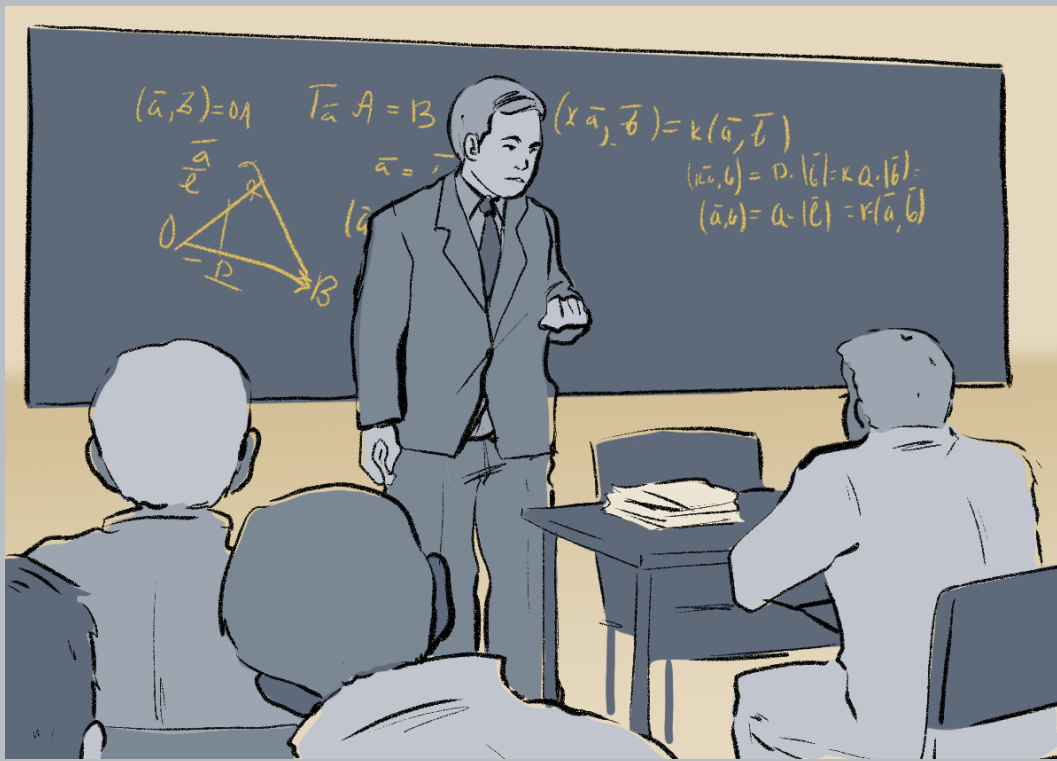




BOURBAKI

Math & Deep Learning en 25 artículos





Esta colección de 25 artículos académicos ha sido seleccionada para acompañar el curso Math & Deep Learning, un programa diseñado para proporcionar al estudiante una formación sólida en los fundamentos matemáticos y computacionales que sustentan las arquitecturas modernas de redes neuronales. El curso busca ir más allá de la aplicación práctica, ofreciendo una comprensión rigurosa de las estructuras algebraicas, los procesos estocásticos y las funciones no lineales que permiten el funcionamiento del aprendizaje profundo.

Los artículos incluidos en esta bibliografía cubren los pilares teóricos y las técnicas centrales tratadas en el curso: redes densas y sus funciones de activación, redes convolucionales vinculadas al análisis de Fourier, modelos recurrentes y su relación con procesos estocásticos, técnicas de optimización para el entrenamiento en GPU, autoencoders combinados con la teoría de los valores extremos, y los fundamentos matemáticos de los modelos generativos y de las Generative Adversarial Networks (GAN), junto con elementos de teoría de juegos. Se ha priorizado literatura que conecte de manera directa el marco matemático con la implementación práctica en Python y con la construcción de modelos robustos y eficientes.

Esta selección busca fortalecer la comprensión conceptual del estudiante y fomentar un enfoque crítico al diseño, calibración y optimización de arquitecturas profundas. Al integrar matemáticas avanzadas con ejemplos reales, la colección permite comprender cómo los principios teóricos se traducen en soluciones aplicadas en visión por computadora, modelado secuencial, compresión de información y generación de datos. Con ello, se aspira a que el participante no solo domine las herramientas, sino que adquiera la capacidad de adaptar, analizar y justificar modelos de deep learning desde sus fundamentos, consolidando una formación integral que articula rigor académico con aplicaciones contemporáneas en inteligencia artificial.

1.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–133.

2.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.

3.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11),

4.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211.

5.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.

6.

Vaswani, A., et al. (2017). Attention is all you need. *NeurIPS*.

7.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks. *Journal of Computational Physics*, 378, 686–707.

8.

Goodfellow, I., et al. (2014). Generative adversarial nets. *NeurIPS*.

9.

Kingma, D. P., & Welling, M. (2014). Auto-Encoding Variational Bayes. *ICLR*.



10.

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166.

11.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.

12.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

13.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *ICML*.

14.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *NeurIPS*.

15.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

16.

Nesterov, Y. (1983). A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, 27(2), 372–376.

17.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR*.

18.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. Wiley.



19.

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. Random House.

20.

Goodfellow, I., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. ICLR.

21.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

22.

Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. JMLR, 17(39), 1–40.

23.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(42), 230–

24.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. CVPR.

25.

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554.



BOURBAKI

ESCUELA DE MATEMÁTICAS

CURIOSIDAD Y SABER

CDMX, MEXICO

Estaremos encantados de explicar
con detalle sobre nuestros cursos

Escríbenos por mail o WhatsApp

info@colegio-bourbaki.com

+52 56 2141 7850

colegio-bourbaki.com

