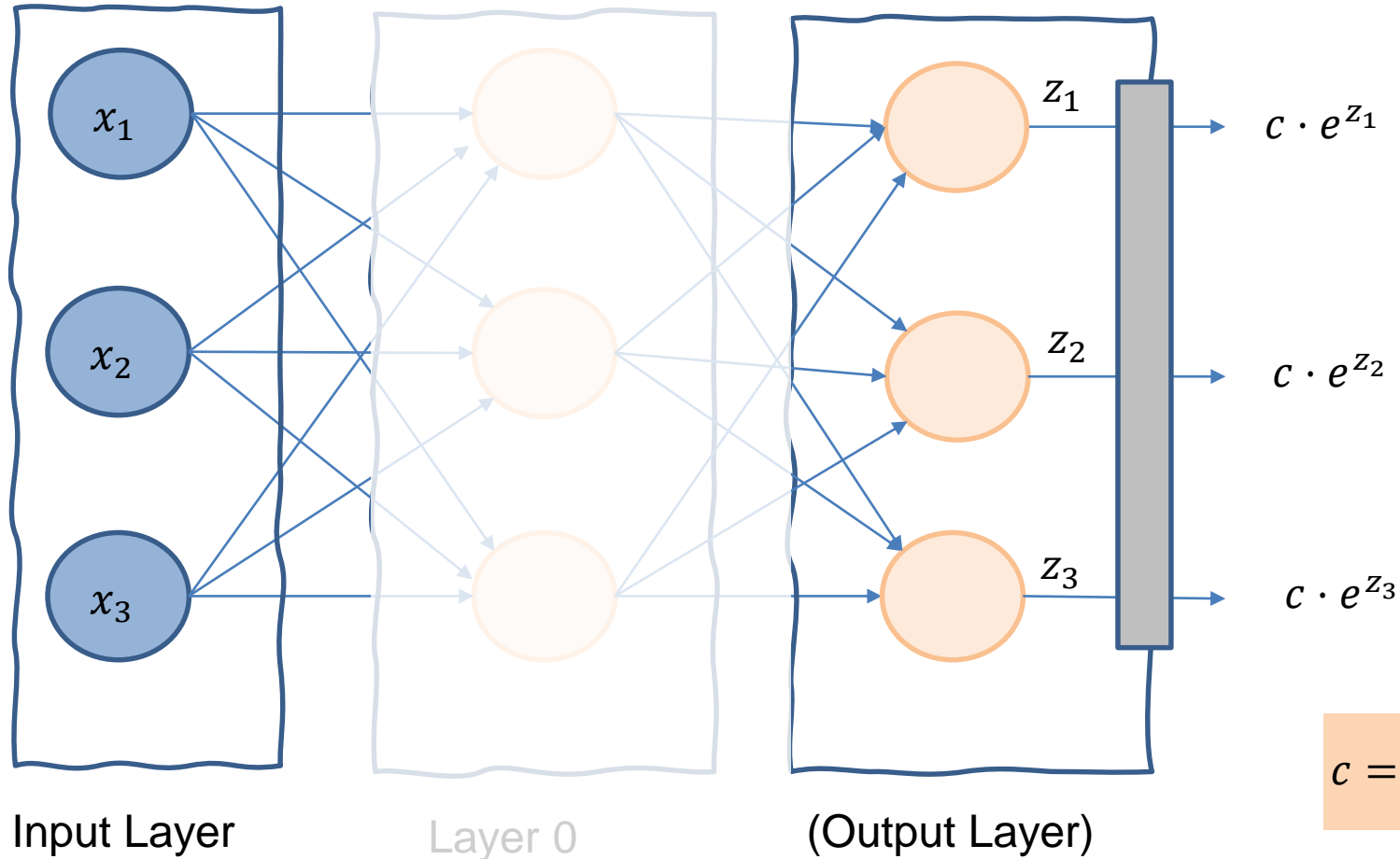


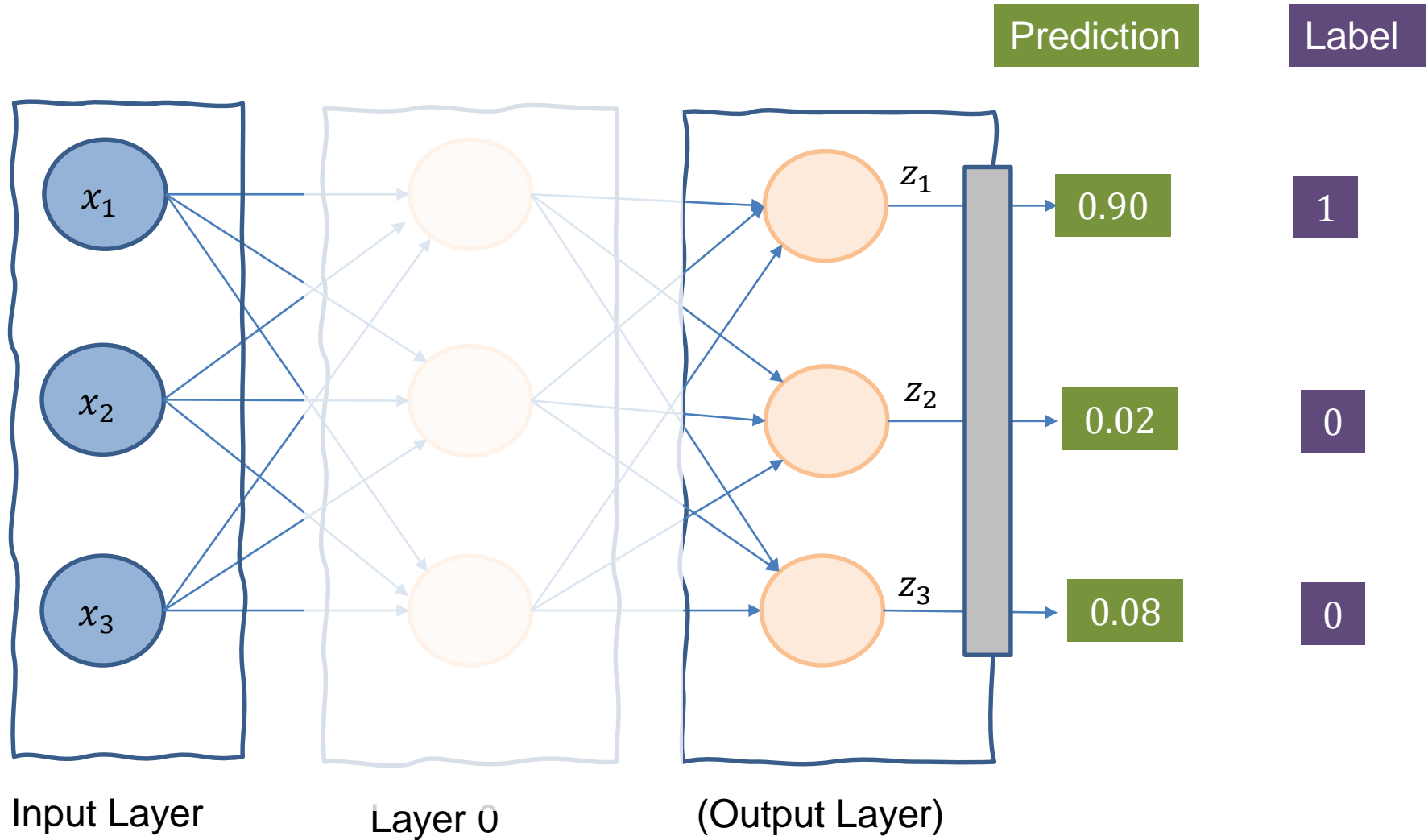


Classification and Softmax

Classification Network with Softmax-Activation(Example)



Prediction vs Label



Labels und Predictions als Zufallsvariablen mit verschiedenen Verteilungen

Zufallsvariable \hat{Y}
mit Verteilung
 $\hat{y}_1, \hat{y}_2, \hat{y}_3$

Zufallsvariable Y
mit Verteilung
 y_1, y_2, y_3

Prediction

Label

0.90

1

0.02

0

0.08

0

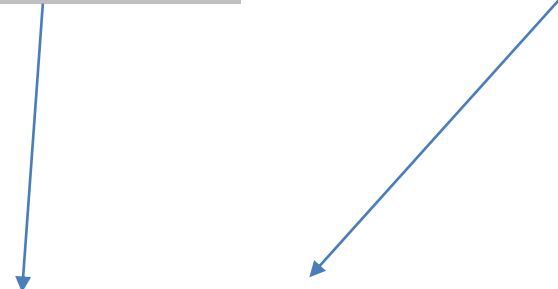
Ende des Stoffs – ab hier optional



Categorical Cross Entropie: Wie groß ist der Abstand der Verteilungen?

Zufallsvariable Y
mit Verteilung
 y_1, y_2, y_3

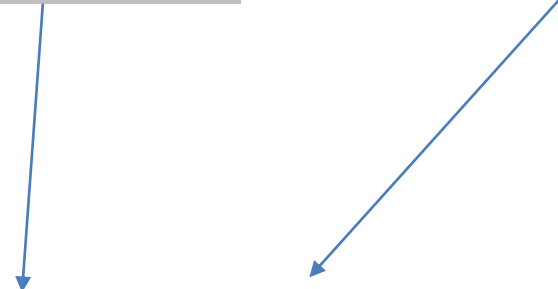
Zufallsvariable \hat{Y}
mit Verteilung
 $\hat{y}_1, \hat{y}_2, \hat{y}_3$

$$CE(Y, \hat{Y}) = - \sum_{i=1}^3 y_i \cdot \log(\hat{y}_i)$$


Categorical Cross Entropie: Wie groß ist der Abstand der Verteilungen?

Zufallsvariable Y
mit Verteilung
 y_1, y_2, y_3

Zufallsvariable \hat{Y}
mit Verteilung
 $\hat{y}_1, \hat{y}_2, \hat{y}_3$

$$CE(Y, \hat{Y}) = - \sum_{i=1}^3 y_i \cdot \log(\hat{y}_i)$$


In MNIST: Nur ein Eintrag von y steht auf 1. Sei i^* der Index dieses Eintrags.

Zufallsvariable Y
mit Verteilung
 y_1, y_2, y_3

Zufallsvariable \hat{Y}
mit Verteilung
 $\hat{y}_1, \hat{y}_2, \hat{y}_3$

$$CE(Y, \hat{Y}) = -1 \cdot \log(\hat{y}_{i^*})$$