LAB1.3 D7047E GROUP 7

Gabriel Röjmyr
gabrjm-0@student.ltu.se
Lulea University of Technology

Benjamin Ådén Fjällman
bendnf-0@student.ltu.se
Lulea University of Technology

Umuthan Ercan
umuerc-2@student.ltu.se
Lulea University of Technology

Aamer Ahmed
aamahm-3@student.ltu.se@student.ltu.se
Lulea University of Technology

Georgios Savvidis
geosav-2@student.ltu.se
Lulea University of Technology

Sushanta Mohapatra
susmoh-3@student.ltu.se
Lulea University of Technology

May 1, 2024

3 Task 1.3

Comparison

Here, it would be best if you did a comparison of both models; you are requested to modify your Chatbot to use the same data as the transformer and answer the following:

- Compare the performance of the two models and explain in which scenarios you would prefer one over the other.
- How did the two models' complexity, accuracy, and efficiency differ? Did one model outperform the other in specific scenarios or tasks? If so, why?
- What insights did you obtain concerning data amount to train? Embeddiutilizedised? Architectural choices made?

1 Answer

The first model might have been less complex, potentially requiring fewer hardware resources, and being more efficient in terms of computational requirements. On the other hand, the second model utilizes advanced transforms which caused an increase in the complexity of the model and demanded higher hardware capabilities, particularly in RAM. This complexity resulted in longer training times due to hardware limitations. Conversely, In scenarios where computational resources are limited or the dataset is relatively small, the first model might be preferred due to its lower complexity and potential for better efficiency. The second model can be preferred if the needed task demands high accuracy and the hardware resources are available to support the computational demands, despite its increased complexity.

Insights gained regarding data amount, embedding utilization, and architectural choices are crucial. For instance, the size of the dataset influences the model's ability to generalize. Larger datasets may benefit from more complex models with higher capacity. Embedding utilization is essential in capturing intricate relationships within the data, but it also adds to the model's complexity. Architectural choices, such as the selection of transforms and optimization algorithms, impact both the model's accuracy and efficiency.

Concerning accuracy and efficiency, the trade-off between the two models becomes apparent. The implementation of the model using transformers, with its higher complexity, might achieve better accuracy but at the cost of longer training times and higher hardware requirements compared to the first model, while potentially less accurate, could offer better efficiency in terms of computational resources and training time.

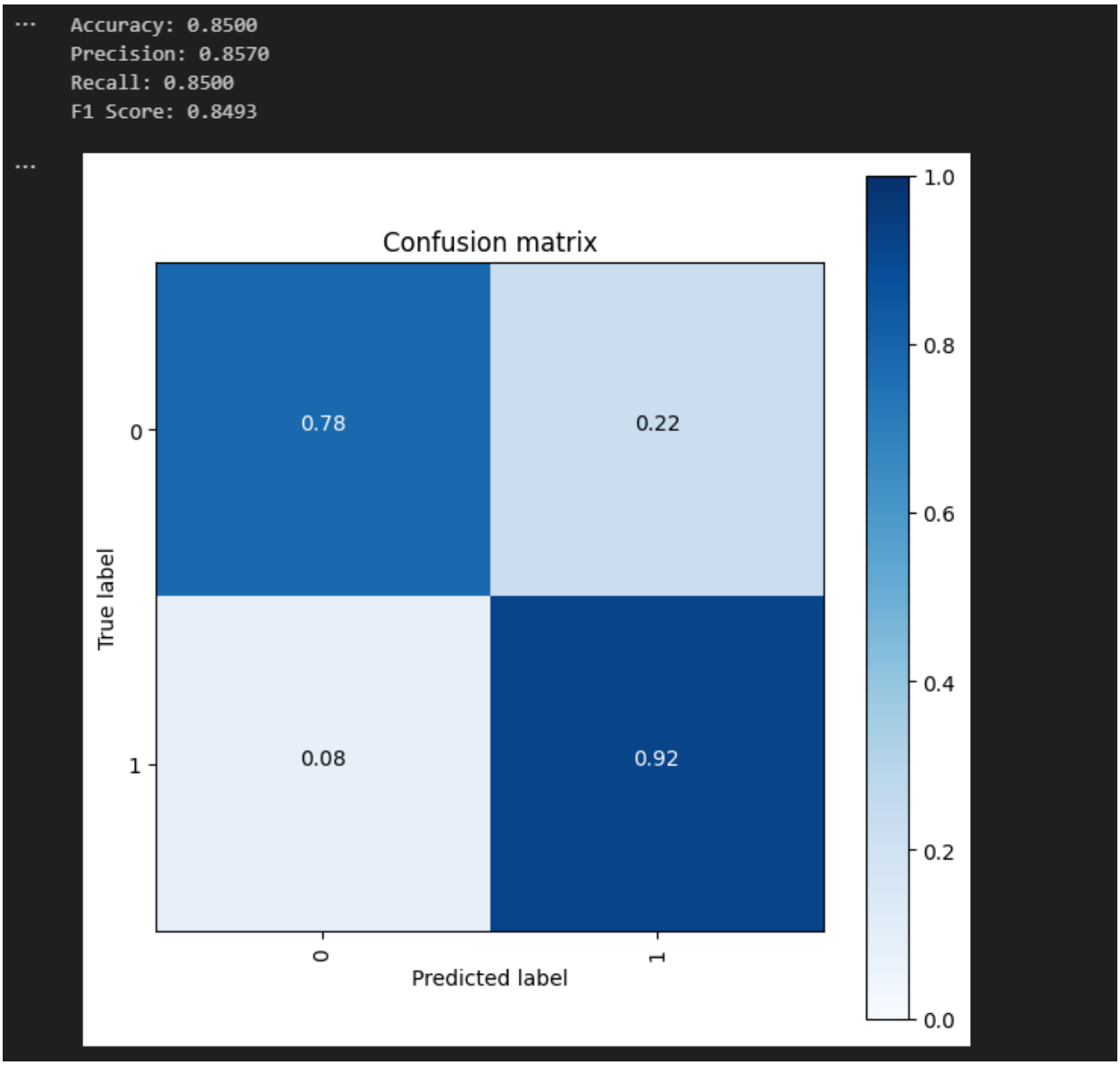

Figure 1: To test our model, we used the SimpleNN network together with a stratified sample from a multiclass dataset from Coursera course reviews (1 to 4 stars). We used around 10K out of 100K stratified samples because of technical resource limitations as using a bigger sample would lead to a much bigger vocabulary and significantly higher memory would be needed.

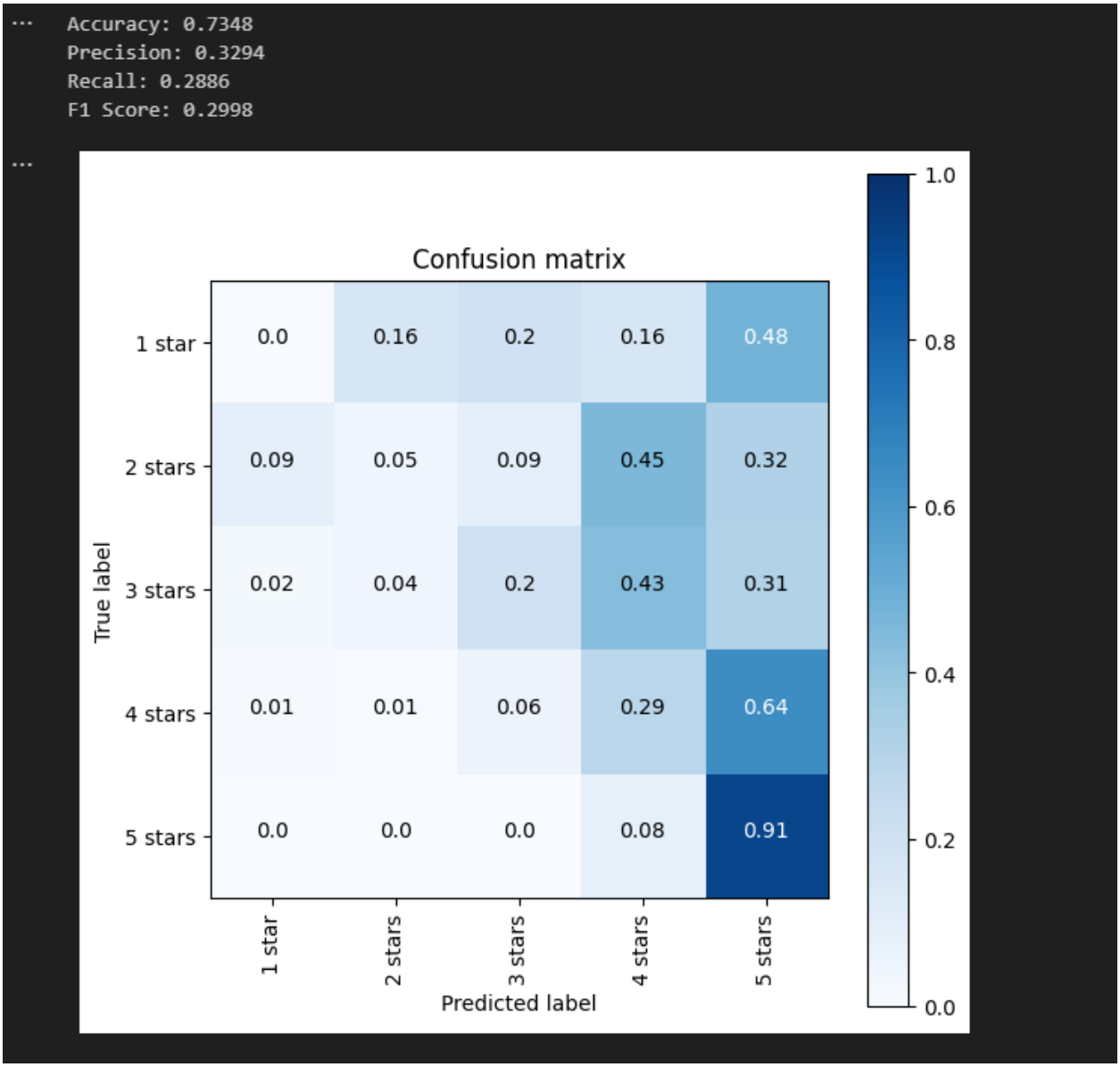

Figure 2: Performance metrics for the model trained on the Cursera reviews dataset (multiclass, 1 to 5 stars) using the SimpleNN model.