

Ime (Keras)	Formula	Opis	Uporaba (dis./zv.)
mean_squared_error (MSE)	$\frac{1}{n} \sum (y_i - \hat{y}_i)^2$	Povprečna kvadratna napaka.	Regresija (zv.)
mean_absolute_error (MAE)	$\frac{1}{n} \sum  y_i - \hat{y}_i $	Povprečna absolutna napaka, manj občutljiva na outlierje.	Regresija (zv.)
binary_crossentropy	$-\frac{1}{n} \sum [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)]$	Verjetnostna izguba za dvo-razredne probleme.	Klasifikacija (dis.)
categorical_crossentropy	$-\sum y_i \log \hat{y}_i$	Večrazredna klasifikacija.	Klasifikacija (dis.)
hinge	$\frac{1}{n} \sum \max(0, 1 - y_i \hat{y}_i)$	Za SVM podobne modele.	Klasifikacija (dis., -1/1 labeli)
Huber	$L_\delta = \begin{cases} \frac{1}{2}(y - \hat{y})^2 &  y - \hat{y}  \leq \delta \\ \delta y - \hat{y}  - \frac{1}{2}\delta^2 & \text{sicer} \end{cases}$	Kombinacija MSE in MAE.	Regresija (zv.)

**Table 1:** Najpogostejše stroškovne funkcije v Kerasu.

Metrika (Keras)	Formula	Opis	Uporaba (dis./zv.)
accuracy	$\frac{\# \text{ pravih}}{\# \text{ vseh}}$	Delež pravih napovedi.	Klasifikacija (dis.)
binary_accuracy	$\frac{1}{n} \sum 1_{[y_i = \hat{y}_i]}$	Točnost pri binarnih problemih.	Binarna klasifikacija (dis.)
categorical_accuracy	$1_{[\arg \max(y_i) = \arg \max(\hat{y}_i)]}$	Točnost pri večrazrednih problemih.	Večrazredna klasifikacija (dis.)
mse	$\frac{1}{n} \sum (y_i - \hat{y}_i)^2$	MSE kot metrika.	Regresija (zv.)
mae	$\frac{1}{n} \sum  y_i - \hat{y}_i $	MAE kot metrika.	Regresija (zv.)
RootMeanSquaredError (RMSE)	$\sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$	Koren MSE, interpretacija v enotah ciljne spremenljivke.	Regresija (zv.)
MeanAbsolutePercentageError (MAPE)	$\frac{100}{n} \sum \left  \frac{y_i - \hat{y}_i}{y_i} \right $	Odstotekna napaka.	Regresija (zv., brez $y_i = 0$ ).

**Table 2:** Najpogostejše metrike v Kerasu.

Ime (Keras)	Formula	Opis	Kdaj uporabiti
SGD	$w_{t+1} = w_t - \eta \nabla L(w_t)$	Osnovni stochastic gradient descent.	Enostavni modeli, baseline.
SGD(momentum)	$v_{t+1} = \mu v_t - \eta \nabla L(w_t)$	Dodatek momentuma za hitrejšo konvergenco.	Globoke mreže.
RMSprop	$w_{t+1} = w_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} g_t$	Normalizira gradient po drseči povprečni kvadratični vrednosti.	RNN, LSTM.
Adagrad	$w_{t+1} = w_t - \frac{\eta}{\sqrt{\sum g_i^2 + \epsilon}} g_t$	Prilagajanje učne stopnje za redke lastnosti.	NLP, priporočila.
Adadelta	–	Razširitev Adagrad (brez hitrega padanja $\eta$ ).	Dolgotrajno učenje.
Adam	$m_t, v_t$ kombinacija momenta in RMSprop	Najbolj uporabljen, robusten.	Prva izbira pri večini problemov.
Nadam	Adam + Nesterov	Hitrejša konvergenca.	Kompleksne mreže.

**Table 3:** Optimizatorji v Kerasu.

Layer	Parametri	Opis	Uporaba
Dense	units, activation	Fully connected sloj.	Klasifikacija, regresija.
Conv1D/2D/3D	filters, kernel_size, strides	Konvolucijski sloji.	Slike, videi, sekvence.
MaxPooling...	pool_size	Zmanjševanje dimenzionalnosti.	CNN arhitekture.
Flatten	–	Preoblikuje vektor.	CNN → Dense.
Dropout	rate	Regularizacija.	Globoke mreže.
BatchNormalization	momentum, epsilon	Normalizacija aktivacij.	Skoraj povsod.
Embedding	input_dim, output_dim	Mapiranje kategorij v vektorje.	NLP, kategorialni podatki.
LSTM/GRU	units, return_sequences	Rekurentni sloji.	Časovne vrste, NLP.

**Table 4:** Najpogostejši layerji v Kerasu.

Funkcija	Formula	Opis	Uporaba
linear	$f(x) = x$	Identiteta.	Regresija.
sigmoid	$f(x) = \frac{1}{1+e^{-x}}$	Izhod $(0, 1)$ .	Binarna klasifikacija.
tanh	$\tanh(x)$	Izhod $(-1, 1)$ .	RNN.
relu	$\max(0, x)$	Najpogostejša aktivacija.	Skrite plasti CNN/MLP.
leaky_relu	$\max(\alpha x, x)$	Odpravi mrtve ReLU.	Globoke mreže.
elu	$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$	Gladkejši od ReLU.	CNN/MLP.
softmax	$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$	Pretvori v verjetnosti.	Večrazredna klasifikacija.

**Table 5:** Aktivacijske funkcije v Kerasu.