Other Deep Learning Algorithms

OTHER DEEP LEARNING ALGORITHMS

أكاديمية طويق
Tuwaiq Academy

Why Explore More Architectures?

What are Neural Networks?

- Building blocks of AI designed to mimic the human brain.
- Consist of layers of interconnected nodes (neurons).

- Basic networks (MLPs, single-layer NNs) have limitations.
- Advanced architectures solve specific problems like **sequential data, image generation, and language understanding.**

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

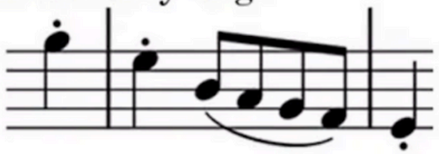
OTHER DEEP LEARNING ALGORITHMS

OTHER DEEP LEARNING ALGORITHMS

أكاديمية طويق
Tuwaiq Academy

Why Explore More Architectures?

Examples of sequence data

Speech recognition		"The quick brown fox jumped over the lazy dog."
Music generation	Ø	
Sentiment classification	"There is nothing to like in this movie."	★☆☆☆☆
DNA sequence analysis	AGCCCCTGTGAGGAACTAG	AGCCCCTGTGAGGAACTAG
Machine translation	Voulez-vous chanter avec moi?	Do you want to sing with me?
Video activity recognition		Running
Name entity recognition	Yesterday, Harry Potter met Hermione Granger.	Yesterday, Harry Potter met Hermione Granger.

Andrew Ng

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

OTHER DEEP LEARNING ALGORITHMS

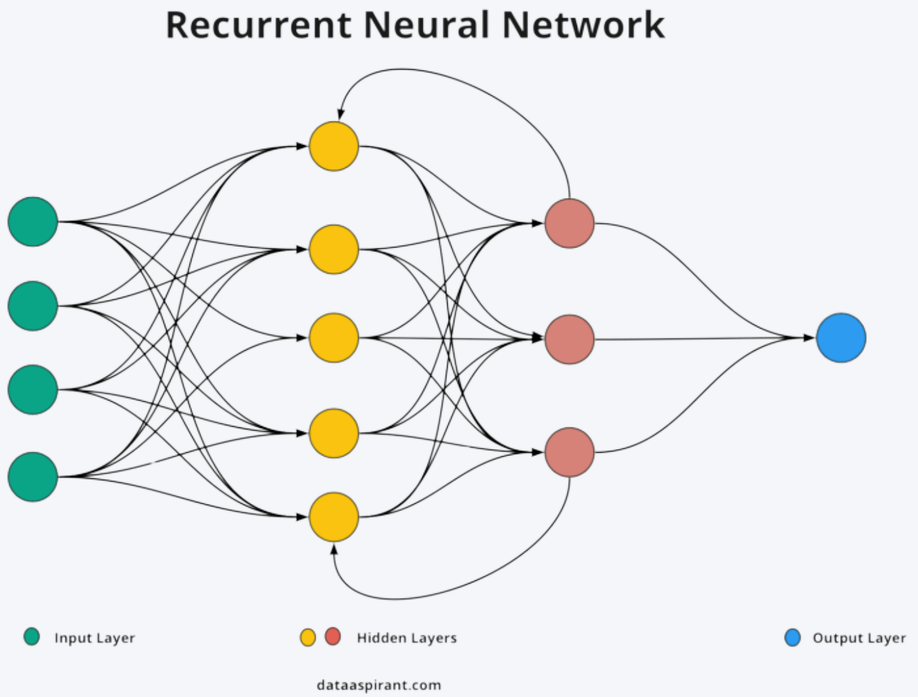
Recurrent Neural Networks (RNNs)

***Purpose**: Handling sequential data (e.g., time-series, language).*

- **How it Works:**
  - Neurons maintain a "memory" of previous inputs via loops.
  - Outputs depend on both current input and past computations.

- **Limitations:**
  - Struggles with long-term dependencies due to vanishing gradients.

- **Applications:**
  - Text generation, speech recognition, and financial modeling.



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

OTHER DEEP LEARNING ALGORITHMS

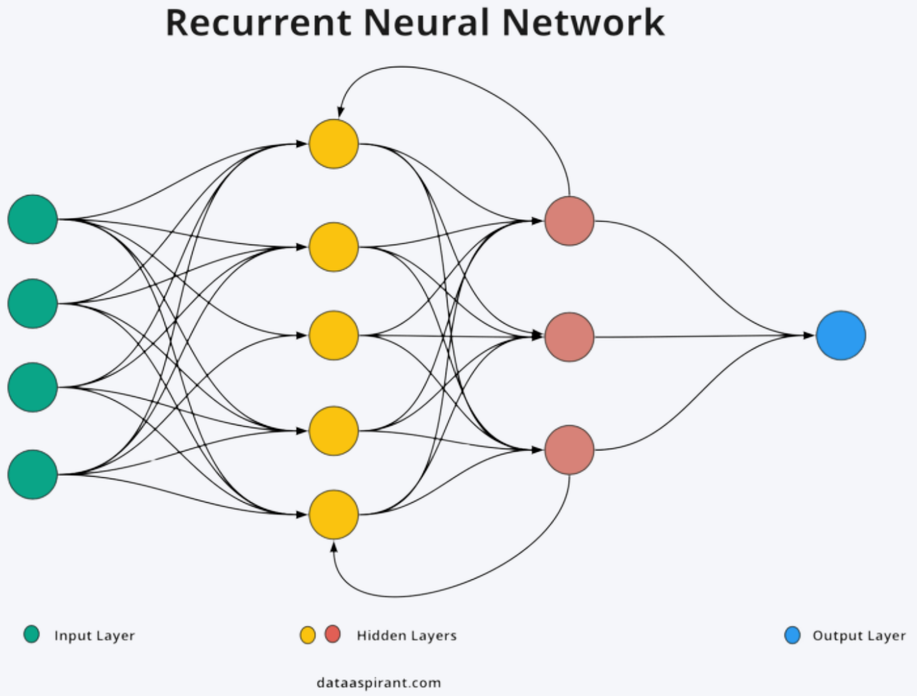
how a Recurrent Neural Network (RNN) works

- **Input Layer**
  - Data is fed into the network one step at a time (e.g., words in a sentence or time-series values).
  - Each input at a time step is processed sequentially.

- **Hidden Layers**
  - These layers process the input and maintain a "memory" by passing information from one time step to the next through loops (recurrent connections).
  - The hidden state captures both the current input and the context from previous time steps.

- **Output Layer**
  - The network produces an output at each time step (e.g., predicting the next word or value).
  - The final output (blue) represents the network's prediction after processing all time steps.

- **Recurrent Connections:**
  - The feedback loops in the hidden layers allow the RNN to "remember" past data, which makes it suitable for sequential data processing.

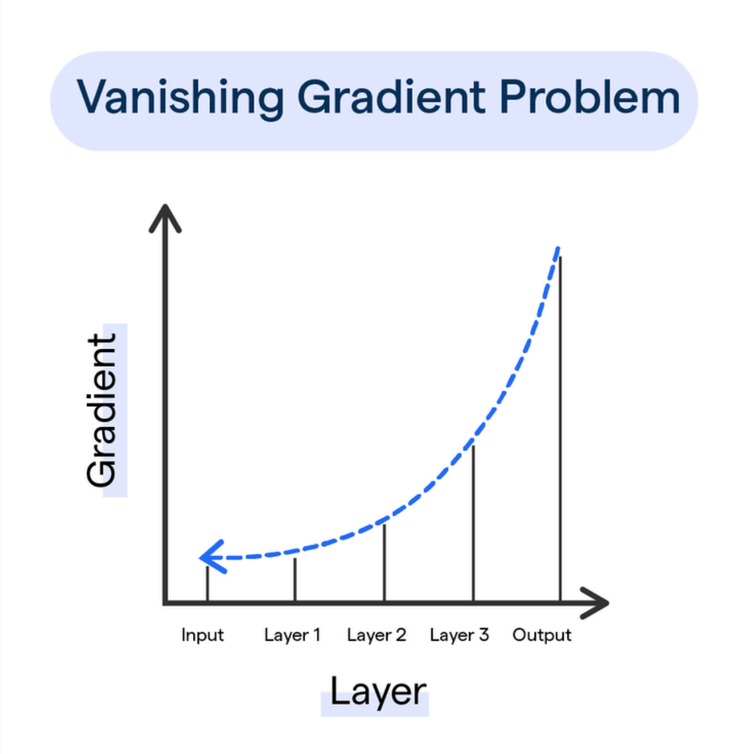


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

OTHER DEEP LEARNING ALGORITHMS

vanishing gradient problem

The **vanishing gradient** problem happens when **gradients** (used to update neural network weights) become very small during **backpropagation**. This makes it hard for the network to learn because the weights in earlier layers are barely updated. It’s common in deep networks like RNNs when processing long sequences, as information from earlier steps "fades away" over time.



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

OTHER DEEP LEARNING ALGORITHMS

Long Short-Term Memory (LSTM) Networks

Why LSTMs?

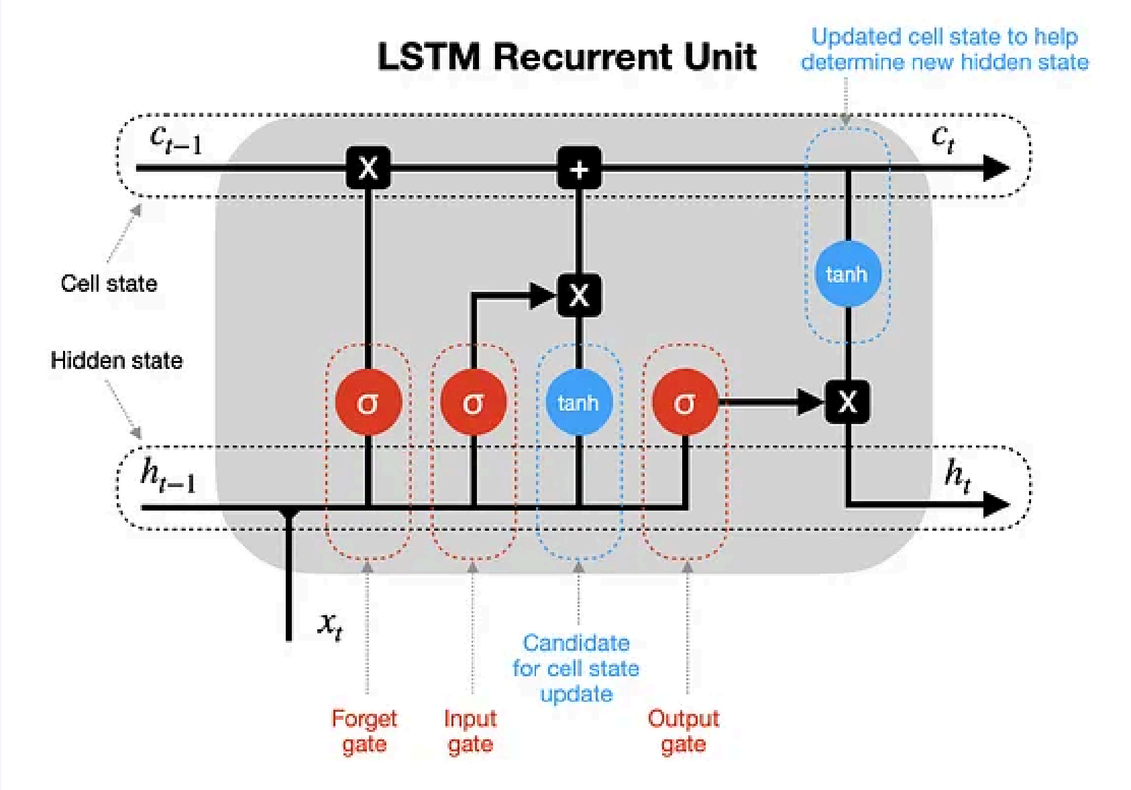
- Overcome RNN's vanishing gradient problem.
- Designed for learning long-term dependencies.

Key Components:

- **Forget Gate:** Decides what to discard from the memory.
- **Input Gate:** Updates the cell state with new information.
- **Output Gate:** Controls the final output based on memory.

Applications:

- Language modeling, video analysis, and predictive analytics.



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

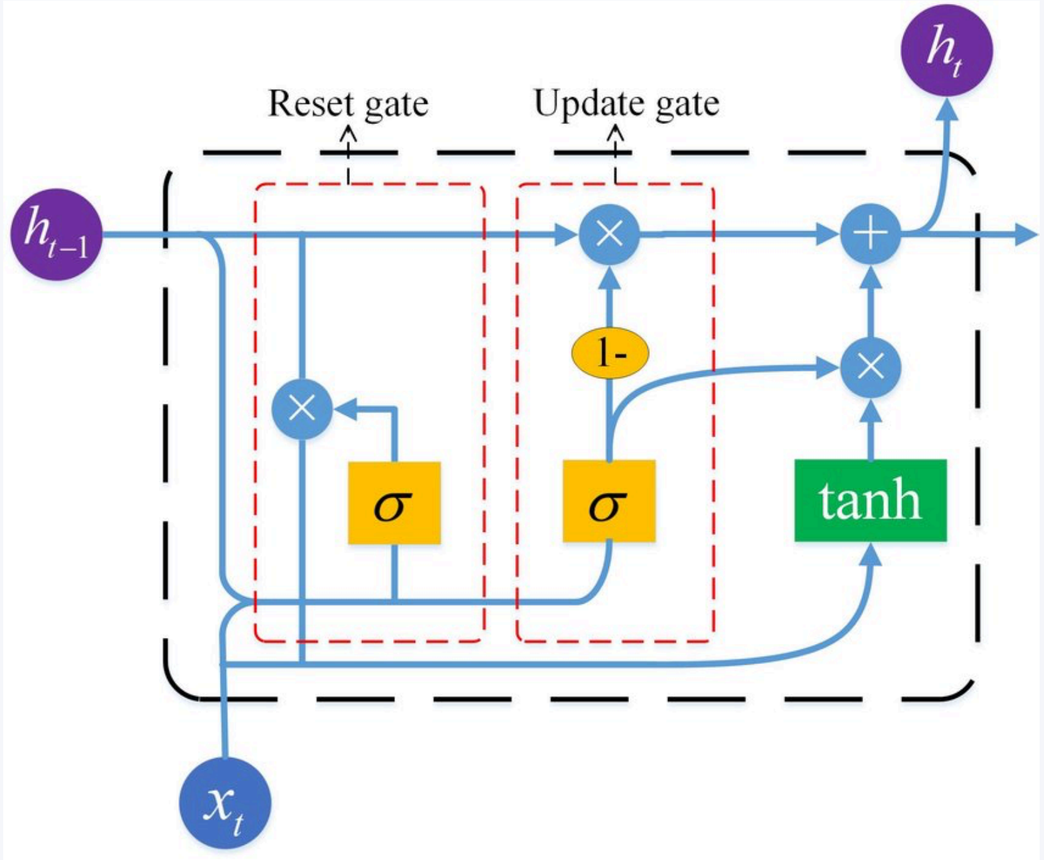
OTHER DEEP LEARNING ALGORITHMS

Gated Recurrent Units (GRUs)

- **Simpler Alternative to LSTMs:**
  - Combines forget and input gates into a single gate.
  - Fewer parameters make it computationally efficient.

- **How it Works:**
  - Update Gate: Determines how much past information to retain.
  - Reset Gate: Controls how much of the past to forget.

- **Applications:**
  - Real-time applications requiring quick computation.



OTHER DEEP LEARNING ALGORITHMS

Encoder-Decoder Architecture (CNN-based)

Converts input data into a compressed representation (Encoder) and reconstructs it into the desired output (Decoder).

- **Encoder**: Uses Convolutional Layers to extract features from input data (e.g., an image). Reduces data dimensions using pooling layers for a compact representation.

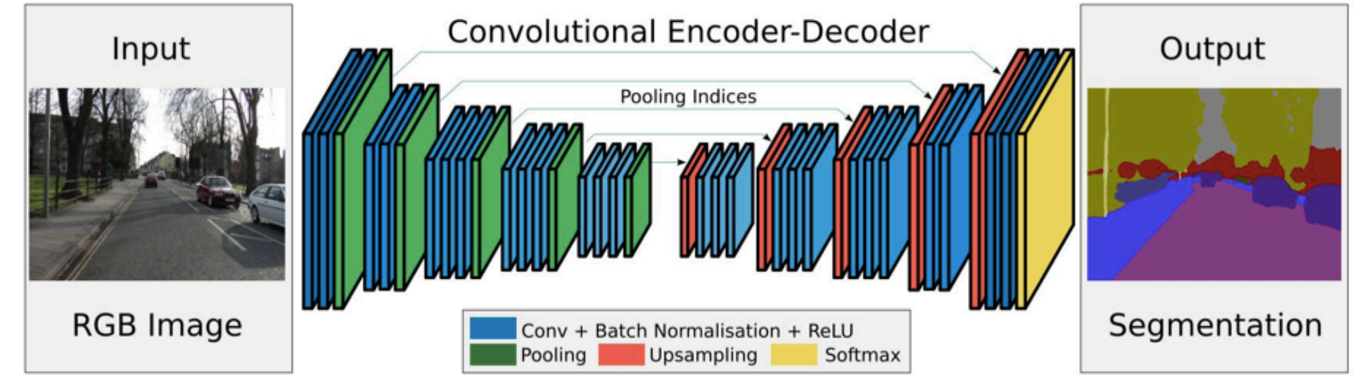
- **Decoder**: Uses Transposed Convolutions to upsample and reconstruct the original data. Rebuilds spatial details layer by layer to match the original input.

Key Features:

- **Encoder**: Captures key features in a smaller, efficient format.
- **Decoder**: Learns how to recreate data from the compressed format.
- **Shared Representation**: Bridges the input and output for seamless reconstruction.

Applications:

- Image-to-Image Translation (e.g., colorization).
- Denoising Autoencoders.
- Generating high-resolution images from compressed data.



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

DEEP LEARNING RESOURCES

أكاديمية طويق
Tuwaiq Academy

Resources

- Types of Neural Networks - GeeksforGeeks : GeeksforGeeks
- 8 Common Types of Neural Networks - Coursera : Coursera
- Types of Neural Networks and their Applications - Analytics Vidhya : Analytics Vidhya
- Arabic books for ML and DL : https://dlarabic.com

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

Trends in Deep Learning Training

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Agenda

- Understanding Overfitting and Underfitting
- Recognizing Overtraining in Deep Learning
- Techniques to Prevent Overfitting
- Early Stopping as a Regularization Method
- Hyperparameter Tuning for Better Models
- Summary and Best Practices

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Overfitting

Overfitting occurs when a model learns the training data too well, including its noise and irrelevant details, which hurts its ability to generalize to new data.

Signs of Overfitting:

- **High accuracy on training data but poor accuracy on validation/testing data.**

Example:

- A neural network memorizes specific training examples instead of learning general patterns.

Solution Approaches:

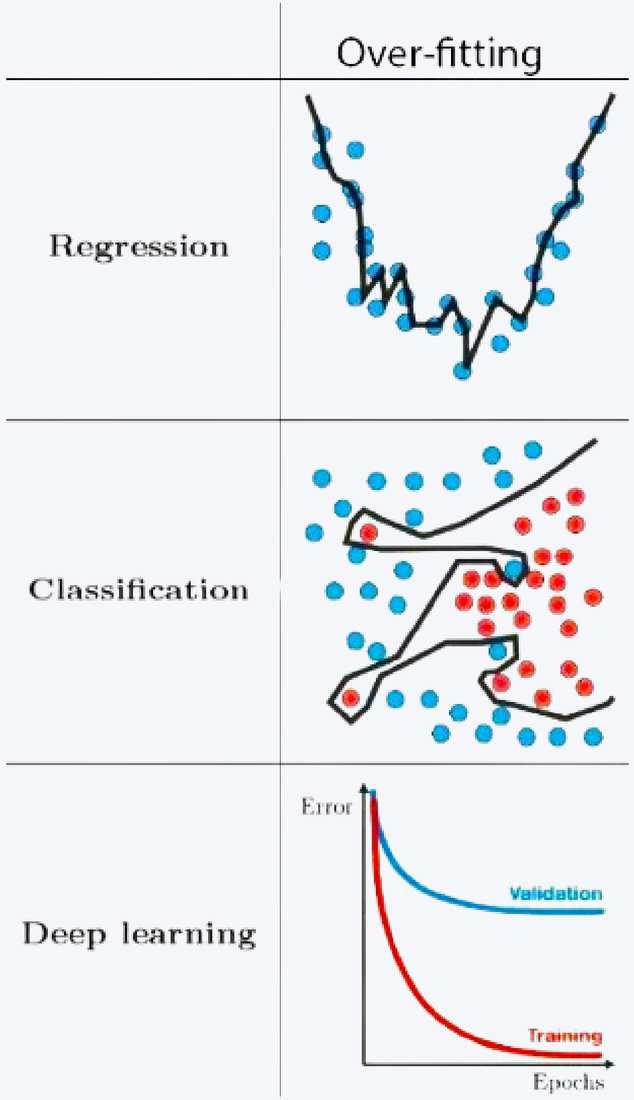
- Use more training data.
- Apply regularization (e.g., L2 or dropout).
- Simplify the model by reducing its complexity.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Overfitting



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Underfitting

Underfitting happens when a model is too simple to capture the underlying patterns in the data, resulting in poor performance on both training and testing data.

Signs of Underfitting:

- Low accuracy on both training and validation/testing data.

Example:

- A shallow neural network trying to classify complex images.

Solution Approaches:

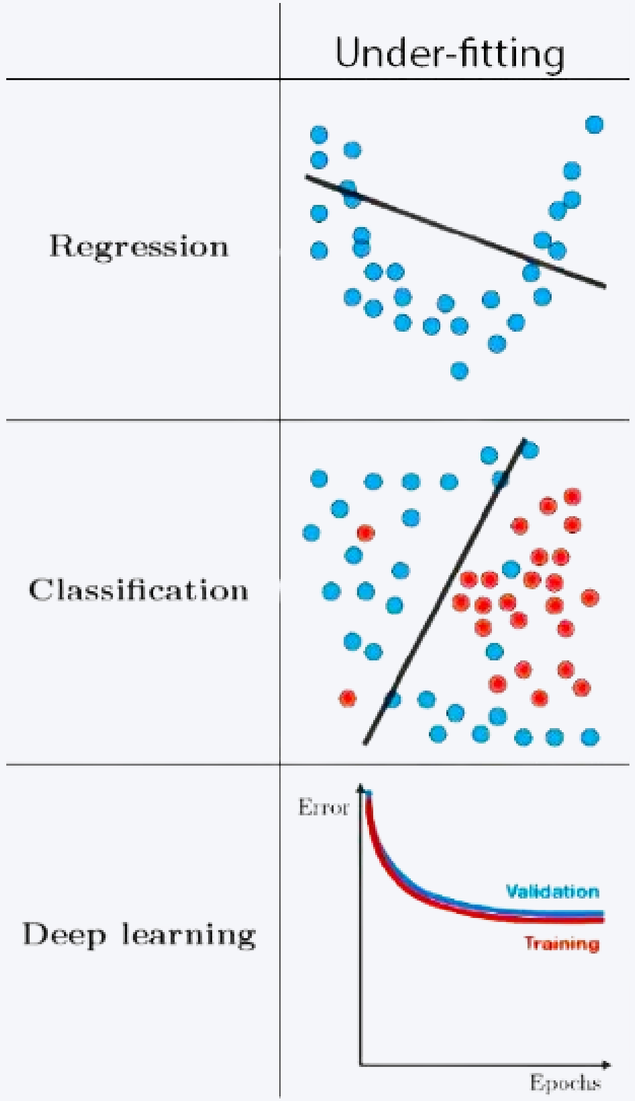
- **Use a more complex model.**
- **Train for more epochs.**
- **Reduce regularization.**

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Overfitting

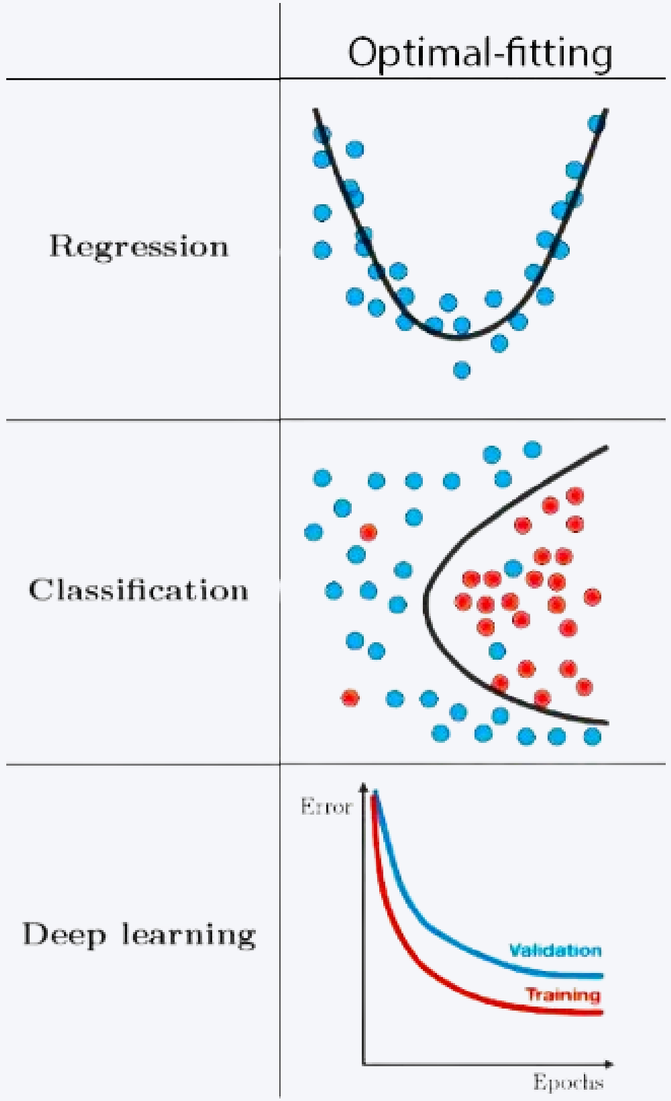


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Optimal Fitting

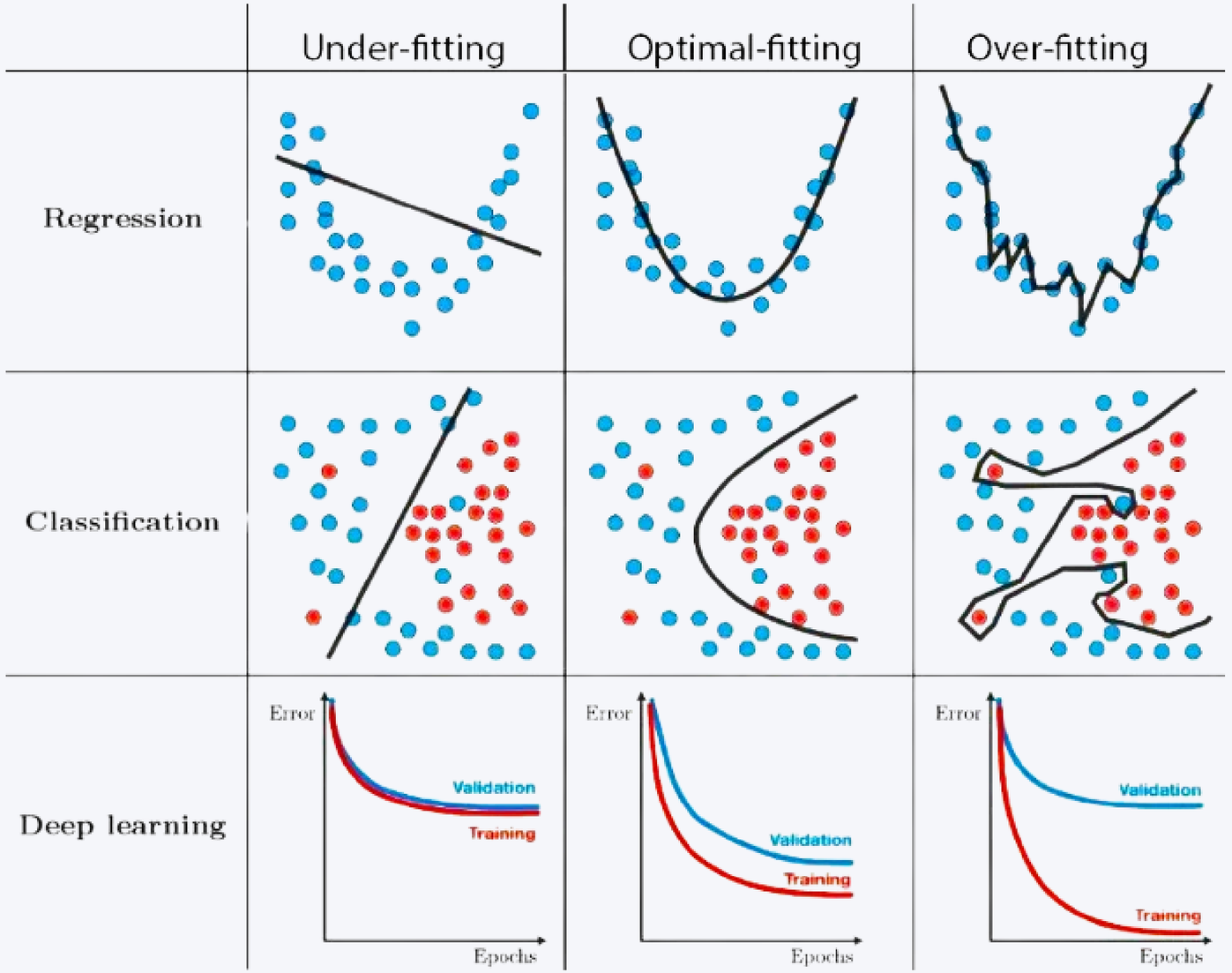


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Putting All Together



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Overtraining in Deep Learning

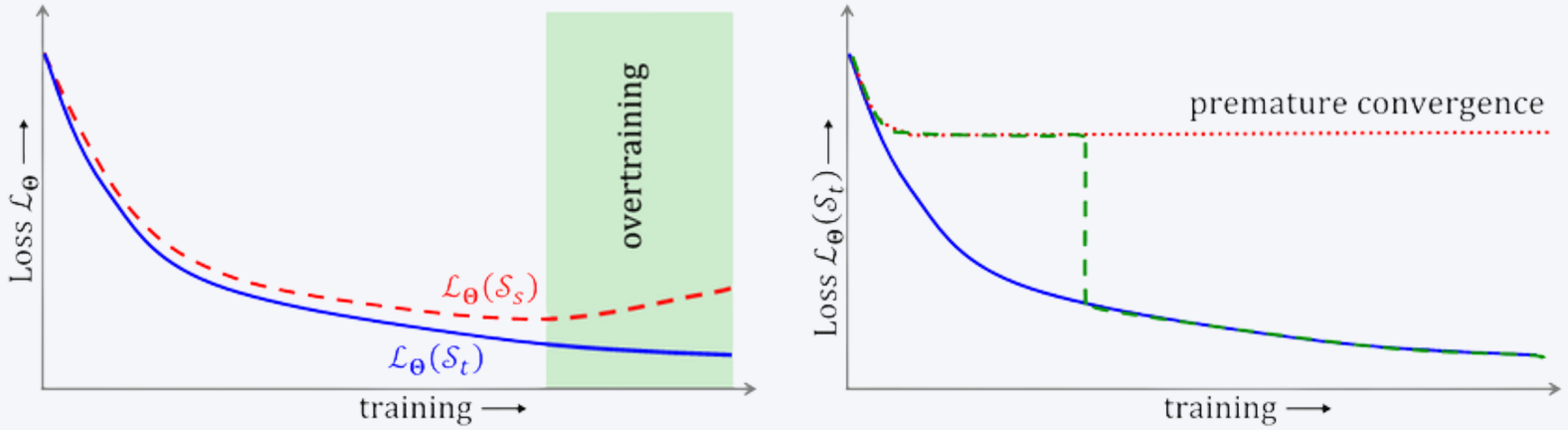
Overtraining is when a model is trained for too many epochs, leading to overfitting and decreased performance on unseen data.

How to Recognize Overtraining:

- Validation accuracy starts to drop after a certain number of epochs.

Prevention:

- Use validation data to monitor performance.
- Apply early stopping.



Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Early Stopping

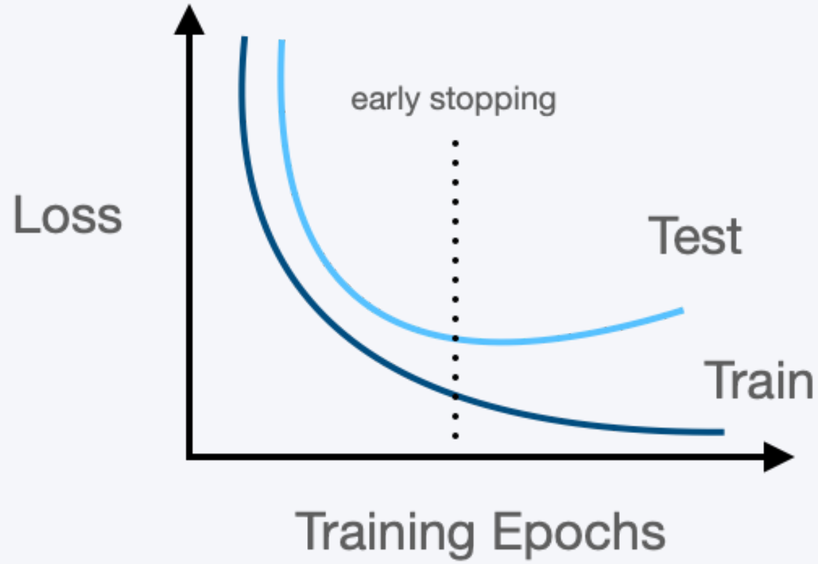
A technique where training is stopped when the model's performance on validation data no longer improves.

How It Works:

1. Track validation loss during training.
2. Stop training if loss does not decrease for a specified number of epochs (patience).

Advantages:

- Prevents overfitting.
- Saves training time.



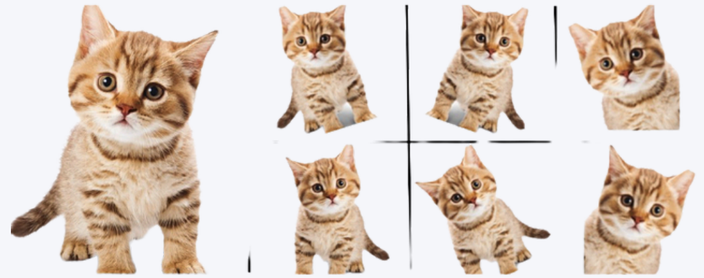
Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

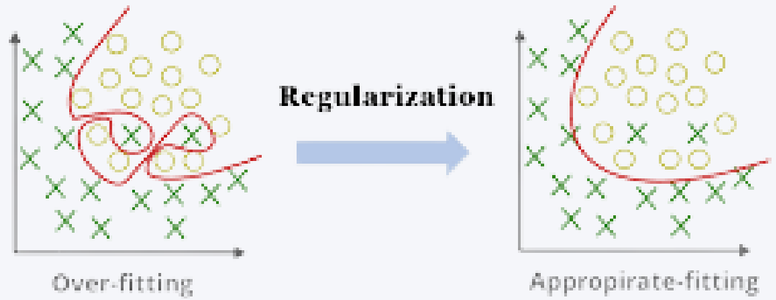
TRENDS IN DEEP LEARNING TRAINING

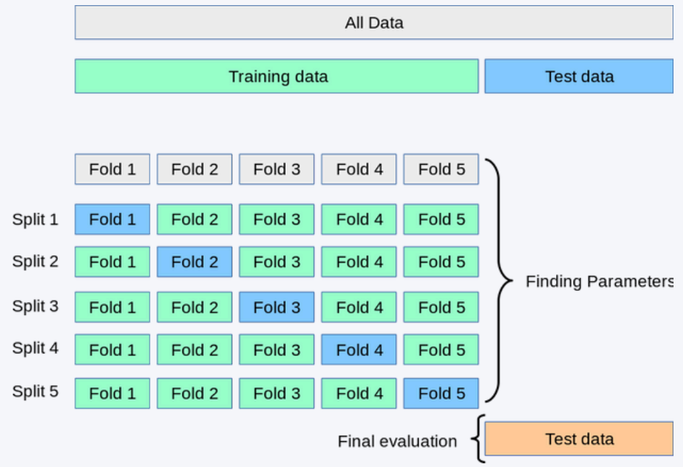
أكاديمية طويق
Tuwaiq Academy

Techniques to Prevent Overfitting

1. **Data Augmentation:**
   - Create new training examples by modifying existing ones (e.g., rotating or flipping images).
2. **Regularization:**
   - **L2 Regularization:** Penalizes large weights to simplify the model.
   - **Dropout**: Randomly drops neurons during training to prevent co-adaptation.
3. **Cross-Validation:**
   - Split data into multiple subsets for training and validation.
4. **Reduce Model Complexity:**
   - Use fewer layers or neurons if the model is too complex for the data.







Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

TRENDS IN DEEP LEARNING TRAINING

أكاديمية طويق
Tuwaiq Academy

Hyperparameter Tuning

Settings that control the learning process, such as learning rate, batch size, and number of layers.

How to Tune Hyperparameters:

- **Grid Search:** Test combinations of hyperparameters systematically.
- **Random Search:** Randomly sample hyperparameter values.
- **Automated Methods:** Use tools like Optuna or Hyperband.

Find the optimal balance between underfitting and overfitting.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine Learning Research 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

أكاديمية طويق
Tuwaiq Academy

Assignemnt