Example of Single Neuron with Activation Function

EXAMPLE OF SINGLE NEURON WITH ACTIVATION FUNCTION

أكاديمية طويق
Tuwaiq Academy

Problem Statement: Binary Classification for Insurance Purchase Prediction

Objective:
Given the age of a person, predict whether the person will buy insurance or not.

Input:

- A single feature: Age (numerical).
-

Output:

- A binary label:
  - 1: Person will buy insurance.
  - 0: Person will not buy insurance.

Age	Buys Insurance
22	0
25	0
47	1
52	0
46	1
56	1

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.

EXAMPLE OF SINGLE NEURON WITH ACTIVATION FUNCTION

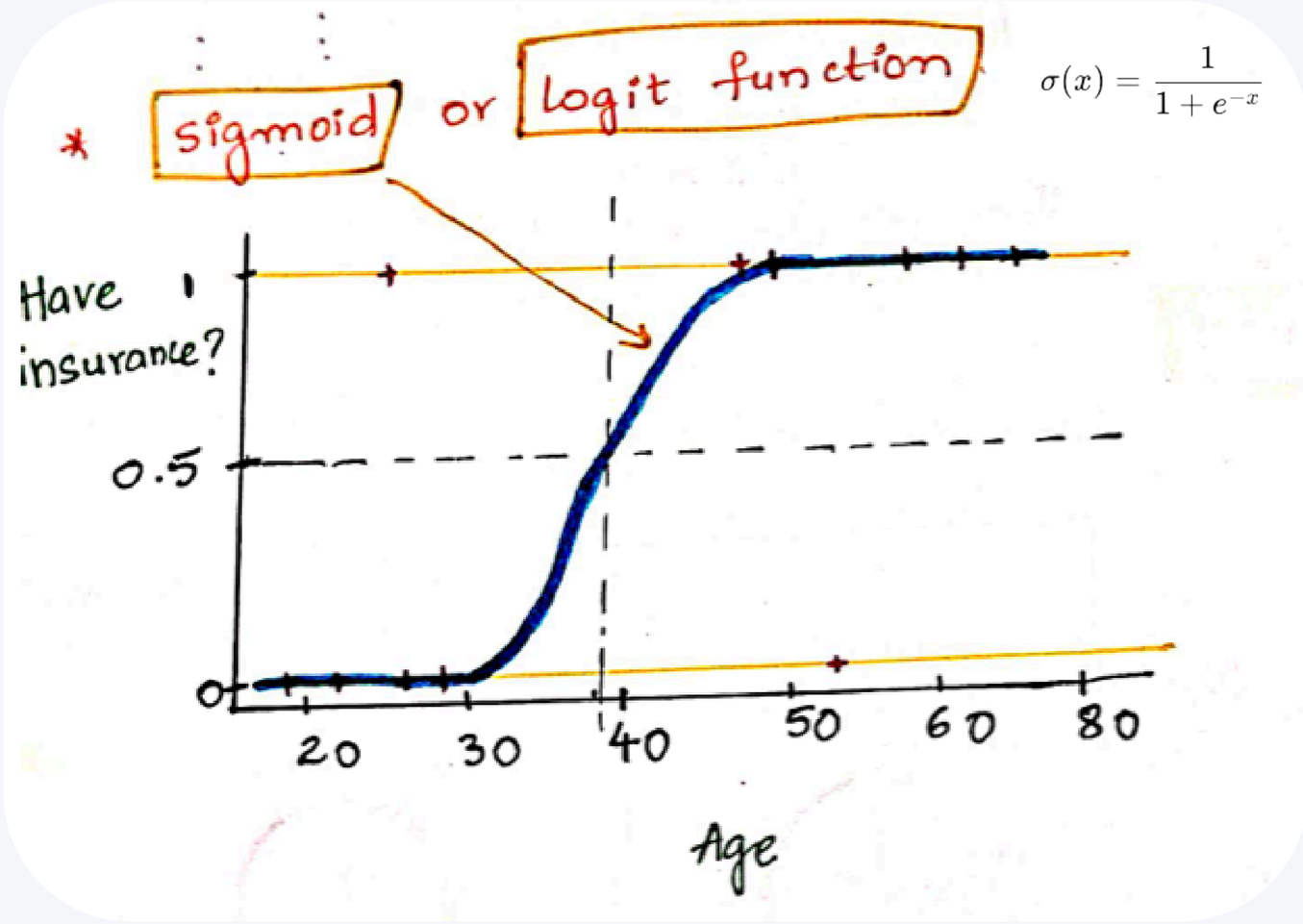
أكاديمية طويق
Tuwaiq Academy

Let’s consider a sigmoid activation function to solve the classification problem

Euler’s Number (e): e≈2.71828.

$$\frac{1}{1+e^{-200}} \approx \frac{1}{1+0} = 1$$

$$\frac{1}{1+e^{200}} \approx \frac{1}{1+\infty} = 0$$



The **Sigmoid** function converts any input into a range between 0 and 1, making it ideal for probability-based outputs

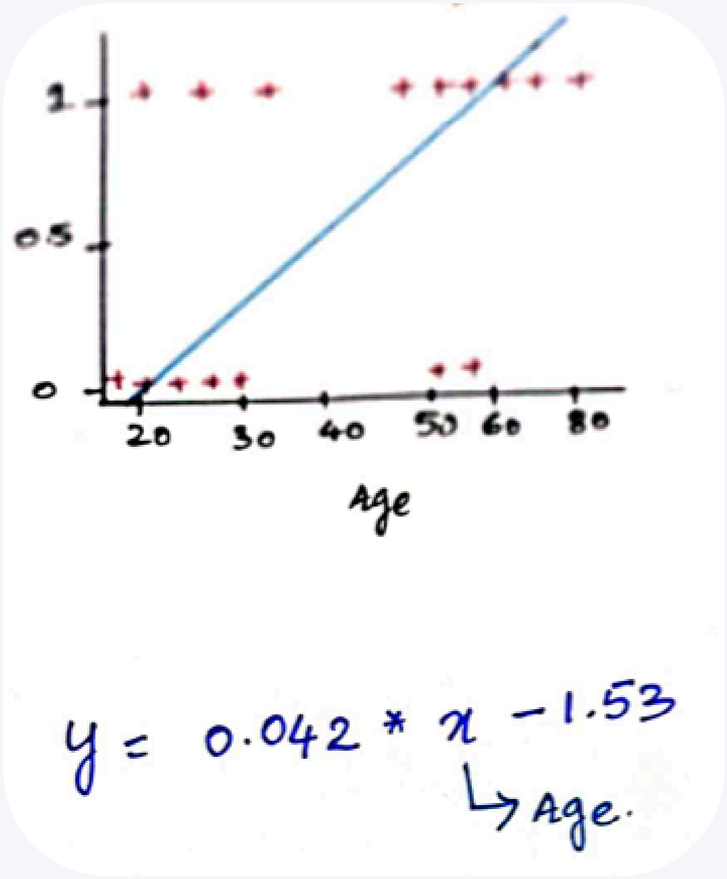
https://media.licdn.com/dms/document/media/v2/D4D1FAQER6f4_6s6s6Q/feedshare-document-pdf-analyzed/B4DZQ7goaBGgAY-/0/1736165331087?e=1737590400&v=beta&t=P-RrkuoV7HpgyEMIUWq3a4JfG2_VAzhE07uZZMukB-Q

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.

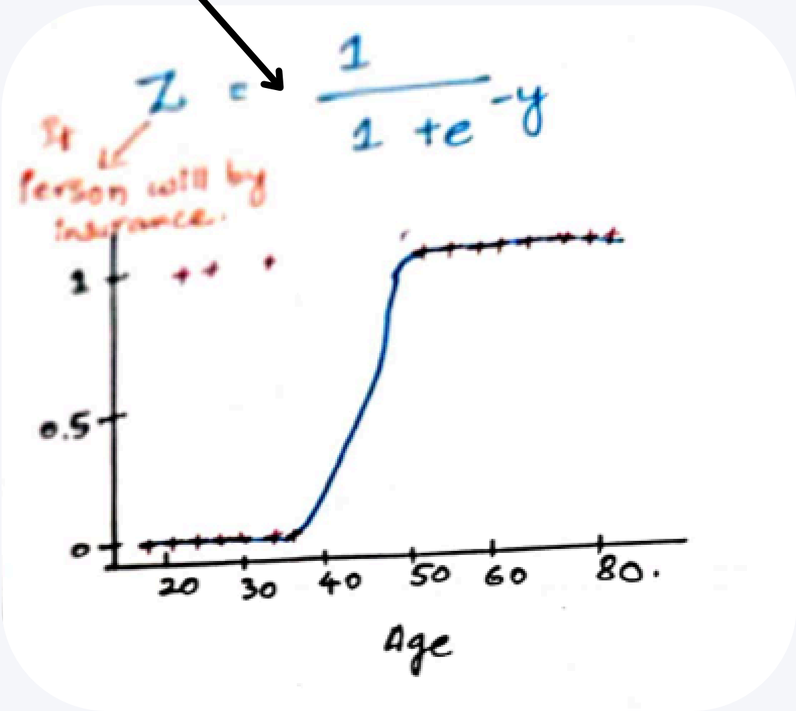
EXAMPLE OF SINGLE NEURON WITH ACTIVATION FUNCTION

Two Steps to solve the problem .. and this is the baisc of NN with signle neuron !!

STEP#1
y = wx+b | Transformation



STEP#2
Z = ... | Activation



https://media.licdn.com/dms/document/media/v2/D4D1FAQER6f4_6s6s6Q/feedshare-document-pdf-analyzed/B4DZQ7goaBGgAY-/0/1736165331087?e=1737590400&v=beta&t=P-RrkuoV7HpgyEMIUWq3a4JfG2_VAzhE07uZZMukB-Q

EXAMPLE OF SINGLE NEURON WITH ACTIVATION FUNCTION

أكاديمية طويق
Tuwaiq Academy

Two Steps to solve the problem .. and this is the baisc of NN with signle neuron !!

1. **Linear Equation Setup**:

$$y = 0.042 \cdot \text{age} - 1.53$$

2. **Sigmoid Activation**:
   The linear output (y) is passed through the sigmoid function:

$$z = \frac{1}{1 + e^{-y}}$$

Threshold Decision:

- $z < 0.5$: Person will **not** buy insurance.
- $z \geq 0.5$: Person **will** buy insurance.

- **For Age = 35**:

$$y = 0.042 \cdot 35 - 1.53 = -0.06$$

$$z = \frac{1}{1 + e^{0.06}} \approx 0.48$$

Decision: **Will not buy insurance**.

- **For Age = 43**:

$$y = 0.042 \cdot 43 - 1.53 = 0.276$$

$$z = \frac{1}{1 + e^{-0.276}} \approx 0.568$$

Decision: **Will buy insurance**.

Generalized Formula for Multiple Features:

For multiple features such as **age**, **income**, and **education**, the linear equation becomes:

$$y = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + b$$

Where:

- w_1, w_2, w_3: Weights for each feature.
- x_1, x_2, x_3: Feature values.
- b: Bias term.

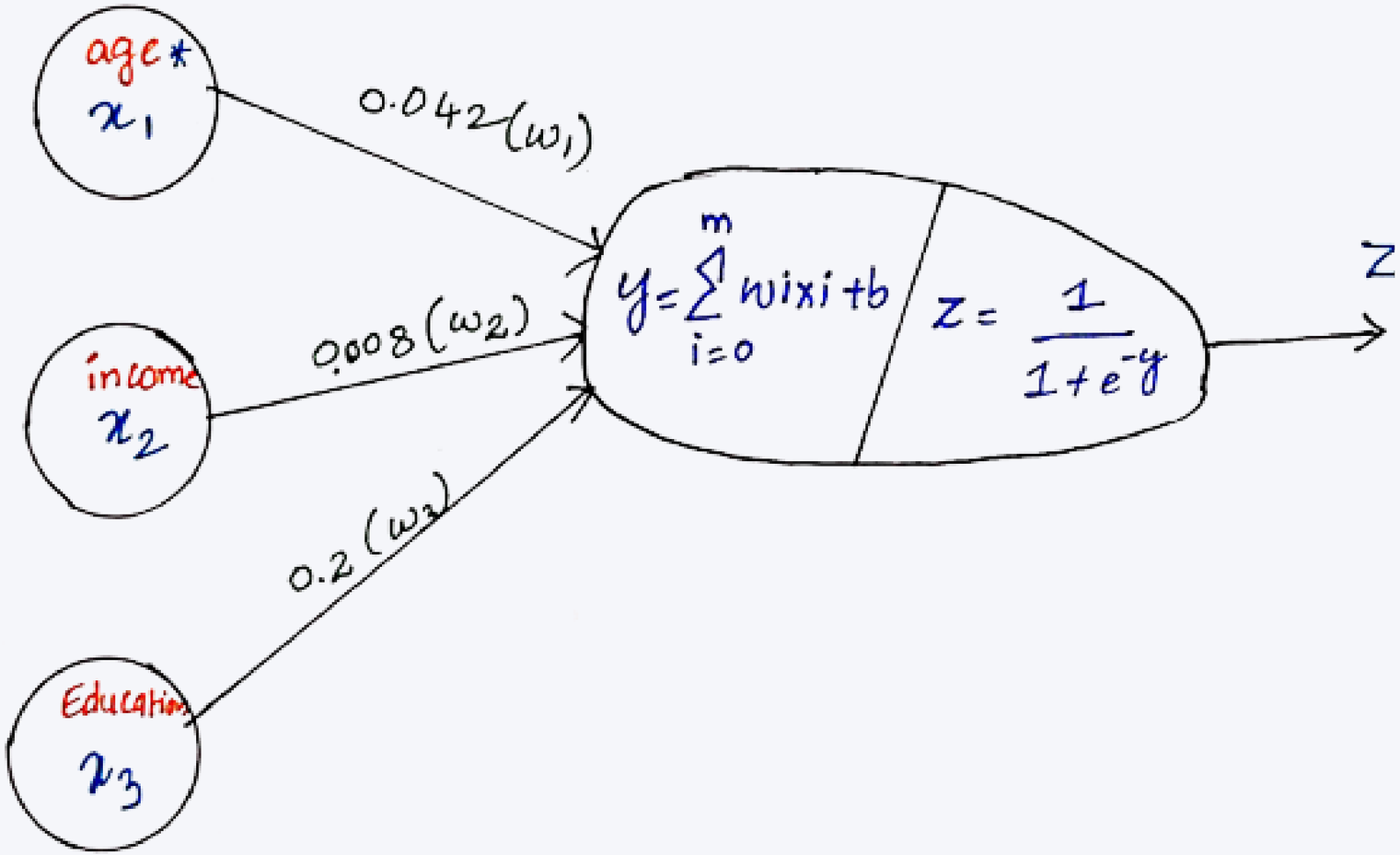
Compact form:

$$y = \sum_{i=1}^{m} (w_i \cdot x_i) + b$$

https://media.licdn.com/dms/document/media/v2/D4D1FAQER6f4_6s6s6Q/feedshare-document-pdf-analyzed/B4DZQ7goaBGgAY-/0/1736165331087?e=1737590400&v=beta&t=P-RrkuoV7HpgyEMIUWq3a4JfG2_VAzhE07uZZMukB-Q

EXAMPLE OF SINGLE NEURON WITH ACTIVATION FUNCTION

Two Steps to solve the problem .. and this is the baisc of NN with signle neuron !!



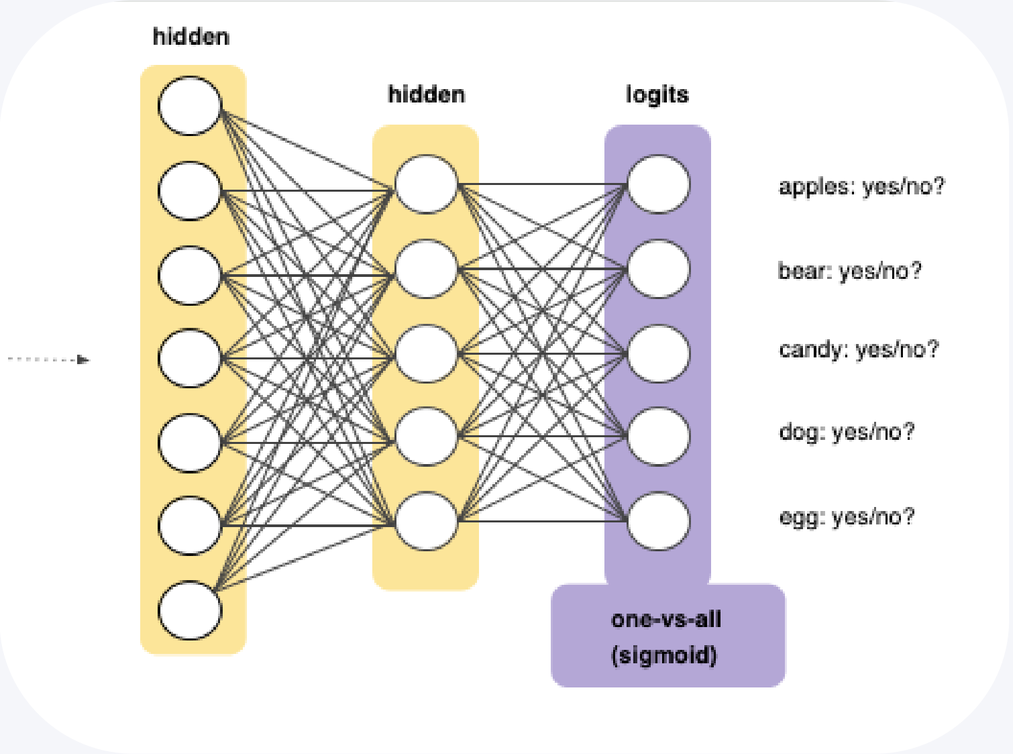
Layer Networks for Multi-Class Classification

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

What About Multi-Class Classification ?

- **We have only considered regression and binary classification problems so far.**
- **How can we get a neural network to perform multiclass classification?**



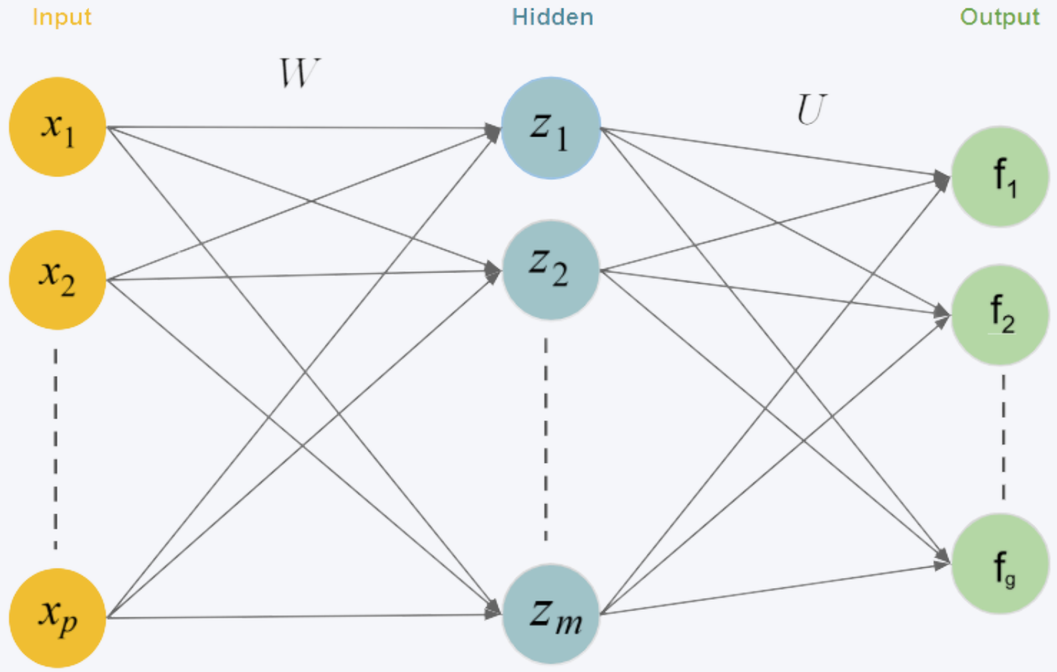
https://www.google.com/url?sa=i&url=https%3A%2F%2Faman.ai%2Fprimers%2Fai%2Fmulticlass-vs-multilabel-classification%2F&psig=AOvVaw0J39FfDxrr6WGkBu7bzvo3&ust=1736691083883000&source=images&cd=vfe&opi=89978449&ved=0CBcQjhxqFwoTCMDA7czs7YoDFQAAAAAdAAAAABAT

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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

How does it work ?

- The first step is to add additional neurons to the output layer.
    - Each neuron in the layer will represent a specific class
    - Number of neurons in the output layer **=** number of classes.



Input features:x=(x1,x2,...,xn)

Hidden neurons: Derived features from the hidden layer are represented as z=(z1,z2,...,zm).

Weights:

- Wj: Weights for connections between input and hidden layer.
- Uk: Weights for connections between hidden layer and output layer.

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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

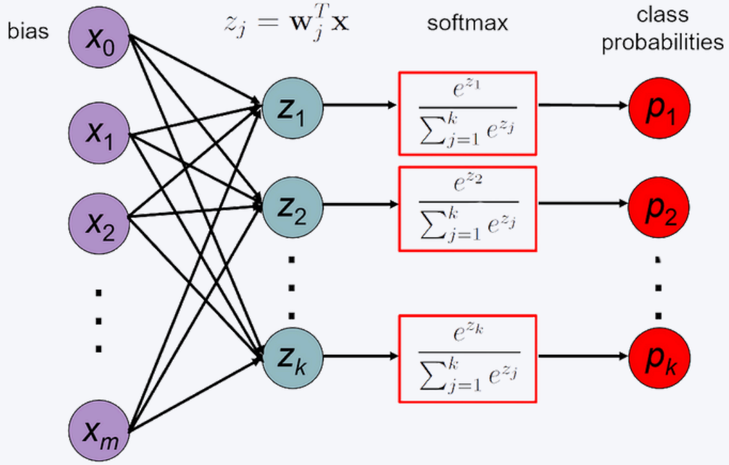
أكاديمية طويق
Tuwaiq Academy

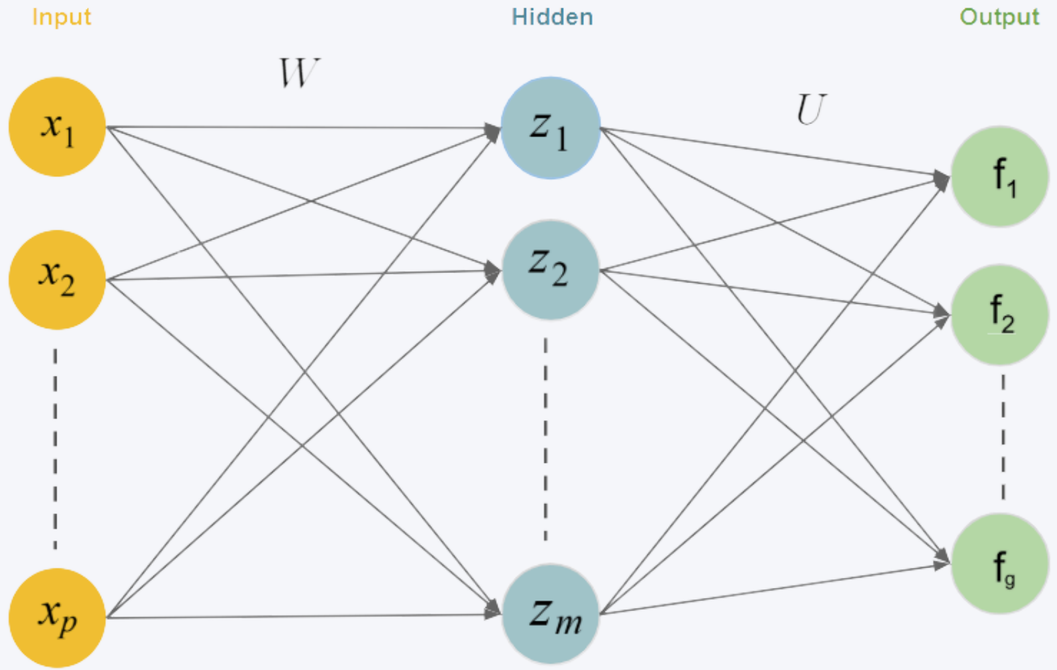
How deos it work ?

The second step is to apply a **softmax activation function** to the output layer.

Softmax provides a **probability distribution** over multiple classes, enabling multi-class predictions.

- Converts the model's output into probabilities for each class.
- Ensures all probabilities sum to 1.



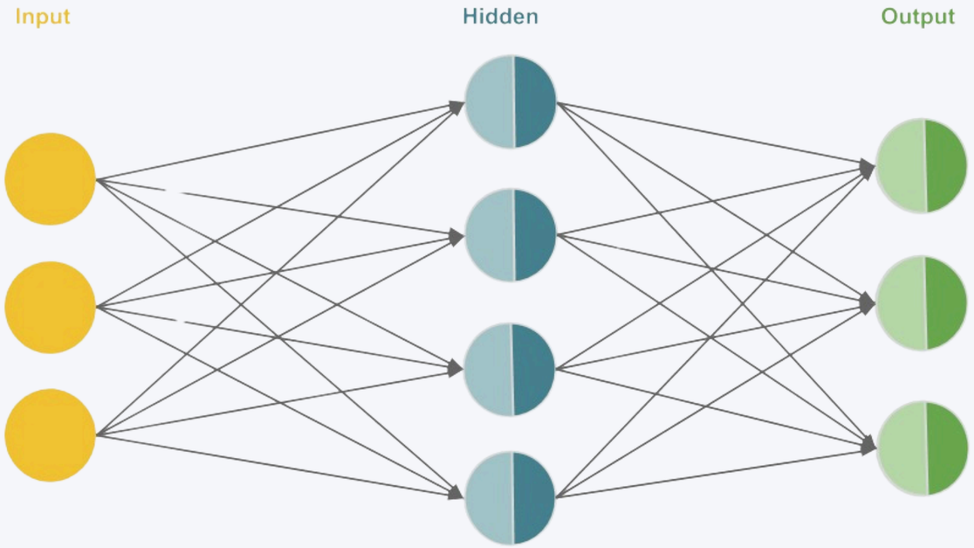


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine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

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).



$$\begin{pmatrix} 3 & -9 & 2 \\ 11 & -2 & 7 \\ -6 & 3 & -4 \\ 6 & -1 & 5 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ -1 \\ 1 \end{pmatrix}$$

W^T b

$$\begin{pmatrix} 3 & -12 & 8 & 1 \\ 2 & -3 & 9 & 1 \\ -5 & 1 & -1 & 7 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ -8 \end{pmatrix}$$

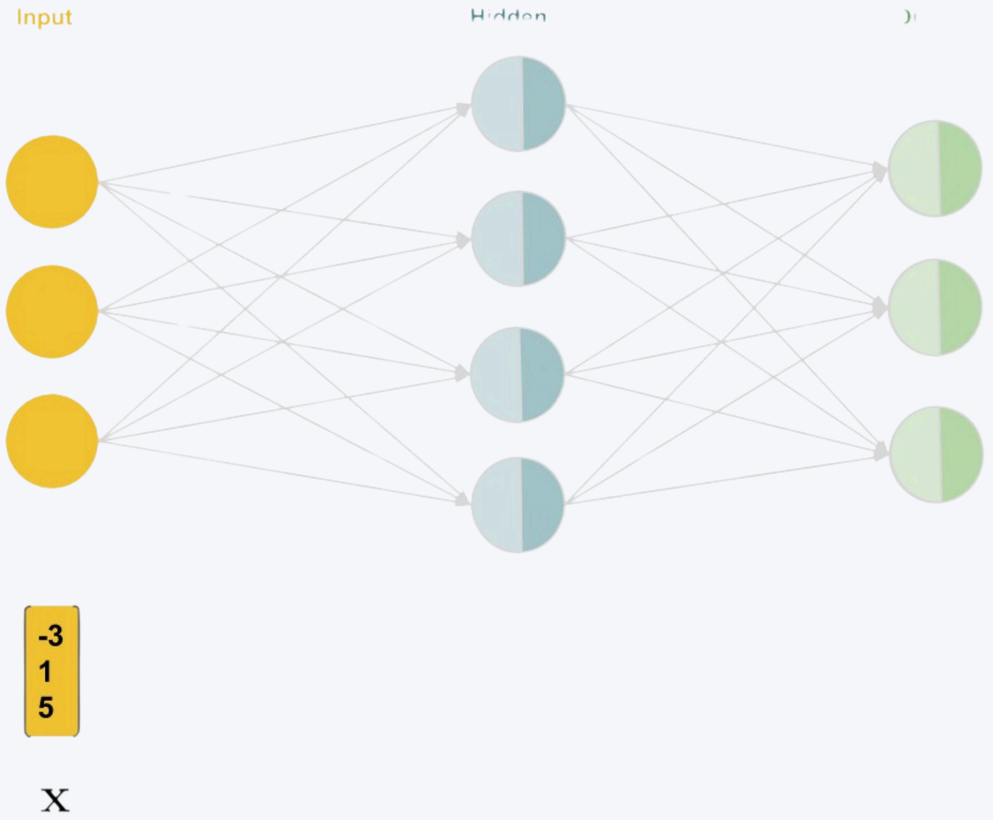
U^T c

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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).

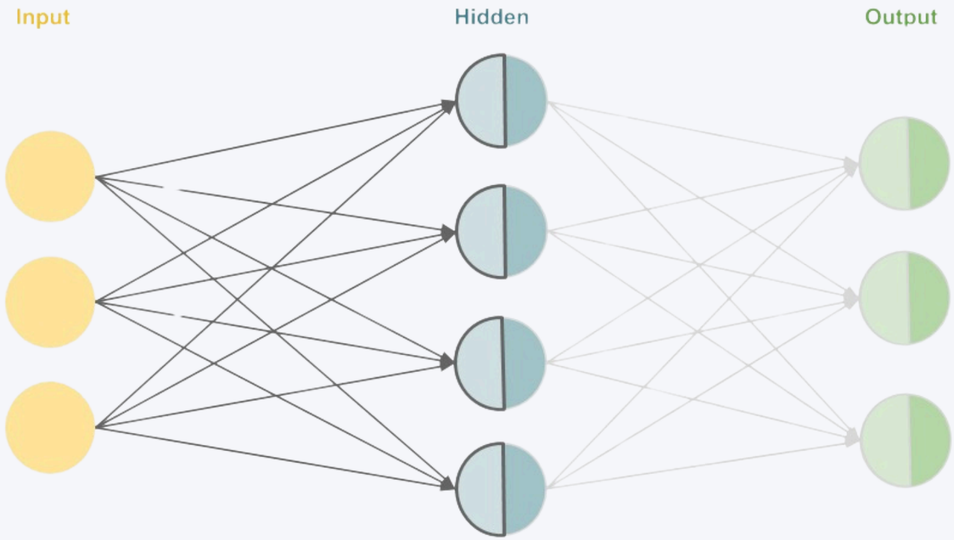


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine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

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).

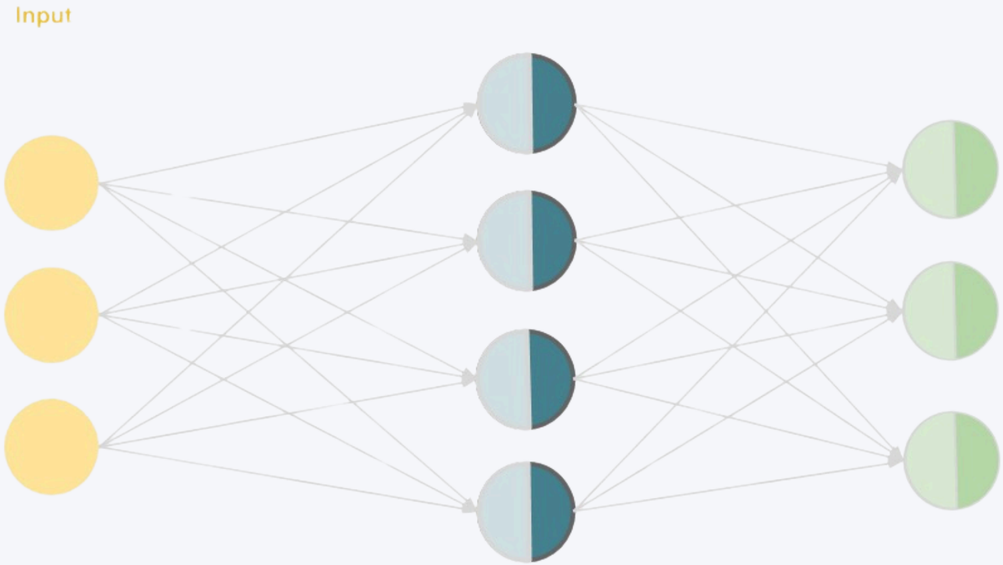


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine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

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).

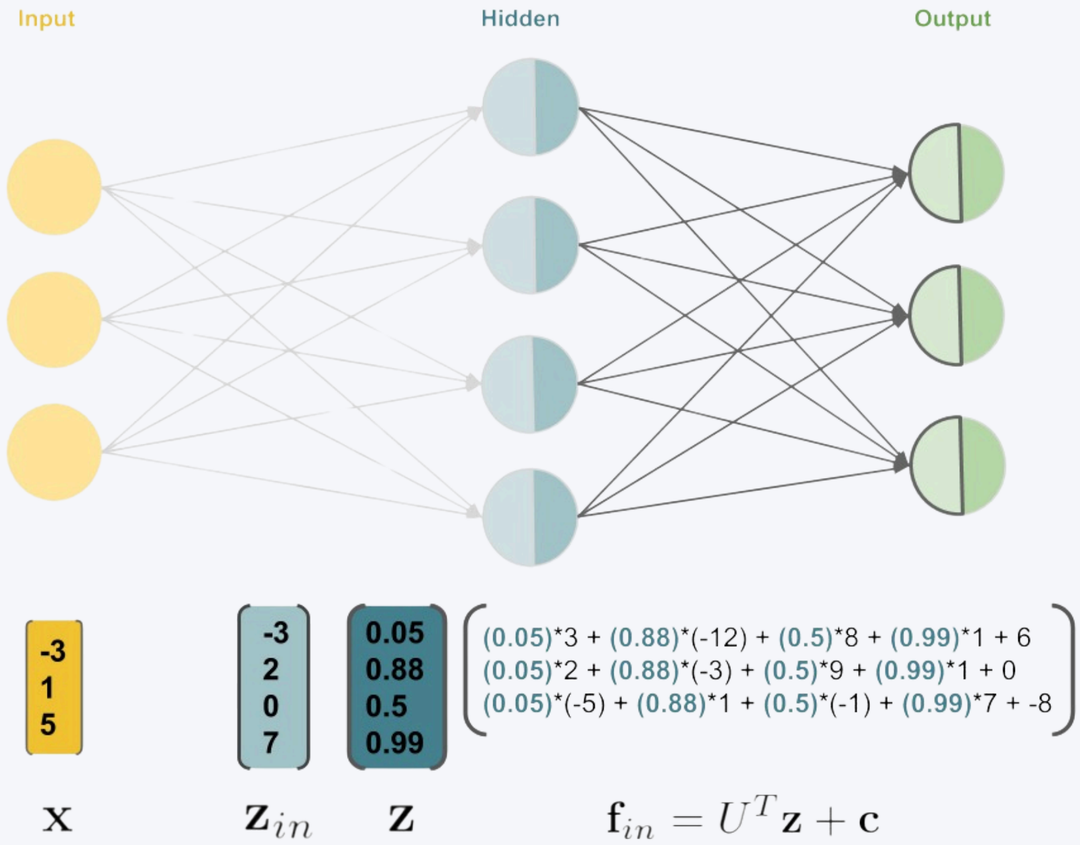


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine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

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

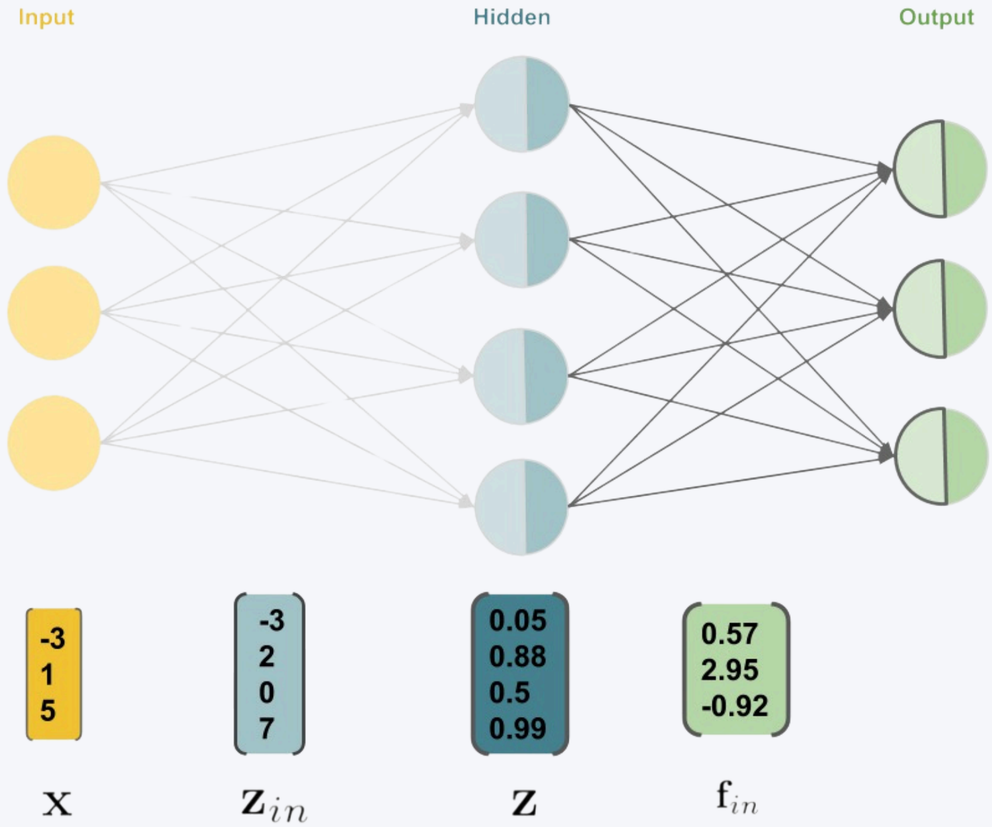
Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).



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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

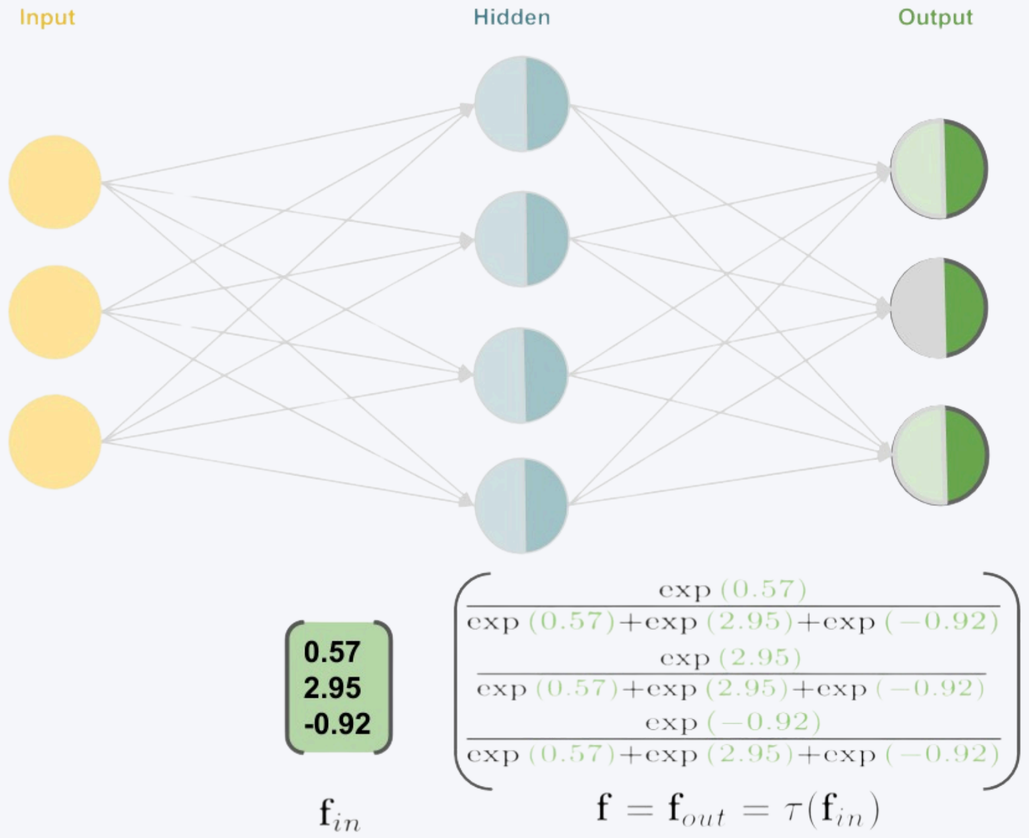
Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).



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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).

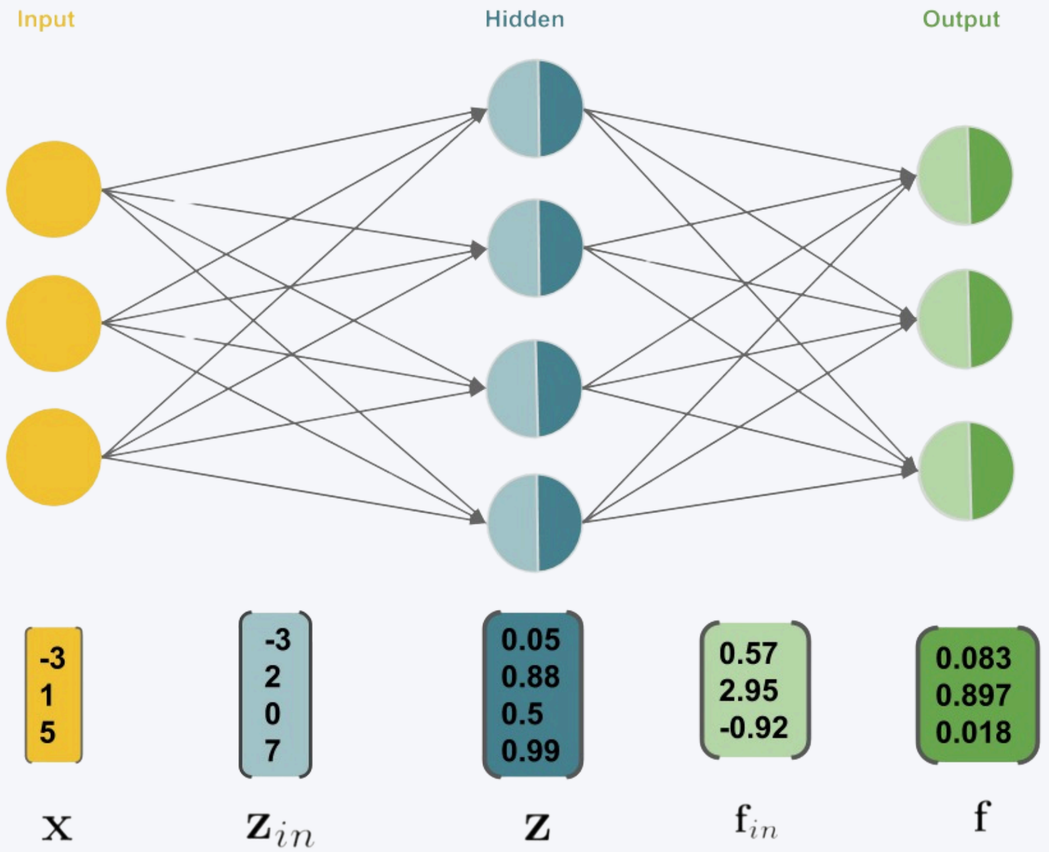


Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop, in Proceedings of Machine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

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Example on MULTI-CLASS CLASSIFICATION
Forward pass (Hidden: Sigmoid, Output: Softmax).



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.

LAYER NETWORKS FOR MULTI-CLASS CLASSIFICATION

أكاديمية طويق
Tuwaiq Academy

Cross-Entropy Loss

Revising: Squared Error:

- Measures the difference between predicted and true values by squaring the error.
- Works well for regression or simple problems but not ideal for classification.
-

Introducing: Cross-Entropy Loss:

- Used in multi-class classification.
- Measures how far the predicted probability is from the correct class.
- Example: If the true class is 3, it compares the predicted probability for class 3 to 1 (perfect match).

$$L = -\sum (\text{True Label}) \times \log(\text{Predicted Probability})$$

Why Cross-Entropy?

- It penalizes wrong predictions more effectively by focusing on the probabilities assigned to the correct class.

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

Let’s Code

Implementing Multi-Class Classification Problem with Keras