

Fundamentos Teóricos de Redes Neuronales

(Seminario de Estadística)

Descripción

Este seminario está diseñado para proporcionar una inmersión en los fundamentos teóricos de las redes neuronales, enfocándose principalmente en la generalización, el paisaje de optimización y la aproximación de funciones. La teoría se complementará con la discusión de experimentos numéricos diseñados para reforzar la intuición y mejorar las habilidades prácticas de los participantes. El seminario está dirigido a estudiantes de posgrado que tengan una sólida formación en estadística y aprendizaje de máquina, y que estén interesados en explorar los desafíos teóricos y prácticos asociados al uso de las redes neuronales.

Prerrequisitos

Para participar de manera efectiva en este seminario, es esencial **haber completado al menos un curso teórico en aprendizaje de máquina** o poseer experiencia previa equivalente. Además, los participantes deben **tener conocimientos básicos de Python** para aprovechar las discusiones de los experimentos numéricos.

Objetivos

Al completar con éxito este seminario, los participantes deberán:

- comprender los fundamentos básicos de la teoría de redes neuronales;
- entender las implicaciones prácticas de estos fundamentos;
- reconocer los desafíos asociados al uso de las redes neuronales.

Temario

1. Introducción a las redes neuronales
 - a. Redes neuronales superficiales y profundas
 - b. Redes neuronales convolucionales y recurrentes
 - c. Autocodificadores y redes generativas
2. Generalización
 - a. Aprendizaje estadístico clásico
 - b. Estabilidad algorítmica y generalización
 - c. Fenómeno del doble descenso
3. Panorama de optimización
 - a. Desafíos de la optimización no convexa
 - b. Mínimos locales, puntos de silla y mínimos globales
 - c. Métodos iterativos y sus propiedades de convergencia
4. Aproximación de funciones
 - a. Propiedad universal de aproximación de funciones
 - b. Resultados de aproximación cuantitativos
 - c. Comparación de redes superficiales y profundas

Bibliografía Básica

- Barron, A.R., 1993. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information Theory*, 39(3), pp.930-945.
- Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep Learning*. MIT Press.

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.
- Hazan, E., 2016. Introduction to Online Convex Optimization. *Foundations and Trends® in Optimization*, 2(3-4), pp.157-325.
- Lee, J.D., Simchowitz, M., Jordan, M.I. and Recht, B., 2016, June. Gradient descent only converges to minimizers. In *Conference on Learning Theory* (pp. 1246-1257). PMLR.
- Liang, S. and Srikant, R., 2017. Why deep neural networks for function approximation?. In *5th International Conference on Learning Representations, ICLR 2017*.
- Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B. and Sutskever, I., 2021. Deep double descent: Where bigger models and more data hurt. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12), p.124003.
- Safran, I. and Shamir, O., 2018, July. Spurious local minima are common in two-layer ReLU neural networks. In *International Conference on Machine Learning* (pp. 4433-4441). PMLR.
- Shalev-Shwartz, S. and Ben-David, S., 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
- Xu, A. and Raginsky, M., 2017. Information-theoretic analysis of generalization capability of learning algorithms. *Advances in Neural Information Processing Systems*, 30.

Acerca del Curso

- Evaluación: tareas (60%) y proyecto final (40%)
- Horario: viernes de 12:00pm a 2:30pm
- Instructor: Mario Diaz, IIMAS - Oficina 115, mario.diaz@sigma.iimas.unam.mx