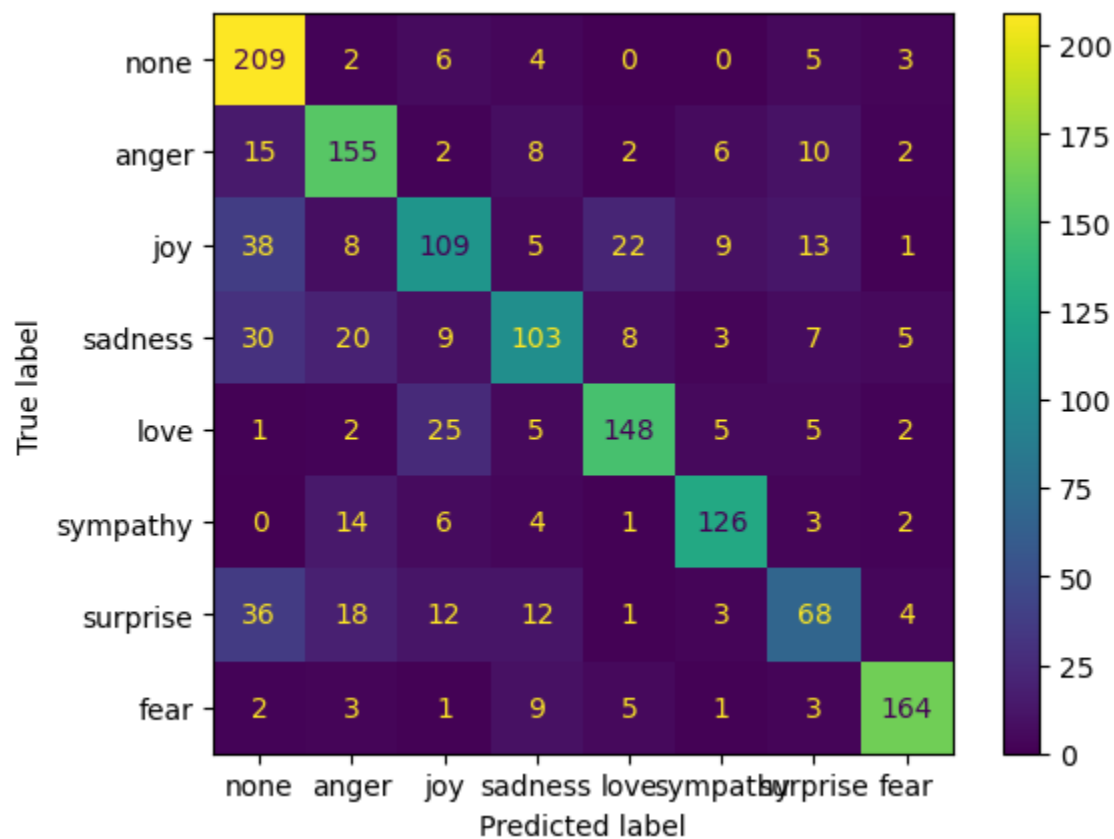
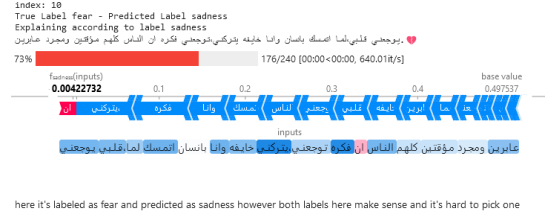
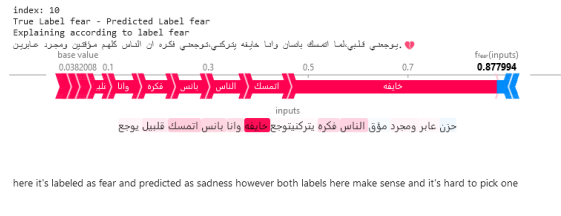


We will be using the light stemming model from now.

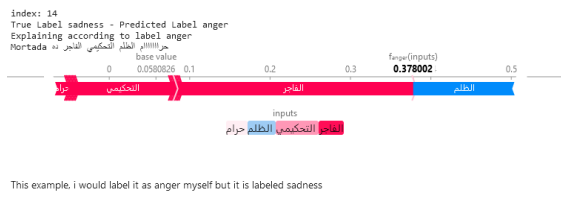
	precision	recall	f1-score	support
none	0.63	0.91	0.75	229
anger	0.70	0.78	0.73	200
joy	0.64	0.53	0.58	205
sadness	0.69	0.56	0.61	185
love	0.79	0.77	0.78	193
sympathy	0.82	0.81	0.82	156
surprise	0.60	0.44	0.51	154
fear	0.90	0.87	0.88	188
accuracy			0.72	1510
macro avg	0.72	0.71	0.71	1510
weighted avg	0.72	0.72	0.71	1510



Light Stemming vs Raw Data:



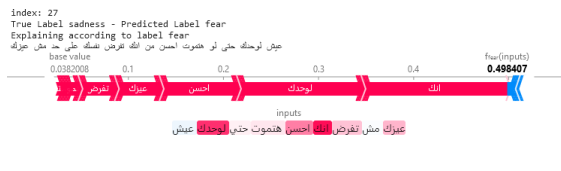
Here the model made the correct prediction with light stemming.



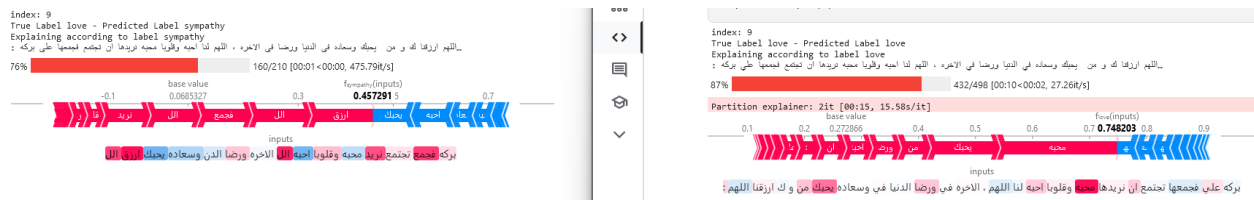
The model stopped using stopwords as explanations.



This was a case which I agreed with the prediction more than the label, but now the model got the label correctly. The explanation is bad too. ***(try the correct word)



This one just got more confusing.



The model got confused in this case, as light stemming loses some of the context, so the mistake is understandable.

Conclusion:

The explanations got better but some of the context was lost.

Shap vs Lime

Interruption:

Some errors were found in light stemming during the comparison of lime and shap.

Raw data:

💔. يوجعني قلبي، لما اتمسك بانسان وانا خايفه **يتركني، توجعني** فكره ان الناس كلهم مؤقتين ومجرد عابرين

Light Stemming:

يُوجع قَلِيل اِتمسك بانس وانا خايفه **يتركنيتوجع** فكره الناس مؤق ومجرد عابر حزن



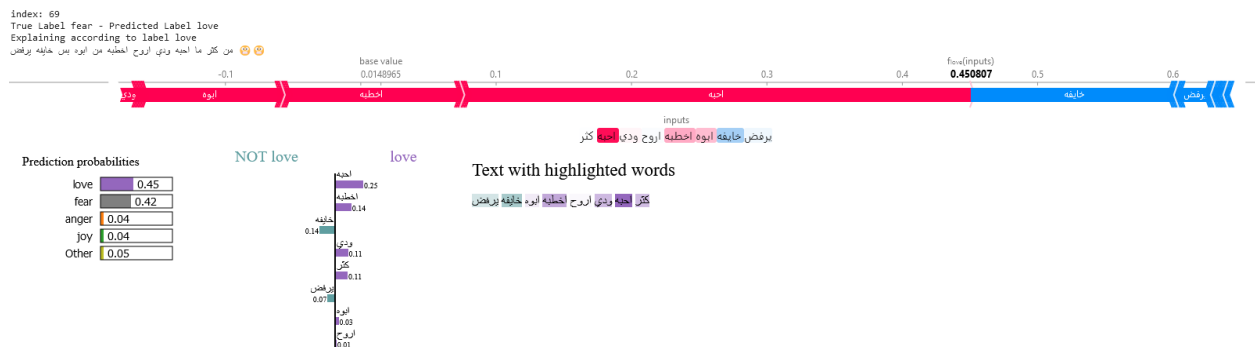
While the main contributor to the label is the same, the other words are assigned different importance in different explainers.



The same words contribute to different labels according to different explainers.



Here, both have mostly the same explanation, but the words are assigned different weights.



Here, LIME gave better insight into the workings of the model by providing the probabilities. We can see that even the model understands that there are mixed feelings.

Interruption (again):

Some errors were found in light stemming during the comparison of lime and shap.

Raw data:

دول **افريفا** مبوزه سمعتنا في الاولمبياد

Light Stemming:

افريف مبوزه سمعت الاولمبياد

There is a spelling mistake here, and some removed stopwords that need to be looked into.

And another one:

Raw data:

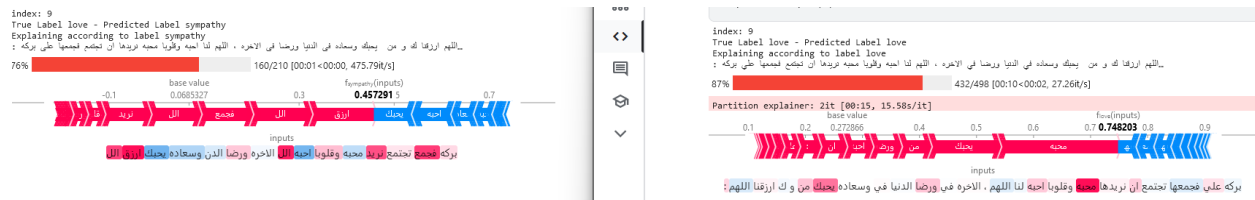
محدث عارف هما تعبوا اد ايه عشان يوصلوا الاولمبياد و هم مش ناقصين تعليقات الجهله في اللعبه لادائهم و لا تعليقات **المصريين**
عموما علي لبسهم

Light Stemming:

محدث عارف تعب اد عشان يوصل الاولمبياد مش ناقص تعليق الجهله اللعبه لادائ تعليق **المصر** **عمو** لبس

There are some removed stopwords that need to be looked into.

Using Cleaned data without stemming:



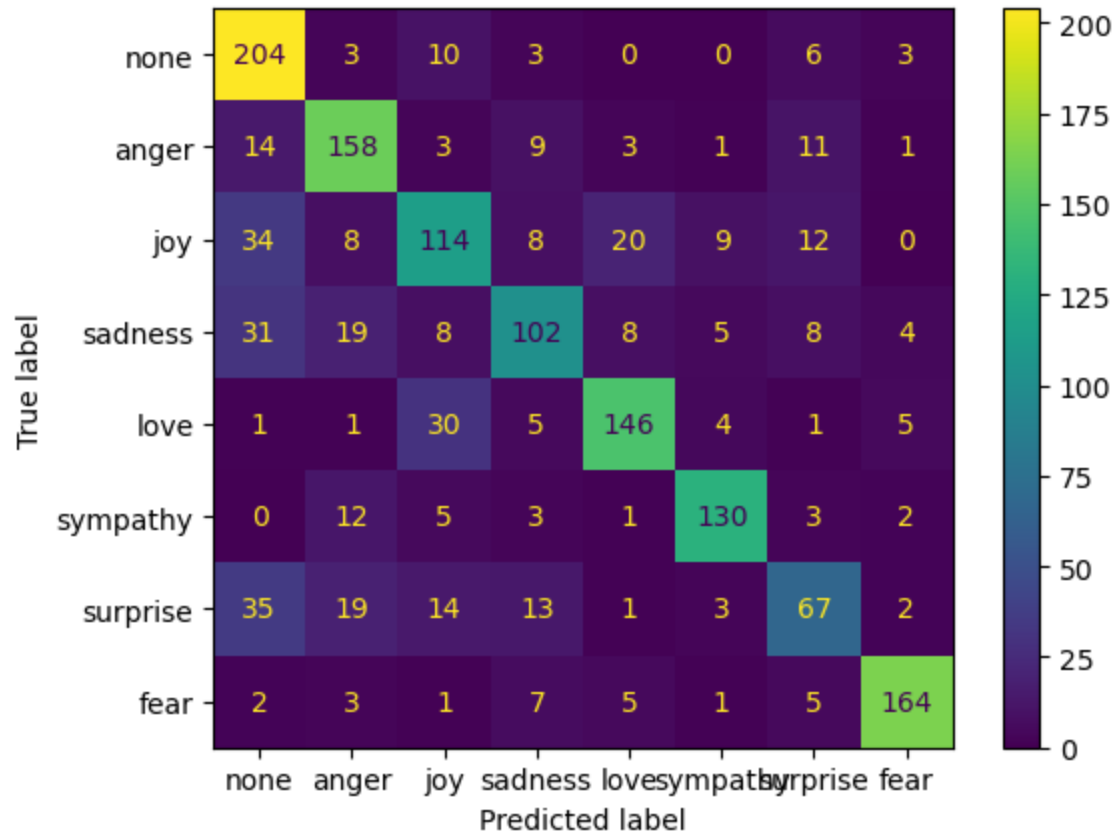
There was this error that I attributed to a loss of context but without stemming, the model got it correctly once more:



So, context can be lost with stemming but it appears that without stemming, it may get better.

Since we will be using cleaned data now, here are the metrics:

	precision	recall	f1-score	support
none	0.64	0.89	0.74	229
anger	0.71	0.79	0.75	200
joy	0.62	0.56	0.58	205
sadness	0.68	0.55	0.61	185
love	0.79	0.76	0.77	193
sympathy	0.85	0.83	0.84	156
surprise	0.59	0.44	0.50	154
fear	0.91	0.87	0.89	188
accuracy			0.72	1510
macro avg	0.72	0.71	0.71	1510
weighted avg	0.72	0.72	0.71	1510

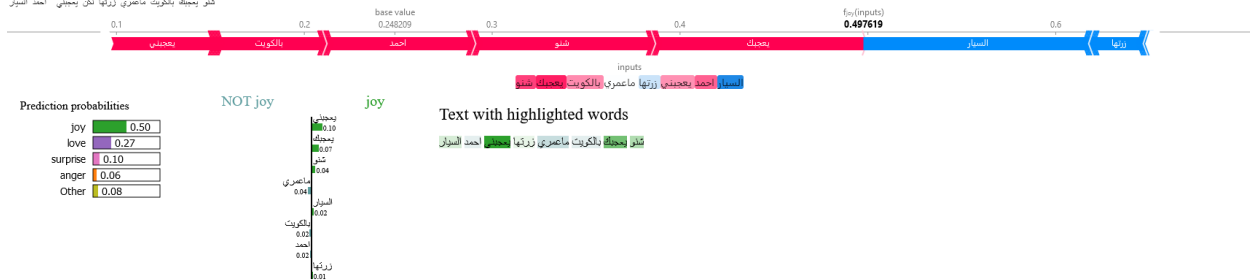


By comparing the confusion matrix with the previous one, we can observe that the box that indicates (true = none, pred = joy) increased from 6 to 10.

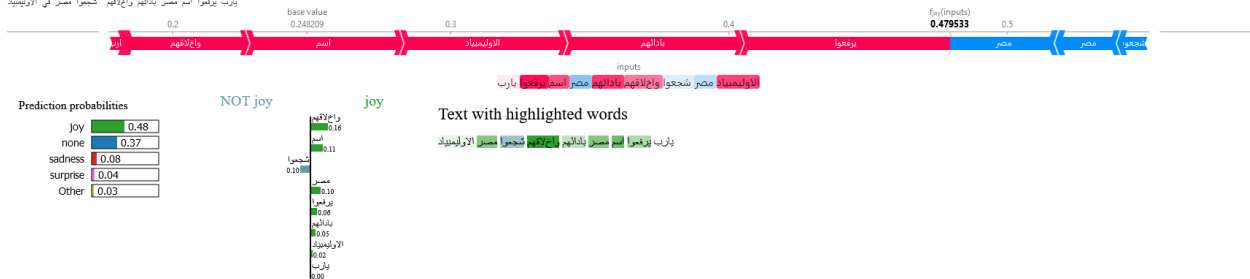
That was despite the fact that context appears to be more complete without stemming so let's examine those 10 samples.



index: 180
True Label none - Predicted Label joy
Explaining according to label joy
خبر يمحيط بالكرت ماعمرى زرتها نكن يمحيطي احمد البوار



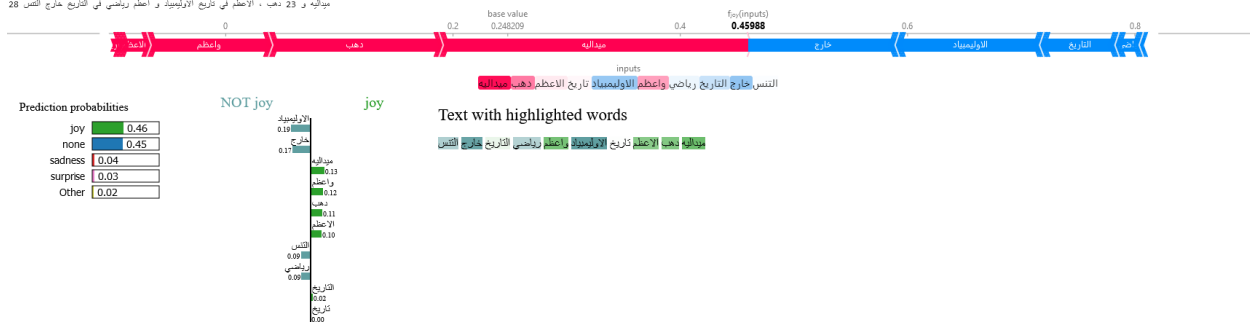
index: 325
True Label none - Predicted Label joy
Explaining according to label joy
بارب برقمنا اسم مضمير بادانهم واجانهم شجروا بارت



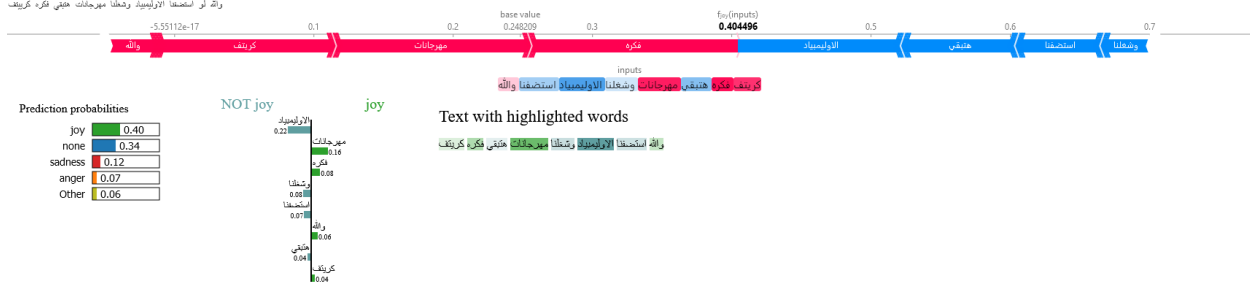
index: 891
True Label none - Predicted Label joy
Explaining according to label joy
احداثات الاوليمبياد 1 احداثات قاع السويس 0



index: 815
True Label none - Predicted Label joy
Explaining according to label joy
ميدالية و 23 ذهب - الاعظم في تاريخ الاوليمبياد و اعظم رياضي في التاريخ خارج التنس



index: 874
True Label none - Predicted Label joy
Explaining according to label joy
واله لو استحق الاوليمبياد وحظا مبرحانات هاني فكره كرتك



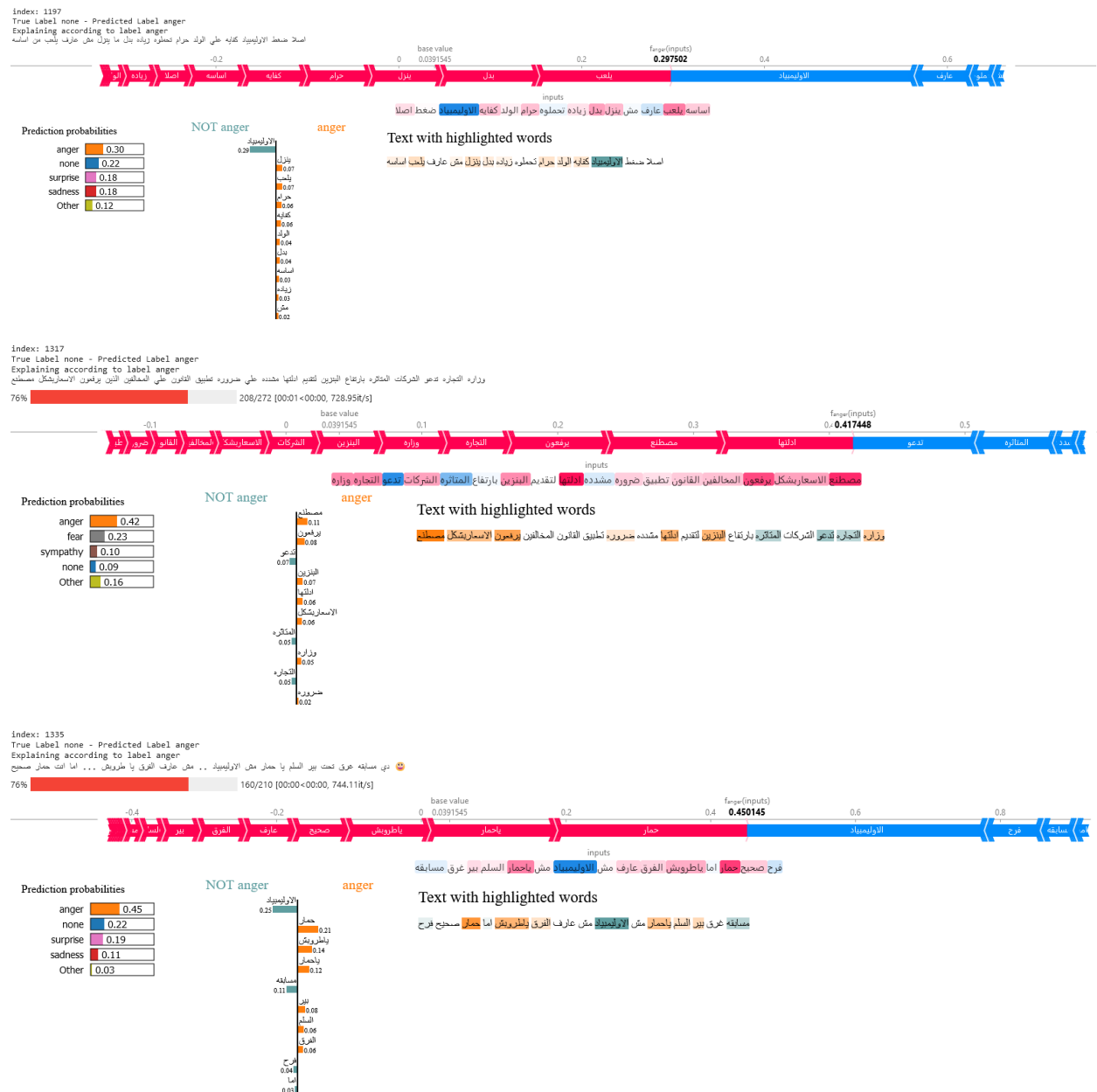


Some examples are understandable why the model predicted them as joy, while others are just confusing to me and the whole context did not help either.

The just context made the model more confused.

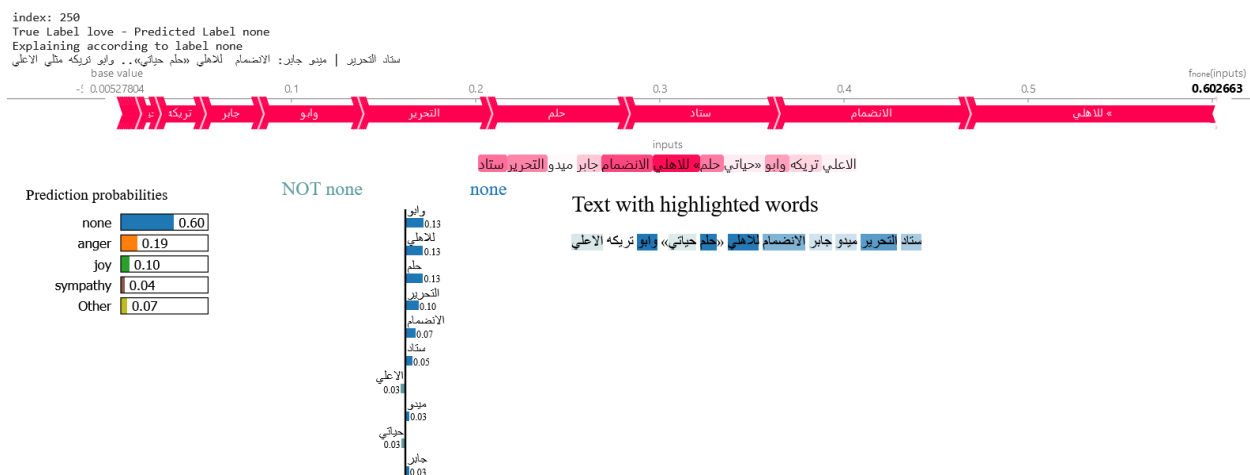
So, let's examine something a bit less confusing.

(true = none, pred = anger, total error = 3)



The first and last example are understandable why the model would label them as anger, and I am inclined to agree with its explanation. Especially the last one, as 3 curse words in one sentence is a convincing argument for why it is predicted anger.

(true = love, pred = none, total error = 1)



Most of the love samples in the dataset are the romantic kind of love, so the problem here is a lack of data in this very specific area.

(true = love, pred = joy, total error = 30)

index: 203

True Label love - Predicted Label joy

Original: احلام العنزي فخر الكويت الف الف مبرووووك بنت قبيلتي وبنت ديرتبيبيي
الديره كل يوم اسم رافع اسمنا بشاره خير ان شاء الله

Cleaned: احلام العنزي فخر الكويت الف الف مبروك بنت قبيلتي وبنت ديرتي كفو عيال الديره يوم اسم رافع اسمنا بشاره خير شاء الله

I would say this one is mixed feelings.

index: 594

True Label love - Predicted Label joy

Original: : الدوله العثمانيه الاسلاميه ستبقي رغم انوف الاعداء ربي احفظ تركيا المسلمه من مكر الماكرين واحفظ كل يلاذ
المسلمين

Cleaned: الدوله العثمانيه الاسلاميه ستبقي رغم انوف الاعداء ربي احفظ تركيا المسلمه مكر الماكرين واحفظ يلاذ المسلمين

This one supports my earlier argument about romantic love and love for one's country.

index: 823

True Label love - Predicted Label joy

Original: لون حياتك بحب وطاعه الله فهي طويله قصيره ... ايمان تقوه

Cleaned: لون حياتك بحب وطاعه الله طويله قصيره ايمان تقوه

Same as the above.

Conclusion on True = love:

It appears that most of them are predicted wrongly due to a few reasons:

1. The difference between romantic and non-romantic love
2. Mixed feelings
3. Misunderstood context
4. Generally confusing such as the following example:

index: 1430

True Label love - Predicted Label joy

Original: الهلال الساعه 7:30 بالتوفيق شباب الاتحاد & الاربعاء ديربي.. اتحاد الشرس

Cleaned: الاربعاء ديربي اتحاد الشرس الهلال الساعه بالتوفيق شباب الاتحاد

(true = surprise, pred = none, total error = 35)

All of them are due to "الاولمبياد"

ALL OF THEM!!!