Q1
Yes
The code is executed with no error.

Q2
Well documented
The code is well documented.

Q3
No

Q4
Yes
The part1 code contains encoding the label, tokenization, training and evaluating the model.
The part2 code contains loading the dataset, initializing, training and evaluating the model.

Q5
No
No bugs.

Q6
The code structure is well organized, the coding is elegant, and they use "assert" to restrict the input, which is a great idea, and they create functions to make the structure straightforward and logical. The way they load the dataframe in part 2 is interesting and instructive.

Q7
They did a great job and nothing need to improve in coding except that they hand in the report in .doc form ☹

Q8
Yes
They trained three models.

Q9
The model trained with 3,000 examples performed the best on the evaluation set. They did not further discuss why this model performs best in the report besides the increased quantity of the dataset.

Q10
Yes
They use bert-base-uncased model,   Pearson Correlation on train set is 98.92%, on test set is 73.61%.

Q11
Yes, they mentioned all questions in green.

Q12

The lab report is well organized and the language is straightforward and logical. They answered the questions with their own thoughts.

Q13
They can resave the report in pdf format, and for the second question, they can discuss further.

Q14
No

Q15
1
Overall, the code structure, the coding, the models' performance, and the report are great and deserve a full mark.