

BIBLIOGRAFÍA / RECURSOS EN INTERNET

Support Vector Machines

- Vapnik, V. (1995). The Nature of Statistical Learning Theory. Springer, New York.
 - En este libro se presentaron los SVM por primera vez.
- James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An Introduction to Statistical Learning, pp. 351-372. Springer.

Redes Neuronales y Deep Learning

- Bengio, Y., Ducharme, R., Vincent, P., Jauvin, C. (2003). A neural probabilistic language model. *Journal of machine learning research* 3 (Feb), 1137–1155.
 - Disponible también en: <https://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf>
- Cho, K., Raiko, T., Ilin, A. (2013). Enhanced gradient for training restricted Boltzmann machines. *Neural Computation*, 25(3):805–831.
 - Describe el Stochastic Gradient Descent, una de las variantes algorítmicas que más han disminuido el tiempo de entrenamiento de las NN
 - Disponible también en: <https://direct.mit.edu/neco/article/25/3/805/7857/Enhanced-Gradient-for-Training-Restricted>
- Chollet, F. (Ed 1, 2017; Ed 2 prevista oct. 2021). Deep Learning with Python. Manning.
 - Libro de referencia. François Chollet es el autor de Keras, capa que ofrece distintos niveles de abstracción sobre tensorflow, uno de los principales framework de Deep Learning.
- Fukushima, K. (1979). Neural network model for a mechanism of pattern recognition unaffected by shift in position - Neocognitron. *Trans. IECE, J62-A(10)*:658–665.
 - El Neocognitron es un antecesor de las Redes Neuronales Convolucionales (CNN)
- Glorot, X., Bengio, Y. (2011). Understanding the difficulty of training deep feedforward neural networks.
 - Artículo de referencia que demuestra cómo una correcta inicialización de los pesos mejora el entrenamiento de la NN.
 - Disponible en: https://www.researchgate.net/publication/215616968_Understanding_the_difficulty_of_training_deep_feedforward_neural_networks
- Goodfellow, I. (2016). NIPS 2016 Tutorial: Generative Adversarial Networks.
 - Tutorial de las Redes Generativas Antagónicas o GAN escrito por el propio autor.
 - Disponible en: <https://dev.arxiv.org/abs/1701.00160v3>
- Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. MIT Press; 2016.

- Libro de referencia. Liderado por Ian Goodfellow, conocido por ser el autor de las GAN
 - Disponible también en <http://www.deeplearningbook.org>
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors.
 - Describe el dropout, que permite disminuir el overfitting y mejorar el entrenamiento
 - Disponible en: <https://arxiv.org/pdf/1207.0580.pdf>
- Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen. Diploma thesis, Institut für Informatik, Lehrstuhl Prof. Brauer, Technische Universität München.
 - Analiza el problema del Vanishing gradient y demuestra que se puede solucionar cambiando la función de activación
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. of the National Academy of Sciences, 79:2554–2558.
 - Describe un primer modelo de Red Neuronal Recurrente
 - Disponible también en: <https://www.pnas.org/content/pnas/79/8/2554.full.pdf>
- Kingma, D. Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR 2015*.
 - Describe el optimizador Adam, que es uno de los más utilizados actualmente
 - Disponible también en: <https://arxiv.org/pdf/1412.6980.pdf>
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation* 1(4), 541–551.
 - Artículo de referencia en el uso de CNN para el conjunto de datos MNIST
 - Disponible también en: <https://direct.mit.edu/neco/article/1/4/541/5515/Backpropagation-Applied-to-Handwritten-Zip-Code>
- McCulloch, W. S. Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133.
 - Describe el primer modelo conceptual de Neurona Artificial y posibles conexiones desde el ámbito de la lógica proposicional
 - Disponible también en: <https://web.csulb.edu/~cwallis/382/readings/482/mcculloch.logical.calculus.ideas.1943.pdf>
- Minsky, M. Papert, S. *Perceptrons*. MIT Press; 1969.
- Ng, A. (2012). Machine Learning. Coursera.
 - MOOC de referencia en Redes Neuronales
 - Disponible en: <https://www.coursera.org/learn/machine-learning>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review* 65(6), 386–408.
 - Describe el modelo perceptron, evolución del modelo de Neurona Artificial
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. Rumelhart, D. E. and McClelland, J. L. (Eds.), *Parallel Distributed Processing*, Vol. 2, 318–362. MIT Press.

- Describe el algoritmo de backpropagation para entrenar las NN
- Disponible también en https://web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVollChapter8.pdf
- Schmidhuber, J. (2014). Deep Learning in Neural Networks: An Overview.
 - Disponible en: <https://arxiv.org/abs/1404.7828>
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *JMLR* 15(1), 1929–1958.
 - Describe el dropout, que permite disminuir el overfitting y mejorar el entrenamiento
 - Disponible también en: <https://jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf>
- Sutton, R. (2019). The bitter lesson.
 - Este post de uno de los padres del Reinforcement Learning demuestra con diversos ejemplos (ajedrez, visión artificial, reconocimiento de voz) que la fuerza bruta de las redes neuronales se impone con frecuencia a modelos específicos.
 - Disponible en: <http://www.incompleteideas.net/InclIdeas/BitterLesson.html>
- Ujjwalkarn (2016). A Quick Introduction to Neural Networks.
 - Disponible en: <https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/>

Procesado de imágenes

- Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks.
 - Artículo de referencia que describe el modelo de Deep Learning AlexNet que ganó la competición ImageNet Large-Scale Visual Recognition Challenge de 2012 (ILSVRC12). Es considerado como la singularidad de las NN, momento en que superaron a otros modelos de Machine Learning
 - Disponible en: https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf
- Lowe, D. G. (1987). Three-Dimensional Object Recognition from Single Two-Dimensional Images. *Artificial Intelligence*, 31, 3 (March 1987), pp. 355–395
 - Disponible también en: <https://www.cs.ubc.ca/~lowe/papers/aij87.pdf>
- Hubel, D. H., Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.*, 148:574–591.
 - Artículo de referencia. Describe un experimento que demostró que la visión de los mamíferos consta de células simples (sensibles a la orientación de bordes) y de complejas (más robustas frente a rotación y traslación).
 - Disponible también en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1363130/pdf/jphysiol01298-0128.pdf>
- Ujjwalkarn (2016). An Intuitive Explanation of Convolutional Neural Networks.

- Disponible en: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>
- Zeiler, M., Fergus, R. (2013). Visualizing and Understanding Convolutional Networks.
 - Disponible en: <https://arxiv.org/abs/1311.2901>

Procesado de lenguaje natural

- Bahdanau, D., Cho, K., Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. ICLR 2015.
 - Artículo de referencia en modelos de atención con LSTM
 - Disponible en: <https://arxiv.org/pdf/1409.0473.pdf>
- Bloem, P. (2019). Transformers from scratch.
 - Disponible en: <http://peterbloem.nl/blog/transformers>
- Britz, D. (2015). Understanding Convolutional Neural Networks for NLP.
 - Disponible en: <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>
- Brownlee, J. (2017). Long Short-Term Memory Networks With Python.
 - Disponible en (previo pago): <https://machinelearningmastery.com/lstms-with-python/>
- Cho, K., Merriënboer, B., Culcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation.
 - Disponible en: <https://arxiv.org/pdf/1406.1078v3.pdf>
- Dimson, T. This word does not exist.
 - Generador de palabras y de su significado, basado en GPT-2.
 - Disponible en: <https://www.thisworddoesnotexist.com>
- Hochreiter, S. Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8):1735–1780.
 - Artículo de referencia. Su impacto ha sido muy posterior: es en gran parte responsable de la irrupción del Deep Learning en el Procesado de Lenguaje Natural y en otras áreas
 - Disponible también en: https://www.researchgate.net/publication/13853244_Long_Short-term_Memory
- Huggingface. Transformers.
 - Web de referencia en Transformers
 - Disponible en: <https://github.com/huggingface/transformers>
- Jurafsky, D., Martin, J. Speech and Language Processing.
 - Libro de referencia en procesado de Voz y Lenguaje Natural
 - Borrador de tercera edición; prevista finalización a finales de 2021.
 - Disponible en: <https://web.stanford.edu/~jurafsky/slp3/>
- Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space.
 - Artículo de referencia. Describe word2vec, que es uno de los modelos responsables de la expansión del Deep Learning en el NLP
 - Disponible en: <https://arxiv.org/abs/1301.3781>

- Olah, C. (2015). Understanding LSTM Networks.
 - Tutorial de referencia para comprender cómo funcionan las LSTM
 - Disponible en: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I. (2017). Attention is all you need. NIPS 2017.
 - Artículo de referencia en modelos de atención (sin LSTM) – Transformers
 - Disponible en: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>

MATERIAL COMPLEMENTARIO

- Fei-Fei Li. ImageNet.
 - Dataset de referencia; fue concebido para servir de banco de pruebas en modelos de clasificación de imágenes. Su versión clásica consta de mil clases y un millón de imágenes
 - Disponible en: <https://image-net.org>
- Russell, S. Norvig, P. Artificial Intelligence: A Modern Approach (2nd Ed.). Prentice Hall. 2002