Moving Towards Open Set Incremental Learning: Readily Discovering New Authors

Justin Leo
Department of Computer Science
University of Colorado
jleo@uccs.edu

Jugal Kalita
Department of Computer Science
University of Colorado
jkalita@uccs.edu

Abstract

The classification of textual data often yields important information. Most classifiers work in a closed world setting where the classifier is trained on a known corpus, and then it is tested on unseen examples that belong to one of the classes seen during training. Despite the usefulness of this design, often there is a need to classify unseen examples that *do not belong* to any of the classes on which the classifier was trained. This paper describes the open set scenario where unseen examples from previously unseen classes are handled while testing. We examine a process of enhanced open set classification with a deep neural network that discovers new classes by clustering the examples identified as belonging to unknown classes, followed by a process of retraining the classifier with newly recognized classes Through this process we move to an incremental learning model where we continuously find and learn from novel classes of data that have been identified automatically. We also develop a new metric that measures multiple attributes of clustering open set data. Multiple experiments across two author attribution data sets show we are able to create an incremental model that produces excellent results.

Introduction

Formal as well as informal textual data are over-abundant in this Internet-connected era of democratized publishing and writing. These textual information sources are in multiple forms such as news articles, electronic books and social media posts. The use of text classification allows us to determine important information about the texts that can often be used to connect to the respective authors, naturally leading to the concept of Authorship Attribution. Authorship Attribution is seen as the process of accurately finding the author of a piece of text based on its stylistic characteristics (Rocha et al. 2016). Authorship Attribution is useful in scenarios such as identification of the author of malicious texts or the analysis of historical works with unknown authors.

Typically, text classification has a few well-established stages. The words in the text corpus are transformed using an embedding algorithm, and a classifier is trained with documents labeled with associated classes. In Authorship Attribution, the text samples tend to be books such as novels, transcribed speeches, or Internet-mediated social media posts, where each sample is labeled with the corresponding author. The trained text classifier is given testing data that is usually unseen text samples from the same set of trained authors. This process describes a closed set approach because the tested samples are associated with the same trained classes. A problem with this process of classification arises if the testing data includes samples from unfamiliar authors. In these cases, the classifier typically and erroneously associates the piece of text with a wrong author—an author on which it was trained. To remedy this problem, a new approach called open set classification has been proposed. Open set classification enables the classifier to discriminate among the known classes, but additionally and importantly, to identify if some test example is not associated with any of the classes on which it was trained (Scheirer et al. 2012).

There has been some recent work on open set classification using convolution neural networks (CNN) and recurrent neural networks (RNN). Prior work on open set classification has often been in areas such as computer vision (Bendale and Boult 2015), speech processing (Dahl et al. 2011), and natural language processing (Higashinaka et al. 2014). In this paper, we utilize open set recognition to identify the presence of test examples from novel classes, and incorporate these new classes to those already known to create an incremental class-learning model.

The rest of the paper is organized as follows. After describing related work in the the next section, we present our approach to identifying new classes and instantiating them. Then, we discuss our evaluation metrics for assessing incremental learning, followed by experimental results using authorship attribution datasets and analysis. We conclude by reiterating our accomplishments and thoughts on future work.

Related Work

We discuss related work in terms of four topics: deep networks for open set classification, metrics for open set classification, open set text classification, and recent proposals to use loss functions for open set classification in the context of computer vision.

Open Set Deep Networks

Using deep neural networks for open set classification often requires a change in the network model. Modern neural net-

works have multiple layers connected in various ways, depending on the classifier architecture being used. Most models eventually include a softmax layer that classifies the data to the known classes, with an associated confidence level or probability for each class. A test example is considered to belong to the class which has the highest probability among all the classes. To adapt this model to the open set scenario, the softmax layer was replaced by a unique layer named the OpenMax layer (Bendale and Boult 2016). This layer estimates the probability of an input being from one of the known classes as well as an "unknown" class, which lumps together all classes unseen during training. Thus, the network is able to recognize examples belonging to unknown classes, enhancing the ability of the closed set classifier it starts with.

Metric for Evaluating Open Set Classification

The process of open set class recognition leads to new challenges during the evaluation process. There are multiple sources of error that could be present including: misclassification of known or unknown classes and determination of novel classes. Bendale and Boult (2015) proposed a metric to evaluate how individual examples are classified. Although they proposed it for use in computer vision, we think it is applicable in author attribution as well.

Deep Open Set Text Classification

Prakhya, Venkataram, and Kalita (2017) modify the single OpenMax layer proposed by (Bendale and Boult 2016) to replace the softmax layer in a multi-layer convolution neural networks with an ensemble of several outlier detectors to obtain high accuracy scores for open set textual classification. The ensemble of classifiers uses a voting model between three different approaches: Mahalanobis Weibull, Local Outlier Factor (Kriegel et al. 2009), and Isolation Forest (Liu, Ting, and Zhou 2008). The average voting method produced results that are more accurate in detecting outliers, making detection of unknown classes better.

Loss Functions for Open Set Classification

A problem that often occurs in open set classification is the classifier labeling known class data as unknown. This problem typically occurs if there are some similar features in the examples of the pre-trained classes and unknown classes encountered during testing. In the context of computer vision, Dhamija, Günther, and Boult (2018) introduce what is called the Entropic Open-Set loss function that increases the entropy of the softmax scores for background training samples and improves the handling of background and unknown inputs. They introduce another loss function called the Objectosphere loss, which further increases softmax entropy and performance by reducing the vector magnitudes of examples of unknown classes in comparison with those from the known classes, lowering the erroneous classification of known class data as unknown. Since this approach squishes the magnitudes of all examples that belong to all unknown classes, it makes later separation of individual unknown classes difficult.
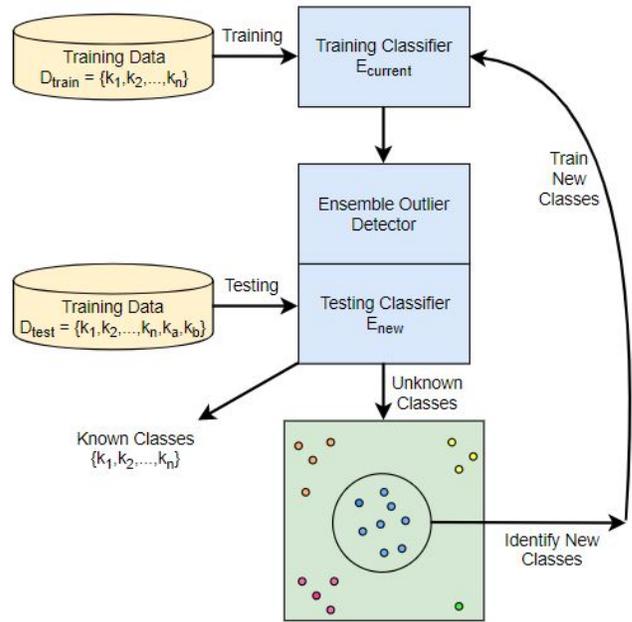
Approach

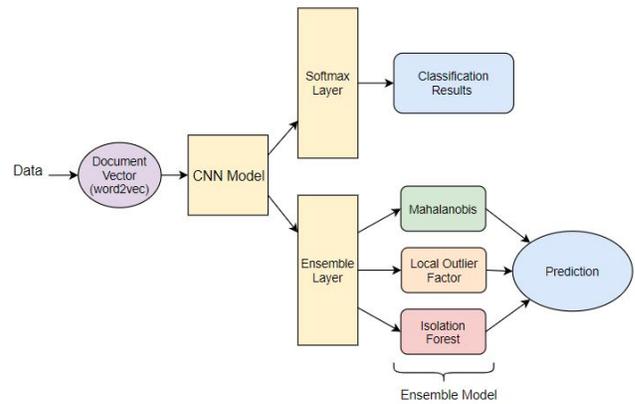Figure 1: Protocol for Open Set Classification and Incremental Class Learning

Figure 2: Ensemble Model and Testing Classifier Diagram. This diagram more clearly describes the 'Ensemble Outlier Detector' component from Figure 1.

This paper explores open set classification and the process of moving towards incremental learning of new classes. The objective is to create a classifier framework that can incrementally learn and expand its knowledge base as additional data is presented as shown in Figure 1. The approach is also outlined in Algorithm 1.

In prior work on open set classification, authors have focused on recognizing test samples as belonging to classes unknown during training. To the best of our knowledge, we

are the first ones to instantiate new classes iteratively, extending prior work to real incremental class learning. We first summarize our approach to provide a easily comprehensible sketch, before moving on to details. We seed our classifier framework by training it with examples from a small number of selected classes. We then expose the trained classifier to a mix of examples from the already-known classes as well unknown classes, during testing. At a certain point, we stop the testing of the current-classifier and cluster all examples recognized as belonging to unknown classes. Clustering allows for the grouping of similar data and visually represents the differences between unique clusters. Our hypothesis is that, if the clustering is good, one or more of the clusters of unknown examples can be thought of as new classes the current-classifier has not seen and these clusters are instantiated as new classes, by making up new unique labels for them. At this point, the current-classifier is updated by retraining it with all examples of the old known classes as well the newly instantiated classes. This process—of training, accumulating of outliers, clustering, and instantiating selected new classes out of the clusters—is repeated repeated a number of times, as long as the error of the entire learning process remains acceptable.

In particular, the classifier is a multi-layer CNN structure for training purposes. During testing, the softmax layer at the very end replaced by an outlier ensemble, following the work of (Prakhya, Venkataram, and Kalita 2017). The outlier detector ensemble consists of a Mahalanobis model, Local Outlier Factor model, and an Isolation Forest model, like (Prakhya, Venkataram, and Kalita 2017). The classifier model, as used in training is shown in Figure 2. Initially the model is created by training a classifier $E_{current}$ with a given k_{seed} number of classes found in the entire training data set $\mathbf{D}$. Then we create a derived dataset $D_{current}^{test}$ for testing the model by mixing examples of $k_{unknown}$ unknown classes with the previously trained k_{seed} classes. We always add k_{new} classes to the number of known classes. Thus, at the end of the ith iteration of class-learning, the classifier knows $k_{seed} + (i-1)k_{new}$ classes. We instantiate "new" classes by choosing dominant clusters, and then retrain the model with these new classes. The classes are then removed from the set of all classes and new ones are selected for the incremental addition.

We experiment with multiple clustering techniques including K-Means (Hartigan and Wong 1979), Birch (Zhang, Ramakrishnan, and Livny 1996), DBScan (Ester et al. 1996), and Spectral (Stella and Shi 2003), to determine the most suitable one for author attribution. We also experiment with various values of the parameters: k_{seed}, $k_{unknown}$ and δ.

Evaluation Methods

Since we use clustering as well as classification in our protocol for incremental classification, we need to evaluate both. Below, we first outline how clusters obtained from examples classified as unknown are evaluated, and then we describe how the incremental classifier is evaluated.

Evaluation of Clustering

There are a variety of clustering algorithms, and we need to choose one that works well in the domain of author attribution. The test samples that are deemed to be outliers are clustered, with the hypothesis that some of these clusters correspond to actual classes in the original dataset. We use the Davies-Bouldin Index as shown in Equation (1) to evaluate clustering (Davies and Bouldin 1979).

$$DB = \frac{1}{n} \sum_{i=1}^{n} max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right) \quad (1)$$

In this formula, n is the number of clusters produced, σ_i is the average distance between the points in cluster i and its centroid, $d(c_i, c_j)$ is the Euclidean distance between the centroids of clusters indexed i and j. Typically lower Davies-Bouldin Index scores indicate better clustering. Another clustering evaluation metric that we use is the V-Measure as shown in Equation (2), which has been widely used in clustering in natural language processing tasks when ground truth is known, i.e., we know samples and the classes they belong to. This metric computes the harmonic mean between homogeneity and completeness (Rosenberg and Hirschberg 2007). Homogeneity measures how close the clustering is such that each cluster contains samples from one class only. Completeness measures how close the clustering is such that samples of a given class are assigned to the same cluster. Typically scores close to 1 indicate better clustering. Here β is a parameter used to weigh between the two components—a higher value of β weighs completeness more heavily over homogeneity, and vice versa.

$$V = \frac{(1 + \beta) * homogeneity * completeness}{\beta * homogeneity + completeness} \quad (2)$$

Evaluation of Open Set Misclassification Error

Assuming there are n known classes, multi-class classification using a classifier $E_n()$, trained on n classes, can be evaluated using the misclassification error:

$$\epsilon_n = \frac{1}{N} \sum_{i=1}^{N} \left[E_n(\mathbf{x}^{(i)}) \neq y^{(i)} \right] \quad (3)$$

where N is the total number of samples in the dataset. When we test the same classifier $E_n()$ in the context of open set classification, we need to keep track of errors due that occur between known and unknown classes. When we test this classifier on N samples from n known classes and N' samples from u unknown classes, we test a total of $N + N'$ samples over $n + u$ classes. The open set classification error ϵ_{OS} for classifier E_n is given as (Bendale and Boult 2015):

$$\epsilon_{OS} = \epsilon_n + \frac{1}{N'} \sum_{j=N+1}^{N'} \left[E_n(\mathbf{x}^{(i)}) \neq unknown \right] \quad (4)$$

Evaluation of Incremental Class Learning Accuracy

For our research we are using clustering in order to obtain new classes after we perform open set recognition. This way

Input: Training Set $\mathbf{D} = \langle \mathbf{x}^{(i)}, y^{(i)} \rangle, i = 1 \cdots N$, samples from all known classes
Output: An incrementally trained classifier E on examples from a number of classes in $\mathbf{D}$

1 $\mathcal{C}_{all} \leftarrow \{C_1, \cdots C_n\}$, set of all known classes
2 $\mathcal{C}_{current}^{train} \leftarrow$ (randomly) pick k_{seed} classes from $\mathcal{C}_{all}$
3 $\mathbf{D}_{current}^{train} \leftarrow \{\langle \mathbf{x}^{(i)}, y^{(i)} \rangle \mid y(i) \in \mathcal{C}_{current}^{train}\}$, samples from classes in $\mathcal{C}_{current}^{train}$
4 **repeat**
5 $\quad$ $\mathcal{C}_{current}^{unknown} \leftarrow$ (randomly) pick $k_{unknown}$ classes from $\mathcal{C}_{all} - \mathcal{C}_{current}^{train}$
6 $\quad$ $\mathbf{D}_{current}^{test} \rightarrow \mathbf{D}_{current}^{train} \bigcup \{\langle \mathbf{x}^{(i)}, y^{(i)} \rangle \mid y^{(i)} \in \mathcal{C}_{current}^{unknown}\}$
7 $\quad$ $E_{current} \leftarrow$ (CNN) classifier trained on $\mathbf{D}_{currrent}^{train}$
8 $\quad$ $\mathbf{O} \leftarrow$ outlier samples detected by ensemble outlier detector when tested on $\mathbf{D}_{current}^{test}$
9 $\quad$ $\mathcal{L} \leftarrow$ set of clusters produced from $\mathbf{O}$ using a selected clustering algorithm
10 $\quad$ $\mathcal{L}_{dominant} \rightarrow$ pick k_{new} dominant clusters from $\mathcal{L}$, call these clusters new classes by making up new labels for them
11 $\quad$ $\mathcal{C}_{current}^{train} \leftarrow \mathcal{C}_{current}^{train} \bigcup \mathcal{L}_{dominant}$, increase the number of "known" classes
12 $\quad$ $\mathbf{D}_{current}^{train} \leftarrow \mathbf{D}_{current}^{train} \bigcup \{\langle x, y \rangle \in L_j \mid L_j \in \mathcal{L}_{dominant}\}$
13 **until** *too low accuracy* **or** *n times*;
14 $E \rightarrow E_{current}$
15 **return** E

Algorithm 1: Algorithm for Incremental Class-Learning

the new data identified for the novel classes can be used to incrementally train the model. For the evaluation of these clusters we present a new metric *ICA* (Incremental Class Accuracy) which takes into account the specific data from an identified cluster and averages calculations of homogeneity, completeness, and unknown identification accuracy of the cluster. We define homogeneity as the ratio of the number of data samples of the predominant class c in the cluster k ($n_{c|k}$) and the total number of values in the cluster (N_k). We define completeness as the ratio of the number of data samples of the predominant class c in the cluster k ($n_{c|k}$) and the total number of tested samples of the same class N_c. We define unknown identification accuracy as ratio of the number of unknown u data samples in the cluster k ($n_{u|k}$) and the total number on values in the cluster N_k. The equation used for ICA assumes only one cluster is being evaluated, but the equation can be adapted for multiple clustering by finding multiple ICA scores for each cluster and averaging.

$$Homogeneity = \frac{max(n_{c|k})}{N_k} \quad (5)$$

$$Completeness = \frac{max(n_{c|k})}{N_c} \quad (6)$$

$$Unknown\ Identification\ Accuracy = \frac{(n_{u|k})}{N_k} \quad (7)$$

$$ICA = \left(\frac{max(n_{c|k})}{N_k} + \frac{max(n_{c|k})}{N_c} + \frac{(n_{u|k})}{N_k} \right) * \frac{1}{3} \quad (8)$$

Other metrics that will be used to determine the performance of the model will be accuracy and F-score, these figures inherently show the accuracy of the classifier as well as novel data detection.

Experiments and Results

In this section we discuss the data sets used, the experiments performed, and the results with analysis.

Datasets

Since our objective is on open set author attribution, we use two datasets each of which contains 50 authors.

- **Victorian Era Literature Data Set** (Gungor 2018): This dataset is a collection of writing excerpts from 50 Victorian authors chosen from the GDELT database. The text has been pre-processed to remove specific words that identify the individual piece of text or author (names, author made words, etc.). Each author has hundreds of unique text pieces with 1000 words each.

- **CCAT-50** (Houvardas and Stamatatos 2006): This data set is a collection of 50 authors each with 50 unique text pieces divided for both training and testing. These texts are collections of corporate and industrial company news stories. This data is a subset of Reuters Corpus Volume 1.

Preliminary Clustering Results

After experimental comparison of the different clustering techniques, we decided to use Spectral Clustering (Stella and Shi 2003) as this typically produces the highest accuracy results as seen in Figure 3 and Figure 4, the clustering evaluation scores are also used for comparison. We use the pre-trained model *word2vec* (Mikolov et al. 2013) to obtain the word embeddings to pass into the multi-layer CNN structure.

Incremental Classification Results

For the first experiment our objective was to see if we could use our method to improve our classification accuracy and to also decide which clustering algorithm would work best. We run both data sets individually with five known training classes and then with ten known training classes, then we introduce three unknown classes during the testing phase for each of the tests. Our results include the comparison with accuracy and F1-Score as found on Table 1; a significant
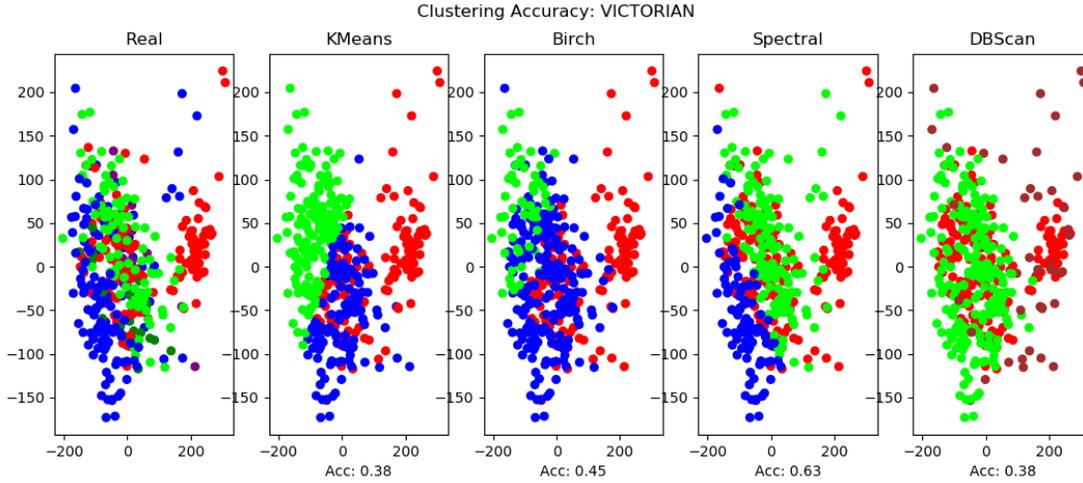
Figure 3: Clustering Plots for Victorian Literature Data with Accuracy Score, 5 Trained Classes and 8 Tested Classes
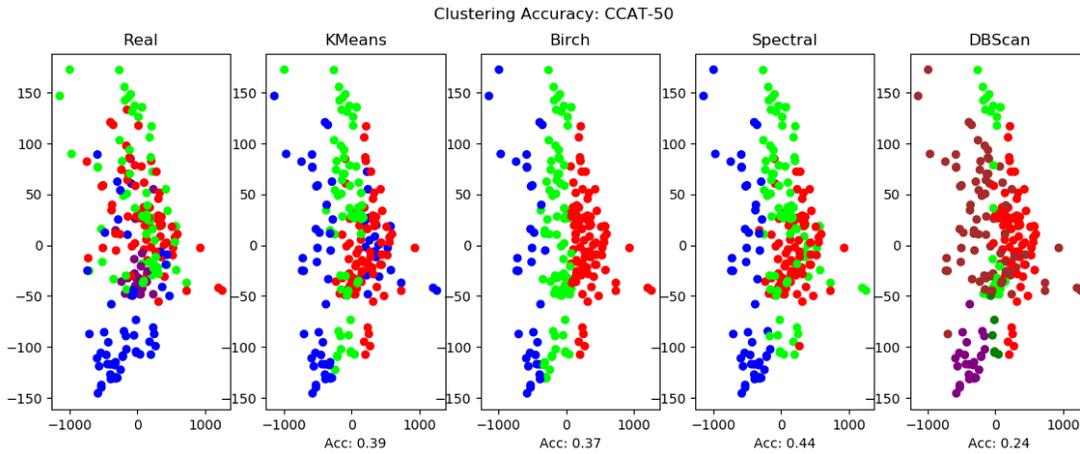

Figure 4: Clustering Plots for CCAT-50 Data with Accuracy Score, 5 Trained Classes and 8 Tested Classes

increase of these values is observed after the classifier is retrained with the identified novel classes. Our clustering evaluation metrics are found on Table 2. V-Measure scores prove to be more useful because the Davies-Bouldin scores do not always indicate the highest accuracy of clustering, this is because the best formed clusters does not necessarily mean higher accuracy. Even though our chosen data sets have not been used for open set classification in prior research we can compare our open set classification scores with the state of the art closed set classification scores. As far as we know, the best classification F1-Score from prior work for the Victorian Literature data set using only few classes is 0.808 (Gungor 2018) and our model has a similar score. Also as far as we know, the best classification accuracy score for the CCAT-50 data set for using only few classes is 86.5% and we are obtaining similar results. Our clustering models seem to have the most error for both data sets (especially the CCAT-50 data), thus presumably better clustering models or would produce greater results.

Dataset	Pre-Trained		Post-OpenSet	
	Acc	F1	Acc	F1
Victorian 5class	56.29%	0.592	85.43%	0.855
CCAT-50 5class	54.75%	0.565	83.00%	0.825
Victorian 10class	61.29%	0.644	71.38%	0.706
CCAT-50 10class	62.50%	0.727	86.77%	0.866

Table 1: Pre-Trained Class Scores and Post-Open Set Classification Scores, Either 5 or 10 initial trained classes and 3 unknown added during testing

For the second experiment we initially train with a fixed amount of classes k_{seed} and then incrementally add a $k_{unknown}$ amount of classes for testing. We repeat this process to demonstrate the model incrementally learns as the learning and open set classification cycle is repeated. We run this test by adding classes for multiple iterations and record the change in the F1-Score for the overall classification and generation of new classes; we attempt to run each test until

Data Set	Davies-Bouldin				V-Measure			
	Vic-5	CCAT-5	Vic-10	CCAT-10	Vic-5	CCAT-5	Vic-10	CCAT-10
K-Means	2.739	2.045	1.989	0.876	0.078	0.039	0.147	0.082
Birch	2.670	2.237	4.193	3.654	0.165	0.075	0.147	0.083
Spectral	4.457	2.550	4.841	0.807	0.319	0.242	0.328	0.258
DBScan	4.031	2.432	4.783	4.349	0.065	0.101	0.158	0.149

Table 2: Davies Bouldin Index and V-Measure Score for Clustering methods evaluated, Either 5 or 10 trained classes and 3 unknown added during testing.

the results drop significantly or until we have reached a max value of classes. From Figure 5 we notice the results of the incremental cycle and we notice that we achieve better results when fewer classes are added at a time. We run tests for adding 1, 2, and 3 classes at a time. We also keep track of the open set error shown in Equation 4; this metric shows error of unknown data identification but not novel class generation. The problem we notice with the experiment is that error will propagate through the process so as error accumulates the results deter. We also notice based on the results from both data sets, adding one class incrementally each iteration has better results because this limits the clustering error. We also notice that the Victorian Literature does worse than the CCAT-50 data and we think this is because of the text samples; the Victorian text includes words with slurs and accent mark symbols and *word2vec* is not pre-trained with these new features. The CCAT-50 data tends to have very distinct authors and the pieces of text tend to also tend to be more unique. Overall based on the results, we notice that most of this error can be attributed to the clustering process.

Form the previous experiments we realized the clustering process tends to have the most variance, this is evident from the low clustering accuracy due to the lack of fully distinct clusters. Thus, there needs to be a way to evaluate the clustering. Using our Incremental Class Accuracy (ICA) metric shown from Equation 5 we will be able to evaluate the clustering in regards to homogeneity, completeness, and unknown identification accuracy. From the previous experiment we also notice that adding one class at a time incrementally tends to produce the best results, so we calculate the ICA score when one class is added and instantiated. The results for both data sets is shown in Table 3. From these results we notice having a fewer amount of initial trained k_{seed} classes produces better results and this is expected as the $k_{unknown}$ classes are more easily identified.

Initial Training	Victorian	CCAT-50
5 Classes	0.687	0.875
10 Classes	0.593	0.754
15 Classes	0.529	0.764
20 Classes	0.387	0.681

Table 3: ICA Scores for 1 added class/cluster evaluation. Scores based on Equation 5.

Conclusion

This research works with open set classification regarding NLP text analysis in the area of Authorship Attribution. The model created will be to determine the originating author for a piece of text based on textual characteristics. We also move towards a novel incremental learning approach where unknown authors are identified and then the data is labeled so the classifier expands on its knowledge. Through this process we expand upon the state of the art implementation by creating a full cycle model by training on given data and then expanding the trained knowledge based on new data found for future testing.

Text based Authorship Attribution can be applied to research involving security and linguistic analysis. Some similar developing work using similar research methods involving image recognition (Rebuffi et al. 2017), this can be applied to facial recognition tasks and video surveillance applications. This model can also be further improved by developing a more precise way of distinguishing different pieces of text. Another method for future research is using backpropagation. Once novel classes are identified, the model should be then able to modify the already trained classifier with the $D_{current}^{train}$ data. Then the model can be tested with the $D_{current}^{test}$ to determine if the model can recognize previously unknown classes. Backpropagation of a neural network requires a fully inter connected set of layers that allow the processing of data through either side of the model (Hecht-Nielsen 1992). This process would save the step of fully retraining the classifier model. A similar approach to this can also be to add new "neurons" to a deep neural network to allow for an extension of a trained model (Draelos et al. 2017). With these new future improvements our model can be further improved and potentially obtain better results.

Acknowledgement

References

Bendale, A., and Boult, T. 2015. Towards open world recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1893–1902.
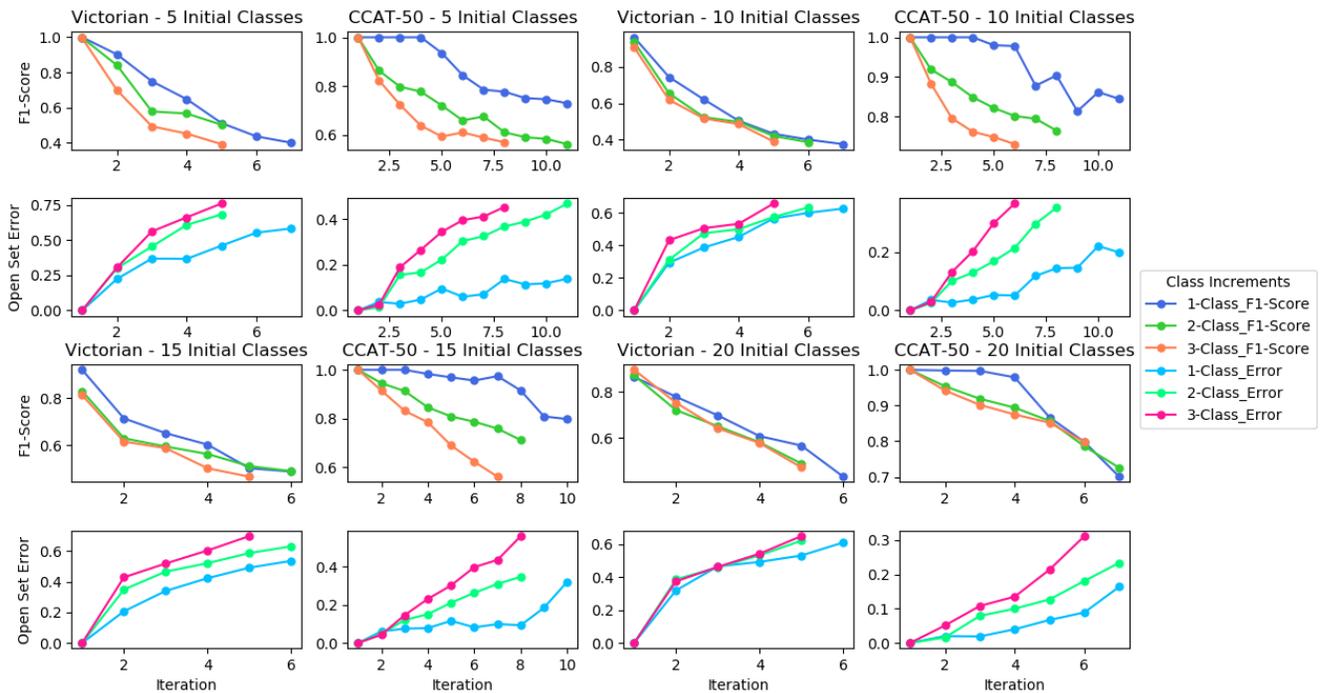
Figure 5: Incremental Learning Plots. Initially trained with 5, 10, 15, and 20 initial classes then tested by incrementally adding 1, 2, and 3 Classes. These plots show the final F1-Scores and Open Set Error from Equation 4.

Bendale, A., and Boult, T. E. 2016. Towards open set deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1563–1572.

Dahl, G. E.; Yu, D.; Deng, L.; and Acero, A. 2011. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Transactions on audio, speech, and language processing* 20(1):30–42.

Davies, D. L., and Bouldin, D. W. 1979. A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence* (2):224–227.

Dhamija, A. R.; Günther, M.; and Boult, T. 2018. Reducing network agnostophobia. In *Advances in Neural Information Processing Systems*, 9157–9168.

Draelos, T. J.; Miner, N. E.; Lamb, C. C.; Cox, J. A.; Vineyard, C. M.; Carlson, K. D.; Severa, W. M.; James, C. D.; and Aimone, J. B. 2017. Neurogenesis deep learning: Extending deep networks to accommodate new classes. In *2017 International Joint Conference on Neural Networks (IJCNN)*, 526–533. IEEE.

Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X.; et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, 226–231.

Gungor, A. 2018. *Benchmarking authorship attribution techniques using over a thousand books by fifty Victorian era novelists*. Ph.D. Dissertation.

Hartigan, J. A., and Wong, M. A. 1979. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28(1):100–108.

Hecht-Nielsen, R. 1992. Theory of the backpropagation neural network. In *Neural networks for perception*. Elsevier. 65–93.

Higashinaka, R.; Imamura, K.; Meguro, T.; Miyazaki, C.; Kobayashi, N.; Sugiyama, H.; Hirano, T.; Makino, T.; and Matsuo, Y. 2014. Towards an open-domain conversational system fully based on natural language processing. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 928–939.

Houvardas, J., and Stamatatos, E. 2006. N-gram feature selection for authorship identification. In *International conference on artificial intelligence: Methodology, systems, and applications*, 77–86. Springer.

Kriegel, H.-P.; Kröger, P.; Schubert, E.; and Zimek, A. 2009. Loop: local outlier probabilities. In *Proceedings of the 18th ACM conference on Information and knowledge management*, 1649–1652. ACM.

Liu, F. T.; Ting, K. M.; and Zhou, Z.-H. 2008. Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, 413–422. IEEE.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; and Dean, J. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 3111–3119.

Prakhya, S.; Venkataram, V.; and Kalita, J. 2017. Open set text classification using convolutional neural networks. In *International Conference on Natural Language Processing, 2017*.

Rebuffi, S.-A.; Kolesnikov, A.; Sperl, G.; and Lampert, C. H. 2017. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2001–2010.

Rocha, A.; Scheirer, W. J.; Forstall, C. W.; Cavalcante, T.; Theophilo, A.; Shen, B.; Carvalho, A. R.; and Stamatatos, E. 2016. Authorship attribution for social media forensics. *IEEE Transactions on Information Forensics and Security* 12(1):5–33.

Rosenberg, A., and Hirschberg, J. 2007. V-measure: A conditional entropy-based external cluster evaluation measure. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, 410–420.

Scheirer, W. J.; de Rezende Rocha, A.; Sapkota, A.; and Boult, T. E. 2012. Toward open set recognition. *IEEE transactions on pattern analysis and machine intelligence* 35(7):1757–1772.

Stella, X. Y., and Shi, J. 2003. Multiclass spectral clustering. In *null*, 313. IEEE.

Zhang, T.; Ramakrishnan, R.; and Livny, M. 1996. Birch: an efficient data clustering method for very large databases. In *ACM Sigmod Record*, volume 25, 103–114. ACM.