# Week 10 — Neural Networks & Deep Learning

Feedforward, RNN, and LSTM for Time Series

Dr Juan Zurita

Economic & Business Forecasting
The University of Edinburgh · School of Economics

# From Linear to Neural

# The Feedforward Neural Network

A single hidden layer:

$$\hat{y} = W_2 \cdot \sigma(W_1\mathbf{x} + \mathbf{b}_1) + b_2$$

- $\sigma$: activation function (ReLU, sigmoid, tanh)
- **Universal approximation theorem**: one hidden layer can approximate any continuous function
- Trained by backpropagation + gradient descent

# Activation Functions

**ReLU**

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

Fast, sparse; default choice.

**Sigmoid**

$\sigma(z) = \frac{1}{1+e^{-z}}$

Output in $(0, 1)$; vanishing gradients.

**Tanh**

$\sigma(z) = \tanh(z)$

Output in $(-1, 1)$; zero-centred.

# Recurrent Neural Networks

# RNNs for Sequences

Standard feedforward networks treat inputs as *unordered*.
RNNs process sequences by maintaining a hidden state:

$$\mathbf{h}_t = \sigma(\mathbf{W}_h \mathbf{h}_{t-1} + \mathbf{W}_x \mathbf{x}_t + \mathbf{b})$$

- Natural fit for time series: respects temporal order
- Problem: vanishing/exploding gradients for long sequences

# LSTM — Long Short-Term Memory

Solves the vanishing gradient problem with gating mechanisms:

- **Forget gate**: what to discard from cell state
- **Input gate**: what new information to store
- **Output gate**: what to output from cell state

# Practical Considerations

# When to Use Neural Networks?

| **Method** | **Strengths** | **Weaknesses** |
|---|---|---|
| ARMA | Interpretable, theory | Linear |
| Exponential Smoothing | Trend + season | Univariate |
| VAR | Multivariate dynamics | Many parameters |
| Ridge/Lasso | Many predictors | Still linear |
| Random Forest | Nonlinear | No uncertainty |
| Neural Networks | Universal approx. | Data-hungry, opaque |

**Rule of thumb:** start simple (ARMA, ETS). Add complexity only if justified by the data.

# Course Summary

1. Time series **components and stationarity**
2. **Probability** and estimation
3. **ARMA**: the linear workhorse
4. **Forecasting**: estimation, selection, evaluation
5. **Exponential smoothing**: trend and season
6. **VAR**: multivariate dynamics
7. **State space**: unobserved components
8. **Dynamic factors**: large panels
9. **Machine learning**: regularisation, trees
10. **Neural networks**: nonlinear frontiers