Sentiment Analysis in Tweets: Exploring Instance-based Transfer Learning for Dataset Enrichment

Eliseu Guimarães[1,2], Daniela Vianna[3,4], Aline Paes[1], Alexandre Plastino[1]

[1] Universidade Federal Fluminense, Brazil
[2] Marinha do Brasil
eliseuguimaraes@id.uff.br {alinepaes,plastino}@ic.uff.br
[3] Universidade Federal do Amazonas, Brazil
[4] Jusbrasil
dvianna@gmail.com

Abstract. Due to the popularity of user-generated content driven by social networks, sentiment analysis has become a very rich and influential research field. A challenging problem in this classification task is curating sufficient labeled data to train a classifier with a good performance. In order to address that issue, a promising strategy is to enrich the dataset of interest with labeled data from other datasets of different domains. However, another issue emerges: how to properly select data from a broad set of datasets to improve the classifier's performance. This manuscript presents instance-based transfer learning strategies to enrich the training set that is initially composed of the labeled target-dataset. Notably, we investigate the benefits of selecting similar and dissimilar instances from a set of source-datasets to transfer them to the target-dataset. Our results show that one of the strategies produces statistically significant performance improvement and that diversity plays an essential role in enhancing performance.

Categories and Subject Descriptors: H.2.8 [**Database Management**]: Database Applications; I.2.7 [**Artificial Intelligence**]: Natural Language Processing

Keywords: machine learning, sentiment analysis, supervised learning, transfer learning, Twitter

1. INTRODUCTION

Sentiment analysis is the computational study of people's opinions, sentiments, emotions, and attitudes [Liu 2020]. With the ever-growing use of social networks, this study field is becoming increasingly important since people are encouraged to share their opinions on various topics. One popular social network is Twitter[1], which is essentially a microblog with short texts – called tweets – in which people share opinions. Despite its simplicity, performing sentiment analysis on Twitter is challenging because of misspelled words, informal language, and lack of context [Martínez-Cámara et al. 2014]. One of the tasks in sentiment analysis is polarity detection, in the case of this study, performed over tweets.

Machine learning is widely used for this purpose, extracting features from tweets and using them to train classifiers. Even though these classifiers are usually trained with data from a specific domain of interest, occasionally there are not enough labeled data (e.g., a rare domain, the prohibitive cost of manually labeling available data, or even low data quality) from this domain and the classifier performance is poor.

In these cases, an instance-based transfer learning approach [Pan and Yang 2010] can increase the

[1]http://www.twitter.com

This research was partially supported by CNPq (under grants 311275/2020-6 and 315750/2021-9) and FAPERJ (under grants E-26/202.914/2019 and E-26/201.139/2022).

training set by selecting instances from another domain (the source domain) to enrich the training set from the target domain, ultimately increasing its predictive power [Guo et al. 2018; Liu et al. 2019; Ruder et al. 2017; Ruder and Plank 2017]. However, most of the previous works require the training of metrics or the partitioning of the source database on smaller subsets. Moreover, previous works did not always deal with sentiment analysis in tweets.

This paper extends the results obtained in [Guimarães et al. 2021][2]. In that paper, we explored three strategies to select data from a set of source-datasets belonging to a wide variety of domains, with the purpose of enriching the training set for polarity detection in a target-dataset. In this scenario, the target-dataset is labeled, but we aim to improve the predictive performance with an enriched training set. In [Guimarães et al. 2021] we address this issue by proposing a random selection, a closeness criterion for selection, and a selection based on diversity. In the latter case, a ratio of 1:1 was used to choose between similar or dissimilar instances. In this work, we conducted new experiments varying this ratio, and the results showed that enriching the training set with dissimilar instances, i.e., increasing the diversity, benefits the performance.

The remaining of this paper is organized as follows. Section 2 presents some related work, while in Section 3 the methodology is described, including a detailed explanation of the experimental procedure. In Section 4, the results of the conducted experiments are shown, and in Section 5, the conclusions are presented, and future research ideas are mentioned.

2. RELATED WORK

Several papers have proposed different approaches to solve the issue of training data selection from one or more source-datasets to train robust classifiers to a target-dataset from a different domain [Guo et al. 2018; Liu et al. 2019; Ruder et al. 2017; Ruder and Plank 2017]. [Guo et al. 2018] used a mixture-of-experts approach with several source-datasets. They considered that each source-dataset is aligned to a distinct region of the target-dataset and a point-to-set metric is learned to weigh the results from the classifiers trained with these source-datasets. The paper concludes that accuracies obtained with that strategy outperformed the results achieved with only one source-dataset or even the union of the source-datasets.

[Liu et al. 2019] proposed a reinforcement learning-based framework that searches for relevant instances and learns better representations for them. The framework has two components. One of them selects data using a distribution vector based on data selection from the previous step. The other component extracts data features, updates the rewards of the distribution vector generation process, and generates the classifier. The authors showed that this approach had the best performance in three out of four target-datasets compared to other studies.

[Ruder et al. 2017] analyzed how the following three factors affect data selection: (i) data representation, (ii) similarity metrics, and (iii) selection level. The authors used the similarity metric most frequently associated with each data representation, and the results showed that using the selection level of subsets of instances can have a better predictive performance than selecting individual instances.

[Ruder and Plank 2017] used Bayesian optimization to learn a similarity metric with the premise that different tasks and domains require different similarity notions. They used six similarity metrics between the datasets, three different data representations, and six diversity metrics on the training dataset. The authors concluded that the concomitant use of both diversity and similarity metrics enhances the predictive performance, overcoming the approaches that use one single metric or randomly select data.

[2]Codes and results of this and the previous work can be found at: `https://github.com/eliseupsg/dataset_enrichment`

Table I. Characteristics of the used datasets.

Dataset	Abreviation	#pos	#neg	% pos	Total
irony	iro	22	43	34%	65
sarcasm	sar	33	38	46%	71
aisopos	ais	159	119	57%	278
SemEval15-Fig	S15	47	274	15%	321
sentiment140	sen	182	177	51%	359
person	per	312	127	71%	439
hobbit	hob	354	168	68%	522
iphone	iph	371	161	70%	532
movie	mov	460	101	82%	561
sanders	san	570	654	47%	1224
Narr	nar	739	488	60%	1227
archeage	arc	724	994	42%	1718
SemEval18	S18	865	994	47%	1859
OMD	OMD	710	1196	37%	1906
HCR	HCR	539	1369	28%	1908
STS-gold	STS	632	1402	31%	2034
SentiStrength	SSt	1340	949	59%	2289
Target-dependent	Tar	1734	1733	50%	3467
Vader	Vad	2897	1299	69%	4196
SemEval13	S13	3183	1195	73%	4378
SemEval17	S17	2375	3972	37%	6347
SemEval16	S16	8893	3323	73%	12216

Our study takes a different approach, using data selection strategies that do not require metric training or splitting the source-datasets into subsets. Additionally, we aimed at sentiment analysis in tweets. Our results are substantiated by an extensive and diverse set of datasets from different domains.

3. METHODOLOGY

This section describes the methodology proposed in this paper to leverage instance-based transfer learning to tweet sentiment analysis. In Subsection 3.1, the datasets utilized are described, and the preprocessing steps used for feature extraction are presented. Subsection 3.2 describes the procedures adopted in the experiments conducted here.

3.1 Datasets and preprocessing

In the conducted assessments, we utilized a set of 22 datasets of tweets in English. Characteristics of these datasets are presented in Table I. We started replacing user mentions and URLs with unique expressions to clean the tweets. Then, tweets were tokenized and lowercased. Features were obtained using word embeddings from a static model [Bravo-Marquez et al. 2016] that has a good performance for sentiment analysis in tweets [Carvalho and Plastino 2021]. For each instance, features were computed using the average of the embeddings of instances' tokens, and, in case a token does not have a corresponding embedding, its embedding was considered a null vector.

3.2 Experimental procedures

In the conducted experiments, we utilized the SVM (Support Vector Machines) algorithm in its scikit-learn [Pedregosa et al. 2011] implementation with the parameter of class weighting set up for balanced mode because of its good performance in sentiment analysis in tweets [Barreto et al. 2021]. Furthermore, as baselines, we considered the accuracy and weighted F_1 values obtained from a stratified 10-fold cross-validation procedure performed over the target-dataset. Each experiment was performed

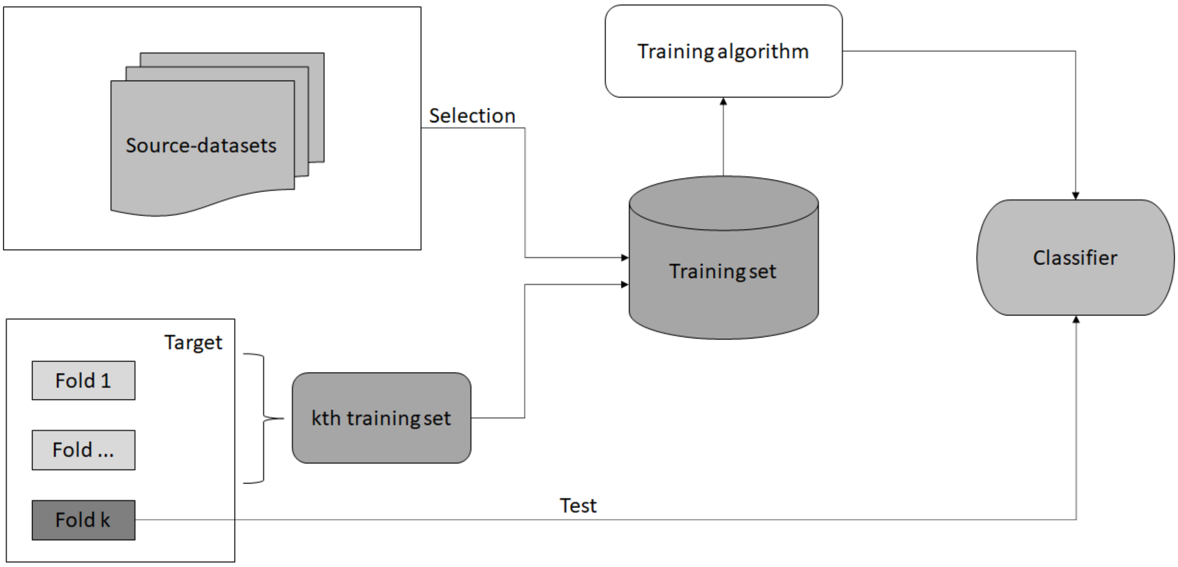

Fig. 1. Schematic of the experimental procedure.

considering each dataset as target-dataset and the remaining 21 datasets as the union of source-datasets.

Figure 1 shows the basic schematic of the experimental procedure used for all the experiments. First of all, the target-dataset is split into the same partitions used for the baseline computation. Then, the training algorithm is fed with the kth training set from the target-dataset, i.e., the nine training partitions, and the selected instances from the union of source-datasets according to a specific criterion and percentage. Details about the selection method are provided ahead in the paper. The training process generates a classifier, which is used to predict the classes for the test partition from the target-dataset and to compute the evaluation metrics. The training and testing procedures are repeated for all the partitions. The evaluation metrics for the employed criterion are computed as the mean of the metrics for all test partitions.

Figure 2 depicts the instances selection method. In Figure 2(a), the selection algorithm receives the difference in the number of elements between the classes of the training set from the target-dataset, the union of the source-datasets, the distances between the instances of the source and the target-datasets, and the selection criterion. Considering that the distances are used for almost all criteria, they were precomputed in order to avoid unnecessary recalculation. Initially, the selection algorithm calculates the number of instances of the minority class needed to balance the target training set. Next, taking into account that these needed instances come from the union of source-datasets, the percentage to be selected is applied in the minority class of the remaining source-dataset. For the majority class in the original target training set, this last quantity is the number of instances to select, while for the minority class, this quantity is added with the amount needed to balancing to define the amount to be selected. Then, the selection algorithm picks up the instances from the union of source-datasets, according to the quantities computed, and considering the criterion for each experiment, as shown ahead. As can be seen in Figure 2(b), these selected instances are added to the original training set forming the balanced training set. For example, consider an initial training set formed by the instances of target-dataset with 16 positive instances and 10 negative instances, the union of source-datasets with 50 positives and 51 negative instances, and 20% as the percentage to select. In the first step, the difference in the training set is calculated, and we conclude that 16-10 = 6 negative instances from the union of source-datasets will be needed to balance the training set. Considering this, there will be 50 positives and 51-6 = 45 negative instances available in the union of source-datasets for us to apply the

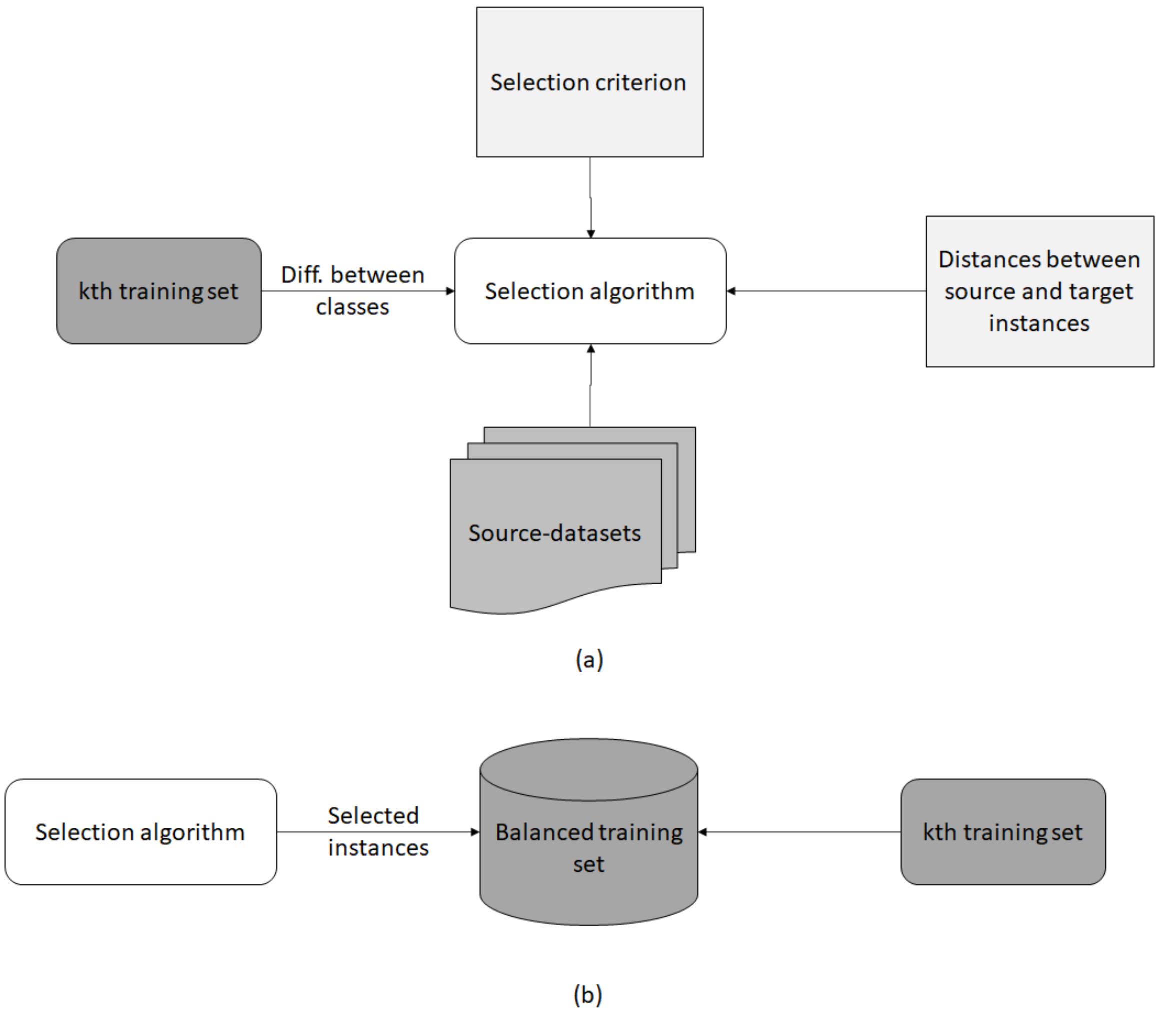

Fig. 2. Schematic of the selection procedure.

percentage. The percentage is applied to the minority class of this remaining source-datasets, then 20% x 45 = 9, and so the quantities of instances to select from the union of the source-datasets are nine positive, and 9+6 = 15 negative. Adding this to the original training set, we have 16+9 = 25 positive and 10+15 = 25 negative, making this a balanced training set.

The first experiment aimed to verify if the enrichment of the training set with the largest possible number of instances from the union of source-datasets improves the performance compared to the baseline. In this experiment, the percentage to select is 100%, and the chosen selection criterion was to select instances at random. This random selection respects the balance in the training set, as described above.

The second experiment aimed to identify if there is a subset of the union of source-datasets that, when added to the target training set, enhances the classifier's performance. For this purpose, three selection methods were used (each one with ten selected percentages): (i) random selection, (ii) closeness selection, and (iii) diversity selection.

The first method is similar to the first experiment but diversifies the selected percentage. The second method considers the assumption that instances that are similar to the original training set

Table II. Accuracies and F_1 obtained with target-dataset (t) and target-dataset+source-dataset (t+s), and their respective gains.

Dataset	Ac_t	Ac_{t+s}	Ac. gain	F_{1-t}	F_{1-t+s}	F_1 gain
iro	0.63	0.77	**1.22**	0.62	0.74	**1.19**
sar	0.69	0.85	**1.22**	0.67	0.85	**1.27**
ais	0.94	0.95	**1.00**	0.94	0.95	**1.00**
S15	0.90	0.76	0.84	0.90	0.78	0.87
sen	0.87	0.87	**1.01**	0.87	0.87	**1.01**
per	0.78	0.82	**1.06**	0.79	0.82	**1.05**
hob	0.89	0.83	0.93	0.89	0.83	0.93
iph	0.79	0.78	0.98	0.80	0.79	0.99
mov	0.83	0.86	**1.05**	0.83	0.87	**1.04**
san	0.83	0.84	**1.00**	0.83	0.84	**1.00**
nar	0.88	0.91	**1.04**	0.88	0.91	**1.04**
arc	0.87	0.85	0.99	0.87	0.85	0.98
S18	0.83	0.83	**1.00**	0.83	0.83	**1.00**
OMD	0.84	0.81	0.96	0.84	0.80	0.96
HCR	0.75	0.78	**1.04**	0.76	0.75	0.99
STS	0.86	0.86	0.99	0.86	0.86	**1.00**
SSt	0.80	0.82	**1.02**	0.80	0.82	**1.02**
Tar	0.83	0.82	0.98	0.83	0.82	0.98
Vad	0.87	0.88	**1.00**	0.88	0.88	**1.00**
S13	0.81	0.84	**1.05**	0.81	0.84	**1.03**
S17	0.88	0.88	**1.01**	0.88	0.88	**1.00**
S16	0.85	0.86	**1.02**	0.85	0.86	**1.02**

can improve the classifier performance. In this way, for each instance in the target training set, this method selects the closest instances in the union of source-datasets that are of the same class as the considered instance in target-dataset. The diversity selection method relies upon the idea that bringing diversity to the training set can augment performance while still selecting closer instances. For this method, seven different ratios between closer instances and farther instances are applied. It also selects only instances from the union of source-datasets with the same corresponding instance class in the target-dataset. In methods (ii) and (iii), the distance between instances is computed using Euclidean distance applied to the average embeddings of the instances.

For comparison purposes, we will compute the gain of the methods for each metric, dividing the performance achieved applying the method by the performance achieved by the baseline. Values of gain greater than 1 indicate that the enriched classifier has a better performance than the classifier trained with only the target-dataset, while values less than 1 show that the classifier has a poorer performance. Gains equal to 1 are obtained for classifiers whose performance remains the same as the original classifier.

4. RESULTS

This section presents the results obtained with the experiments described in Section 3. Table II presents the results for the first experiment. It shows the accuracies and the weighted F_1 obtained when the training set consists only of the target-dataset (columns Ac_t and F_{1-t}), and when the training set is composed of the target-dataset enriched with the union of the source-datasets (columns Ac_{t+s} and F_{1-t+s}). Columns "Ac. gain" and "F_1 gain" present the gains of accuracy and F_1, i.e., the results of the division of the columns Ac_{t+s} and F_{1-t+s} by the columns Ac_t and F_{1-t}, respectively. The boldfaced values represent gains greater than or equal to 1, that is, values that indicate the performance using the union of source-datasets aggregated to target-dataset outperformed or maintained the result with only the target-dataset. For both accuracy and F_1, it occurred for 15 datasets. For most datasets, the gain was close to 1, meaning that the difference in performance was irrelevant.

Tables III to XII present the results of the second experiment. Table III presents values of accuracy and F_1 obtained when the training set is composed of instances from target-dataset and instances from the union of source-datasets randomly selected. The selection size is defined in percentage and varies from 0.0% to 100.0%. This last percentage is equivalent to the first experiment. One important

Table III. Accuracies and F_1 obtained with a random selection of percentages from source-dataset associated with target-dataset.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	0.65	0.66	0.66	0.72	0.77	0.74	0.75	0.75	0.77	0.77	0.64	0.64	0.64	0.71	0.74	0.71	0.73	0.73	0.74	0.74
sar	0.69	0.79	0.76	0.75	0.75	0.82	0.78	0.86	0.85	0.85	0.67	0.78	0.75	0.74	0.73	0.81	0.77	0.85	0.84	0.85
ais	0.93	0.92	0.92	0.91	0.92	0.92	0.94	0.94	0.94	0.95	0.93	0.92	0.92	0.91	0.92	0.92	0.94	0.94	0.94	0.95
S15	0.88	0.88	0.88	0.87	0.87	0.85	0.83	0.79	0.76	0.76	0.89	0.88	0.88	0.87	0.87	0.85	0.84	0.80	0.78	0.78
sen	0.87	0.87	0.87	0.87	0.88	0.88	0.89	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.88	0.88	0.89	0.87	0.87	0.87
per	0.77	0.79	0.80	0.80	0.82	0.81	0.81	0.82	0.82	0.82	0.77	0.79	0.80	0.80	0.82	0.81	0.81	0.82	0.82	0.82
hob	0.88	0.88	0.88	0.86	0.87	0.84	0.83	0.83	0.83	0.83	0.88	0.88	0.88	0.86	0.86	0.84	0.83	0.83	0.82	0.83
iph	0.80	0.80	0.81	0.81	0.79	0.78	0.79	0.78	0.78	0.78	0.80	0.80	0.82	0.82	0.80	0.79	0.80	0.79	0.79	0.79
mov	0.84	0.87	0.86	0.85	0.84	0.86	0.85	0.86	0.86	0.86	0.83	0.86	0.85	0.85	0.84	0.86	0.85	0.86	0.86	0.87
san	0.84	0.84	0.84	0.84	0.83	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.83	0.83	0.83	0.83	0.83	0.84
nar	0.88	0.89	0.89	0.89	0.89	0.90	0.90	0.91	0.91	0.91	0.88	0.89	0.89	0.89	0.90	0.90	0.90	0.91	0.91	0.91
arc	0.87	0.87	0.87	0.87	0.87	0.86	0.85	0.86	0.86	0.85	0.87	0.87	0.87	0.87	0.86	0.86	0.85	0.86	0.86	0.85
S18	0.83	0.83	0.83	0.84	0.84	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.84	0.83	0.83	0.83	0.83	0.83
OMD	0.83	0.83	0.83	0.83	0.83	0.84	0.82	0.82	0.81	0.81	0.83	0.82	0.83	0.83	0.83	0.83	0.82	0.81	0.81	0.80
HCR	0.79	0.79	0.79	0.80	0.80	0.79	0.79	0.79	0.79	0.78	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.75
STS	0.86	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.86	0.86
SSt	0.81	0.81	0.81	0.81	0.81	0.81	0.82	0.81	0.82	0.82	0.81	0.81	0.81	0.81	0.81	0.81	0.82	0.81	0.82	0.82
Tar	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.82	0.82	0.82	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.82	0.82	0.82
Vad	0.87	0.87	0.88	0.88	0.88	0.88	0.88	0.87	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
S13	0.82	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.85	0.84	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.84
S17	0.88	0.88	0.88	0.88	0.88	0.88	0.89	0.89	0.89	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.89	0.88	0.88	0.88
S16	0.85	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.85	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86

point to observe is that the selection of 0.0% can select some instances of the union of source-datasets to balance the training set.

Table IV presents the gains obtained considering the random selection method; that is, values shown in this table are values from Table III divided by their respective baselines. The results are in bold when the gain is greater than or equal to 1. The last four rows of the table show, for each selected percentage, the average F_1 gain (Avg gain), the number of F_1 gains that are greater than or equal to 1 (#Gain), how many times the percentage produced the best performance, in terms of F_1, between all analyzed percentages (#Best), and the ranking of the percentage, i.e., the mean position of the percentage considering the ten evaluated percentages. The best results for each one of these evaluations are boldfaced. Considering that the ranking better represents the performance of a strategy, the best performance occurs with the selection of 5.0% (ranking 3.64) of the union of source-datasets. This percentage has the best results in terms of Avg gain (1.02), tied with 10.0%, 40.0%, 80.0%, and 100.%, and in terms of #Gain (18 gains), tied with 0.5%, 1.0%, and 2.5%. Despite the fact that 80.0% has the best performance for #Best (9 times vs. 6 times for 5.0%), the other metrics allow the choice of 5.0% as the best percentage for this selection method.

Table V presents the results of accuracy and F_1 gains for the strategy that selects instances from the union of source-datasets according to the closeness criterion. Gains greater than or equal to 1 are boldfaced and the best results for the four evaluations are at the bottom of the table. This highlighting is done for all the tables. The best ranking (2.14) for this strategy is obtained when 20.0% of the union of source-datasets are selected. This percentage has the best results in terms of Avg gain (1.03), and #Best (13 times), and its performance for #Gain (17 gains) is close to the best performance (19 gains for the selected percentage of 10.0%).

Table VI presents the accuracy and F_1 gains obtained with the strategy of selecting instances from the union of source-datasets that are closer to and farther from every instance in the target training set in a ratio of 1:1. For this criterion, the best-selected percentage is 10.0%, which obtained a ranking of 2.86. Considering the Avg gain evaluation, this percentage achieved the best result (1.03), tied with the selection of 20.0%. In terms of #Gain, percentages 1.0% and 2.5% obtained 19 gains that outperformed the baseline, but the result of 10.0% (18 gains) is very close, the same occurring with #Best that reaches 10 times, while the best result is 11 times (40.0%).

Table VII presents the results obtained with the selection of the closest and the farthest instances in a ratio of 2:1. The assessment of Avg gain yields a tie (1.02) with seven out of 10 percentages (2.5%, 5.0%, 10.0%, 20.0%, 40.0%, 80.0%, and 100.0%). In terms of #Gain, the performance achieved varies only between 16 and 18, a very tight margin. Even though the selection of 40.0% produces a #Best of 10, selecting 20.0% (#Best of 7) has the lowest ranking value (2.91). For this reason, this percentage is considered the best result for this ratio.

Table VIII shows the results for the strategy with a ratio closest:farthest of 3:1. For this ratio, selecting 40.0% from the union of source-datasets produces the best results in terms of Avg gain (1.03, tied with 20.0%), #Gain (18 gains, tied with 0.0%), #Best (11 times), and ranking (2.91). Considering these results, the selection of 40.0% can be, undoubtedly, chosen as the best percentage for this diversity criterion with a ratio of 3:1.

Table IX shows the result for the diversity criterion with a ratio of 4:1. In the same way that happened with the previous ratio, in this case the same percentage (20.0%) reached the best results in all evaluation topics. It obtained the best Avg gain (1.03), #Gain (18 gains), tied with 0.0% and 40.0%, #Best (9 times), tied with selecting 40.0%, and ranking (2.68). Then, this percentage is the best choice for the diversity strategy with a ratio of 4:1.

The ratios were inverted after increasing the first element of the ratio, adding the closest elements in quantity larger than the farthest elements. Table X presents the results obtained with the diversity criterion for the selection with a ratio of 1:2. Two percentages (20.0% and 40.0%) reached the best performance in terms of Avg gain (1.03). The greater #Gain (20 gains) occurred with the selection of 0.5%, and the greater #Best (10) occurred selecting 40.0%, which also had the best result in the ranking (3.09). This percentage achieved 17 gains in #Gain, which is a little far from the best result, but it is still a good performance.

The diversity criterion with a ratio of 1:3 is presented in Table XI. As can be seen in the table, two percentages (10.0% and 20.0%) had achieved the best performance for Avg gain (1.03), while three obtained the best result for #Gain (19 gains): 0.5%, 2.5%, and 5.0%. Two percentages (2.5% and 20.0%) reached the greater #Best (10 times). Taking into account the ranking, the best percentage to select is 20.0%, with a ranking of 2.36. This percentage also has a good performance in terms of #Gain (18 gains).

Table XII presents the results when selecting the closest and the farthest instances in a ratio of 1:4.

Table IV. Accuracy and F_1 gains obtained with a random selection of percentages from source-dataset associated with target-dataset.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.03**	**1.06**	**1.05**	**1.15**	**1.22**	**1.17**	**1.20**	**1.20**	**1.22**	**1.22**	**1.02**	**1.03**	**1.03**	**1.13**	**1.18**	**1.13**	**1.17**	**1.17**	**1.19**	**1.19**
sar	**1.00**	**1.14**	**1.10**	**1.08**	**1.08**	**1.18**	**1.12**	**1.24**	**1.22**	**1.22**	**1.01**	**1.18**	**1.13**	**1.11**	**1.10**	**1.22**	**1.15**	**1.28**	**1.27**	**1.27**
ais	0.98	0.98	0.98	0.96	0.97	0.98	**1.00**	0.99	**1.00**	**1.00**	0.99	0.98	0.98	0.96	0.97	0.98	**1.00**	0.99	**1.00**	**1.00**
S15	0.98	0.98	0.98	0.96	0.97	0.94	0.92	0.87	0.84	0.84	0.99	0.98	0.98	0.96	0.97	0.95	0.93	0.89	0.87	0.87
sen	**1.00**	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.03**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.03**	**1.00**	**1.00**	**1.01**
per	0.99	**1.02**	**1.03**	**1.03**	**1.06**	**1.04**	**1.04**	**1.06**	**1.06**	**1.06**	0.98	**1.00**	**1.02**	**1.02**	**1.04**	**1.03**	**1.03**	**1.05**	**1.04**	**1.05**
hob	0.98	0.99	0.98	0.97	0.97	0.95	0.94	0.94	0.93	0.93	0.99	0.99	0.98	0.97	0.97	0.94	0.93	0.93	0.92	0.93
iph	**1.01**	**1.01**	**1.03**	**1.03**	**1.00**	0.99	**1.00**	0.99	0.99	0.98	**1.00**	**1.01**	**1.03**	**1.03**	**1.00**	0.99	**1.00**	0.99	0.99	0.99
mov	**1.02**	**1.05**	**1.05**	**1.03**	**1.02**	**1.04**	**1.03**	**1.05**	**1.04**	**1.05**	**1.00**	**1.03**	**1.02**	**1.02**	**1.01**	**1.03**	**1.02**	**1.04**	**1.04**	**1.04**
san	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
nar	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**
arc	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	0.98	0.99	0.99	0.99	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	0.98	0.99	0.99	0.98
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
OMD	0.99	0.99	0.99	0.99	**1.00**	**1.00**	0.98	0.98	0.97	0.96	0.99	0.98	0.99	0.99	0.99	0.99	0.98	0.97	0.97	0.96
HCR	**1.04**	**1.05**	**1.05**	**1.06**	**1.06**	**1.05**	**1.05**	**1.05**	**1.04**	**1.04**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	0.99	0.99
STS	0.99	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	0.99	0.99	0.99	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
SSt	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.98	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.98	0.98
Vad	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
S13	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**
										Avg gain	1.00	1.01	1.01	1.01	**1.02**	**1.02**	1.01	**1.02**	**1.02**	**1.02**
										#Gain	16	**18**	**18**	**18**	**18**	17	17	15	15	15
										#Best	7	6	7	7	6	4	6	8	**9**	8
										Ranking	4.91	4.00	3.73	4.00	**3.64**	3.86	4.14	3.95	4.23	4.36

Table V. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer to target-dataset.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.15**	**1.20**	**1.18**	**1.18**	**1.15**	**1.20**	**1.23**	**1.20**	**1.25**	**1.33**	**1.15**	**1.19**	**1.15**	**1.16**	**1.13**	**1.18**	**1.22**	**1.17**	**1.23**	**1.32**
sar	**1.00**	**1.06**	**1.08**	**1.12**	**1.12**	**1.12**	**1.18**	**1.16**	**1.18**	**1.22**	**1.01**	**1.08**	**1.11**	**1.16**	**1.16**	**1.16**	**1.23**	**1.21**	**1.21**	**1.26**
ais	0.98	0.98	0.98	0.98	0.98	**1.00**	0.99	0.97	0.97	0.98	0.98	0.99	0.98	0.98	0.98	**1.00**	0.99	0.97	0.97	0.98
S15	0.96	0.97	0.96	0.94	0.94	0.92	0.90	0.89	0.84	0.81	0.97	0.98	0.96	0.95	0.95	0.93	0.92	0.91	0.87	0.85
sen	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	0.99	0.98	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	0.99	0.98
per	**1.01**	**1.01**	**1.01**	**1.03**	**1.03**	**1.07**	**1.05**	**1.04**	**1.06**	**1.07**	**1.00**	**1.00**	**1.00**	**1.02**	**1.03**	**1.06**	**1.04**	**1.03**	**1.05**	**1.05**
hob	0.99	0.99	0.98	0.97	0.97	0.96	0.96	0.95	0.92	0.92	0.99	0.99	0.98	0.97	0.97	0.95	0.96	0.94	0.92	0.91
iph	0.99	**1.01**	**1.01**	**1.01**	**1.03**	**1.03**	**1.04**	**1.03**	0.99	0.98	0.99	**1.01**	**1.01**	**1.01**	**1.03**	**1.02**	**1.04**	**1.03**	0.99	0.98
mov	**1.00**	**1.01**	**1.02**	**1.03**	**1.03**	**1.05**	**1.05**	**1.04**	**1.05**	**1.05**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.04**	**1.04**	**1.04**	**1.04**	**1.04**
san	**1.01**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	0.99	0.97	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	0.99	0.97
nar	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.03**	**1.03**	**1.01**	**1.02**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.03**	**1.02**	**1.01**	**1.02**
arc	**1.00**	**1.00**	0.99	0.99	**1.00**	**1.00**	0.99	0.99	0.97	0.97	**1.00**	**1.00**	0.99	0.99	**1.00**	**1.00**	0.99	0.99	0.97	0.97
S18	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**
OMD	0.97	0.97	0.98	0.99	0.99	0.98	0.98	0.97	0.96	0.95	0.97	0.97	0.98	0.99	0.99	0.98	0.98	0.96	0.95	0.94
HCR	**1.02**	**1.03**	**1.02**	**1.04**	**1.04**	**1.04**	**1.04**	**1.05**	**1.04**	**1.04**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
STS	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	0.99	0.98	0.97	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	0.99	0.98	0.97
SSt	**1.00**	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.98	0.98	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.98	0.98	0.98
Vad	0.99	0.99	0.99	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
S13	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**
										Avg gain	1.00	1.01	1.01	1.01	1.02	1.02	**1.03**	1.01	1.01	1.01
										#Gain	15	17	16	16	18	**19**	17	15	12	12
										#Best	5	6	3	5	8	9	**13**	7	5	7
										Ranking	5.32	4.59	5.18	3.91	3.14	2.86	**2.14**	4.09	5.36	5.05

The best percentage for this strategy is 5.0% with a ranking of 3.00, and the best results in terms of Avg gain (1.03), tied with 10.0% and 20.0%. This percentage presented good results in terms of #Gain (19 gains, while the best result is 20 gains for 0.5% and 1.0%) and #Best (8 times, with the best performance obtained by 80.0% with nine times).

As we can observe from the tables above, generally, the gains are slightly greater than 1. Nevertheless, two datasets have higher gains for some strategy combinations and selected percentages: irony and sarcasm. These two are the smallest datasets, with just a few tens of instances.

Table XIII compares the best results for all the strategies, concluding the evaluation. Once again, gains greater than or equal to 1 are boldfaced, and the best result for each evaluation is at the bottom

Table VI. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 1:1.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.03**	**1.13**	**1.20**	**1.20**	**1.28**	**1.30**	**1.23**	**1.25**	**1.25**	**1.25**	**1.00**	**1.12**	**1.21**	**1.21**	**1.27**	**1.30**	**1.23**	**1.24**	**1.21**	**1.21**
sar	0.96	**1.02**	**1.10**	**1.10**	**1.10**	**1.12**	**1.20**	**1.14**	**1.20**	**1.24**	0.97	**1.04**	**1.12**	**1.13**	**1.14**	**1.15**	**1.25**	**1.18**	**1.25**	**1.29**
ais	0.99	0.98	0.99	0.98	0.98	0.99	0.99	0.99	**1.00**	**1.00**	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	**1.00**	**1.00**
S15	**1.00**	0.99	0.98	0.95	0.96	0.94	0.91	0.88	0.84	0.82	0.99	0.99	0.98	0.96	0.97	0.95	0.92	0.90	0.87	0.85
sen	**1.01**	**1.00**	**1.03**	**1.01**	**1.02**	**1.03**	**1.01**	**1.03**	**1.02**	**1.01**	**1.01**	**1.00**	**1.03**	**1.01**	**1.02**	**1.03**	**1.01**	**1.03**	**1.02**	**1.01**
per	**1.04**	**1.01**	**1.02**	**1.03**	**1.03**	**1.05**	**1.06**	**1.05**	**1.06**	**1.06**	**1.03**	**1.00**	**1.01**	**1.02**	**1.02**	**1.04**	**1.05**	**1.04**	**1.04**	**1.05**
hob	0.98	0.99	**1.00**	0.98	0.96	0.98	0.95	0.96	0.93	0.92	0.98	0.99	**1.00**	0.98	0.96	0.97	0.95	0.96	0.93	0.90
iph	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	0.99	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**
mov	**1.03**	**1.02**	**1.01**	**1.03**	**1.04**	**1.04**	**1.05**	**1.05**	**1.06**	**1.06**	**1.02**	**1.01**	**1.00**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.04**
san	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**
nar	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**
arc	**1.00**	0.99	**1.00**	**1.00**	0.99	0.99	**1.00**	0.99	0.99	0.99	**1.00**	0.99	**1.00**	**1.00**	0.99	0.99	**1.00**	0.99	0.99	0.99
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
OMD	0.98	0.99	0.99	**1.00**	0.99	**1.00**	0.98	0.98	0.97	0.96	0.98	0.99	0.99	**1.00**	0.99	**1.00**	0.98	0.97	0.96	0.96
HCR	**1.03**	**1.04**	**1.04**	**1.03**	**1.03**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
STS	**1.01**	**1.01**	**1.00**	**1.01**	**1.02**	**1.02**	**1.01**	**1.01**	0.99	0.98	**1.00**	**1.01**	**1.00**	**1.01**	**1.02**	**1.02**	**1.01**	**1.01**	**1.00**	0.99
SSt	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
Tar	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	0.99	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	0.99	0.99	0.98
Vad	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.00**
S13	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.05**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
										Avg gain	1.00	1.01	1.02	1.02	1.02	**1.03**	**1.03**	1.02	1.02	1.02
										#Gain	16	17	**19**	**19**	17	18	18	16	17	16
										#Best	5	4	6	6	4	10	9	**11**	9	8
										Ranking	5.00	5.09	4.36	4.09	3.68	**2.86**	3.14	2.95	3.68	4.59

Table VII. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 2:1.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.13**	**1.18**	**1.13**	**1.23**	**1.18**	**1.23**	**1.25**	**1.25**	**1.27**	**1.25**	**1.12**	**1.16**	**1.12**	**1.22**	**1.17**	**1.22**	**1.25**	**1.24**	**1.25**	**1.21**
sar	0.98	**1.06**	**1.08**	**1.08**	**1.10**	**1.14**	**1.16**	**1.16**	**1.18**	**1.18**	0.99	**1.08**	**1.11**	**1.11**	**1.13**	**1.19**	**1.20**	**1.21**	**1.23**	**1.22**
ais	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.99	**1.00**	**1.00**	0.99	0.99	0.98	0.98	0.99	0.99	0.99	0.99	**1.00**	**1.00**
S15	**1.00**	0.99	0.98	0.95	0.93	0.93	0.90	0.87	0.83	0.80	**1.01**	0.99	0.98	0.96	0.94	0.94	0.91	0.89	0.87	0.84
sen	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.02**
per	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.05**	**1.07**	**1.06**	**1.06**	**1.06**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.04**	**1.06**	**1.04**	**1.04**	**1.05**
hob	0.99	0.99	0.98	0.97	0.97	0.97	0.95	0.95	0.94	0.92	0.99	0.99	0.99	0.97	0.97	0.96	0.95	0.94	0.93	0.91
iph	0.99	**1.00**	**1.01**	**1.01**	**1.02**	**1.04**	**1.02**	**1.02**	**1.01**	**1.01**	0.98	0.99	**1.01**	**1.01**	**1.02**	**1.04**	**1.02**	**1.02**	**1.01**	**1.01**
mov	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.05**	**1.05**	**1.06**	**1.06**	**1.06**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.04**	**1.04**	**1.05**	**1.04**	**1.05**
san	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
nar	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.03**	**1.03**	**1.05**	**1.04**	**1.04**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.05**	**1.04**	**1.04**
arc	**1.00**	**1.00**	0.99	**1.00**	0.99	0.99	0.99	0.99	0.99	0.99	**1.00**	**1.00**	0.99	**1.00**	0.99	0.99	0.99	0.99	0.99	0.99
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
OMD	0.98	0.98	0.98	0.99	**1.00**	**1.00**	0.98	0.98	0.97	0.97	0.98	0.98	0.99	0.99	**1.00**	**1.00**	0.98	0.97	0.96	0.96
HCR	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**
STS	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	0.99	0.98
SSt	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.99	0.98
Vad	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**
S13	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.05**	**1.05**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**
										Avg gain	1.01	1.01	1.01	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
										#Gain	17	17	17	**18**	**18**	**18**	17	16	16	16
										#Best	4	3	2	4	4	7	7	**10**	7	7
										Ranking	5.27	5.32	5.05	4.18	4.00	2.95	**2.91**	3.05	3.86	4.18

of the table. Considering the Avg gain, there is a tie (1.03) between seven out of the nine strategies, and the other two are close to this value of gain (1.02). This closeness in the results also occurs with the #Gain, with a variation from 17 (3 strategies: Closeness, Div. 2:1, and Div. 1:2) to 19 (Div. 1:4). The evaluation of #Best shows a different scenario: while the strategy Div. 1:4 achieves 10 times the best position, some strategies only reach the best result three times (Random and Div. 2:1). Observing the ranking, one strategy is way better than others – Div. 1:3 has a mean position of 2.73, while the next better-ranked strategy (Div. 1:1) has only a position of 3.41. Taking into account that the Div. 1:3 strategy also reaches the best Avg gain and has a slightly worse result for #Gain and #Best, compared to the best results for these two evaluations, this strategy, with its corresponding selected percentage of 20.0%, is the best we found in our research.

Table VIII. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 3:1.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.10**	**1.12**	**1.13**	**1.20**	**1.21**	**1.23**	**1.30**	**1.25**	**1.27**	**1.25**	**1.10**	**1.13**	**1.12**	**1.21**	**1.20**	**1.23**	**1.30**	**1.24**	**1.25**	**1.21**
sar	**1.00**	**1.12**	**1.08**	**1.06**	**1.12**	**1.14**	**1.14**	**1.20**	**1.20**	**1.20**	**1.01**	**1.15**	**1.11**	**1.09**	**1.16**	**1.19**	**1.18**	**1.25**	**1.25**	**1.25**
ais	0.99	0.98	0.98	0.98	0.98	0.99	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.98	0.98	0.99	**1.00**	**1.00**	**1.00**	**1.00**
S15	**1.00**	0.99	0.97	0.94	0.93	0.93	0.91	0.87	0.83	0.82	**1.01**	0.99	0.98	0.95	0.94	0.94	0.92	0.90	0.87	0.85
sen	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.01**	**1.03**	**1.02**	**1.02**	**1.02**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.01**	**1.03**	**1.02**	**1.02**	**1.02**
per	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.06**	**1.07**	**1.05**	**1.06**	**1.08**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.05**	**1.06**	**1.04**	**1.05**	**1.06**
hob	0.99	0.98	0.98	0.97	0.98	0.97	0.97	0.95	0.95	0.92	0.99	0.98	0.98	0.97	0.98	0.96	0.96	0.94	0.94	0.91
iph	0.99	0.99	**1.00**	**1.02**	**1.02**	**1.02**	**1.03**	**1.01**	**1.01**	**1.00**	0.99	0.99	**1.00**	**1.02**	**1.02**	**1.02**	**1.03**	**1.01**	**1.01**	**1.00**
mov	**1.01**	**1.00**	**1.00**	**1.02**	**1.02**	**1.05**	**1.05**	**1.05**	**1.05**	**1.06**	**1.01**	**1.00**	0.99	**1.01**	**1.01**	**1.04**	**1.03**	**1.04**	**1.04**	**1.05**
san	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**
nar	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**
arc	**1.00**	**1.00**	0.99	0.99	0.99	**1.00**	**1.00**	**1.00**	0.99	0.98	**1.00**	**1.00**	0.99	0.99	0.99	0.99	0.99	**1.00**	0.99	0.98
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**
OMD	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	0.97	0.96	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	0.97	0.96
HCR	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
STS	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.01**	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.01**	0.99	0.98
SSt	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.00**	**1.00**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.99
Vad	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**
S13	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.05**	**1.05**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**
										Avg gain	1.01	1.01	1.01	1.01	1.02	1.02	**1.03**	**1.03**	1.02	1.02
										#Gain	**18**	17	16	17	17	17	17	**18**	16	16
										#Best	5	3	2	4	4	3	10	**11**	8	10
										Ranking	5.14	4.73	5.05	4.09	3.95	3.82	2.95	**2.91**	3.68	4.27

Table IX. Accuracy and F_1 gains obtained with selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 4:1.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.05**	**1.15**	**1.18**	**1.18**	**1.23**	**1.23**	**1.25**	**1.25**	**1.25**	**1.25**	**1.05**	**1.14**	**1.16**	**1.18**	**1.23**	**1.23**	**1.25**	**1.24**	**1.21**	**1.21**
sar	**1.00**	**1.06**	**1.10**	**1.08**	**1.14**	**1.14**	**1.14**	**1.18**	**1.18**	**1.20**	**1.01**	**1.09**	**1.13**	**1.11**	**1.18**	**1.19**	**1.18**	**1.23**	**1.23**	**1.25**
ais	0.98	0.97	0.99	0.98	0.98	0.99	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.97	0.99	0.98	0.98	0.99	**1.00**	**1.00**	**1.00**	**1.00**
S15	0.99	0.99	0.98	0.94	0.92	0.93	0.91	0.88	0.84	0.81	**1.00**	0.99	0.98	0.94	0.93	0.94	0.92	0.90	0.87	0.85
sen	**1.01**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.01**	**1.00**	**1.02**	**1.01**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.01**	**1.00**	**1.02**
per	**1.00**	**1.01**	**1.02**	**1.02**	**1.03**	**1.06**	**1.06**	**1.05**	**1.06**	**1.07**	**1.00**	**1.00**	**1.01**	**1.01**	**1.02**	**1.05**	**1.05**	**1.03**	**1.05**	**1.06**
hob	0.99	0.98	0.98	0.97	0.98	0.96	0.96	0.95	0.94	0.92	0.99	0.99	0.98	0.97	0.97	0.96	0.96	0.94	0.94	0.91
iph	0.99	0.99	**1.01**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.01**	**1.01**	0.99	0.99	**1.01**	**1.01**	**1.02**	**1.03**	**1.03**	**1.03**	**1.01**	**1.01**
mov	**1.01**	**1.01**	**1.00**	**1.02**	**1.03**	**1.05**	**1.05**	**1.05**	**1.05**	**1.06**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.04**	**1.04**	**1.04**	**1.04**	**1.05**
san	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
nar	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**
arc	**1.00**	**1.00**	0.99	0.99	0.99	0.99	**1.00**	**1.00**	0.99	0.98	**1.00**	**1.00**	0.99	0.99	0.99	0.99	0.99	**1.00**	0.99	0.98
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.02**	**1.01**	**1.01**	**1.00**
OMD	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.98	0.97	0.96	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.97	0.96	0.96
HCR	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**
STS	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	0.99	0.98
SSt	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.00**	**1.00**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99
Vad	**1.00**	**1.00**	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**
S13	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.04**	**1.04**	**1.05**	**1.04**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**
										Avg gain	1.00	1.01	1.01	1.01	1.02	1.02	**1.03**	1.02	1.02	1.02
										#Gain	**18**	17	17	17	17	17	**18**	**18**	16	16
										#Best	5	3	2	5	5	5	**9**	**9**	7	8
										Ranking	4.95	5.27	4.73	4.32	3.64	3.09	**2.68**	3.23	4.05	4.45

Table X. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 1:2.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.05**	**1.08**	**1.23**	**1.21**	**1.22**	**1.25**	**1.23**	**1.30**	**1.25**	**1.22**	0.91	**1.03**	**1.24**	**1.21**	**1.22**	**1.24**	**1.22**	**1.29**	**1.21**	**1.19**
sar	0.98	**1.06**	**1.06**	**1.08**	**1.14**	**1.16**	**1.18**	**1.16**	**1.24**	**1.26**	0.99	**1.08**	**1.08**	**1.11**	**1.18**	**1.20**	**1.23**	**1.21**	**1.29**	**1.31**
ais	0.98	0.98	0.99	0.98	0.98	0.97	0.99	0.99	**1.00**	**1.00**	0.98	0.99	0.99	0.98	0.98	0.97	0.99	0.99	**1.00**	**1.00**
S15	**1.00**	**1.00**	**1.00**	0.98	0.97	0.93	0.92	0.87	0.84	0.83	0.99	**1.00**	0.99	0.98	0.97	0.93	0.93	0.89	0.88	0.86
sen	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**
per	**1.03**	**1.02**	**1.02**	**1.04**	**1.03**	**1.03**	**1.04**	**1.06**	**1.06**	**1.06**	**1.01**	**1.01**	**1.01**	**1.03**	**1.02**	**1.03**	**1.03**	**1.05**	**1.05**	**1.04**
hob	0.99	**1.00**	0.98	0.98	0.98	0.97	0.97	0.95	0.93	0.92	0.99	**1.00**	0.98	0.98	0.97	0.97	0.96	0.95	0.92	0.91
iph	**1.01**	**1.01**	**1.00**	**1.00**	**1.02**	**1.03**	**1.02**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	0.99	**1.00**	**1.02**	**1.03**	**1.02**	**1.01**	**1.01**	**1.00**
mov	**1.03**	**1.04**	**1.03**	**1.03**	**1.03**	**1.02**	**1.06**	**1.06**	**1.06**	**1.05**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.01**	**1.05**	**1.05**	**1.05**	**1.03**
san	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**
nar	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.04**	**1.04**	**1.04**
arc	**1.00**	0.99	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.99	**1.00**	0.99	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99	0.99	0.99
S18	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
OMD	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	0.99	0.99	0.98	0.97	0.97	0.99	**1.00**	**1.00**	**1.00**	**1.01**	0.99	0.99	0.98	0.97	0.96
HCR	**1.04**	**1.03**	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.06**	**1.05**	**1.05**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	**1.00**	**1.02**	**1.01**	**1.00**
STS	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.00**	0.99	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.00**	**1.00**
SSt	**1.00**	**1.01**	**1.00**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.00**	**1.01**	**1.00**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98
Vad	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**
S13	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.05**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.03**	**1.04**	**1.03**	**1.04**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
										Avg gain	1.00	1.01	1.02	1.02	1.02	1.02	**1.03**	**1.03**	1.02	1.02
										#Gain	16	**20**	18	19	19	16	17	17	17	17
										#Best	4	6	6	6	8	6	6	**10**	9	8
										Ranking	5.05	4.14	3.82	3.55	3.14	4.36	3.45	**3.09**	3.86	4.77

We performed statistical tests to verify our findings. At first, we applied the Friedman test over the F_1 values of the best strategies, displayed in Table XIII, and found a p-value of 0.146, showing that these strategies have no statistically significant difference. Despite this table presenting the F_1 gains, we performed the tests over the F_1 values. Later on, we used the Friedman test over the F_1 values of the strategies and the baseline, and the result was a p-value of 0.010, indicating that at least one strategy presents a statistically significant difference compared to the baseline. Then, we used the Nemenyi test, and, considering that we are comparing nine different strategies against the baseline, we performed a Bonferroni correction for the p-value. Then, the null hypothesis should be rejected for a p-value fewer than $0.05/9 \approx 0.006$, and we discovered that the Div. 1:3 - 20% strategy is the only one with a p-value (0.001) fewer than the corrected value when compared to the baseline, confirming this is the best strategy for the dataset enrichment.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we investigated if using data from a set of source-datasets can improve the performance of classification models for a labeled target-dataset, in the context of sentiment analysis in tweets. For this purpose, two experiments were conducted: the first one aggregating the maximum amount of instances from the union of source-datasets to the training set of the classifier, and the second one proposing strategies for selecting instances from the union of source-datasets according to three selection methods: (i) random selection, (ii) closeness selection, and (iii) diversity selection. For all experiments, balancing the training set was assured. The generated models were tested in partitions of the target-dataset via a cross-validation procedure. The results were compared to the performance of the classifier trained only with the target-dataset, without any balancing procedure, in terms of accuracy and weighted F_1 using a gain computation – a division between the metrics values obtained selecting from the union of the source-datasets with the target-dataset and the metrics using just the target-dataset.

The first experiment's results show that using a set of source-datasets to enrich the training set produces performance gains for most of target-datasets, both in terms of accuracy and F_1. However, the gain was not so high for most of the target-datasets, indicating that some selection strategies could be useful.

The results of the second experiment revealed that some combinations of method and percentage presented superior performance. The first criterion used was to select instances from the union of source-datasets at random. The best results were achieved with a selection of 5.0%. Using a closeness criterion, i.e., selecting instances from the union of source-datasets closer to the instances in target-dataset, the best results were obtained by selecting 20.0% of the instances. Then, criteria that consider diversity, selecting closer and farther instances, were tested varying the closest:farthest ratio. For a ratio 1:1, selecting 10.0% reached the best results, while a selected percentage of 20.0% performed the best results for the ratio 2:1. When the tested ratio was 3:1, the percentage that obtained better results was 40.0%, and for the 4:1 ratio, it occurs when selecting 20.0%. The ratios that select more farther instances than closer instances (namely, 1:2, 1:3, 1:4) achieved the best results for the percentages

Table XI. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 1:3.

Dataset	Accuracy										F_1									
	Selected percentages										Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.05**	**1.12**	**1.16**	**1.20**	**1.22**	**1.28**	**1.33**	**1.25**	**1.25**	**1.22**	0.85	**1.03**	**1.15**	**1.21**	**1.19**	**1.27**	**1.32**	**1.23**	**1.21**	**1.19**
sar	0.98	**1.04**	**1.10**	**1.08**	**1.12**	**1.18**	**1.20**	**1.20**	**1.24**	**1.26**	0.97	**1.05**	**1.12**	**1.10**	**1.16**	**1.23**	**1.25**	**1.25**	**1.29**	**1.31**
ais	0.98	0.99	0.99	0.98	0.98	0.98	0.99	0.99	**1.00**	**1.00**	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	**1.00**	**1.00**
S15	**1.01**	**1.01**	**1.01**	**1.00**	0.99	0.97	0.93	0.89	0.84	0.82	0.99	**1.00**	0.99	0.99	0.99	0.96	0.94	0.90	0.87	0.86
sen	**1.01**	**1.01**	**1.02**	**1.03**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.02**	**1.03**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**
per	**1.02**	**1.03**	**1.02**	**1.04**	**1.05**	**1.04**	**1.04**	**1.06**	**1.06**	**1.05**	**1.00**	**1.02**	**1.01**	**1.03**	**1.04**	**1.03**	**1.03**	**1.05**	**1.04**	**1.03**
hob	0.99	0.99	0.98	0.99	0.98	0.97	0.96	0.95	0.93	0.93	0.99	0.99	0.98	0.99	0.97	0.96	0.96	0.95	0.92	0.92
iph	**1.02**	**1.02**	**1.00**	**1.02**	**1.01**	**1.02**	**1.02**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.02**	**1.01**	**1.02**	**1.02**	**1.01**	**1.00**	**1.00**
mov	**1.04**	**1.03**	**1.03**	**1.03**	**1.03**	**1.03**	**1.06**	**1.05**	**1.05**	**1.04**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.05**	**1.04**	**1.04**	**1.03**
san	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**
nar	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.04**	**1.03**	**1.04**	**1.04**	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.03**	**1.04**	**1.04**
arc	**1.00**	**1.00**	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	**1.00**	**1.00**	0.99	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	0.99
S18	0.99	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.99	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99
OMD	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	0.99	0.99	0.98	0.98	0.97	0.99	0.99	**1.00**	**1.01**	**1.01**	0.99	0.99	0.98	0.97	0.96
HCR	**1.04**	**1.04**	**1.04**	**1.03**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.05**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
STS	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.00**	0.99	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.00**	0.99
SSt	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.01**	**1.02**	**1.02**	**1.03**	**1.03**	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.01**	**1.02**	**1.02**	**1.03**	**1.03**
Tar	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.98	0.98	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	0.98	0.98
Vad	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	0.99	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**
S13	**1.03**	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.05**	**1.04**	**1.04**	**1.02**	**1.02**	**1.02**	**1.03**	**1.02**	**1.03**	**1.03**	**1.04**	**1.03**	**1.03**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
										Avg gain	0.99	1.01	1.02	1.02	1.02	**1.03**	**1.03**	1.02	1.02	1.02
										#Gain	15	**19**	18	**19**	**19**	18	18	16	17	15
										#Best	6	8	4	**10**	6	7	**10**	4	5	6
										Ranking	5.41	4.09	4.36	2.55	3.41	3.05	**2.36**	4.23	4.27	5.27

Table XII. Accuracy and F_1 gains obtained with a selection of percentages from source-dataset that are closer and farther to target-dataset in a ratio 1:4.

Dataset	Accuracy									
	Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	**1.05**	**1.13**	**1.15**	**1.20**	**1.30**	**1.28**	**1.33**	**1.28**	**1.22**	**1.22**
sar	0.98	**1.00**	**1.00**	**1.10**	**1.12**	**1.20**	**1.14**	**1.20**	**1.24**	**1.28**
ais	0.98	0.99	**1.00**	0.99	0.99	0.97	0.99	0.98	**1.00**	**1.00**
S15	0.99	**1.01**	**1.01**	**1.00**	0.98	0.96	0.93	0.90	0.85	0.83
sen	**1.01**	**1.00**	**1.03**	**1.02**	**1.00**	**1.01**	**1.02**	**1.01**	**1.01**	**1.01**
per	**1.01**	**1.03**	**1.04**	**1.04**	**1.05**	**1.03**	**1.03**	**1.05**	**1.05**	**1.04**
hob	**1.00**	**1.00**	0.98	0.99	0.99	0.96	0.97	0.95	0.93	0.92
iph	**1.02**	**1.03**	**1.00**	**1.02**	**1.01**	**1.01**	**1.02**	**1.01**	0.99	**1.00**
mov	**1.04**	**1.04**	**1.04**	**1.03**	**1.04**	**1.03**	**1.05**	**1.06**	**1.06**	**1.05**
san	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.02**	**1.01**
nar	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**
arc	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	0.98
S18	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99	**1.00**	**1.00**	0.99
OMD	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	0.98	0.97
HCR	**1.05**	**1.04**	**1.04**	**1.04**	**1.04**	**1.04**	**1.05**	**1.05**	**1.05**	**1.04**
STS	**1.01**	**1.01**	**1.02**	**1.02**	**1.03**	**1.02**	**1.03**	**1.01**	**1.00**	0.99
SSt	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**
Tar	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	0.98
Vad	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**
S13	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**	**1.04**	**1.04**	**1.04**	**1.05**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**

Dataset	F_1									
	Selected percentages									
	0.0	0.5	1.0	2.5	5.0	10.0	20.0	40.0	80.0	100.0
iro	0.85	0.98	**1.03**	**1.20**	**1.27**	**1.27**	**1.32**	**1.25**	**1.19**	**1.19**
sar	0.97	**1.01**	**1.01**	**1.13**	**1.16**	**1.25**	**1.18**	**1.25**	**1.29**	**1.33**
ais	0.98	0.99	**1.00**	0.99	0.99	0.97	0.99	0.98	**1.00**	**1.00**
S15	0.96	**1.00**	0.99	0.99	0.97	0.96	0.94	0.91	0.88	0.87
sen	**1.01**	**1.00**	**1.03**	**1.02**	**1.00**	**1.01**	**1.02**	**1.01**	**1.01**	**1.01**
per	0.98	**1.02**	**1.03**	**1.03**	**1.04**	**1.02**	**1.03**	**1.04**	**1.04**	**1.03**
hob	**1.00**	**1.00**	0.98	0.99	0.99	0.96	0.97	0.95	0.93	0.92
iph	**1.01**	**1.02**	**1.00**	**1.02**	**1.01**	**1.01**	**1.02**	**1.01**	0.99	**1.00**
mov	**1.00**	**1.01**	**1.02**	**1.01**	**1.02**	**1.02**	**1.04**	**1.05**	**1.05**	**1.04**
san	**1.00**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.02**	**1.01**
nar	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.03**	**1.03**	**1.03**	**1.04**	**1.04**
arc	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	0.99	**1.00**	0.99	0.98	0.98
S18	0.99	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.99
OMD	0.99	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	0.97	0.97
HCR	**1.01**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**	0.99
STS	**1.00**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.01**	**1.00**	0.99
SSt	**1.00**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.03**
Tar	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	0.99	0.98	0.98
Vad	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**
S13	**1.02**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.03**	**1.04**	**1.03**	**1.04**
S17	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
S16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
Avg gain	0.99	1.01	1.01	1.02	**1.03**	**1.03**	**1.03**	1.02	1.02	1.02
#Gain	15	**20**	**20**	19	19	18	18	16	16	14
#Best	4	6	6	7	8	4	7	7	**9**	7
Ranking	6.00	4.36	3.95	3.09	**3.00**	3.73	3.09	4.05	4.41	5.36

40.0%, 20.0%, and 5.0%, respectively.

Analyzing the results from the second experiment, some interesting information can be observed. First of all, the best results for each criterion were achieved for percentages not greater than 40.0%, indicating that the performance can be degraded by noisy instances at some point. However, it is important noticing that the best performance is reached by a criterion that selects more farther instances than closer instances (the 1:4 ratio), which points out that diversity plays an essential role in the training set.

The comparison between the best results for all criteria shows that all of them achieved good performance, especially for the #Gain evaluation. Despite that, only one strategy (Div. 1:3 - 20%) presented statistically significant results compared to the baseline. One more interesting point is that

Table XIII. Comparison between best percentages according F_1 gain in all criteria.

Dataset	Criterion								
	Random (5.0%)	Closeness (20.0%)	Div. 1:1 (10.0%)	Div. 2:1 (20.0%)	Div. 3:1 (40.0%)	Div. 4:1 (20.0%)	Div. 1:2 (40.0%)	Div. 1:3 (20.0%)	Div. 1:4 (5.0%)
irony	**1.18**	**1.22**	**1.30**	**1.25**	**1.24**	**1.25**	**1.29**	**1.32**	**1.27**
sarcasm	**1.10**	**1.23**	**1.15**	**1.20**	**1.25**	**1.18**	**1.21**	**1.25**	**1.16**
ntua	0.97	0.99	0.99	0.99	**1.00**	**1.00**	0.99	0.99	0.99
SemEval15-Task11	0.97	0.92	0.95	0.91	0.90	0.92	0.89	0.94	0.97
sentiment140	**1.02**	**1.02**	**1.03**	**1.02**	**1.02**	**1.03**	**1.02**	**1.02**	**1.00**
person	**1.04**	**1.04**	**1.04**	**1.06**	**1.04**	**1.05**	**1.05**	**1.03**	**1.04**
hobbit	0.97	0.96	0.97	0.95	0.94	0.96	0.95	0.96	0.99
iphone	**1.00**	**1.04**	**1.01**	**1.02**	**1.01**	**1.03**	**1.01**	**1.02**	**1.01**
movie	**1.01**	**1.04**	**1.03**	**1.04**	**1.04**	**1.04**	**1.05**	**1.05**	**1.02**
sanders	**1.00**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.01**	**1.01**	**1.01**
Narr-KDML-2012	**1.02**	**1.03**	**1.02**	**1.03**	**1.04**	**1.03**	**1.04**	**1.03**	**1.02**
archeage	**1.00**	0.99	0.99	0.99	**1.00**	0.99	0.99	**1.00**	**1.01**
SemEval18	**1.01**	**1.01**	**1.00**	**1.01**	**1.01**	**1.02**	**1.00**	**1.00**	**1.00**
debate08	0.99	0.98	**1.00**	0.98	0.98	0.98	0.98	0.99	**1.00**
HCR	**1.01**	**1.00**	**1.00**	**1.01**	**1.01**	**1.00**	**1.02**	**1.00**	**1.00**
STS-gold	**1.01**	**1.01**	**1.02**	**1.01**	**1.01**	**1.01**	**1.01**	**1.02**	**1.02**
SentiStrength	**1.01**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
Target-dependent	**1.00**	**1.00**	**1.01**	**1.00**	0.99	**1.00**	**1.00**	**1.00**	**1.01**
VADER	**1.00**	**1.00**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.02**
SemEval13	**1.03**	**1.02**	**1.02**	**1.03**	**1.03**	**1.03**	**1.04**	**1.03**	**1.03**
SemEval17-test	**1.01**	**1.01**	**1.01**	**1.00**	**1.01**	**1.00**	**1.00**	**1.01**	**1.00**
SemEval16	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**	**1.01**
Avg gain	1.02	**1.03**	**1.03**	1.02	**1.03**	**1.03**	**1.03**	**1.03**	**1.03**
#Gain	18	17	18	17	18	18	17	18	**19**
#Best	3	5	8	3	6	5	7	8	**10**
Ranking	4.64	3.55	3.41	3.73	3.59	3.45	3.50	**2.73**	3.45

for irony and sarcasm, the smallest datasets, the results obtained were superior to the achieved for the 20 others. This may indicate that a smaller dataset can benefit most from this kind of instance-based transfer learning. Or even, the fact that these two datasets were from known, challenging domains to classify can show that enrichment is more critical for them than for other domains.

Future work may include the use of other metrics for selecting instances. In this work, we chose a simple representation for the datasets, but embeddings of sentences can be used to search for better results in training set enrichment. Furthermore, future research can focus on why some datasets results are so different from the others.

REFERENCES

Barreto, S., Moura, R., Carvalho, J., Paes, A., and Plastino, A. Sentiment analysis in tweets: an assessment study from classical to modern text representation models. *CoRR* vol. abs/2105.14373, 2021.

Bravo-Marquez, F., Frank, E., Mohammad, S. M., and Pfahringer, B. Determining word-emotion associations from tweets by multi-label classification. In *Proceedings of the 2016 IEEE/WIC/ACM Int. Conf. on Web Intelligence (WI)*. IEEE, pp. 536–539, 2016.

Carvalho, J. and Plastino, A. On the evaluation and combination of state-of-the-art features in twitter sentiment analysis. *Artificial Intelligence Review* vol. 54, pp. 1887–1936, 03, 2021.

Guimarães, E., Vianna, D., Paes, A., and Plastino, A. Enriching datasets for sentiment analysis in tweets with instance selection. In *Anais do IX Symposium on Knowledge Discovery, Mining and Learning*. SBC, Porto Alegre, RS, Brasil, pp. 73–80, 2021.

Guo, J., Shah, D., and Barzilay, R. Multi-source domain adaptation with mixture of experts. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. ACL, pp. 4694–4703, 2018.

Liu, B. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Studies in Natural Language Processing. Cambridge University Press, 2020.

Liu, M., Song, Y., Zou, H., and Zhang, T. Reinforced training data selection for domain adaptation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. ACL, pp. 1957–1968, 2019.

Martínez-Cámara, E., Martín-Valdivia, M., López, L., and Montejo-Ráez, A. Sentiment analysis in twitter. *Natural Language Engineering* vol. 20, pp. 1–28, 01, 2014.

Pan, S. J. and Yang, Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22 (10): 1345–1359, 2010.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* vol. 12, pp. 2825–2830, 2011.

Ruder, S., Ghaffari, P., and Breslin, J. G. Data selection strategies for multi-domain sentiment analysis. *CoRR* vol. abs/1702.02426, 2017.

Ruder, S. and Plank, B. Learning to select data for transfer learning with Bayesian Optimization. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. ACL, pp. 372–382, 2017.