Homework 2: Aspect-Based Sentiment Analysis

Michele Conti
conti.1599133@studenti.uniroma1.it

6 Figures and tables

All figures and tables have been intentionally placed at the end of the document, in this section, so that they don't affect the maximum limit of three pages for the text.
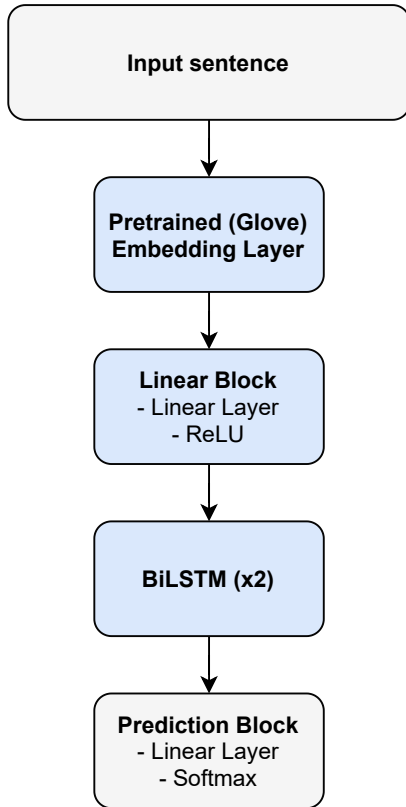

Figure 1: Baseline model (M0) architecture.

Figure 2: Final model (M3) architecture.

Sentiment	Train	Dev
positive	2605	546
negative	1364	307
neutral	877	216
conflict	111	25

Table 1: Polarities support for targets (task A+B) in the training and development set (i.e., Restaurants + Laptops).
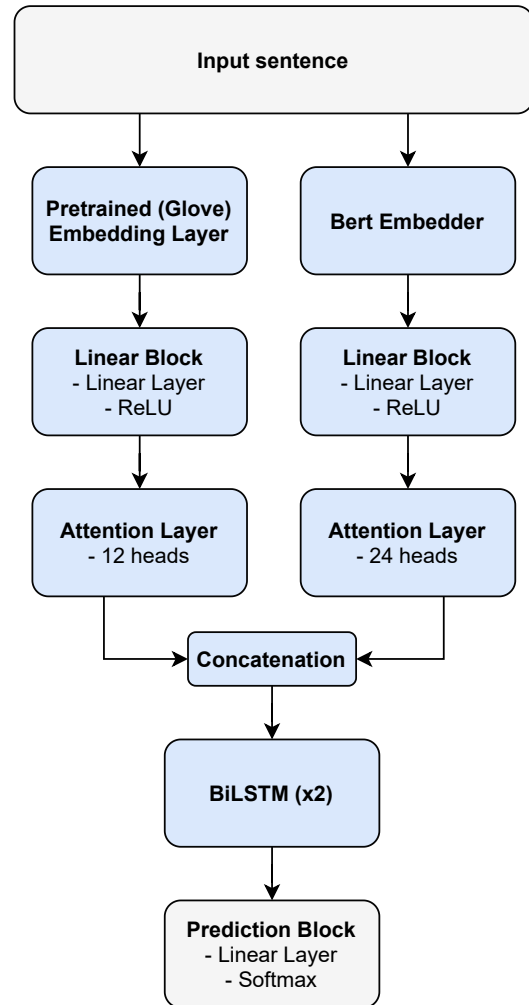
Category	Train	Dev
anecdotes/miscellaneous	941	191
price	268	53
food	1008	224
ambience	355	76
service	478	119

Table 2: Categories support (task C+D) in the training and development set (i.e., Restaurants).

Sentiment	Train	Dev
positive	1803	376
negative	672	167
neutral	411	89
conflict	164	31

Table 3: Polarities support for categories (task C+D) in the training and development set (i.e., Restaurants).

Figure 3: Task A+B: Baseline (M0) loss on the training and development set.
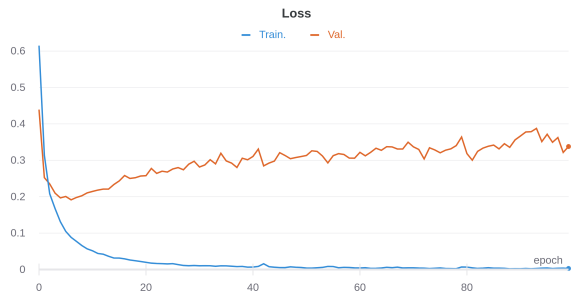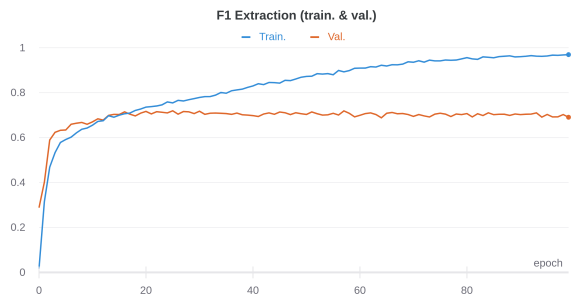

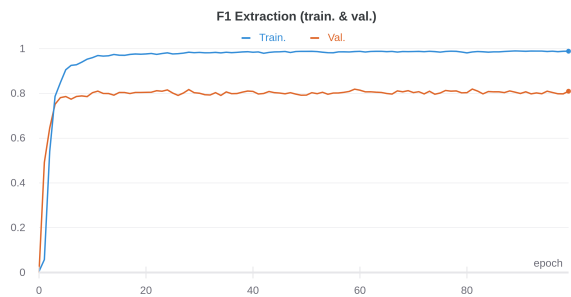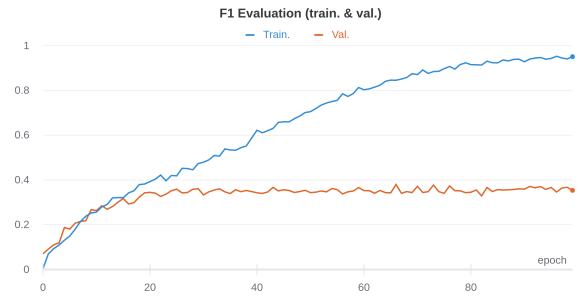Figure 7: Task A+B: Baseline (M0) macro f1 evaluation score on the training and development set.

Figure 4: Task A+B: Final model (M3) loss on the training and development set.

Figure 8: Task A+B: Final model (M3) macro f1 evaluation score on the training and development set.

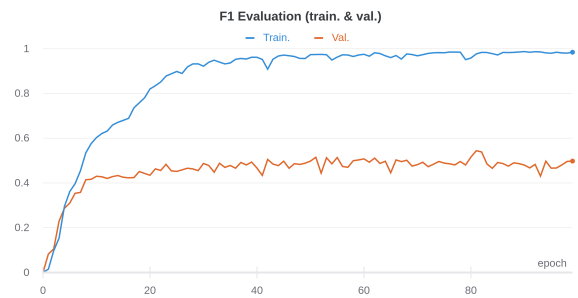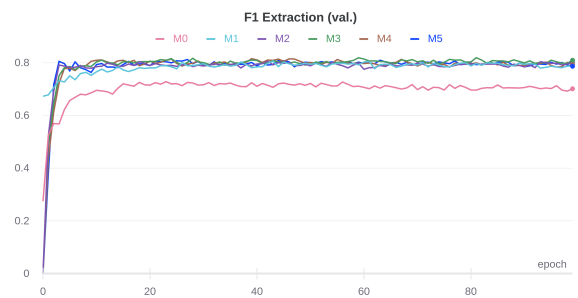Figure 5: Task A+B: Baseline (M0) macro f1 extraction score on the training and development set.

Figure 9: Task A+B: Comparative plot of f1 extraction score for the 6 model variations (M0, M1, M2, M3, M4, M5).

Figure 6: Task A+B: Final model (M3) macro f1 extraction score on the training and development set.
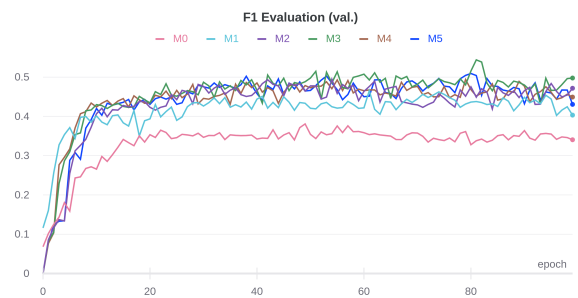

Figure 10: Task A+B: Comparative plot of f1 evaluation score for the 6 model variations (M0, M1, M2, M3, M4, M5).
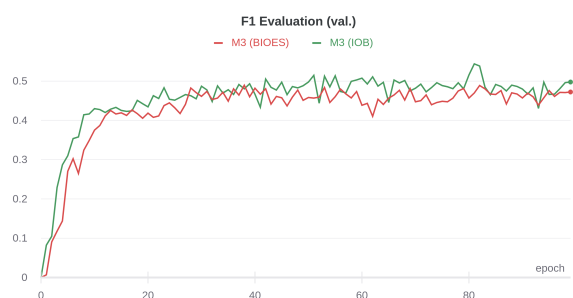
Figure 11: Task A+B: Comparative plot of f1 extraction score for the best model (M3) using two different tagging schemas (IOB, BIOES).
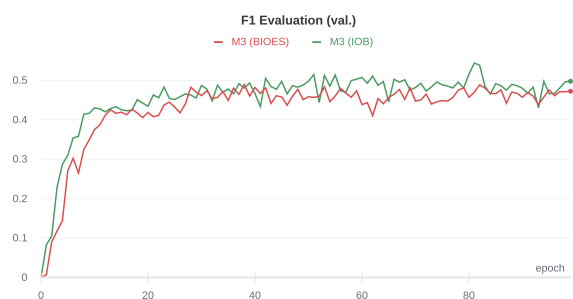

Figure 12: Task A+B: Comparative plot of f1 evaluation score for the best model (M3) using two different tagging schemas (IOB, BIOES).
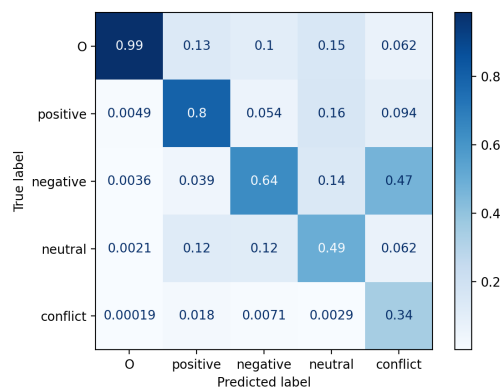

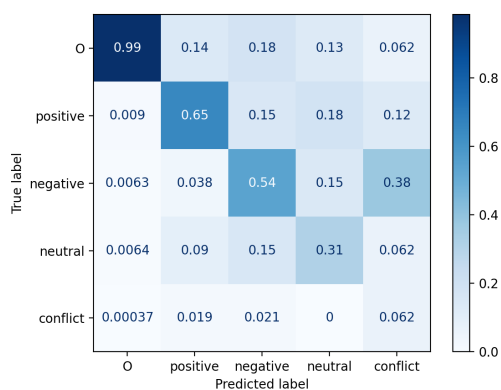Figure 14: Task A+B: Normalized confusion matrix of the best model (M3).

Figure 13: Task A+B: Normalized confusion matrix of the baseline model (M0).