimientos de computación importantes se están popularizando ya en la industria.

El libro acaba con un capítulo de clausura donde me permito hacer unas reflexiones sobre el tema y su impacto. Este va acompañado de un repositorio de código en el Github donde el lector puede encontrar los ejemplos presentados en el libro. Muchos de los ejemplos usados están inspirados en los ejemplos que François Chollet ha compartido en el github de su último libro, *Deep Learning with Python*¹. ¡Gracias François!

Requisitos para seguir el libro

Como hemos mencionado con frecuencia, esta obra pretende ser una introducción; por ello, no es necesario que el lector sea un experto en Python, solo ha de tener, evidentemente, conocimientos de programación y un poco de 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 primeros conceptos sobre el tema. Solo se suponen los conocimientos básicos en matemáticas que cualquier estudiante de la rama científico-técnica. A partir de ellos, a lo largo de los capítulos se repasan muy brevemente los conceptos más importantes de *Machine Learning* que se puedan requerir.

Asumimos que el lector, antes de empezar a leer, tiene instalado Python y el paquete de Keras en su ordenador, portátil o su instancia de máquina virtual en el Cloud. Es importante remarcar que todos los códigos que se proponen al lector para que los pruebe en su ordenador pueden ser ejecutados en una plataforma

¹ Véase Github Companion Jupyter notebooks for the book “Deep Learning with Python” by François Chollet Accesible en: <https://github.com/fchollet/deep-learning-with-python-notebooks>

que solo tenga CPU; no es un requisito tener una GPU, pero sí que es cierto que la segunda parte del libro se proponen códigos que pueden requerir entrenamientos de redes que en plataformas con solo CPU pueden tardar mucho tiempo. En este caso, y si el lector le apetece, les proponemos que cancelen el proceso de cálculo de los parámetros y se descarguen una versión ya entrenada que contendrá los mismos valores que el lector llegaría a conseguir en su ordenador.

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

¿Por qué Keras?

Keras² es la librería recomendada para *beginners*, puesto que su curva de aprendizaje es muy suave en comparación con otras, a la vez que es, sin duda, una de las herramientas para implementar redes neuronales de mayor popularidad en el momento después de TensorFlow.

Keras es una librería de Python que proporciona, de una manera sencilla, la creación de una gran gama de modelos de *Deep learning* usando como *backend* otras librerías como TensorFlow, Theano o CNTK. Fue desarrollado y es mantenido por François Chollet³, ingeniero de Google, y su código ha sido liberado bajo la licencia permisiva del MIT.

Personalmente, valoro la austeridad y simplicidad que presenta este modelo de programación, sin adornos y maximizando la legibilidad; permite expresar redes neuronales de una manera muy modular, considerando un modelo como una secuencia (o un grafo si se trata de modelos más avanzados). Por último, pero no menos importante, creo que es un gran acierto haberse decantado por usar el

²Véase más en las páginas de documentación de Keras disponibles en: <https://keras.io>

³ Remito aquí la cuenta de Twitter del creador: <https://twitter.com/fchollet>, que verán que es una persona muy activa en twitter.

lenguaje de programación Python.

Por todo ello, he considerado usar Keras en este libro, a pesar de que ya les avanzo que si desean realizar un proyecto para ponerlo en producción deben considerar seriamente usar TensorFlow⁴ y todo su ecosistema de utilidades una vez hayan acabado con este libro.

⁴ Véase la página de documentación de TensorFlow disponible en: <http://www.tensorflow.org>