

# Explainable AI and the brain

Characterizing heterogeneity in diverse clinical cohorts

---

Esten H. Leonardsen

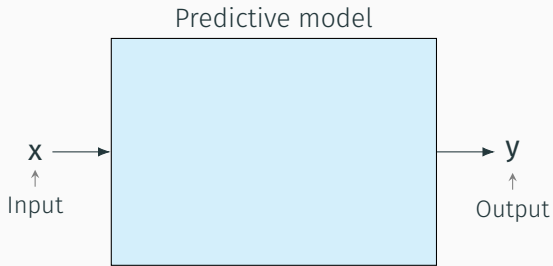
24.10.24



UNIVERSITETET  
I OSLO

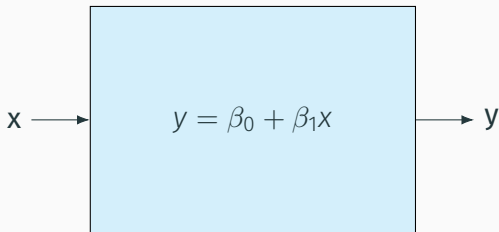
1. Theoretical background
  - a. Artificial neural networks
  - b. Explainable AI
2. Explainable AI and dementia
3. Explainable AI and brain age

# Artificial neural networks



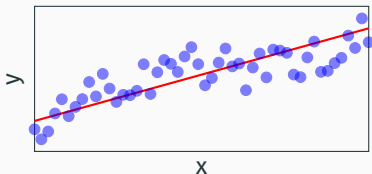
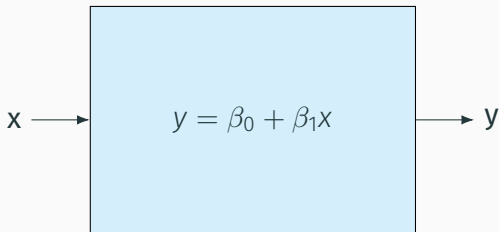


Linear regression model

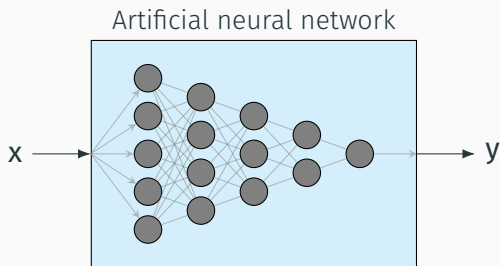




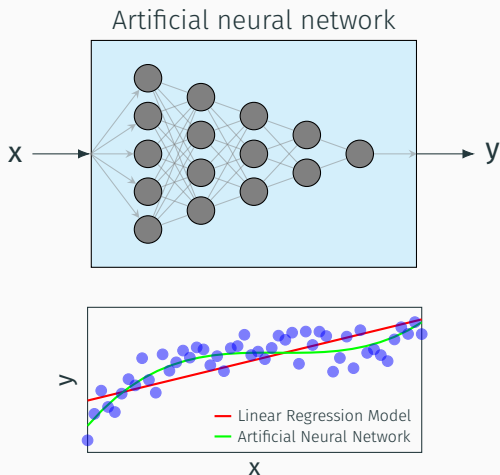
Linear regression model



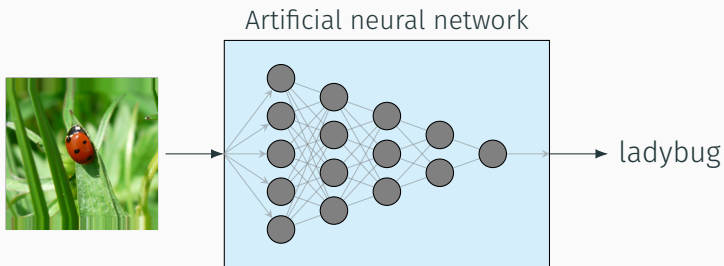
# Artificial neural networks



# Artificial neural networks

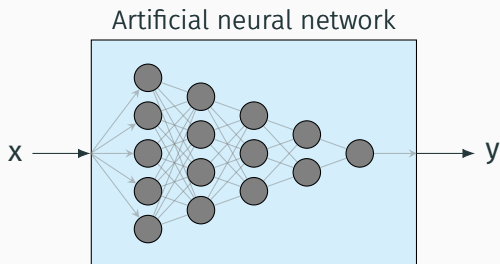


# Artificial neural networks

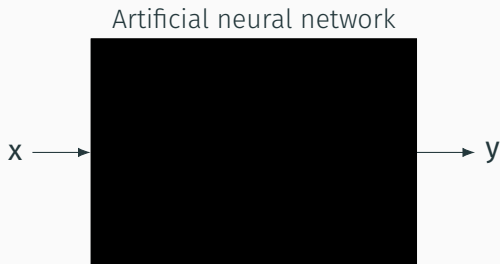




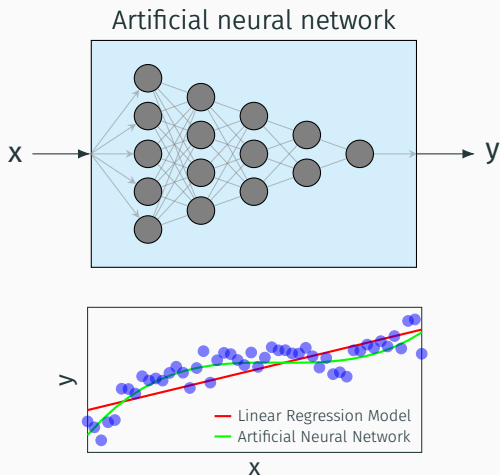
# Artificial neural networks: Interpretability



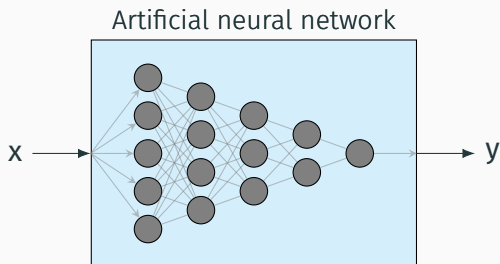
# Artificial neural networks: Interpretability



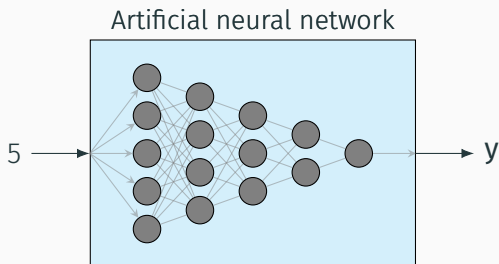
# Artificial neural networks: Interpretability



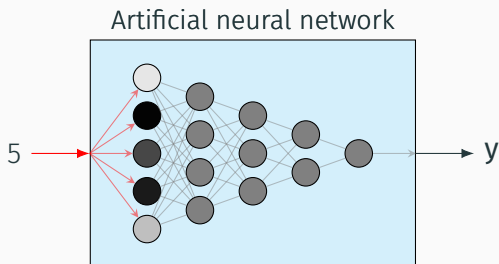
# Artificial neural networks: Interpretability



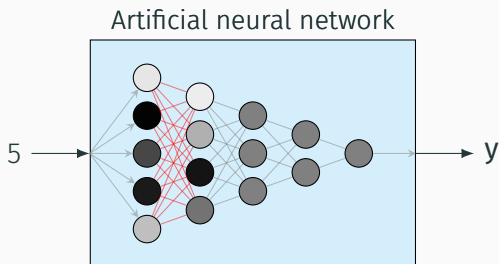
# Artificial neural networks: Interpretability



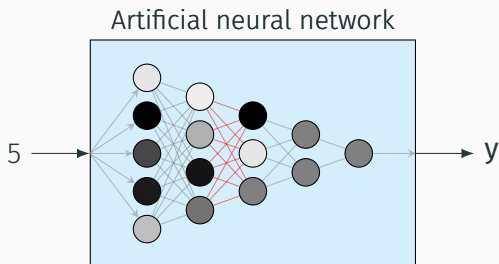
# Artificial neural networks: Interpretability



# Artificial neural networks: Interpretability

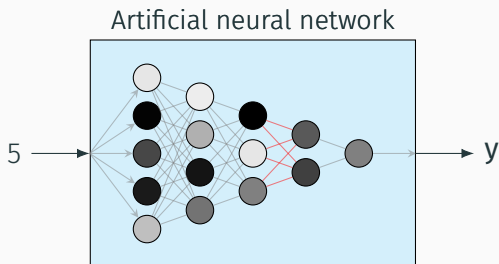


# Artificial neural networks: Interpretability

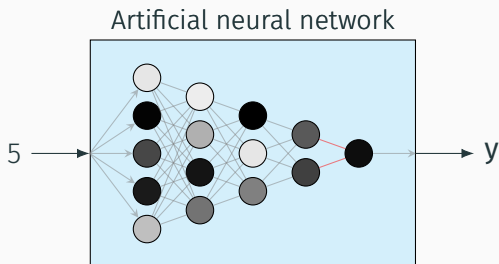




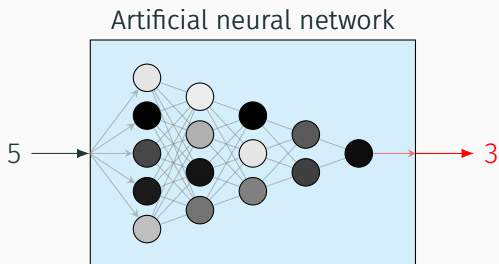
# Artificial neural networks: Interpretability



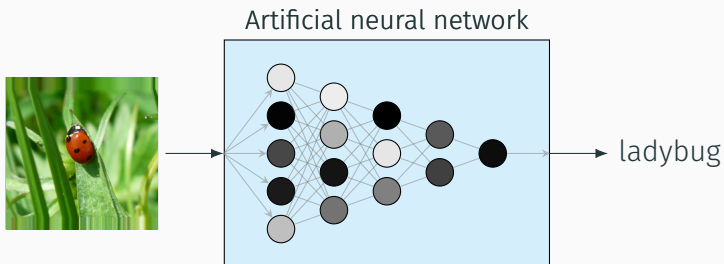
# Artificial neural networks: Interpretability



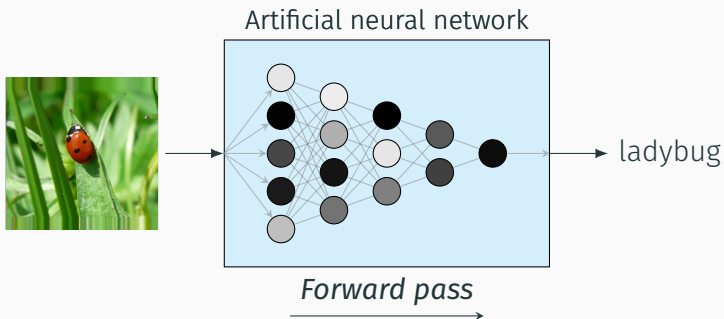
# Artificial neural networks: Interpretability



# Artificial neural networks: Interpretability



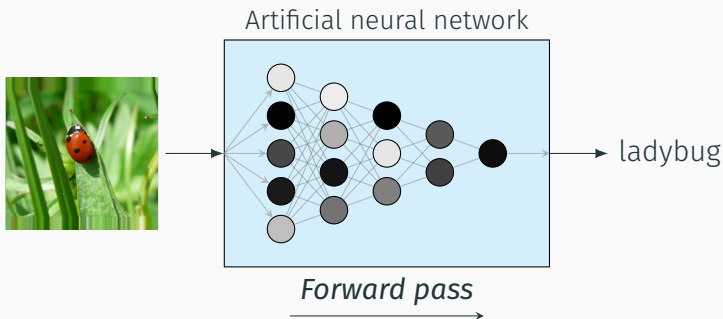
# Artificial neural networks: Interpretability



$$n_j^i = f\left(\sum_{k=0}^n w_{jk}^i n_k^{i-1}\right)$$



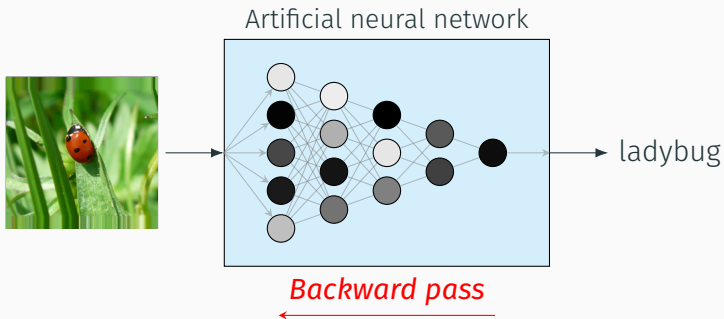
# Artificial neural networks: Explainability



$$n_j^i = f(\sum_k w_{jk}^i n_k^{(i-1)})$$



# Artificial neural networks: Explainability



$$R_j^{i-1} = \sum_k \frac{n_j^{(i-1)} w_{jk}^{(i-1)}}{\sum_l n_l^{(i-1)} w_{lk}^{(i-1)}} R_k^i$$



# Artificial neural networks: Explainability

