

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Base de datos de dígitos manuscritos
 - Del 0 al 9 escritos a mano por diferentes personas
 - Clasificación: 10 clases
 - 60.000 imágenes de entrenamiento
 - 10.000 imágenes de prueba
 - Imágenes en b/n de 28x28 píxeles
- Página web oficial:
<http://yann.lecun.com/exdb/mnist/>

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - M A Belyaev and A A Velichko, Classification of handwritten digits using the Hopfield network, 2020, IOP Conference Series: Materials Science and Engineering, 862 052048

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Preprocesado:
 - Método de *deskewing*
 - Ajusta el desplazamiento lineal y angular de la imagen
 - Se basa en la distribución del brillo de los píxeles
 - En la imagen procesada, el dígito está situado en el centro de la imagen y tiene una inclinación normal



(a)



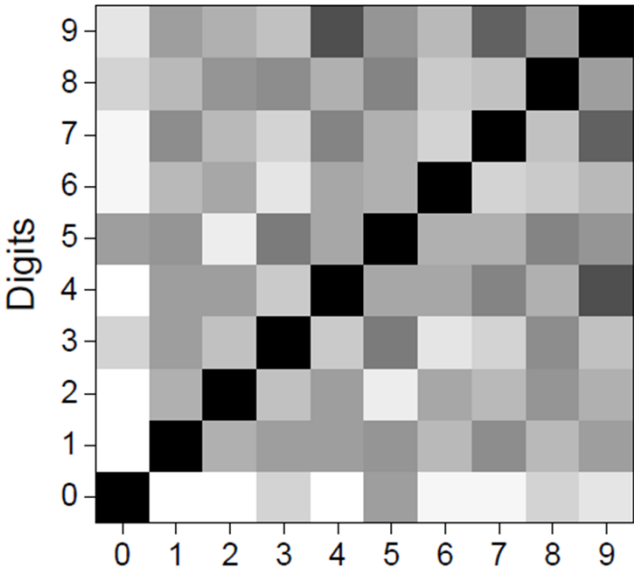
(b)

Figure 1. Image of handwritten digits from the MNIST database before (a) and after (b) processing by the deskewing method.

REDES DE HOPFIELD

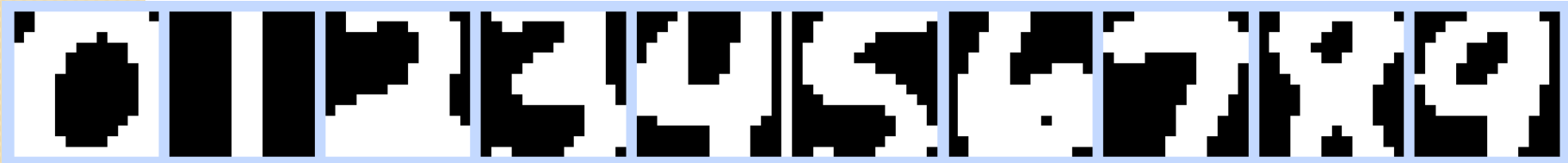
- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Se promedian todas las imágenes del conjunto de entrenamiento para cada dígito
 - 6000 imágenes/dígito
 - Se aplica un umbral a estas 10 imágenes
 - Se convierten a binarias
 - $28 \times 28 = 784$ valores binarios (784 neuronas)
 - Una red de Hopfield podrá almacenar aproximadamente 108 patrones ($\sim 0,138 \times 784$)
 - Los patrones binarios resultantes se utilizan luego como entradas para la red Hopfield
 - Cada patrón se almacena como un estado atractor en la red

- Clasificación con Redes de Hopfield



REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Para aumentar la distancia:
 - Se elimina de los patrones el espacio en blanco que rodea al dígito
 - Se convierten a resoluciones más pequeñas
 - 14x14
 - Patrones finales utilizados para calcular la matriz de pesos:



REDES DE HOPFIELD

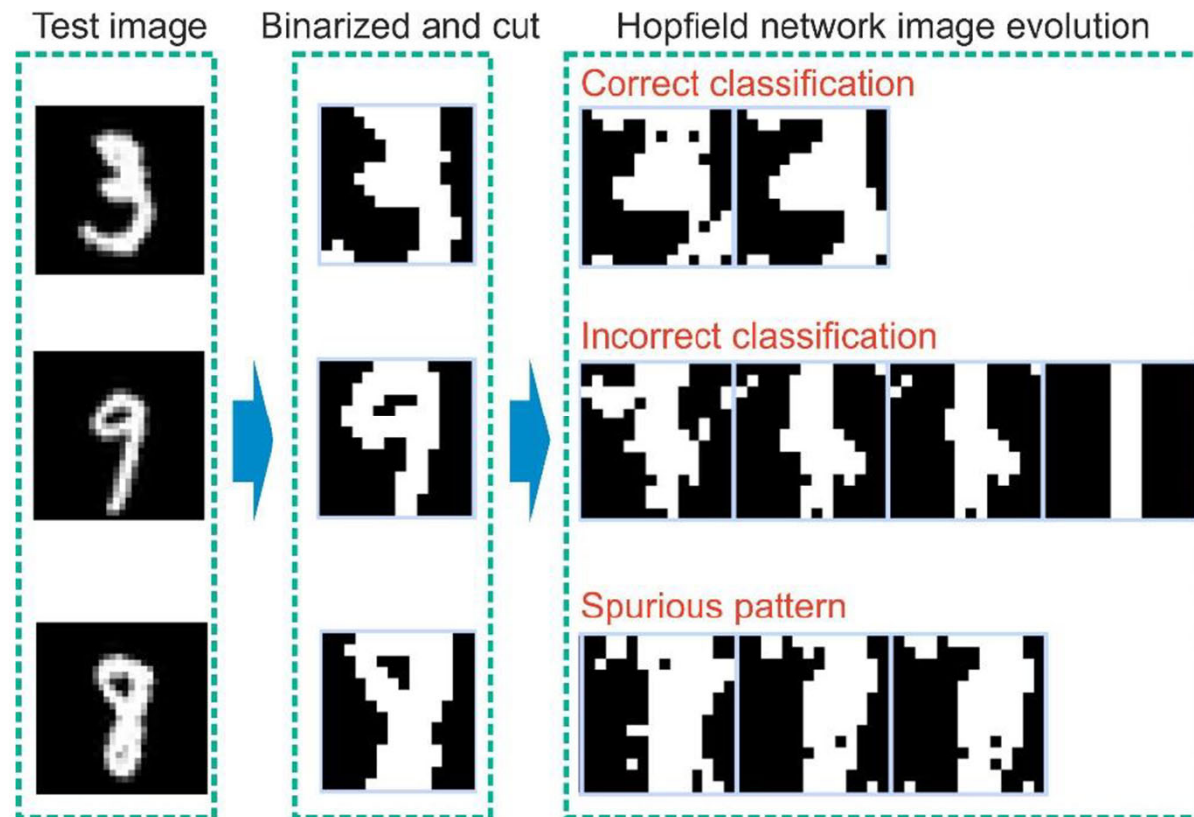
- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Entrenamiento: Método Storkey
 - Aumentar la capacidad de la memoria asociativa
 - Método iterativo

$$\mathbf{W}_i = \mathbf{W}_{i-1} + \frac{1}{N} \mathbf{X}_{\text{mem}_i} \cdot \mathbf{X}_{\text{mem}_i}^T - \frac{1}{N} \cdot \mathbf{X}_{\text{mem}_i} \cdot (\mathbf{W}_{i-1} \cdot \mathbf{X}_{\text{mem}_i})^T - \frac{1}{N} \cdot (\mathbf{W}_{i-1} \cdot \mathbf{X}_{\text{mem}_i}) \cdot \mathbf{X}_{\text{mem}_i}^T$$

- Busca modificar los pesos de la red para que los patrones de entrenamiento sean atractores estables
 - Que los estados espurios sean repelidos
 - Esto contribuye a mejorar la capacidad de almacenamiento y recuperación de patrones

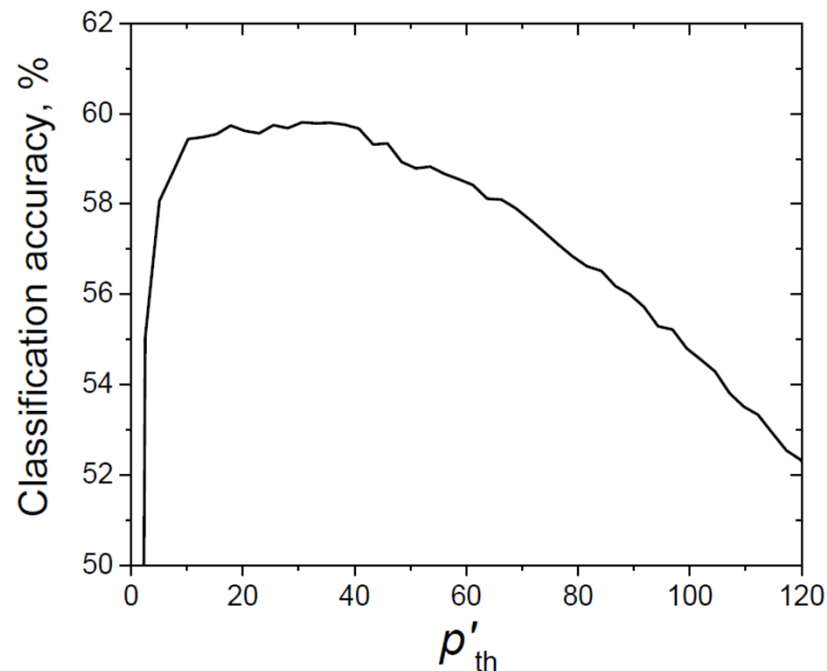
REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados:



REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados:
 - Precisión en la clasificación: 59.8%
 - Dependencia del umbral



REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Entrenamiento:

Digit	Percent of correct classification, %	Percent of incorrect classification, %	The most common erroneous result	Percent of spurious patterns, %
0	69	5.49	6	25.5
1	90.2	0.35	7	9.49
2	73.7	4.44	1	21.9
3	67.7	8.72	2	23.5
4	0	16.1	6	83.9
5	52.4	9.04	6	38.5
6	80.4	3.94	1	15.7
7	69.6	4.86	2	25.5
8	57.3	13.3	3	29.3
9	29.9	25.4	7	44.8

REDES DE HOPFIELD

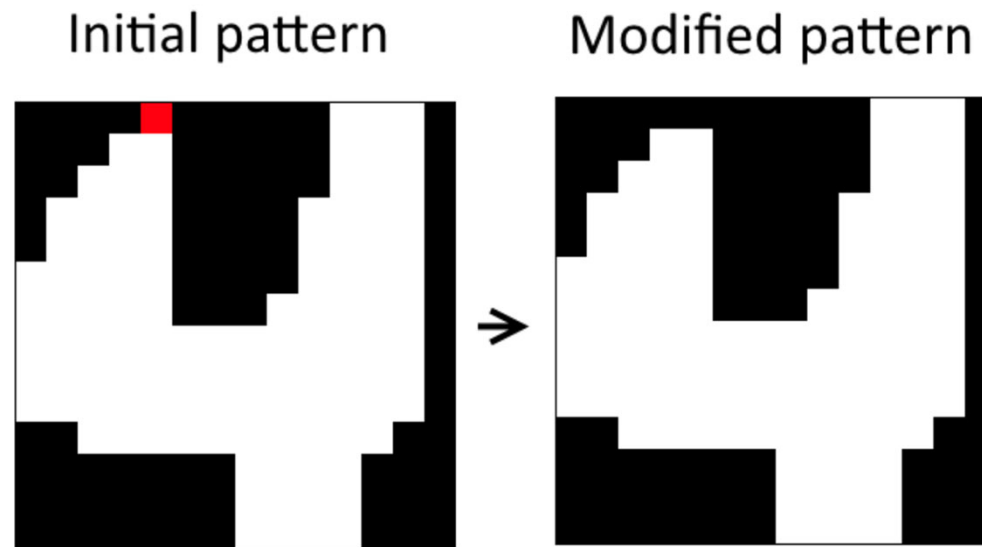
- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Entrenamiento:
 - Los dígitos 1, 2, 6 y 7 tienen las precisiones más altas
 - Muy probablemente, sus patrones son fuertes atractores
 - Esta suposición confirma el hecho de que estos dígitos son el resultado erróneo más probable de la clasificación de otros dígitos
 - El dígito 1 es el que se reconoce con mayor precisión
 - Esto se debe, probablemente, a la pequeña variabilidad de su forma de escritura

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Entrenamiento:
 - Los dígitos 9 y 4 tienen la menor distancia de Hamming relativa media con respecto a otros dígitos
 - Tienen el mayor porcentaje de clasificación incorrecta
 - El dígito 4 nunca se ha clasificado correctamente, y la mayoría de los resultados erróneos están relacionados con patrones espurios (83.9%)
 - El patrón espurio más común difiere del patrón memorizado en sólo 1 píxel
 - Esto significa que ambos patrones son atractores de la red, pero el patrón espurio es un atractor más fuerte, y en muchos casos la red converge hacia él

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Por este motivo, se sustituye el patrón original del dígito 4



- Se vuelve a calcular la matriz de pesos W

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Entrenamiento:

Digit	Percent of correct classification, %	Percent of incorrect classification, %	The most common erroneous result	Percent of spurious patterns, %
0	68.6	5.49	6	25.9
1	90.8	0.35	7	8.87
2	73.5	4.84	1	21.7
3	66.9	8.33	2	24.8
4	15.8	16.3	6	67.9
5	52.8	8.57	6	38.6
6	80.7	3.94	1	15.4
7	69.5	5.89	2	24.6
8	56.7	13.5	3	29.9
9	27.2	31.1	7	41.7

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Entrenamiento:
 - La precisión de la clasificación del dígito 4 aumentó al 15,8%
 - Sin embargo, la precisión de la clasificación del dígito "9" descendió del 29.9% al 27,2%
 - Debido a un aumento de la similitud de los patrones de los dígitos "4" y "9"
 - La precisión de la clasificación de la red aumenta al 61%

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Test:
 - Muestra de test: 10000 imágenes
 - Precisión en test: 61,5%
 - Similar a la anterior
 - Distribución también similar a la tabla anterior
 - La precisión de clasificación de los dígitos 4 y 9
 - 4: 16,9%
 - 9: 27,6%

REDES DE HOPFIELD

- Ejemplo: Base de datos MNIST
 - Clasificación con Redes de Hopfield
 - Resultados: Test:
 - En general, precisión muy inferior a otras técnicas
 - MLP
 - Deep Learning
 - La precisión de la clasificación de la red se puede mejorar mediante el preprocesamiento de los datos de entrada y la selección correcta de patrones memorizados
 - Los patrones espurios introducen un error significativo

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield continuas
 - Redes de Hopfield con asimetría
 - Redes de Hopfield con autoconexiones
 - Redes de Hopfield con dinámica estocástica
 - Redes de Hopfield con aprendizaje
 - Otras variantes importantes:
 - Redes de Hopfield con ruido
 - Redes de Hopfield con umbrales adaptativos
 - Redes de Hopfield con múltiples capas
 - Redes de Hopfield con aprendizaje supervisado

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield continuas (1/3)
 - Permiten valores de activación neuronales reales en lugar de binarios (0/1)
 - Esto aumenta drásticamente la complejidad de los patrones que la red puede aprender y recuperar
 - Ofrecen mayor flexibilidad y precisión en la representación de patrones

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield continuas (2/3)
 - La capacidad de almacenamiento de la red es muy superior que en las redes binarias
 - La cantidad de patrones distintos que puede recordar y recuperar es muy superior
 - Las redes binarias son propensas a mínimos "falsos"
 - Si los recuerdos aprendidos por una red binaria son demasiado similares, o si se aprenden demasiados vectores de patrones, la red corre el riesgo de converger en un recuerdo intermedio
 - Alguna combinación de patrones aprendidos
 - Es decir, la red no podrá discriminar entre patrones y se volverá inútil

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield continuas (3/3)
 - Se utilizan en aplicaciones como el procesamiento de imágenes y la optimización
 - Referencias:
 - Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81(10), 3088-3092.
 - Tank, D. W., & Hopfield, J. J. (1986). Simple neural optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Transactions on Circuits and Systems*, 33(11), 1088-1100.

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield con asimetría
 - Las conexiones entre neuronas no son necesariamente simétricas
 - Permiten la formación de patrones más complejos y la discriminación entre patrones similares
 - Se utilizan en aplicaciones como el reconocimiento de patrones y la clasificación
 - Referencias:
 - Peretto, P. (1988). Collective properties of neural networks with asymmetric interactions. *Journal of Physics A: Mathematical and General*, 21(11), 2855-2862.

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield con autoconexiones
 - Las neuronas pueden tener conexiones consigo mismas
 - Permiten la implementación de funciones de memoria a corto plazo y el auto-reforzamiento de patrones
 - Se utilizan en aplicaciones como el aprendizaje asociativo y la predicción
 - Referencias:
 - Wang, D., & Buhmann, J. M. (1993). Learning in recurrent Hopfield networks with self-connections. *Neural Networks*, 6(5), 747-762

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield con dinámica estocástica
 - Incorporan ruido en la dinámica de la red
 - Permiten la exploración de diferentes soluciones y la evasión de mínimos locales
 - Se utilizan en aplicaciones como la optimización combinatoria y la búsqueda de soluciones creativas
 - Referencias:
 - Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6(6), 721-741

REDES DE HOPFIELD

- Variantes más importantes:
 - Redes de Hopfield con aprendizaje
 - Los pesos de las conexiones se pueden modificar de forma dinámica en función de la experiencia
 - Permiten la adaptación de la red a nuevos patrones y la mejora del rendimiento con el tiempo
 - Se utilizan en aplicaciones como el aprendizaje automático y la robótica
 - Referencias:
 - Williams, R. J., & Zipser, D. (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1(2), 270-280

REDES DE HOPFIELD

- Variantes más importantes:
 - Otras variantes importantes:
 - Redes de Hopfield con ruido
 - Permiten la presencia de ruido en el proceso de almacenamiento y recuperación de patrones
 - Se aumenta la robustez de la red a datos corruptos o incompletos
 - Redes de Hopfield con umbrales adaptativos
 - Ajustan dinámicamente los umbrales de activación de las neuronas en función de la entrada
 - Se mejora la capacidad de la red para almacenar y recuperar patrones con diferentes características

REDES DE HOPFIELD

- Variantes más importantes:
 - Otras variantes importantes:
 - Redes de Hopfield con múltiples capas
 - Permiten una mayor complejidad en la representación de patrones y relaciones entre ellos
 - Similar a las redes neuronales profundas
 - Pero con la ventaja de la eficiencia computacional y la interpretabilidad de las redes de Hopfield
 - Redes de Hopfield con aprendizaje supervisado
 - Combinan el almacenamiento asociativo de las redes de Hopfield con la capacidad de aprendizaje de las técnicas de aprendizaje supervisado
 - Permiten la clasificación y recuperación de patrones con mayor precisión

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Ramsauer, H., Schäfl, B., Lehner, J., Seidl, P., Widrich, M., Gruber, L., Holzleitner, M., Adler, T., Kreil, D., Kopp, M. K., Klambauer, G., Brandstetter, J., & Hochreiter, S. (2021). Hopfield Networks is All You Need. arXiv preprint arXiv:2008.02217.
 - Avance más importante en el ámbito de las redes Hopfield continuas
 - Conexión con arquitecturas profundas

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - El artículo propone un marco novedoso denominado «Deep Hopfield Networks» que combina los puntos fuertes del aprendizaje profundo y las redes Hopfield
 - Utiliza redes de Hopfield **continuas** como una capa dentro de una **estructura profunda**
 - La red Hopfield se encarga del almacenamiento y recuperación de patrones
 - Las otras capas de la red profunda se encargan de tareas como la extracción de características y la clasificación

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Ideas clave en ese trabajo:
 - Estados continuos
 - Valores de activación reales para las neuronas
 - Aumenta la flexibilidad y precisión
 - Regla de actualización eficiente
 - Introduce una nueva regla de actualización derivada del mecanismo de atención de los Transformers
 - Equivalente al mecanismo de atención de los Transformers
 - Esto permite actualizaciones rápidas y estables
 - Arquitectura profunda
 - Se pueden integrar en arquitecturas de aprendizaje profundo

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Dentro del aprendizaje profundo, las capas de Hopfield permiten nuevas formas de capa
 - Más allá de las capas totalmente conectadas, convolucionales o recurrentes
 - Proporcionan mecanismos de agrupación, memoria, asociación y atención
 - Mejoraron el estado del arte en la clasificación del repertorio inmunitario y los problemas de aprendizaje de múltiples instancias

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Características de las redes de Hopfield:
 - Red de Hopfield moderna con estados continuos
 - Esta nueva red Hopfield puede almacenar muchos patrones de forma exponencial
 - Recupera el patrón con una actualización y tiene errores de recuperación exponencialmente pequeños
 - Tiene tres tipos de mínimos de energía: punto fijo global, estados metaestables y puntos fijos
 - Nueva regla de actualización
 - Equivalente al mecanismo de atención utilizado en los Transformers

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Principales ventajas de utilizar redes de Hopfield en una estructura profunda:
 - Almacenamiento masivo
 - Puede almacenar una cantidad de patrones muy superior a las redes clásicas
 - Ideales para tareas que requieren almacenar grandes cantidades de datos
 - Recuperación rápida
 - Recupera patrones con una sola actualización
 - Ideales para tareas que requieren un acceso rápido a la información

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Principales ventajas de utilizar redes de Hopfield en una estructura profunda:
 - Robustez al ruido
 - Ideales para tareas que se trabajan con datos corruptos o incompletos
 - Integración con aprendizaje profundo
 - Se puede integrar en arquitecturas de aprendizaje profundo como capas dedicadas al almacenamiento, acceso y procesamiento de datos

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Hasta ahora, el aprendizaje profundo tenía dificultades con los conjuntos de datos pequeños
 - Sin embargo, las redes de Hopfield son prometedoras para manejar conjuntos de datos pequeños
 - Pueden almacenar los puntos de datos de entrenamiento o sus representaciones para realizar métodos basados en similitud, vecino más cercano o cuantificación vectorial de aprendizaje

REDES DE HOPFIELD

- Redes de Hopfield continuas y profundas
 - Resultados
 - 75 datasets “pequeños” del UCI
 - Menos de 1000 muestras
 - Comparación con la mayoría de técnicas de AA

Method	avg. rank diff.	<i>p</i> -value
Hopfield (ours)	− 3.92	—
SVM	−3.23	0.15
SNN	−2.85	0.10
RandomForest	−2.79	0.05
...
Stacking	8.73	1.2e−11

Table 2: Results on 75 small datasets of the UCI benchmarks given as difference to average rank.