Assignment 1

1. **Top 5 bigrams before smoothing and after each of the 2 selected smoothing techniques along with their probabilities BEFORE applying emotion component.**
    - **Top 5 bigrams based on probability:**

```
Top 5 Bigrams with Probabilities:
('href', 'http'): 1.0
('tychelle', 'to'): 1.0
('hang', 'out'): 1.0
('nonexistent', 'social'): 1.0
('alex', 'and'): 1.0
```

    - **Top 5 bigrams based on probability with laplace smoothing:**

```
Top 5 Bigrams with Probabilities:
('i', 'feel'): 0.11049603820688159
('feel', 'like'): 0.0350976507217662
('i', 'am'): 0.03191142950179095
('that', 'i'): 0.02655721873491067
('and', 'i'): 0.02311382468382032
```

    - **Top 5 bigrams based on probability with kneser-ney smoothing:**

```
Top 5 Bigrams with Probabilities:
('href', 'http'): 0.9680013095432968
('i', 'feel'): 0.2702267942802305
('here', 'i'): 0.15810056697272282
('link', 'href'): 0.10006547716483874
('count', 'link'): 0.05013095432967751
```

2. **Reasoning for method used for including emotion component:**

The function calculate_all_bigram_probabilities_with_emotions, computes bigram probabilities for a given corpus along with inclusion of an emotional component for each word.
In the bigram_probability_with_emotion method, a base probability for the bigram is determined using the underlying bigram model.
To enhance this probability with emotional context, the method extracts emotion scores for each word using the get_emotion_score function, which uses the utils.py file.
The emotion scores for both words of a bigram are then averaged and added to the base probability. So basically for each bigram, we generate an array as a value where there are the sum of averages for each emotion and the base probability.
By averaging and blending emotion scores into the probability, the model makes sure to grasp how emotions may influence the likelihood of a specific bigram occurrence.

3. **2 generated samples for each emotion, for a total of 12 generated samples**

```
Samples for sadness:
i forget to suffer i hurt when ever make
i forget for losing again and suffer from new limitations which dying
Samples for joy:
i remember moments you feel faithful feel gorgeous colors
i perform a smart consumer too cheap so content and
Samples for love:
i love caring people caring people caring eh
im supporting walmart because again its warmth sense for
Samples for anger:
i write it things here that as usual at
i stared up between my bum are heading for and was
Samples for fear:
i could almost weird considering there if id just surreal confusion and
i could almost weird considering everything and suddenly startled me
Samples for surprise:
i wonder why ive gone and curious why should reiterate
i wonder are surprised us no resemblance to
```

4. **Accuracy and macro F1 scores obtained from extrinsic evaluation**

```
Fitting 5 folds for each of 60 candidates, totalling 300 fits
Model saved as best_model.pkl
Classification Report for {'svc__C': [1, 10, 100, 1000, 5000],
              precision    recall  f1-score   support

       anger       0.57      0.30      0.35        50
        fear       0.79      0.52      0.63        50
         joy       0.58      0.76      0.66        50
        love       0.81      1.00      0.88        50
     sadness       0.63      0.84      0.72        50
    surprise       0.88      0.90      0.89        50

    accuracy                           0.71       300
   macro avg       0.71      0.71      0.69       300
weighted avg       0.71      0.71      0.69       300
```

5. **For each emotion, pick 1 of the generated sample and reason why it is generated according to its corresponding emotion.**

   Emotion: sadness
   Sample: i forget for losing again and suffer from new limitations which dying
   Reason: words like ‘losing’, ‘suffer’, ‘limitation’ and ‘dying’ show sadness

   Emotion: joy
   Sample: i remember moments you feel faithful feel gorgeous colors
   Reason: words like ‘faithful’, ‘moments’ and ‘gorgeous’ show joy

   Emotion: love
   Sample: i love caring people caring people caring
   Reason: words like ‘caring’ and ‘love’ show love

Emotion: anger
Sample: i stared up between my bum are heading for and was
Reason: word like ‘stared’ show anger

Emotion: fear
Sample: i could almost weird considering everything and suddenly startled me
Reason: words like ‘weird and ‘suddenly’ ‘startled’ show fear

Emotion: surprise
Sample: i wonder are surprised us no resemblance to
Reason: words like ‘wonder and ‘surprised’ show surprise

6. **A credit statement reflecting the contribution of each member of the group for the assignment.**
   Task 1 - Kuber
   Task 1 (tokenisation) - Priya
   Task 2 (1st 3) - Gurmehak
   Task 2 (Extrinsic evaluation) - Yuvraj