

Diskussion Performance

Vergleich der Ergebnisse der Modelle.

by Alexander Messner, Erik Pfandler

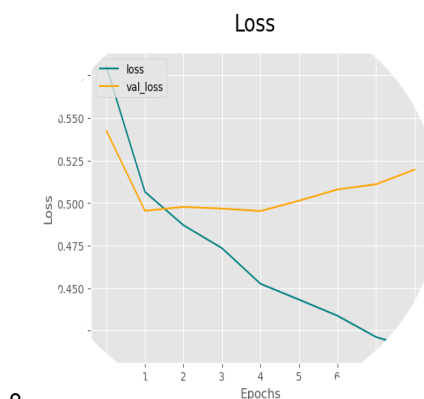
Modellstrukturen

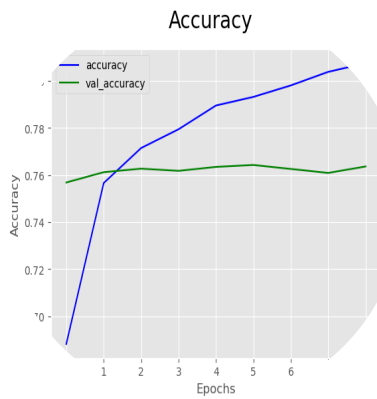
Alle Modelle verwenden das selbe Data-Preprocessing.

Modell 1

100000 Datensätze (50/50), Batch-Size 128, Epochen 15

- Schichten:
 - Embedding-Schicht: 5000 Vokabulargrößen, 75 output.
 - LSTM: 64 Neuronen, dropout=0.5, recurrent_dropout=0.2
 - Dropout(0.4)
 - Dense(64): Aktivierung relu
 - BatchNormalization
 - Dropout(0.5)
 - Dense(1): Aktivierung sigmoid
- Callbacks:
 - EarlyStopping(monitor='val_accuracy',patience = 3 ,mode='max')
 - ReduceLROnPlateau(monitor='val_loss', factor=0.5, patience=2, min_lr=0.00001)
- Performance:





o

o Test Evaluation: accuracy: 0.7632 - loss: 0.5214

- Analyse: Die Trainingsperformances verbessern sich stetig, Die Validierungs-Performance bleibt aber stetig, das ist ein Zeichen von Overfitting.

Modell 2

60000 Datensätze (50/50), Batch-Size 128, Epochen 7

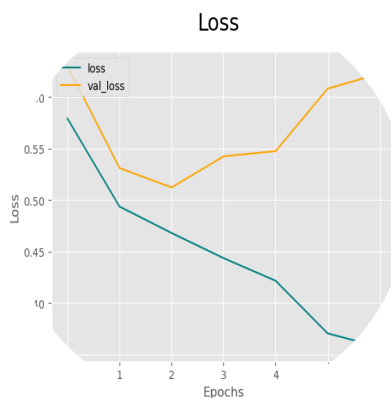
- Schichten:

- o Embedding-Schicht: 5000 Vokabulargrößen, 75 output.
- o LSTM: 64 Neuronen, dropout=0.2, recurrent_dropout=0.2, return_sequences=True
- o LSTM: 64 Neuronen, dropout=0.3, recurrent_dropout=0.2,
- o Dropout(0.2)
- o Dense(64): Aktivierung relu
- o BatchNormalization
- o Dropout(0.3)
- o Dense(1): Aktivierung sigmoid

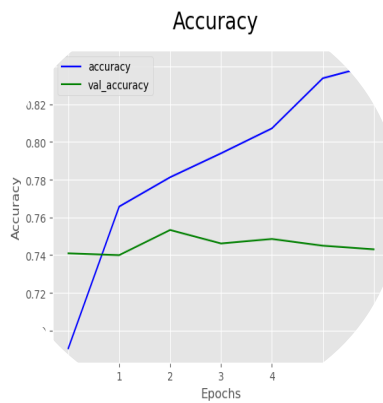
- Callbacks:

- o EarlyStopping(monitor='val_accuracy',patience = 10 ,mode='max')
- o ReduceLROnPlateau(monitor='val_loss', factor=0.3, patience=2)

- Performance:



o



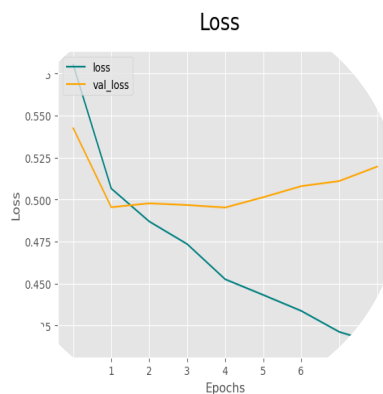
- o
- o Test Evaluation: accuracy: 0.7378 - loss: 0.6140

- Analyse: Die Trainingsperformancen verbessern sich stetig, Die Validierungs-Performance bleibt stetig bzw. der Loss steigt nach Epoche 3 sogar an , das ist ein Zeichen von Overfitting.

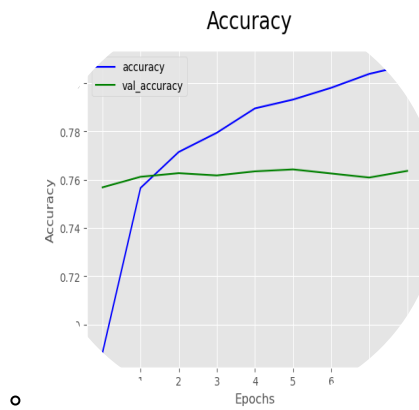
Modell 3

100000 Datensätze (50/50), Batch-Size 128, Epochen 15

- Schichten:
 - o Embedding-Schicht: 5000 Vokabulargrößen, 75 output.
 - o LSTM: 64 Neuronen, dropout=0.5, recurrent_dropout=0.2,
 - o Dropout(0.4)
 - o Dense(64): Aktivierung relu
 - o BatchNormalization
 - o Dropout(0.5)
 - o Dense(1): Aktivierung sigmoid
- Callbacks:
 - o EarlyStopping(monitor='val_accuracy',patience = 3 ,mode='max')
 - o ReduceLROnPlateau(monitor='val_loss', factor=0.5, patience=2, min_lr=0.00001)
- Performance:



- o



o Test Evaluation: accuracy: accuracy: 0.7632 - loss: 0.5214

- Analyse: Die Trainingsperformancen verbessern sich stetig, Die Validierungs-Performance bleibt stetig über den gesamten Zeitraum. Overfitting ist vorhanden, es muss mehr generalisiert werden.

Diskussion

Modellstruktur und Datenmenge

- Modell 1 und Modell 3 nutzen 100.000 Datensätze, während Modell 2 nur 60.000 Datensätze verarbeitet. Dies spiegelt sich in einer besseren Testgenauigkeit (0.7632 vs. 0.7378) bei den Modellen 1 und 3 wider.
- Modell 2 integriert eine zusätzliche LSTM-Schicht, was die Modellkomplexität erhöht, jedoch zu stärkerem Overfitting führt.

Overfitting und Regularisierung

Alle Modelle zeigen Anzeichen von Overfitting, besonders Modell 2, bei dem der Validierungsverlust nach Epoche 3 deutlich ansteigt. Höhere Dropout-Werte (0.4–0.5 in Modell 1 und 3) wirken Overfitting besser entgegen als die niedrigeren Werte in Modell 2 (0.2–0.3). Overfitting ist jedoch bei jedem Modell vorhanden.

Leistung und Effizienz

Modell 1 und Modell 3 liefern nahezu identische Ergebnisse mit der höchsten Genauigkeit (0.7632) und dem niedrigsten Verlust (0.5214). Sie sind damit die effizientesten Modelle. Modell 2 schneidet schlechter ab, was auf die geringere Datenmenge und die ineffektive Regularisierung zurückzuführen ist.