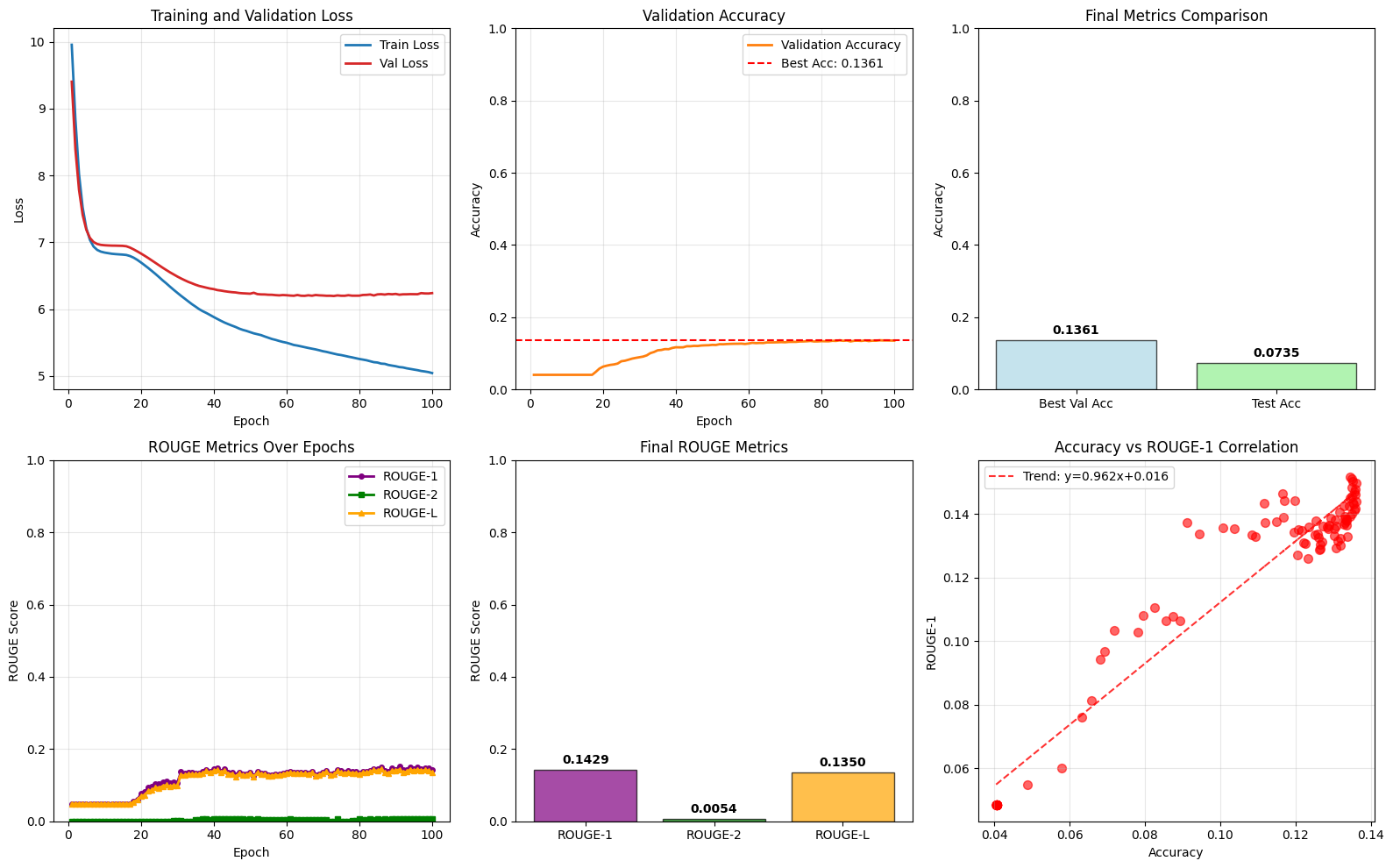
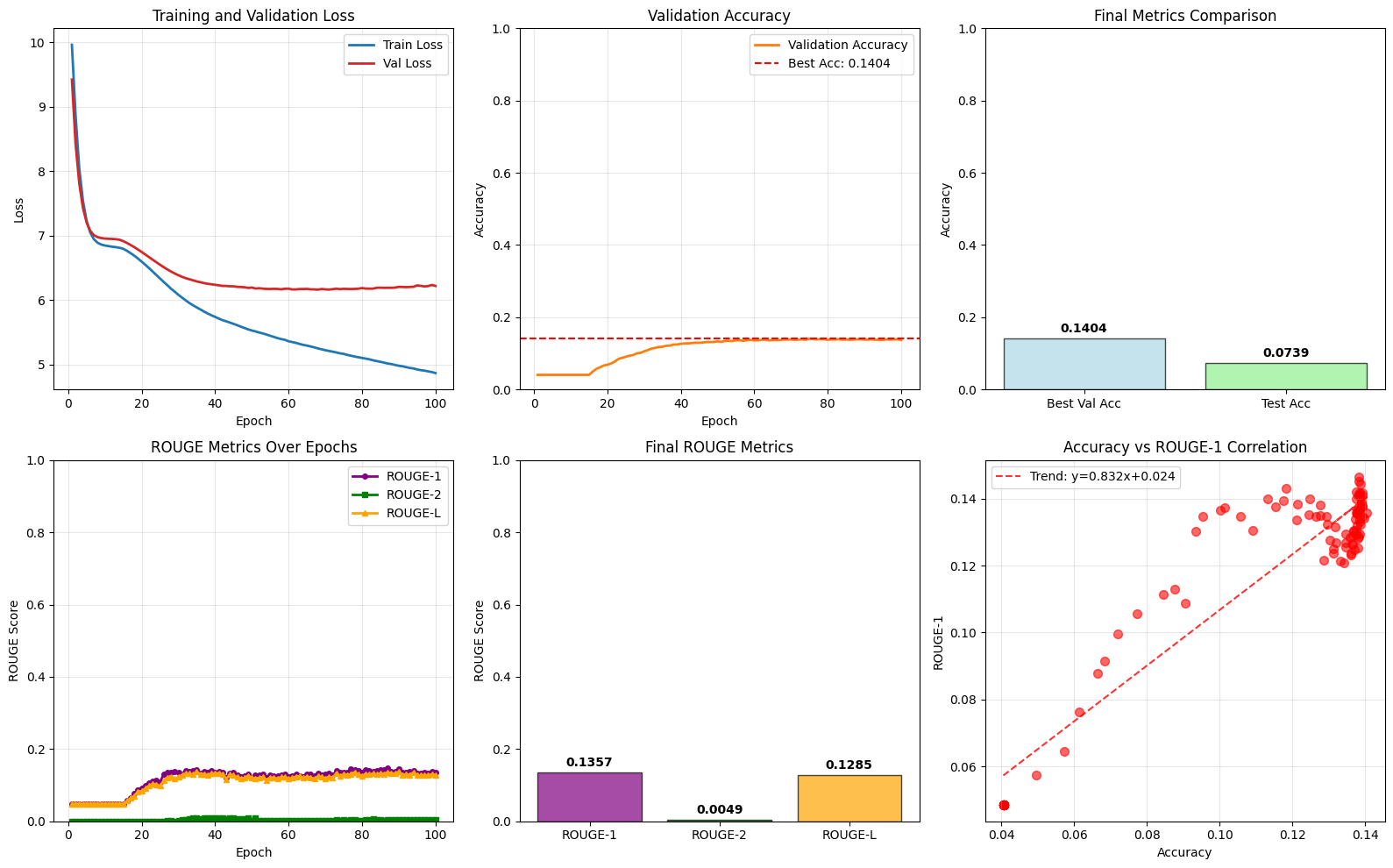
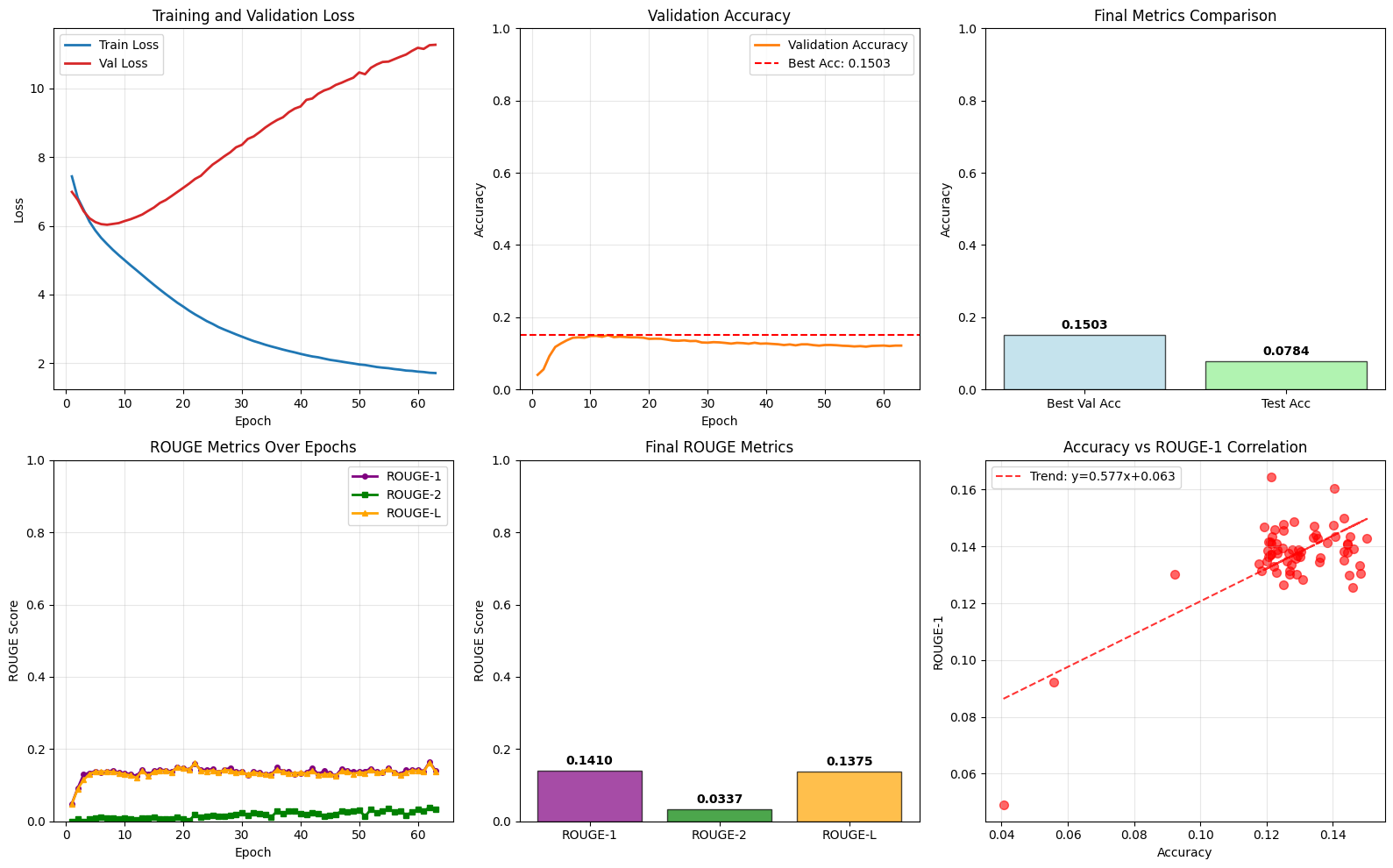
Emb dim 150, nn.Dropout(0.5)



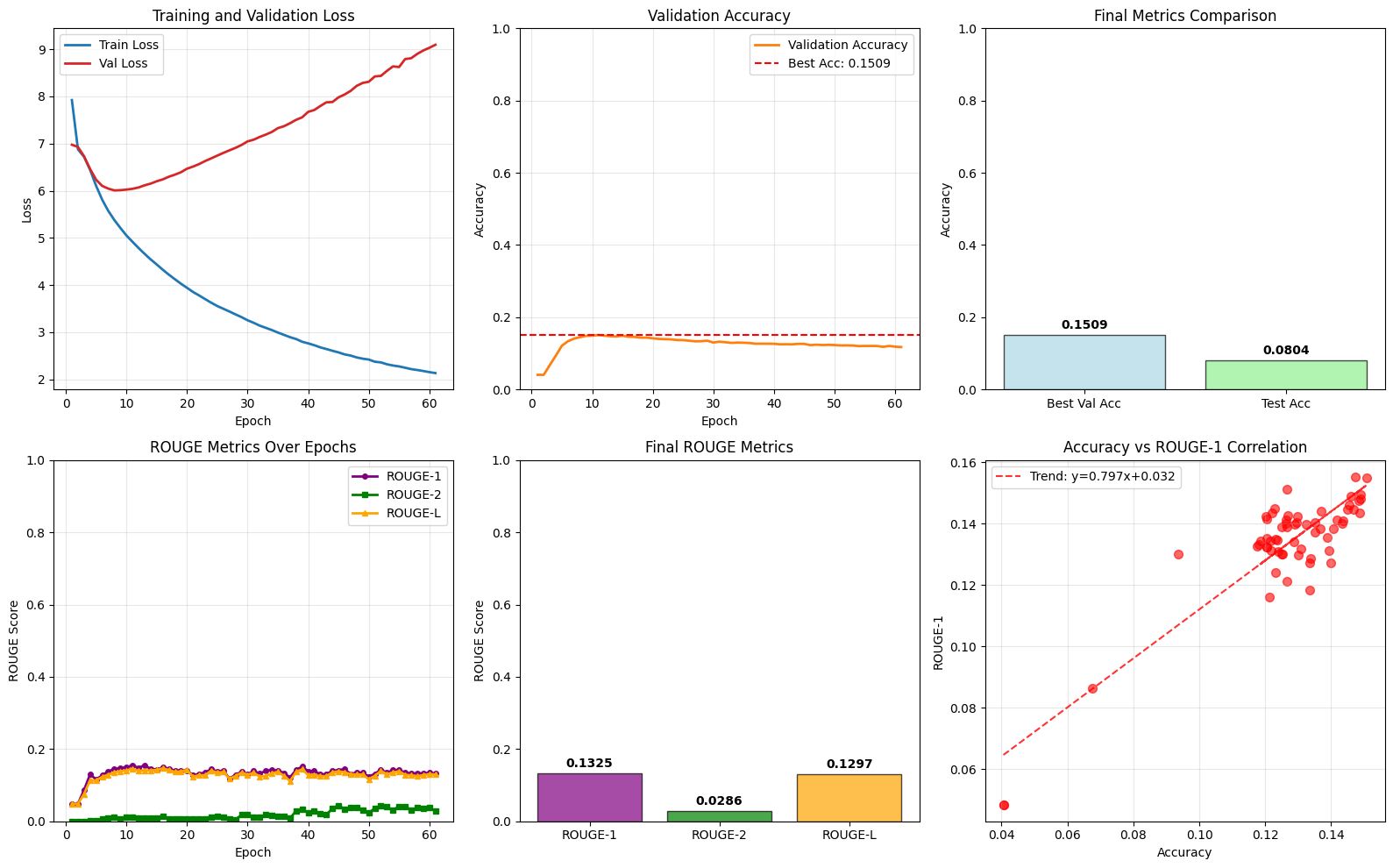
Emb\_dim 300, lr=1e-4 (0,0001) , nn.Dropout(0.5)



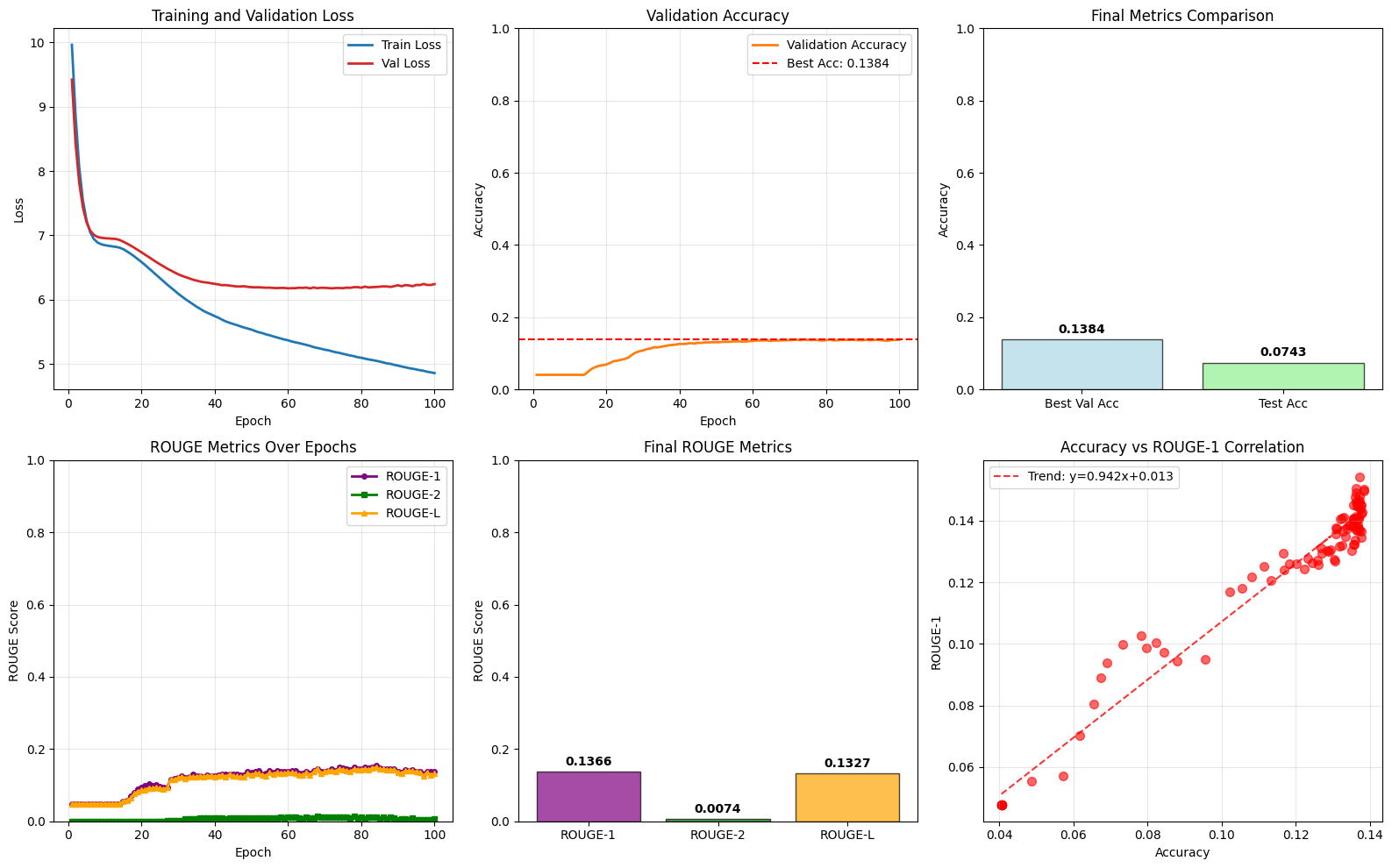
Emb\_dim 300, lr=3e-3 высокая скорость обучения, weight\_decay=0.01 , nn.Dropout(0.5)



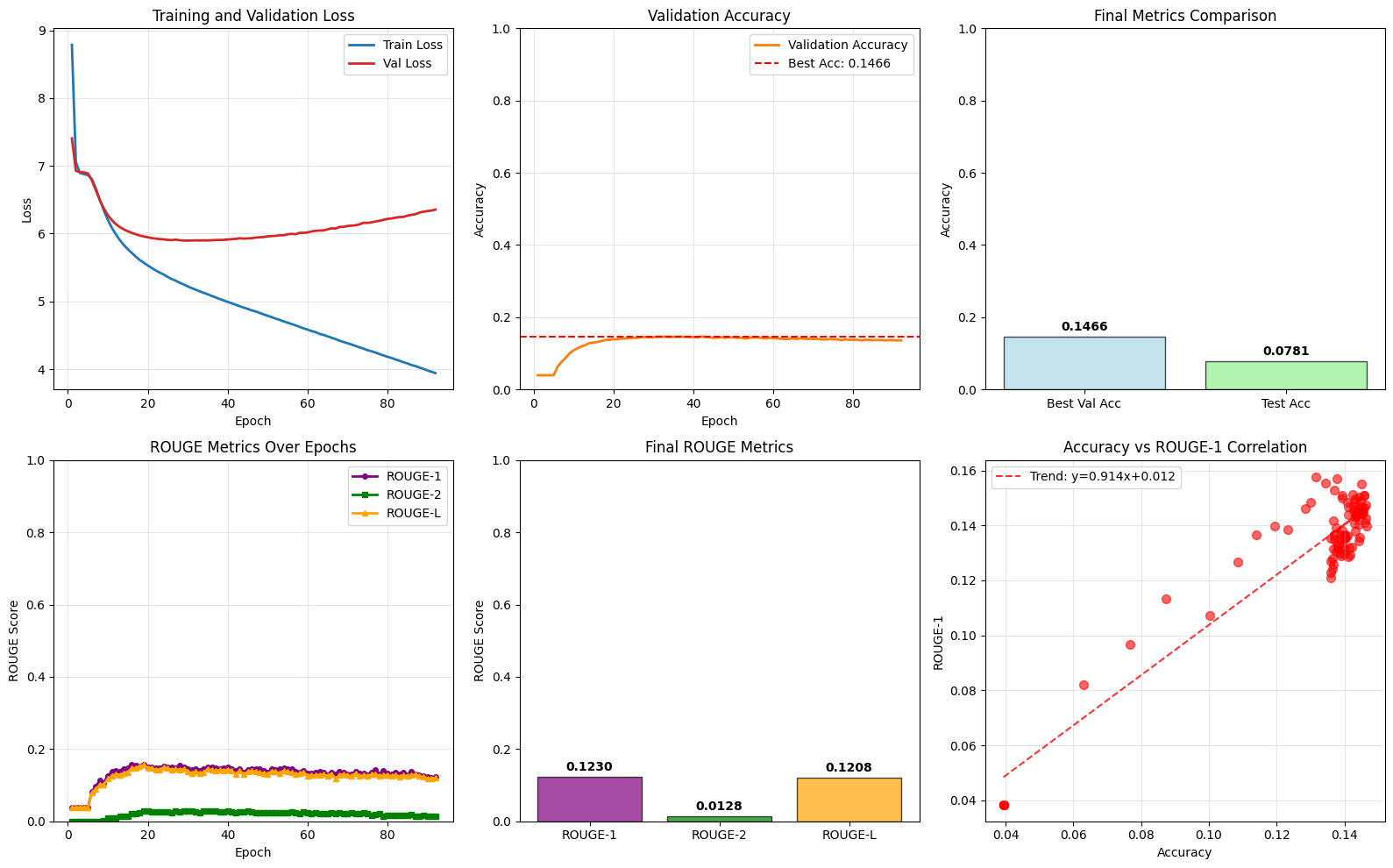
Emb\_dim 300, lr= 1e-3, weight\_decay=0.001 , nn.Dropout(0.5)



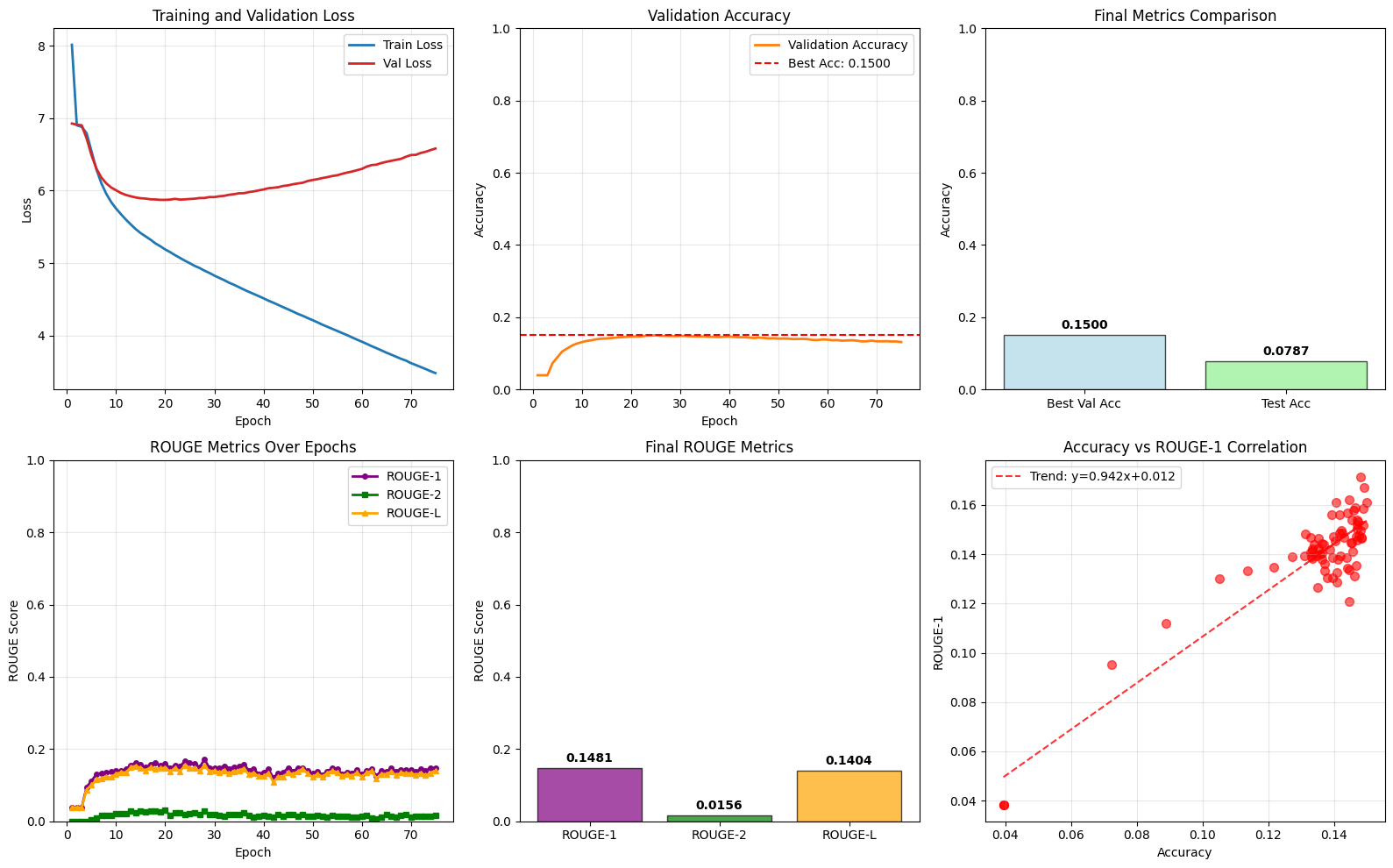
Emb\_dim 300, lr= 1e-4, weight\_decay=0.001 , nn.Dropout(0.5)



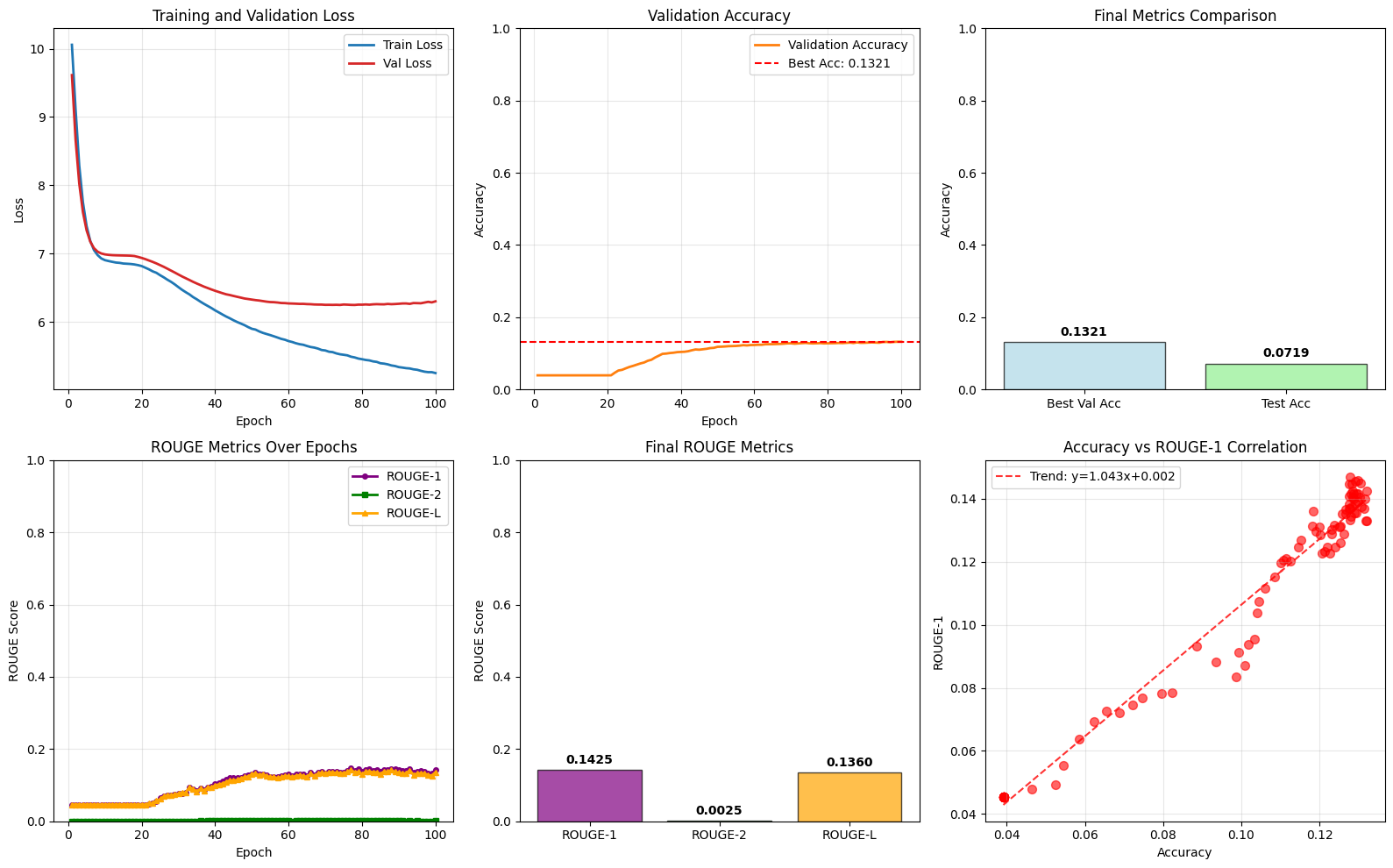
1mln sample Emb\_dim 300, lr= 1e-4, weight\_decay=0.001, 'hidden\_dim': 256 , nn.Dropout(0.5)



, nn.Dropout(0.5)



Maxlen 20



Maxlen 10

'MAX\_LEN': 10,

    'num\_epoch':100,

    'limit': 500000,

    'num\_layers':2,

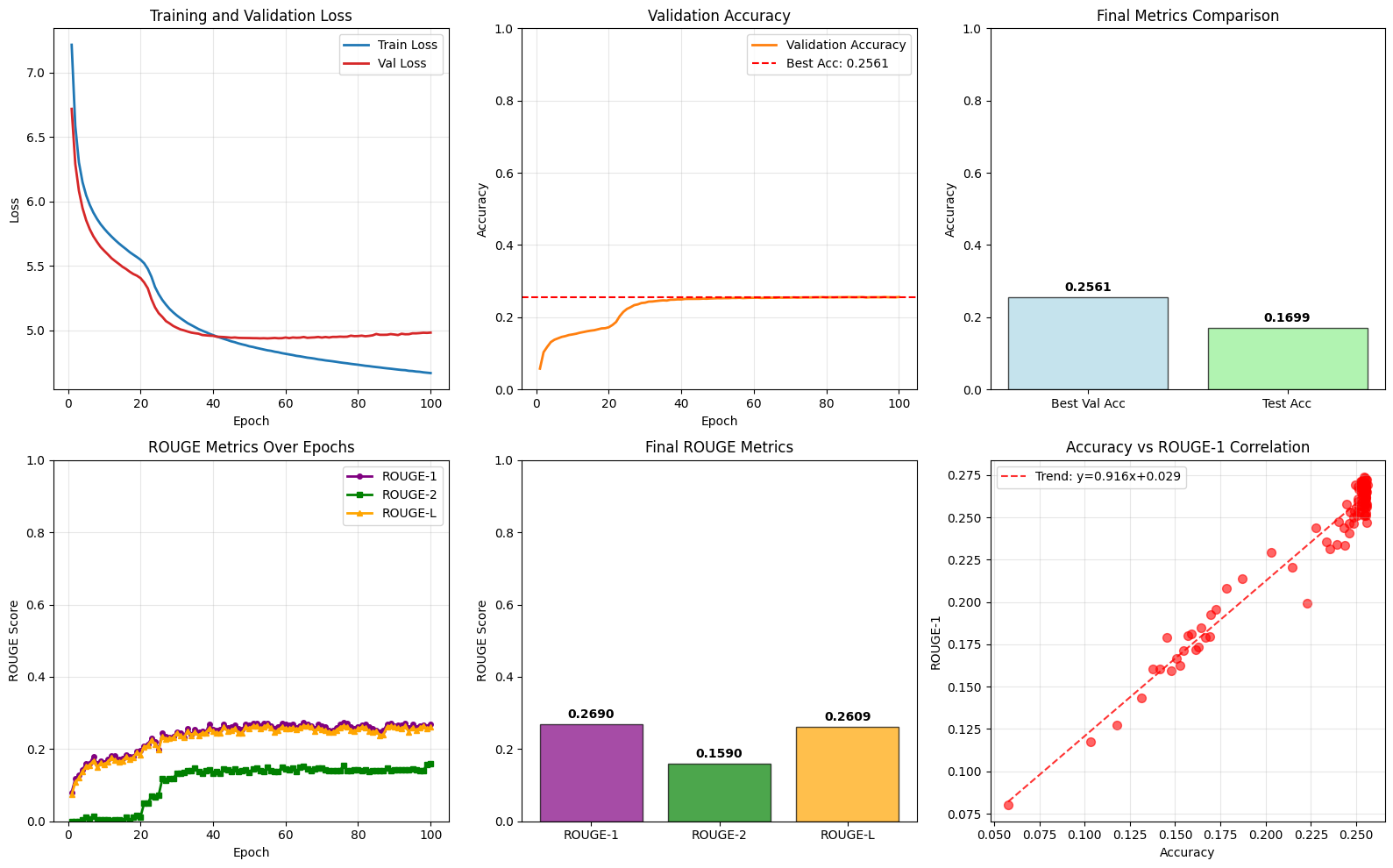
    'hidden\_dim': 128, #128#256#512

    'emb\_dim': 300,

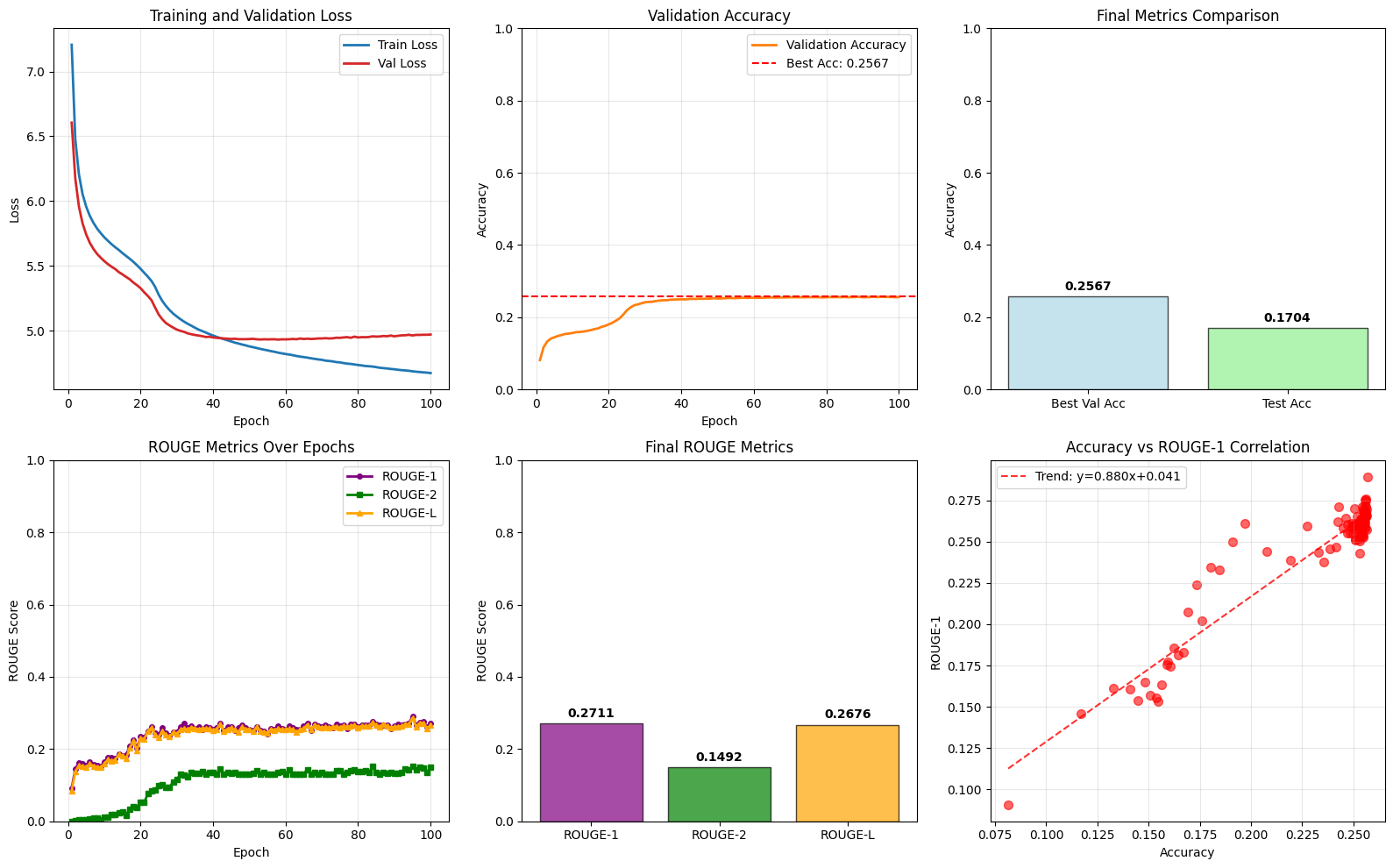
    'batch\_size': 256,

    'rnn\_type': "LSTM",

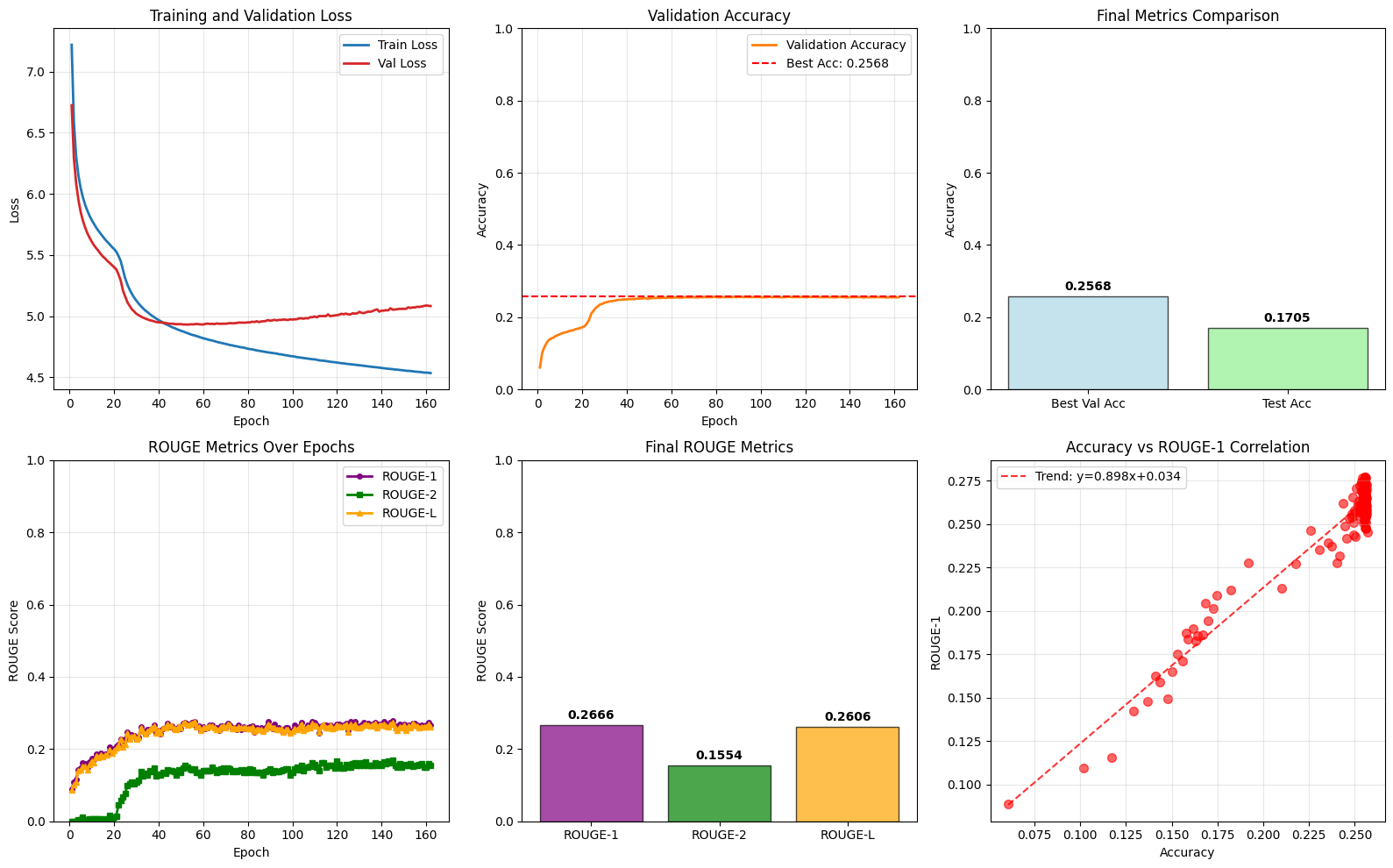
lr=1e-4, weight\_decay=0.05



GRU



LSTM 200 эпох, выход на 162 (patience =50)



Самый последний график – без градиент клиппинга

lr=3e-3

'MAX\_LEN': 10,

    'num\_epoch':200,

    'limit': 1000000,

    'num\_layers':2,

    'hidden\_dim': 128,

    'emb\_dim': 300,

    'batch\_size': 256,

    'rnn\_type': "LSTM"

