Over-sampling

SMOTE **×** (val\_acc only 0.0095)

SMOTEENN **×** (val\_acc only 0.0050)

Under-sampling

TomekLinks **O** (val\_acc: 0.4939 epoch4)

Cluster Centroids **O** (val\_acc: 0.4939 epoch4)

Assign class -weight

Why Conv layer can speed up training time?

Conv layer is helpful because it reduces the features via filters defined. Therefore, we do not need to use all of the high dimensional features but the filtered ones that speed the training up.

Is oversampling different from class weighting?

Data-oversampling and class weighting are equivalent. Copying the samples of a class 3X is equivalent to assigning a 3X weight to the class. However, the weighting is better from storage and computational point of view since it avoids working with a larger data-set.

What is the difference between Keras' MaxPooling1D and GlobalMaxPooling1D functions?

<https://stackoverflow.com/questions/43728235/what-is-the-difference-between-keras-maxpooling1d-and-globalmaxpooling1d-functi>

What is the difference between Keras' recurrent dropout & dropout?

<https://stackoverflow.com/questions/44924690/keras-the-difference-between-lstm-dropout-and-lstm-recurrent-dropout>

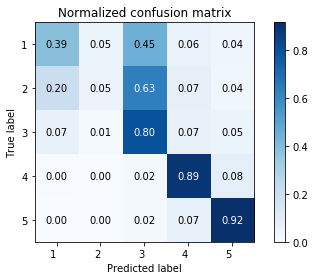
What is the difference between LSTM’s dropout input parameter & external dropout layer?

<https://stackoverflow.com/questions/50720670/using-dropout-with-keras-and-lstm-gru-cell>

Embedding trainable = false

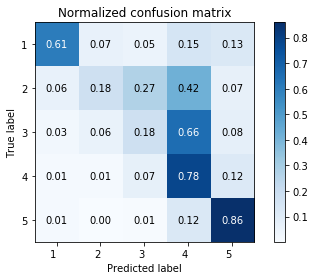
Ratings 1,2,3= cons; rating 4 = combine, ratings 5 = pros, Epoch = 10, normal (val highest 78% !!!)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.42 | 0.39 | 0.40 | 793 |
| 2 | 0.43 | 0.05 | 0.09 | 1075 |
| 3 | 0.63 | 0.80 | 0.70 | 2537 |
| 4 | 0.86 | 0.89 | 0.88 | 4423 |
| 5 | 0.88 | 0.92 | 0.90 | 4678 |
| Micro avg | 0.79 | 0.79 | 0.79 | 13506 |
| Macro avg | 0.64 | 0.61 | 0.60 | 13506 |
| Weighted avg | 0.77 | 0.79 | 0.76 | 13506 |



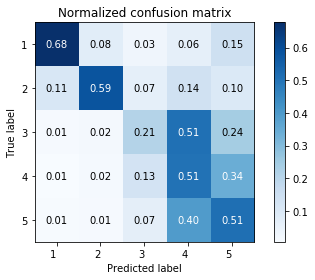
Ratings 1= cons; rating 2,3,4 = combine, ratings 5 = pros, Epoch = 10, normal (val 64%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.66 | 0.61 | 0.63 | 793 |
| 2 | 0.40 | 0.18 | 0.24 | 1075 |
| 3 | 0.40 | 0.18 | 0.25 | 2537 |
| 4 | 0.55 | 0.78 | 0.65 | 4423 |
| 5 | 0.82 | 0.86 | 0.84 | 4678 |
| Micro avg | 0.64 | 0.64 | 0.64 | 13506 |
| Macro avg | 0.57 | 0.52 | 0.52 | 13506 |
| Weighted avg | 0.61 | 0.64 | 0.61 | 13506 |



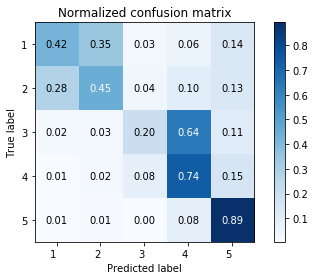
Ratings 1= cons; rating 2 = combine, ratings 3,4,5 = pros, Epoch = 10, normal (val 47%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.70 | 0.68 | 0.69 | 793 |
| 2 | 0.73 | 0.59 | 0.65 | 1075 |
| 3 | 0.35 | 0.21 | 0.26 | 2537 |
| 4 | 0.40 | 0.51 | 0.45 | 4423 |
| 5 | 0.50 | 0.51 | 0.51 | 4678 |
| Micro avg | 0.47 | 0.47 | 0.47 | 13506 |
| Macro avg | 0.54 | 0.50 | 0.51 | 13506 |
| Weighted avg | 0.47 | 0.47 | 0.46 | 13506 |



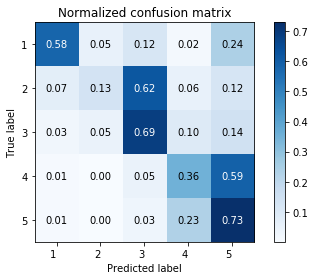
Ratings 1,2 = cons; rating 3,4 = combine, ratings 5 = pros, Epoch = 10, normal (val 65%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.45 | 0.42 | 0.44 | 793 |
| 2 | 0.49 | 0.45 | 0.47 | 1075 |
| 3 | 0.54 | 0.20 | 0.29 | 2537 |
| 4 | 0.60 | 0.74 | 0.66 | 4423 |
| 5 | 0.78 | 0.89 | 0.83 | 4678 |
| Micro avg | 0.65 | 0.65 | 0.65 | 13506 |
| Macro avg | 0.57 | 0.54 | 0.54 | 13506 |
| Weighted avg | 0.63 | 0.65 | 0.62 | 13506 |



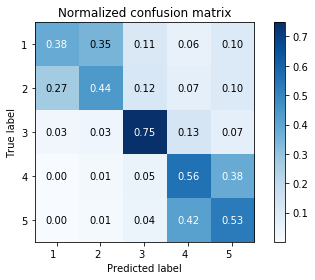
Ratings 1 = cons; rating 2,3 = combine, ratings 4,5 = pros, Epoch = 10, normal (val 54%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.69 | 0.58 | 0.63 | 793 |
| 2 | 0.46 | 0.13 | 0.21 | 1075 |
| 3 | 0.61 | 0.69 | 0.65 | 2537 |
| 4 | 0.53 | 0.36 | 0.43 | 4423 |
| 5 | 0.51 | 0.73 | 0.60 | 4678 |
| Micro avg | 0.54 | 0.54 | 0.54 | 13506 |
| Macro avg | 0.56 | 0.50 | 0.50 | 13506 |
| Weighted avg | 0.54 | 0.54 | 0.52 | 13506 |



Ratings 1,2 = cons; rating 3 = combine, ratings 4,5 = pros, Epoch = 10, normal (val 56%)

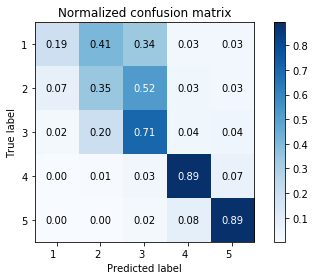
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.45 | 0.38 | 0.41 | 793 |
| 2 | 0.50 | 0.44 | 0.47 | 1075 |
| 3 | 0.76 | 0.75 | 0.75 | 2537 |
| 4 | 0.51 | 0.56 | 0.53 | 4423 |
| 5 | 0.55 | 0.53 | 0.54 | 4678 |
| Micro avg | 0.56 | 0.56 | 0.56 | 13506 |
| Macro avg | 0.55 | 0.53 | 0.54 | 13506 |
| Weighted avg | 0.56 | 0.56 | 0.56 | 13506 |



Embedding trainable = true

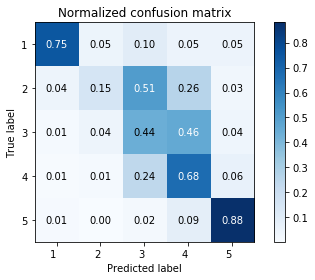
Ratings 1,2,3= cons; rating 4 = combine, ratings 5 = pros, Epoch = 10, normal (val 81%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.48 | 0.19 | 0.28 | 793 |
| 2 | 0.30 | 0.35 | 0.32 | 1075 |
| 3 | 0.63 | 0.71 | 0.66 | 2537 |
| 4 | 0.88 | 0.89 | 0.89 | 4423 |
| 5 | 0.91 | 0.89 | 0.90 | 4678 |
| Micro avg | 0.77 | 0.77 | 0.77 | 13506 |
| Macro avg | 0.64 | 0.61 | 0.61 | 13506 |
| Weighted avg | 0.77 | 0.77 | 0.77 | 13506 |



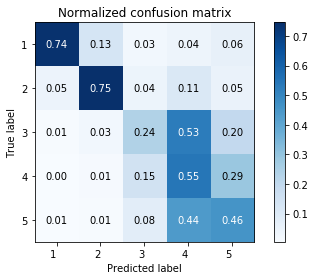
Ratings 1= cons; rating 2,3,4 = combine, ratings 5 = pros, Epoch = 10, normal (val 66%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.78 | 0.75 | 0.76 | 793 |
| 2 | 0.43 | 0.15 | 0.22 | 1075 |
| 3 | 0.38 | 0.44 | 0.41 | 2537 |
| 4 | 0.61 | 0.68 | 0.64 | 4423 |
| 5 | 0.90 | 0.88 | 0.89 | 4678 |
| Micro avg | 0.66 | 0.66 | 0.66 | 13506 |
| Macro avg | 0.62 | 0.58 | 0.59 | 13506 |
| Weighted avg | 0.66 | 0.66 | 0.66 | 13506 |



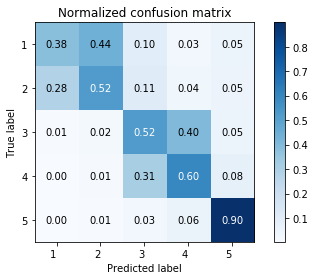
Ratings 1= cons; rating 2 = combine, ratings 3,4,5 = pros, Epoch = 10, normal (val 48%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.82 | 0.74 | 0.78 | 793 |
| 2 | 0.74 | 0.75 | 0.75 | 1075 |
| 3 | 0.36 | 0.24 | 0.29 | 2537 |
| 4 | 0.41 | 0.55 | 0.47 | 4423 |
| 5 | 0.53 | 0.46 | 0.49 | 4678 |
| Micro avg | 0.49 | 0.49 | 0.49 | 13506 |
| Macro avg | 0.57 | 0.55 | 0.55 | 13506 |
| Weighted avg | 0.49 | 0.49 | 0.48 | 13506 |



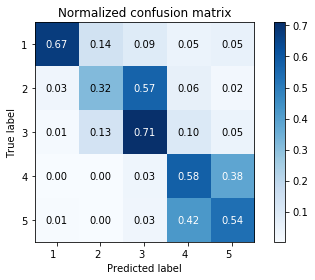
Ratings 1,2 = cons; rating 3,4 = combine, ratings 5 = pros, Epoch = 10, normal (val 67%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.46 | 0.38 | 0.42 | 793 |
| 2 | 0.53 | 0.52 | 0.53 | 1075 |
| 3 | 0.44 | 0.52 | 0.48 | 2537 |
| 4 | 0.66 | 0.60 | 0.63 | 4423 |
| 5 | 0.88 | 0.90 | 0.89 | 4678 |
| Micro avg | 0.67 | 0.67 | 0.67 | 13506 |
| Macro avg | 0.60 | 0.58 | 0.59 | 13506 |
| Weighted avg | 0.67 | 0.67 | 0.67 | 13506 |



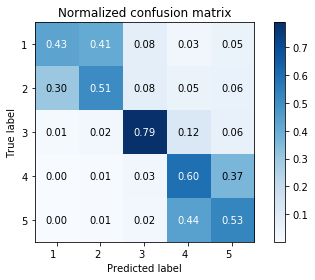
Ratings 1 = cons; rating 2,3 = combine, ratings 4,5 = pros, Epoch = 10, normal (val 57%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.84 | 0.67 | 0.75 | 793 |
| 2 | 0.42 | 0.32 | 0.36 | 1075 |
| 3 | 0.65 | 0.71 | 0.68 | 2537 |
| 4 | 0.52 | 0.58 | 0.55 | 4423 |
| 5 | 0.57 | 0.54 | 0.55 | 4678 |
| Micro avg | 0.57 | 0.57 | 0.57 | 13506 |
| Macro avg | 0.60 | 0.56 | 0.58 | 13506 |
| Weighted avg | 0.57 | 0.57 | 0.57 | 13506 |



Ratings 1,2 = cons; rating 3 = combine, ratings 4,5 = pros, Epoch = 10, normal (val 59%)

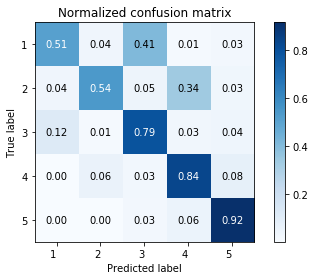
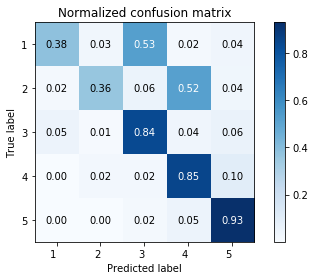
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.48 | 0.43 | 0.45 | 793 |
| 2 | 0.55 | 0.51 | 0.53 | 1075 |
| 3 | 0.85 | 0.79 | 0.82 | 2537 |
| 4 | 0.52 | 0.60 | 0.56 | 4423 |
| 5 | 0.57 | 0.53 | 0.55 | 4678 |
| Micro avg | 0.59 | 0.59 | 0.59 | 13506 |
| Macro avg | 0.59 | 0.57 | 0.58 | 13506 |
| Weighted avg | 0.60 | 0.59 | 0.59 | 13506 |



**Special case, trainable = false (because true positive rate of rating 2 is too low when pair together with 3, model keep predicting rating 2’s review to be rating 3)**

Ratings 1,3= cons; rating 2,4 = combine, ratings 5 = pros, Epoch = 10, normal (80~81% !!!)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.62 | 0.38 | 0.47 | 720 |
| 2 | 0.71 | 0.36 | 0.48 | 1058 |
| 3 | 0.77 | 0.84 | 0.80 | 2544 |
| 4 | 0.81 | 0.85 | 0.83 | 4530 |
| 5 | 0.87 | 0.93 | 0.90 | 4654 |
| Micro | 0.81 | 0.81 | 0.81 | 13506 |
| Macro | 0.76 | 0.67 | 0.70 | 13506 |
| Weighted | 0.80 | 0.81 | 0.80 | 13506 |

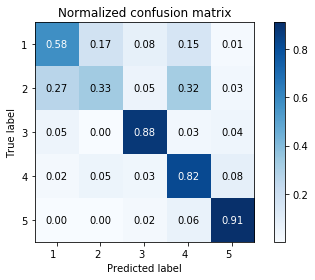
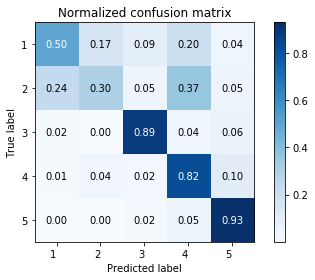


1. 0.8082
2. 0.8174
3. 0.8129
4. 0.7985
5. 0.8049
6. 0.8101
7. 0.7918
8. 0.8036
9. 0.8166
10. 0.8131 (看2nd matrix, 这个combination的1和2的true positive rate有时会高过它们偏向的rating的TPR)

Average = 0.8077

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.54 | 0.50 | 0.52 | 781 |
| 2 | 0.49 | 0.30 | 0.37 | 1049 |
| 3 | 0.87 | 0.89 | 0.88 | 2490 |
| 4 | 0.81 | 0.82 | 0.82 | 4493 |
| 5 | 0.86 | 0.93 | 0.90 | 4693 |
| Micro | 0.81 | 0.81 | 0.81 | 13506 |
| Macro | 0.72 | 0.69 | 0.70 | 13506 |
| Weighted | 0.80 | 0.81 | 0.80 | 13506 |

Ratings 3= cons; rating 1,2,4 = combine, ratings 5 = pros, Epoch = 10, normal (80~81% !!!) (应该不值得因为1和2两个一起分散掉4的TPR)

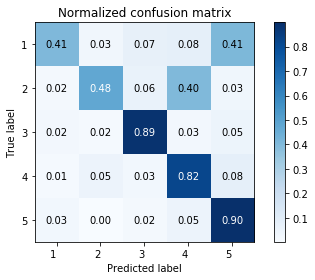
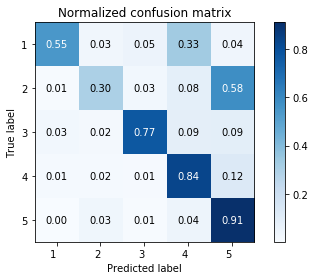


1. 0.8118
2. 0.8067
3. 0.7856
4. 0.8068
5. 0.8051
6. 0.8082
7. 0.8113
8. 0.7991
9. 0.8038
10. 0.8076 (看2nd matrix, 这个combination的1和2的true positive rate有时会高过它们偏向的rating的TPR)

Average = 0.8046

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.76 | 0.55 | 0.64 | 775 |
| 2 | 0.51 | 0.30 | 0.38 | 1060 |
| 3 | 0.93 | 0.77 | 0.84 | 2564 |
| 4 | 0.83 | 0.84 | 0.83 | 4499 |
| 5 | 0.74 | 0.91 | 0.82 | 4608 |
| Micro | 0.79 | 0.79 | 0.79 | 13506 |
| Macro | 0.75 | 0.67 | 0.70 | 13506 |
| Weighted | 0.79 | 0.79 | 0.78 | 13506 |

Ratings 3= cons; rating 2,4 = combine, ratings 1,5 = pros, Epoch = 10, normal (80~81% !!!)

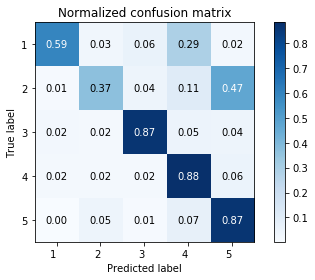
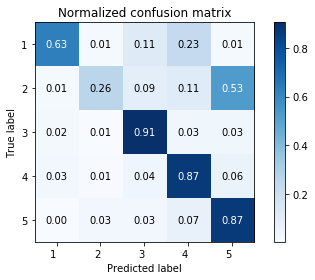


1. 0.8027
2. 0.8070
3. 0.8047
4. 0.8096
5. 0.8042
6. 0.7919
7. 0.8084
8. 0.7995
9. 0.8049
10. 0.8048 (看2nd matrix, 这个combination的1和2的true positive rate有时会高过/等于它们偏向的rating的TPR)

Average = 0.8038

Ratings 3= cons; rating 1,4 = combine, ratings 2,5 = pros, Epoch = 10, normal (80~81% !!!)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.71 | 0.63 | 0.67 | 777 |
| 2 | 0.60 | 0.26 | 0.36 | 1078 |
| 3 | 0.82 | 0.91 | 0.86 | 2559 |
| 4 | 0.85 | 0.87 | 0.86 | 4515 |
| 5 | 0.81 | 0.87 | 0.84 | 4577 |
| Micro | 0.81 | 0.81 | 0.81 | 13506 |
| Macro | 0.76 | 0.71 | 0.72 | 13506 |
| Weighted | 0.80 | 0.81 | 0.80 | 13506 |

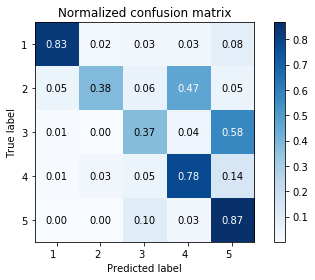


1. 0.8101
2. 0.8140
3. 0.8159
4. 0.8085
5. 0.8076
6. 0.8129
7. 0.8086
8. 0.8173 (看2nd matrix, 这个combination的1和2的true positive rate有时会高过它们偏向的rating的TPR)
9. 0.8079
10. 0.8122

Average = 0.8115

Ratings 1= cons; rating 2,4 = combine, ratings 3, 5 = pros, Epoch = 10, normal (79% !!!) （不值得因为1的rating很少，不值得去牺牲3的accuracy提高1的accuracy）

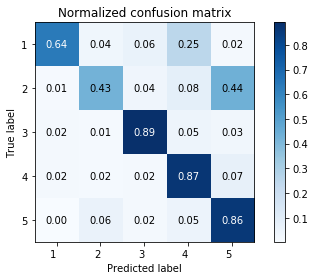
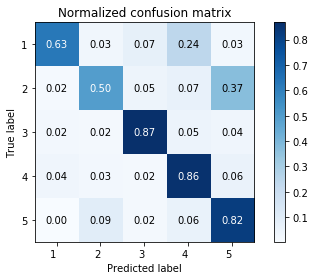
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.86 | 0.83 | 0.85 | 776 |
| 2 | 0.72 | 0.38 | 0.50 | 1087 |
| 3 | 0.56 | 0.37 | 0.45 | 2541 |
| 4 | 0.82 | 0.78 | 0.80 | 4439 |
| 5 | 0.65 | 0.87 | 0.74 | 4663 |
| Micro avg | 0.70 | 0.70 | 0.70 | 13506 |
| Macro avg | 0.72 | 0.65 | 0.67 | 13506 |
| Weighted avg | 0.70 | 0.70 | 0.69 | 13506 |



Embedding trainable = True

Ratings 3= cons; rating 1,4 = combine, ratings 2,5 = pros, Epoch = 10, normal (81~82% !!!)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| rating | precision | recall | F1 | support |
| 1 | 0.73 | 0.64 | 0.68 | 711 |
| 2 | 0.51 | 0.43 | 0.46 | 1002 |
| 3 | 0.89 | 0.89 | 0.89 | 2619 |
| 4 | 0.86 | 0.87 | 0.87 | 4563 |
| 5 | 0.82 | 0.86 | 0.84 | 4611 |
| Micro | 0.83 | 0.83 | 0.83 | 13506 |
| Macro | 0.76 | 0.74 | 0.75 | 13506 |
| Weighted | 0.82 | 0.83 | 0.82 | 13506 |

1. 0.8223
2. 0.8113
3. 0.8167
4. 0.8227
5. 0.8151
6. 0.8226
7. 0.8256
8. 0.8173 (看2nd matrix, 这个combination的1和2的true positive rate有时会高过它们偏向的rating的TPR)
9. 0.8196
10. 0.8213

Average = 0.8195