LECTURE 07 | المحاضرة 07

Deep Learning

التعلم العميق

Deep Neural Networks, ReLU, CNN, RNN, and Power-System Applications

الشبكات العصبية العميقة وReLU وCNN وRNN وتطبيقاتها في نظم الطاقة

Prepared by | إعداد: Dr.-Ing. Aouss Gabash

🎯 Learning Objectives

- Understand what makes a neural network deep.
- Distinguish shallow and deep networks.
- Explain hierarchical feature learning.
- Understand ReLU and the vanishing-gradient problem.
- Understand the basic ideas of CNNs and RNNs.
- Identify deep-learning applications in power systems.

🎯 أهداف التعلم

- فهم المقصود بالشبكة العصبية العميقة.
- التمييز بين الشبكات الضحلة والعميقة.
- فهم التعلم الهرمي للميزات.
- فهم ReLU ومشكلة تلاشي التدرج.
- فهم الفكرة الأساسية لـ CNN وRNN.
- التعرف على تطبيقات التعلم العميق في نظم الطاقة.

1. What is Deep Learning?

Deep Learning is a branch of Machine Learning based mainly on neural networks containing multiple hidden layers.

> Deep networks learn increasingly abstract representations as information moves through successive layers.

1. ما هو التعلم العميق؟

التعلم العميق هو فرع من التعلم الآلي يعتمد بصورة أساسية على الشبكات العصبية متعددة الطبقات المخفية.

> تتعلم الشبكات العميقة تمثيلات أكثر تعقيدًا وتجريدًا كلما انتقلت المعلومات بين الطبقات.

2. Shallow vs Deep Networks

Shallow Network
Usually one hidden layer or only a few layers.

Deep Network
Multiple hidden layers arranged hierarchically.

2. الشبكات الضحلة والعميقة

الشبكة الضحلة
تحتوي عادةً على طبقة مخفية واحدة أو عدد قليل من الطبقات.

الشبكة العميقة
تحتوي على عدة طبقات مخفية مرتبة بصورة هرمية.

3. Hierarchical Feature Learning

Each layer transforms the representation produced by the previous layer.

Raw Input → Simple Features → Combined Features →

High-Level Representation → Prediction

Depth allows complex features to be built from simpler ones.

3. التعلم الهرمي للميزات

تحول كل طبقة تمثيل الطبقة السابقة إلى تمثيل جديد.

تبني الطبقات المتقدمة ميزات معقدة انطلاقًا من ميزات أبسط.

4. Mathematical Form

For layer l:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}$$

$$a^{(l)} = \varphi(z^{(l)})$$

- $W^{(l)}$: weight matrix
- $b^{(l)}$: bias vector
- $a^{(l-1)}$: previous-layer output
- φ: activation function

4. الصيغة الرياضية

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}$$

$$a^{(l)} = \varphi(z^{(l)})$$

تستقبل كل طبقة خرج الطبقة السابقة وتطبق عليه تحويلًا خطيًا ثم دالة تنشيط.

5. ReLU Activation

$$ReLU(z) = \max(0, z)$$

If z≤0
ReLU(z)=0

If z>0
ReLU(z)=z

Its derivative is:

$$ReLU'(z) = 0 for z < 0, \quad \text{and} \quad 1 for z > 0$$

ReLU is computationally simple and often supports more stable training in deep networks.

5. دالة **ReLU**

$$ReLU(z) = \max(0, z)$$

تعطي صفرًا للقيم السالبة، وتعيد القيمة نفسها عندما تكون موجبة.

تمتاز ReLU بالبساطة الحسابية وتساعد غالبًا على تدريب الشبكات العميقة.

6. Vanishing Gradients

Backpropagation multiplies derivatives across many layers.

$$\partial E/\partial W^{(1)} = \partial E/\partial a^{(L)} \cdot \partial a^{(L)}/\partial a^{(L-1)} \cdot \ldots \cdot \partial a^{(1)}/\partial W^{(1)}$$

If many derivatives are smaller than one, their product may become extremely small.

When gradients vanish, early layers learn very slowly.

6. تلاشي التدرج

يضرب الانتشار الخلفي مشتقات كثيرة عبر الطبقات.

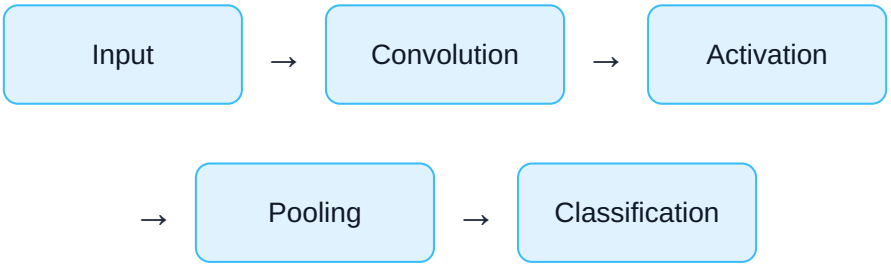
إذا كانت هذه المشتقات أصغر من الواحد، فقد يصبح حاصل ضربها صغيرًا جدًا.

عند تلاشي التدرج تتعلم الطبقات الأولى ببطء شديد.

7. Convolutional Neural Networks

CNNs are specialized for extracting local patterns from structured arrays such as images, time-frequency maps, and spatial measurements.

$$FeatureMap = Input * Kernel$$



7. الشبكات الالتفافية CNN

تستخدم CNN لاستخراج الأنماط المحلية من الصور أو الخرائط الزمنية-الترددية أو القياسات المكانية.

تتعلم المرشحات الالتفافية تلقائيًا الميزات المهمة من البيانات.

8. Recurrent Neural Networks

RNNs are designed for sequential data and maintain information from previous time steps.

$$h_t = \varphi(W_x x_t + W_h h_{t-1} + b)$$

$$y_t = W_y h_t + c$$

The hidden state h_t acts as a memory of earlier inputs.

8. الشبكات المتكررة RNN

صممت RNN للبيانات المتسلسلة، وتستخدم حالة مخفية للاحتفاظ بمعلومات من الأزمنة السابقة.

$$h_t = \varphi(W_x x_t + W_h h_{t-1} + b)$$

9. LSTM and Long-Term Dependencies

Long Short-Term Memory networks extend RNNs using gates that control what information is stored, forgotten, and exposed.

Forget Gate
Controls discarded information.

Input Gate
Controls new stored information.

Output Gate
Controls the exposed state.

9. شبكات LSTM

توسع LSTM بنية RNN باستخدام بوابات تتحكم في المعلومات التي يتم الاحتفاظ بها أو نسيانها أو إخراجها.

10. Training a Deep Network

Mini-Batch → Forward Pass → Loss → Backpropagation → Optimizer

$$\theta_{k+1} = \theta_k - \eta \nabla \theta J(\theta_k)$$

Common optimizers include SGD, Momentum, RMSProp, and Adam.

10. تدريب الشبكة العميقة

دفعة بيانات ← انتشار أمامي ← حساب الخطأ ← انتشار خلفي ← تحديث المعاملات

$$\theta_{k+1} = \theta_k - \eta \nabla \theta J(\theta_k)$$

11. Regularization

Dropout
Randomly disables neurons during training.

L2 Regularization
Adds a weight penalty to the loss.

$$J_r = J + \lambda \|W\|^2$$

Early Stopping
Stops training when validation performance worsens.

Data Augmentation
Creates additional training examples.

11. الحد من فرط المطابقة

تشمل الطرائق Dropout وL2 والإيقاف المبكر وزيادة البيانات.

$$J_r = J + \lambda \|W\|^2$$

12. Power-System Applications

Load Forecasting
Learn nonlinear temporal demand patterns.

Renewable Forecasting
Predict PV and wind generation.

Fault Diagnosis
Classify faults from measurements or signal maps.

State Estimation
Estimate system states from incomplete measurements.

Battery Management
Estimate SOC, SOH, and degradation.

Cybersecurity
Detect abnormal and malicious behavior.

12. تطبيقات نظم الطاقة

- التنبؤ بالحمل والطاقة المتجددة.
- تشخيص الأعطال.
- تقدير حالة النظام.
- تقدير SOC وSOH للبطاريات.
- كشف الهجمات والسلوك غير الطبيعي.

13. Example: Sequence-Based Load Forecasting

An RNN or LSTM can receive a sequence of previous loads:

$$[P_{t-23}, P_{t-22}, \ldots, P_t] \longrightarrow \text{Network} \longrightarrow \hat{P}_{t+1}$$

Additional inputs may include temperature, hour, weekday, and holiday indicators.

13. مثال: التنبؤ المتسلسل بالحمل

يمكن لـ RNN أو LSTM استقبال سلسلة من قيم الحمل السابقة والتنبؤ بالحمل في الساعة القادمة.

$$[P_{t-23}, \ldots, P_t] \rightarrow \text{الشبكة} \rightarrow \hat{P}_{t+1}$$

14. Advantages and Limitations

Advantages

- Automatic feature learning.
- Strong nonlinear modeling.
- Good performance with large datasets.

Limitations

- High data and computing requirements.
- Complex tuning.
- Limited interpretability.

14. المزايا والقيود

المزايا

- تعلم الميزات تلقائيًا.
- تمثيل العلاقات غير الخطية.
- أداء قوي مع البيانات الكبيرة.

القيود

- الحاجة إلى بيانات وحوسبة كبيرة.
- صعوبة ضبط النموذج.
- ضعف قابلية التفسير.

15. Summary

- Deep networks contain multiple hidden layers.
- Depth supports hierarchical feature learning.
- ReLU is widely used in deep networks.
- Vanishing gradients can slow early-layer learning.
- CNNs learn local spatial patterns.
- RNNs and LSTMs model sequential data.
- Deep Learning supports forecasting, diagnosis, estimation, and monitoring in power systems.

15. الخلاصة

- تحتوي الشبكات العميقة على عدة طبقات مخفية.
- يسمح العمق بتعلم ميزات هرمية.
- تستخدم ReLU على نطاق واسع.
- قد يبطئ تلاشي التدرج تعلم الطبقات الأولى.
- تتعلم CNN الأنماط المكانية المحلية.

- تعالج RNN وLSTM البيانات المتسلسلة.
- يدعم التعلم العميق التنبؤ والتشخيص والتقدير والمراقبة في نظم الطاقة.

📝 Review Questions

1. What makes a neural network deep?
2. How does hierarchical feature learning work?
3. Write the equations of one deep-network layer.
4. What is ReLU?
5. What is the vanishing-gradient problem?
6. What are CNNs used for?
7. What are RNNs and LSTMs used for?
8. Name four power-system applications.

📝 أسئلة المراجعة

1. متى تسمى الشبكة العصبية عميقة؟
2. كيف يعمل التعلم الهرمي للميزات؟
3. اكتب معادلات طبقة واحدة.
4. ما هي ReLU؟
5. ما مشكلة تلاشي التدرج؟
6. في ماذا تستخدم CNN؟
7. في ماذا تستخدم RNN وLSTM؟
8. اذكر أربعة تطبيقات في نظم الطاقة.

References | المراجع

المراجع العامة المستخدمة في هذا المقرر

[1] A. Gabash, *Flexible Optimal Operations of Energy Supply Networks with Renewable Energy Generation and Battery Storage*. Saarbrücken, Germany: Südwestdeutscher Verlag für Hochschulschriften, 2014.

[2] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th Global ed. Harlow, U.K.: Pearson Education Limited, 2022.

[3] A. Gabash, “Energy Market Transition and Climate Change: A Review of TSOs-DSOs C+++ Framework from 1800 to Present,” *Energies*, vol. 16, no. 17, Art. no. 6139, 2023, doi: 10.3390/en16176139.

[4] A. Gabash, *AI Applications in Electrical Power Systems*, Version 1.0, 2026. [Online]. Available: https://aoussgabash.com

Publication Information | معلومات النشر

Document	Lecture 07 · Deep Learning
Course	AI Applications in Electrical Power Systems
Author	Dr.-Ing. Aouss Gabash
Version	1.0
Publication year	2026
Official website	https://aoussgabash.com
Citation style	IEEE

Recommended citation:

Gabash, A. (2026). Deep Learning. AI Applications in Electrical Power Systems, Lecture 07, Version 1.0. https://aoussgabash.com

المرجع المقترح:

أوس غباش، 2026، محاضرة 07، تطبيقات الذكاء الاصطناعي في أنظمة الطاقة الكهربائية، الإصدار 1.0، https://aoussgabash.com

© 2026 Dr.-Ing. Aouss Gabash. Educational use with attribution to the author and official website.