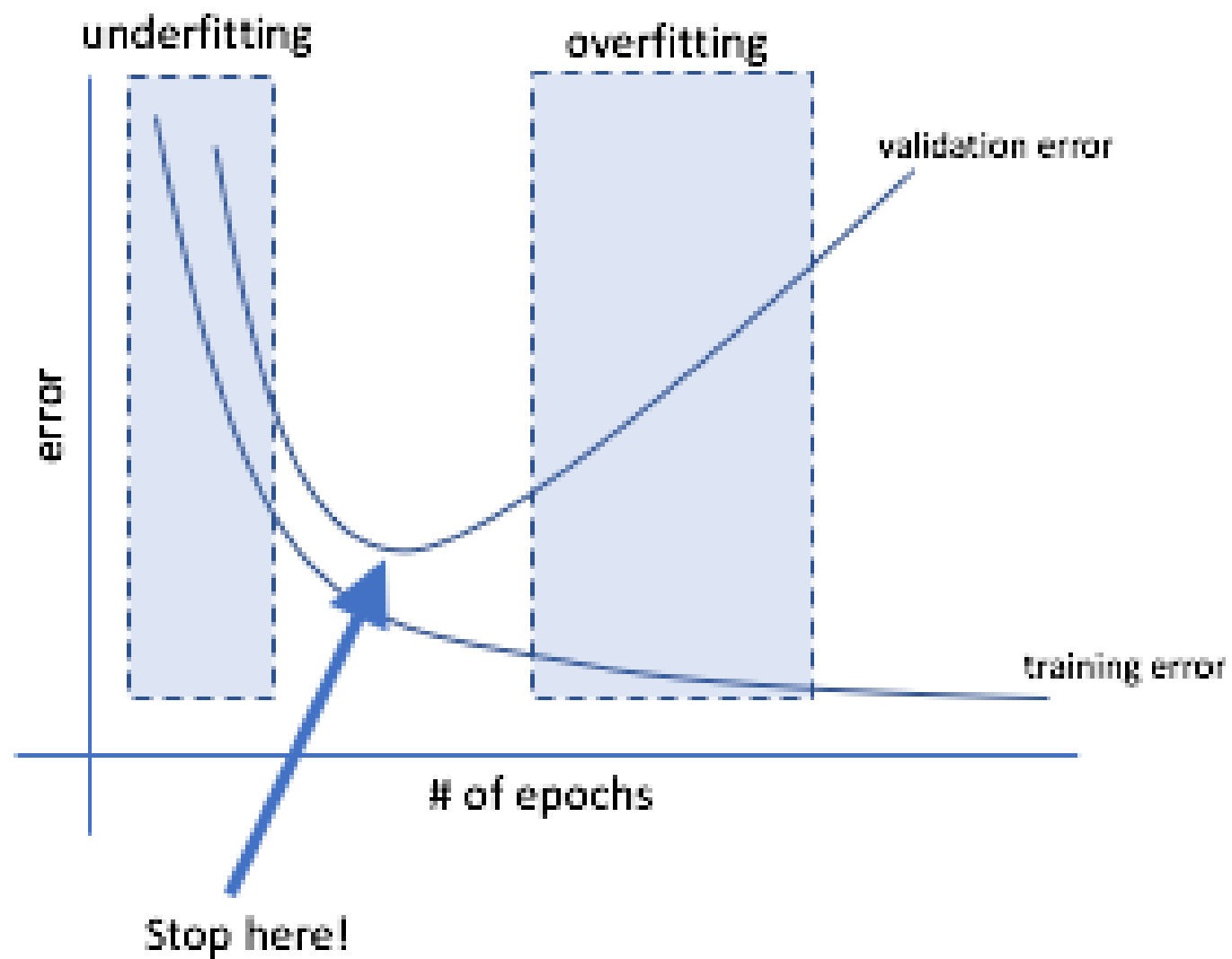


# Jak długo trenować sieć?

Weronika Hryniewska

Zbyt słabe  
czy  
zbyt silne  
dopasowanie modelu?

	Underfitting	Just right	Overfitting
Symptoms	<ul style="list-style-type: none"> <li>• High training error</li> <li>• Training error close to test error</li> <li>• High bias</li> </ul>	<ul style="list-style-type: none"> <li>• Training error slightly lower than test error</li> </ul>	<ul style="list-style-type: none"> <li>• Very low training error</li> <li>• Training error much lower than test error</li> <li>• High variance</li> </ul>
Regression illustration			
Classification illustration			
Deep learning illustration			
Possible remedies	<ul style="list-style-type: none"> <li>• Complexify model</li> <li>• Add more features</li> <li>• Train longer</li> </ul>		<ul style="list-style-type: none"> <li>• Perform regularization</li> <li>• Get more data</li> </ul>



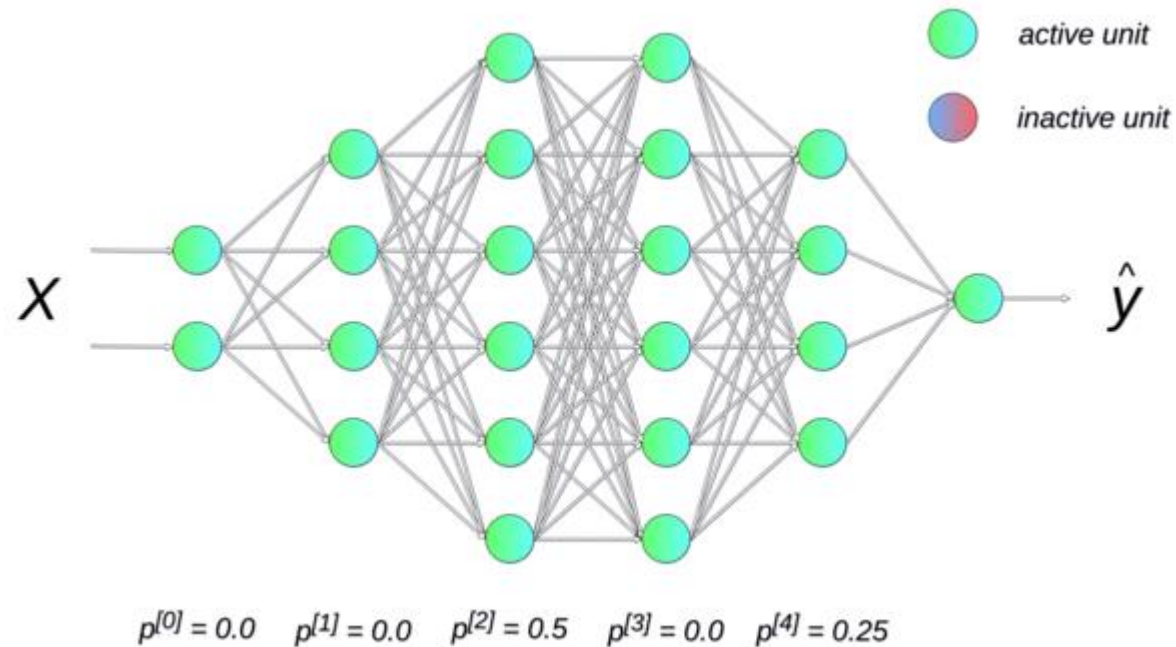
# Sposoby radzenia sobie z za małym modelem

- dodać więcej warstw
- zwiększyć rozmiar warstw
- zwiększyć liczbę epok trenowania

# Sposoby radzenia sobie z za dużym modelem

- Więcej danych treningowych
  - Sztuczne powiększenie zbioru danych
- Regularyzacja
  - redukcja rozmiaru sieci
  - dodawanie regularyzacji wag
  - porzucanie – technika dropout
- Wczesne zatrzymanie trenowania

# Mechanizm porzucania



# Regularyzacja L1 i L2

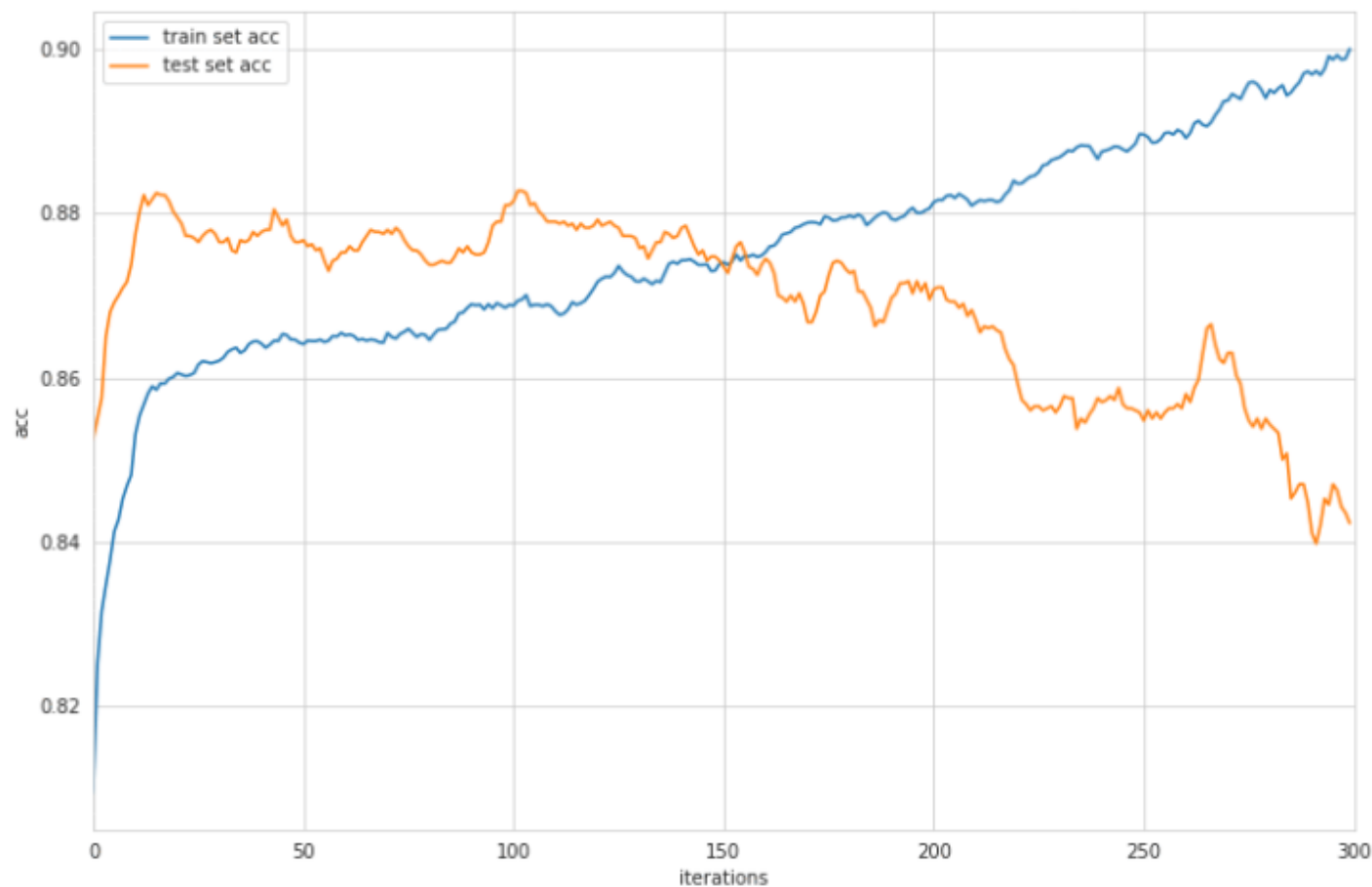
*Funkcja  
błędu*

$$\Sigma(y - \hat{y})^2 + \lambda(|m_1| + |m_2| + \dots + |m_n|)$$

$$\Sigma(y - \hat{y})^2 + \lambda(m_1^2 + m_2^2 + \dots + m_n^2)$$

L1 Regularization	L2 Regularization
1. L1 penalizes sum of absolute values of weights.	1. L2 penalizes sum of square values of weights.
2. L1 generates model that is simple and interpretable.	2. L2 regularization is able to learn complex data patterns.
3. L1 is robust to outliers.	3. L2 is not robust to outliers.

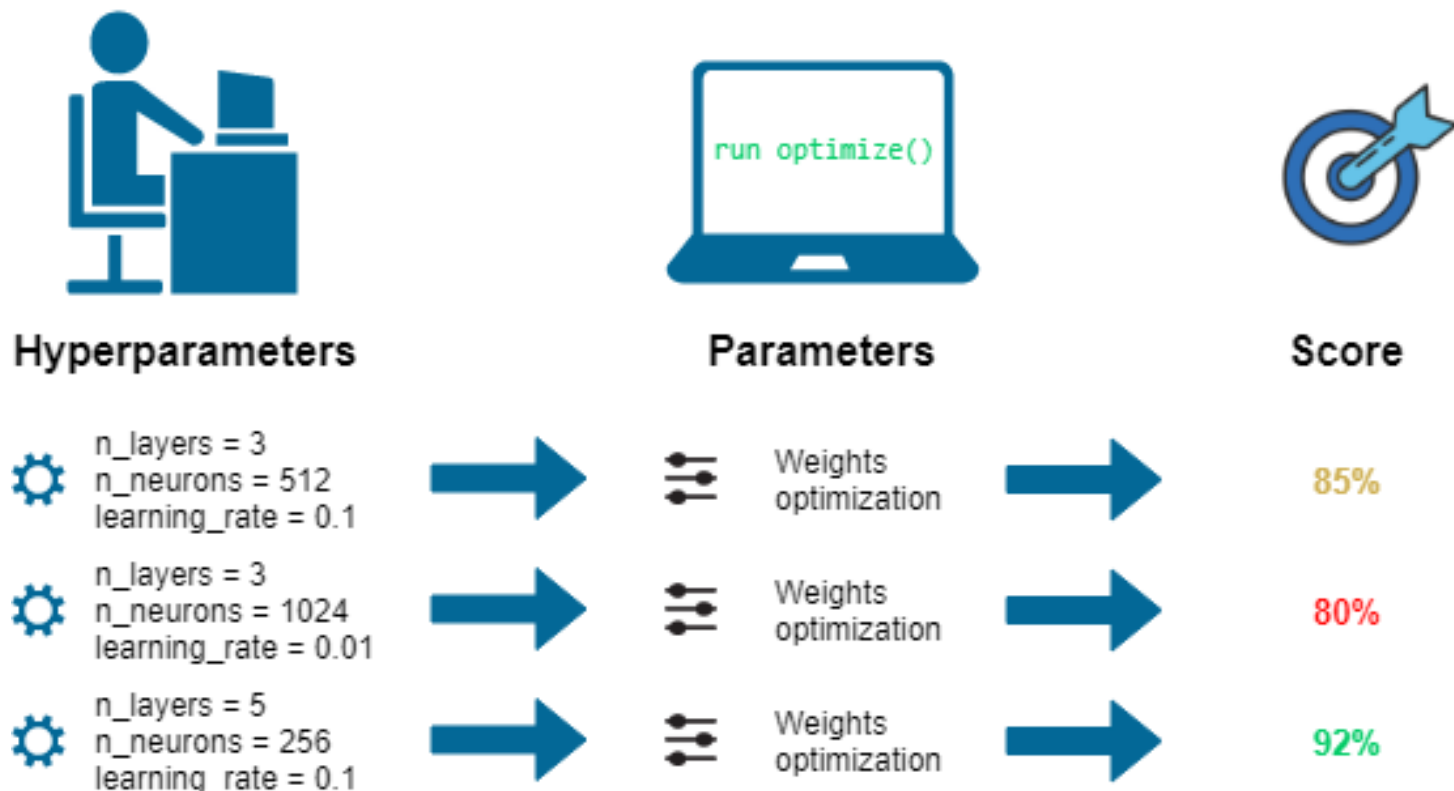
# Wczesne zatrzymanie trenowania



```
EarlyStopping(  
    monitor='val_accuracy',  
    min_delta=1,  
    patience=2,  
    mode='max',  
    restore_best_weights=True,  
)
```



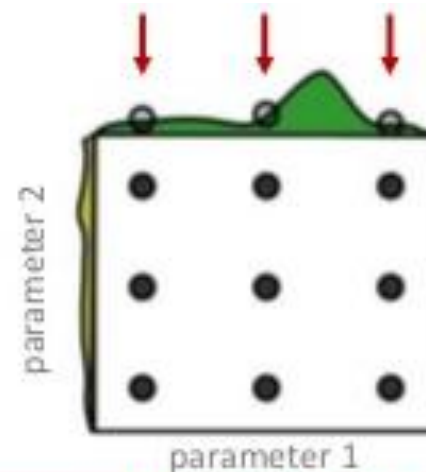
# Dostrajanie hiperparametrów modelu



# Dostrajanie hiperparametrów modelu

## Grid search

1. Select values for each hyperparameter to test
2. Try ALL combinations



## Random search

- Varies important hyperparameters more
- More efficient at model tuning

