

Prefacio

Tecnológicamente hablando, nos encontramos en un momento crucial. Saber diseñar y programar tecnología es posiblemente la habilidad que determine la calidad de vida en el futuro. Muchos ya vienen advirtiéndolo la crudeza del «gap tecnológico». No solo habrá que tomar el rol de consumidor de tecnología, sino, en el nivel que sea, el de productor. Y en estas estamos ahora, en una oportunidad para mucha gente, o en un reto, según se mire.

Lo que sí es seguro es que todas las disciplinas académicas se verán impactadas y quien más, quien menos, estará invitado, porque las cosas están más fáciles que antaño. Por un lado, los lenguajes de alto nivel, en especial los lenguajes funcionales y los más orientados a la ciencia del dato, están dando la posibilidad a muchos perfiles de aplicar distintos algoritmos a diversas disciplinas o, simplemente, a participar en proyectos tecnológicos. Por otro lado, las nuevas formas de almacenamiento y computación, habitualmente considerados en el marco del *big-data*, están evitando a esos programadores el tener que enfrentarse a implementar directamente estructuras para manejar datos masivos y computar operaciones sobre ellos. Estos dos condicionantes hacen que estemos asistiendo a una explosión de tecnología, acaso ciertamente desordenada y asimétrica, pero potencialmente muy fructífera. «Ite, inflamate Omnia».

A la vanguardia de toda esta tecnología se encuentran hoy los algoritmos de aprendizaje automático, en especial las redes neuronales artificiales de última generación, consideradas como aprendizaje profundo. De entre ellas, las redes neuronales recurrentes (RNN-LSTM) y los «transformers», a los que se dedica este libro, cobran especial importancia en el procesamiento y producción de secuencias, y en especial de frases y textos.

Existen muchos vídeos y páginas web dedicadas a enseñar a programar, con pocas líneas de código, un modelo basado en

RNN-LSTM o en bloques transformer. También existe la posibilidad de integrar nuestro código con modelos preentrenados con simples peticiones a servicios (por ejemplo, ELMo, BERT, GPT-2, GPT-3, GPT-4, RoBERTa, LaMDA y PaLM, chatGPT, BART). Sin embargo, independientemente de este ímpetu codificador que podamos tener, los problemas que resuelven con esas líneas de código (en ocasiones extremadamente pocas) y los algoritmos que estas despliegan por debajo, pueden ser descritos de una manera conceptual y formal en toda su amplitud. Incluso en ausencia de máquina, estos algoritmos podrían ser idealmente resueltos computando operaciones en la arena de una playa. Esta consigna guiará el libro. Aunque se propone código de ejemplo, las RNN-LSTM y los transformers son explicadas de una manera agnóstica a código. Se pretende que los que están aprendiendo ahora a usarlas en código de alto nivel, también sepan entenderlas de manera exhaustiva, en todas sus partes. No ser un mero ejecutor de unas líneas. Esto ayuda a diseñar mejores entornos y, no menos importante, a explicarlos y comunicarlos. No nos quedemos con la magia y solo con la magia. Comprendamos, queremos comprender. Porque..., aun con el código apuntalado y funcionando, ¿cuántas veces hemos sentido cierto miedo por no comprender las cajas negras de los algoritmos que corremos?

El propósito de este libro es justo ese. Contar los aspectos importantes de las RNN y los transformers como a nosotros mismos nos hubiera gustado que nos las contasen. Ayudar a hacer ese camino. En el libro se tocan formalmente una gran diversidad de temas relevantes que tienen que formar parte del argot y del bagaje intelectual de los que usen redes neuronales recurrentes y transformers, especialmente en procesamiento del lenguaje. El planteamiento es explicarlas de manera pormenorizada sin dar nada por sabido. Si hemos tenido que poner y explicar una misma fórmula en tres sitios distintos, lo hemos hecho. El fin justifica el papel. Entre otras cosas, y sin afán

de exhaustividad, hemos tratado temas molles de las RNN-LSTM y los transformers, como el valor añadido que aportan sobre otro tipo de redes, los mecanismos de memoria LSTM, el mecanismo de retropropagación clásico así como el llamado «a través del tiempo», las distintas topologías de las RNNs y sus tareas a resolver, los hiperparámetros, los algoritmos de optimización del gradiente descendiente, los mecanismos de atención, los bloques transformer y sus funcionalidades e incluso las diferencias en cuanto a recurrencia de una arquitectura Transformer y una RNN-LSTM.

Además, tanto el lenguaje del libro como el tratamiento de la argumentación está sesgado hacia la ciencia cognitiva. Es importante entender la posibilidad que ofrecen todas estas técnicas de homologar comportamiento artificial con comportamiento humano, y reflexionar en cuanto a esta plausibilidad. Esto es útil no solo para crear mejores algoritmos y, por ende, mejores sistemas, sino para crear modelos cognitivos del lenguaje y de la generación de significado altamente explicativos. Por eso también hemos incluido apartados con valoraciones de este tipo.

Como recomendación, invitamos al lector a tomar el libro por dónde más necesite, «por do más pecado haya». Tenemos la confianza de que, al final, la necesidad de ir completando lagunas lleve a la satisfacción del baño placentero, y de esa forma, tener la sensación de no dejar demasiadas cajas negras que abrir, y poder hacer el camino, acaso hacia Ítaca*.

* Todos los recursos gráficos incluidos en esta obra han sido generados por el autor, a excepción de los gráficos 6, 7, 8, 9, 10 y 11, en los que la elaboración del autor es parcial a partir de otras imágenes con licencia CC BY (ver atribución en la mención a estos gráficos) y la Figura 1, obtenida de Visual Phenomena & Optical Illusions de Michael Bach: <https://michaelbach.de/ot/cog-Dalmatian/index.html>. © 2002-20 M.Bach y a su vez de Gregory R (2001) The Medawar Lecture 2001 Knowledge for vision: vision for knowledge. *Phil Trans Biol Sci*.