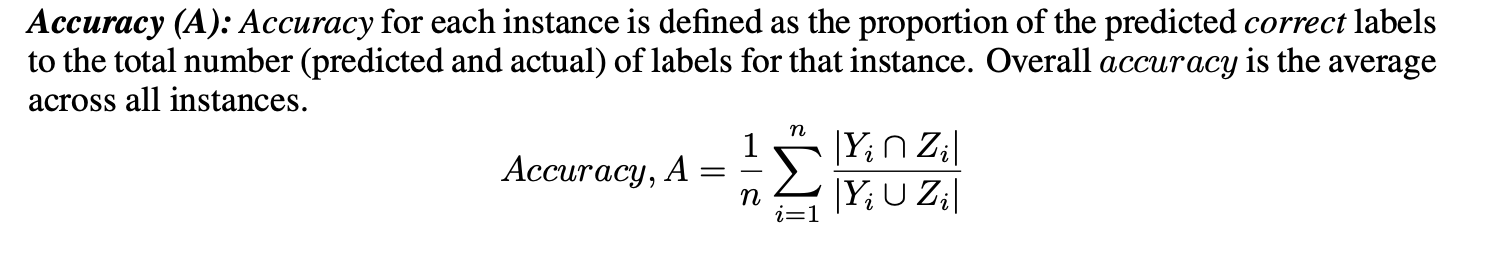
**Measuring Accuracy:**

Because we are working on a multi-label classification task, we cannot measure the accuracy of our model simply by using accuracy\_score from scikit-learn.

In order to calculate the accuracy of our model, we compared the sets of labels between true labels and predicted labels for each datapoint to compute its similarity and then took the average of the similarities among those data points.

We used this formula to compute the accuracy of our classification:



<https://www.researchgate.net/profile/Mohammad_Sorower/publication/266888594_A_Literature_Survey_on_Algorithms_for_Multi-label_Learning/links/58d1864392851cf4f8f4b72a/A-Literature-Survey-on-Algorithms-for-Multi-label-Learning.pdf>

We used this to compute the similarity between two sets of labels:

<https://www.geeksforgeeks.org/python-percentage-similarity-of-lists/>

**Result / Outputs:**

**Training Accuracy and Testing Accuracy:**

**Training accuracy: 99.94526546250684 %**

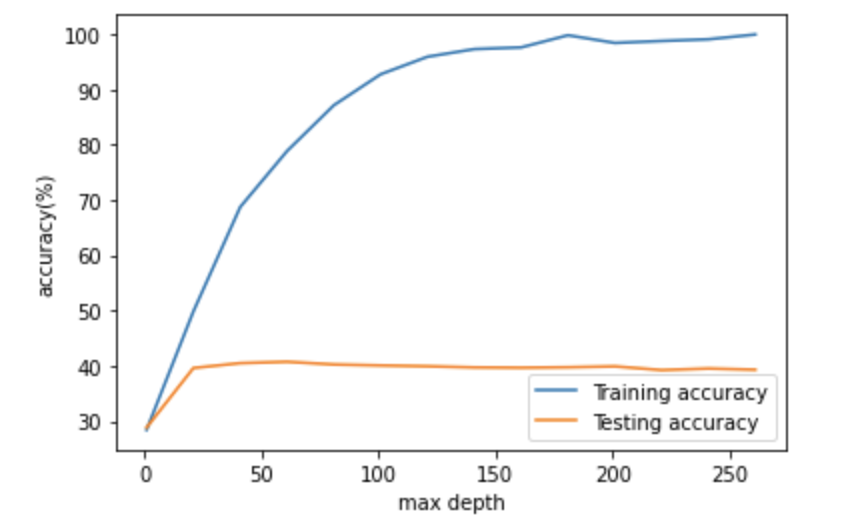
**Testing accuracy: 39.35805180166717 %**

**Comparison between true labels and predicted labels, and its similarity:**

***Table length=1827***

| ***idx*** | ***true\_label*** | ***predict*** | ***accuracy(%)*** |
| --- | --- | --- | --- |
| ***int64*** | ***object*** | ***object*** | ***float64*** |
| ***708*** | ***['animal', 'bad', 'bar', 'behavior', 'into', 'miscellaneous', 'partying', 'walks']*** | ***['animal', 'bad', 'bar', 'behavior', 'into', 'miscellaneous', 'partying', 'walks']*** | ***100.0*** |
| ***47*** | ***['animal', 'insults', 'lawyer', 'miscellaneous', 'work']*** | ***['dirty', 'doctor', 'lawyer', 'men', 'miscellaneous', 'women', 'work']*** | ***33.33333333333333*** |
| ***3995*** | ***['god', 'military', 'miscellaneous', 'news', 'police', 'politics']*** | ***['miscellaneous', 'nationality', 'news', 'politics', 'work']*** | ***37.5*** |
| ***1513*** | ***['blonde', 'dirty', 'men', 'miscellaneous', 'women']*** | ***['blonde', 'men', 'miscellaneous', 'women']*** | ***80.0*** |
| ***3686*** | ***['fat', 'good', 'insults', 'lookin', 'mama', 'miscellaneous', 'yo']*** | ***['fat', 'good', 'insults', 'lookin', 'mama', 'miscellaneous', 'yo']*** | ***100.0*** |
| ***2835*** | ***['dirty', 'doctor', 'men', 'miscellaneous', 'women']*** | ***['animal', 'miscellaneous']*** | ***16.666666666666664*** |
| ***2492*** | ***['celebrity', 'culture', 'dirty', 'lines', 'miscellaneous', 'pick', 'pop', 'up']*** | ***['celebrity', 'culture', 'dirty', 'lines', 'miscellaneous', 'pick', 'pop', 'up']*** | ***100.0*** |
| ***4192*** | ***['dirty', 'god', 'men', 'miscellaneous', 'women']*** | ***['gross', 'miscellaneous']*** | ***16.666666666666664*** |
| ***6865*** | ***['dirty', 'good', 'lookin', 'men', 'miscellaneous', 'money', 'women']*** | ***['dirty', 'marriage', 'men', 'miscellaneous', 'women']*** | ***50.0*** |
| ***6354*** | ***['athletes', 'men', 'miscellaneous', 'sports', 'women']*** | ***['miscellaneous', 'nationality']*** | ***16.666666666666664*** |
| ***...*** | ***...*** | ***...*** | ***...*** |
| ***6538*** | ***['good', 'lookin', 'miscellaneous']*** | ***['doctor', 'miscellaneous']*** | ***25.0*** |
| ***1533*** | ***['blue', 'celebrity', 'collar', 'culture', 'god', 'miscellaneous', 'news', 'politics', 'pop']*** | ***['athletes', 'bad', 'behavior', 'miscellaneous', 'nationality', 'news', 'partying', 'politics', 'sports']*** | ***20.0*** |
| ***2455*** | ***['dirty', 'marriage', 'men', 'miscellaneous', 'women']*** | ***['animal', 'marriage', 'men', 'miscellaneous', 'women']*** | ***66.66666666666666*** |
| ***3656*** | ***['fat', 'good', 'insults', 'lookin', 'mama', 'miscellaneous', 'yo']*** | ***['insults', 'mama', 'miscellaneous', 'yo']*** | ***57.14285714285714*** |
| ***5471*** | ***['athletes', 'kids', 'miscellaneous', 'nationality', 'sports']*** | ***['celebrity', 'culture', 'dark', 'humor', 'marriage', 'men', 'military', 'miscellaneous', 'money', 'police', 'pop', 'women']*** | ***6.25*** |
| ***4855*** | ***['good', 'insults', 'lookin', 'miscellaneous']*** | ***['miscellaneous', 'nationality', 'news', 'politics']*** | ***14.285714285714285*** |
| ***1880*** | ***['bad', 'behavior', 'dark', 'doctor', 'humor', 'marriage', 'men', 'military', 'miscellaneous', 'partying', 'police', 'women']*** | ***['marriage', 'men', 'military', 'miscellaneous', 'money', 'nationality', 'police', 'women']*** | ***42.857142857142854*** |
| ***8473*** | ***['car', 'miscellaneous', 'travel']*** | ***['athletes', 'miscellaneous', 'sports']*** | ***20.0*** |
| ***5967*** | ***['marriage', 'men', 'miscellaneous', 'women']*** | ***['dark', 'humor', 'kids', 'marriage', 'men', 'military', 'miscellaneous', 'police', 'women']*** | ***44.44444444444444*** |
| ***7971*** | ***['celebrity', 'culture', 'miscellaneous', 'pop', 'work']*** | ***['gross', 'miscellaneous']*** | ***16.666666666666664*** |

**Mean accuracy of classification against the max depth set for the Decision Tree:**



**Future Tasks:**

The testing accuracy for our model is very low while the training accuracy for our model goes up as the max depth set for the tree increases. One of the possible factors for this is overfitting. We will work on improvement of this model by reducing the likelihood of overfitting as much as possible.

To improve the accuracy:

* Will work on Pruning to avoid overfitting
* Could also try pretrain word vector (GloVe)