2018 International Conference on Promising Electronic Technologies

Bias Detection of Palestinian/Israeli Conflict in Western Media

A Sentiment Analysis Experimental Study

Dr. Wael F. Al-Sarraj
Faculty of Information Technology
Islamic University of Gaza Gaza, Palestine
wsarraj@gmail.com

Eng. Heba M. Lubbad
Faculty of Information Technology
Islamic University of Gaza Gaza, Palestine
hebalubbad@gmail.com

***Abstract*—The online mass media plays a critical role in influencing the public opinion about controversial political events. Bias in press reports and articles to some ideological or political sides is common and opposites the neutrality nature of press and media. Bias can take different aspects and ways. One of the main aspects of press bias is using mislead terms and vocabularies. In summer 2014, Western media, news and press agencies covered Israeli war on Gaza. In general, Palestinian people complain that there is a notable bias in western media with the Israeli story and opinion and vice versa. In this research paper we report a text mining experimental study, that's have conducted on western media analysis to identify patterns in the press orientation and further in the media bias towards side to another. We have followed the text mining techniques and machine learning in an effort to detect the bias in news agencies. We have crawled news articles form seven major outlets in the western media. Then we have made preprocessing to convert them into useful structured form, building sentiment classifiers that be able to predict articles bias. In addition, we have compared three of supervised machine learning algorithms used in sentiment classification associated with different number of grams, where we have found that SVM with bio-gram gave the better outperformed outputs, with performance metrics are 91.76% accuracy, 88.33% recall and f-measure 91.46%.**

Keywords*— media bias, text mining sentiment analysis, machine learning, news domain, SVM, accuracy, recall and f-measure.

I. Introduction

The fourth power term called on news media because of its pivotal role in formation and influencing on public perception about current event, through manipulation of ideational information's to disseminate and reinforce ideology. Media outlet journalists or columnists express their opinions indirectly [2 ,1]. In most news articles journalists use a clear language to give an impression of objectivity, and state their opinions indirectly by embedding statements in a more complex discourse or argument structure [3].

During July 2014 Palestinian and Israeli had fought unequal military battle force continue for 51 day can be described a deadly and destructive for Palestinian people more than 2200 people were killed, and 11,000 were wounded, 2500 were lost their homes and become homeless ,for Israel people 64 soldiers and 6 civilians were killed, 720 were wounded, dependance on the Euro-Mediterranean Observatory for Human Rights comprehensive report[1].

Stand to reason in wartime reporting current events by journalists, would aim to exercise the utmost caution to make certain they report these charges thoroughly and fairly [4].Western media coverages accused be Pro-Israeli bias through supportive and adoption of the Israeli narrative in their news articles. Bias automatic detection in news domain through employment sentiment analysis for news reports and articles consider hard process according to varity, words similarities, words ambiguity and large amount of text to select biased pattern.

Many researcher's investigate with different text analytic techniques to discover news bias pattern as sentiment analysis by embedding machine learning algorithms with text mining techniques [3, 5], semantic similarity based on feature extraction from larges text and estimates of the relatedness between words and concepts [6], building graph model based on iterative algorithms computes bias to predict bias in news [7].

In this paper we explore news bias in western online news agencies covered Palestinian-Israeli 2014 conflict using text analysis technique. Our methodology based on using subjective sentiment analysis with different supervised machine learning, that classify bias polarity of news. Therefore, embedding text mining techniques which provide us to obtain biased patterns.

This paper is divided into five sections where section 2 reviews the related work, section 3 illustrates the research methodology, section 4 represents and discuss the results, section 5 Evaluation based on research results, and section 6 concludes the paper.

II. Related Works

Sentiment analysis aiming to determine the attitude (sense, emotion, opinion etc.) of a speaker or a writer with respect to a

[1] https://euromedrights.org

978-1-5386-5697-6/18/$31.00 ©2018 IEEE
DOI 10.1109/ICPET.2018.00024

specified topic has become a major area of interest in the field of NLP [3]. It is the computational study of people's opinions, appraisals, attitudes, and emotions toward entities. For this, there is need for data mining as well as text mining techniques [8]. Finding political bias in the published articles is a hard problem because of the huge text corpus from the articles and sparse hyperlink information [7]. News articles have received much less attention, although news bias across different news sources has been discussed by a few and some initial efforts have concentrated on sentiment analysis in the news area [9].

For example, [10] choose a supervised method to detect the opinion of news, by considering the score of the sentiment words and statistics analysis for some features such as the polarity of the title, the length of the news, content and structure features of the news. They select the good features from positive news and the negative news to make the classifier of sentiment analysis. Their proposed approach is self-growth graph algorithm used in sentiment detection beside integrating the lexicon-based methods and the supervised machine learning methods. Their results show C4.5 rule decision-tree method with self-growth algorithm is much better than SVM with self-growth.

Another example for news sentiment anlysis is [11],focus on under- resources languages by taken Norwegian ploitical news domain. Authors propose an unsupervised machine learning approach that takes sentiment-bearing Co-Occurring Terms (COTs) and negation words as features. Three different classifiers were evaluated J48, Random Forest, and Naïve Bayes. Their experiment yield precison over 70% performed by NB,concluded usage of negation words did not contribute will in classification task.

Also in another example [3], researchers compared four supervised machine learning algorithms of Naïve Bayes, Maximum Entropy, SVM and the character based N-Gram Language Model for sentiment classification of Turkish political news. From their empirical findings using different features, all the approaches reached accuracies of 65% to 77, SVM performs worse than other algorithms, for Sentiment Classification of short text SVM performs better than Naïve Bayes and Maximum Entropy, and for relative performance, the N-gram based character Language Model and Maximum Entropy performed better than Naïve Bayes and SVM.

Researchers at [12] make study for media bias classification to predict the American for Democratic Action (ADA) score of the source agency. They have 63% accuracy for political alignment by SVM polynomial classifier, that recorded under RapidMiner[2] tools which performed more reliable than python environment.

Quotes was extracted from news articles for news sentiment detaction by [13], authors argue that quotes is more subjective than the other parts of news articles.They find out smaller window sizes performs better than computing the overall sentiment of texts where the entities are mentioned.Thier classification proccess on using differrent sentiment dictionaries are WordNet, SentiWordNet, JRC Tonality and MicroWN, on a window of 6 words, where 82% accuracy obtioned by JRC Tonality and MicroWN.

Also researchers in [7] exploit quotes in articles , they implement graph approach predicting political Republican vs. Democratic bias of American news websites and political blogsusing the phrases quoted in their text. They argue their algorithm predicts labels better than supervised classification approach with accuracy 80.63%, where SVM recorded accuracy 73.23% that's indicating the advantage of using the graph structure in predicting bias.

[2] https://rapidminer.com

III. METHODOLOGY

Our experimental study focused to detect news bias through exploitation of sentiment analysis for unstructured text content, moreover embedded text mining techniques with machine learning algorithms to have acceptable performance for our model. We organize our implementation as illustrated by figure 1:

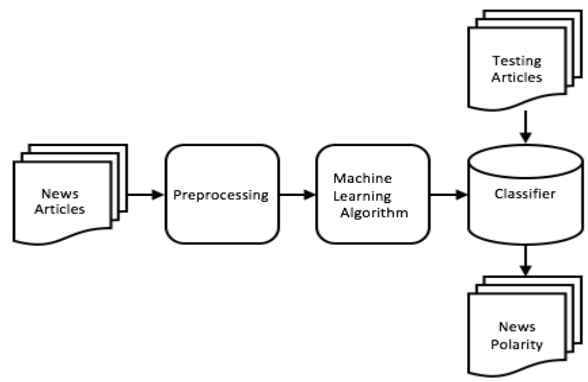

Figure 1:Methodology phases

A. News Articles

The first phase of our implementation involves news articles gathering from different famous online news outlets, namely Aljazeera America, ABC News and Washington Post represents American outlets, Aljazeera English and BBC represents British outlets, in addition to Global news and CTV represents Canadian outlets, from 1th July till 30th August 2014 include period of 50 day of Gaza war. Articles extracted using Outwit Hub[3], that extracts information elements from web pages and organizes them into usable collections. Every extracted article has tile, article, author, data and URL link written to excel file. As illustrated in Table I, statistics of extracted articles for each outlet, each outlet has splitted by 20% for training and 80% for testing model.

B. Preprocessing

At second phase we need to optimize the structure of articles and convert it from unstructured to more structured suitable for applying text mining algorithms. To process the document of articles to term vector of string features implement the following steps on data:

- Tokenization: split the text of articles, where streams of texts are broken into tokens.
- Filter tokens: filter the number of char for terms by length be in range (4:25).
- Transform: to reduce number of terms normalize letters to lowercase.
- Stop word removal: remove stop words from document to reduces the dimensionality of term

space, that does not give the meaning of the documents [14].

- N-Gram: a set of adjacent following items extracted from text, where similar words will have a high quantity of n-grams [14].
- Weighting: Indexing each article in document to terms vector of string features to reduce complexity and easier to handle, terms have weight factor for by Term Frequency–Inverse Document Frequency(TF-IDF), which reveals that a word is how important to a document.

Table I: Dataset statistics

Agency	*# Training*	*#Testing*	*# Total*
Aljazeera America	17	65	*81*
Washington post	27	95	*122*
ABC News	8	34	*42*
Aljazeera English	55	212	*267*
BBC	43	170	*213*
Global News	11	30	*41*
CTV	10	23	*33*
Total	*171*	*564*	*799*

C. Machine Learning Algorithms

The third phase, contians training part of sentiment classifiers till be able to predict the biased sentiment polarity of articles. We used 171 articles manually labled by expert for two polarites Pro-Palestinian and Pro-Israeli to feed the classifiers with training data. At our experiment , we choose three different supervised machine learning algorithms are Naïve Bias, SVM and Logisstic Regrssion SVM associated with unigrams , bi-grams and 3-Grams.

- *Naïve Bias Classifier:* a simple probabilistic classifier based on applying Bayes' theorem, characterized a high-bias, low-variance, and computationally inexpensive. Used in text classification first by equation 1it's find the probability of term assigned to class c, x term (feature), i position of term in document vector, then by equation 2 assign the term to the class c depending to occurence at all document.

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(x_i \mid c_j) \tag{1}$$

$$\hat{P}(w_i \mid c_j) = \frac{count(w_i, c_j)}{\sum_{w \in V} count(w, c_j)} \tag{2}$$

- *SVM:* is classifier formally defined by a separating hyperplane. At training part SVM classifier finds a linear separating hyperplane with the maximal margin in this higher dimensional space between the vectors of the two classes, where at testing part classifer be able to assign the target label based ondocument hyperplane where fall in. We used LibSVM with linear kernel which recommended for large scale linear classification.

- *Logisstic Regression SVM:*is logistic regression algorithm based on *SVM.* It is based on the solution of a dual problem using ideas similar to those of the Sequential Minimal Optimization algorithm for SVM f solving the optimization problem [15]. We choose this algorithm as it fast, robust, and scales good to large size data.

IV. EXPERMENTS

A. Experiment Setup

In order to evaluate our methodology, we make multible experiments divided for two-parts training classifiers under different N grams and then valiadate them by testing articles of outlets under trained classifiers. We used 171 articles for training classifier, which form 20% of articles total, where 86 articles labeled Pro-Israeli and 85 articles labeled Pro-Palestinian manually by expert of speech discourse. All sentiment classifications in our experiments are set up using the Rapid Miner framework, that consider integrated environment for machine learning, data mining, text mining and predictive analytics. Collected dataset was stored into excel file to fed for Rapid miner. Through preprocessing phase for dataset we use different n-Grams are unigram, bigram and 3 Gram. For training dataset, we get 9945 features for unigram,62835 for bigram and 131146 for 3-Gram.Then apply cross validation with 10 folds under different supervised algorism are SVM, Naïve bias and SVM Logistic Regression, for each fold data sampled by stratified sampling. Finally apply the trained classifier at testing data to extract the sentiments polarity for each article, where polarity with major confident value assigned as sentiment for article. We can see a sample of the most weighted features with SVM under 3 Grams at Table II, cease fire term mentioned Pro-Israeli as " violation of the ceasefire by Gazan armed militants groups", mentioned nutrality as " restore the cease-fire", rocket fire term mentioned Pro-Israeli as " civilians killed by rockets fired from Gaza", tunnels term mentioned Pro-Isreal as " destroy a sophisticated network of cross-border tunnels", hospital term mentiond Pro-Plaestinain as " a school in Maghazi, central Gaza, sheltering people" and militants term mentiond Pro-Israeli as " Palestinian militants in the Gaza Strip to infiltrate southern Israel".

Table II: Most Weighted Features by SVM with 3 Gram

Feature	*Weight*
Cease,	0.52
cease_fire	0.50
fire	0.56
gaza_strip	0.44
hamas	0.92
hospital	0.47
israeli	0.57

militants	0.60
rocket	0.59
rocket_fire	0.42
school	0.55
soldiers	0.41
tunnel	0.80

B. Evaluation

We investigate to evaluate our experimented classifiers through four metrics are accuracy, precision, recall and F-measure, these metrics associated with four terms are:

- True Positive (TP): for cases when actual and predicted target class are true.
- True Negative (TF): for cases when actual and predicted target class are false.
- False Positive (FP): for cases when actual class false and the predicted class are true, at these cases classifies predicted incorrectly with positive label.
- False Negative(FN): for cases when actual class true and the predicted class are false, at these cases classifies predicted incorrectly with negative label.

Accuracy: the ratio of correct prediction's made by classifier over all prediction's. as shown in equation 3.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{3}$$

Precision: express the most relevant true positive, its defined as ratio number of true positive over number true positive and true negatives as shown at equation 4.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

Recall: express the actual positive labeled successfully by the classifier, it's measured as ratio true positive over true positive and false negative as shown at equation 5.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

F-measure: express the precise of classifier, its measured as the mean of precision and recall as shown at equation 6.

$$F_1 = 2 * \frac{precision * recall}{precision + recall} \tag{6}$$

C. Evaluation Comparison

In our experiment study; we choose three supervised machine learning under different N-Gram and investigate through making comparison between them as illustrated by figure 9 under machine learning performance metrics. For accuracy superiority had been 91.76% for SVM with Unigram, although small differences exist for SVM between bio and 3 Gram. We suppose the SVM with the least feature number executes more well.

For precision superiority had been 100% with Logistic Regression SVM with unigram, Bio-Gram and 3-Gram, because FN was with 0 value, which means not any of training Pro-Israeli articles assigned by classifier false as Pro-Palestinian, but recall recorded the least values with LR-SVM with all N-Gram, because of the high number of Pro-Palestinian articles predicted as Pro-Israeli, consequently the worst F- measure recorded by LR-SVM algorithm. We noticed that with LR-SVM with 3-Gram the recall value improved consequently affected on F-measure.

We think the SVM algorithm was the more one stable with height metric values through all N-Gram. Also Naïve Bias executed well with stability, and its improve more well every time number of N- Gram value increase.

D. Results

We adopt the SVM and Naïve Bias results, because of their stability behavior through evaluation metrics, especially convergent values with unigram and Bio-Gram for both SVM and Naïve Bias, we select SVM with bio-gram outputs to discuss bias at testing articles for outlets. Figures [2-8] summarized the results of our conducted sentiment analysis experiments for outlets, where its contain quantitative information about each outlet polarity bias. ABC News agency have eight trained articles classified by expert as Pro-Israeli, at testing part we can notice dominant behavior is Pro-Israeli sentiment at figure 2, SVM with Bio-Gram performed with testing articles 11 Pro-Palestinian and 23 Pro-Israeli. Trained articles from BBC news are 34 Pro-Israeli and 9 Pro-Palestinian, figure 8 shows performed output of testing articles, where 56 articles are Pro-Palestinian and 114 articles are Pro-Israeli. Washington Post testing data classified by expert 26 articles are Pro-Israeli and 1 article Pro-Palestinian, testing articles classified 37 articles are Pro-Palestinian and 58 Pro-Israeli as figure 5 shows. CTV Canadian agency have 8 Pro-Israeli articles and 2 articles Pro-Palestinian at training dataset, testing dataset classified 18 articles Pro-Israeli and 5 articles Pro-Palestinian. Global News training dataset are 11 articles Pro-Israeli, testing articles classified 9 articles Pro-Palestinian and 21 articles Pro-Israeli as figure 6 shows. Aljazeera English outlet have 55 training articles are Pro-Palestinian, testing articles classified 203 articles Pro-Palestinian and 5 articles Pro-Israeli as figure 4 shows. Aljazeera America have 17 articles Pro-Palestinian at training dataset, testing articles classified 64 articles Pro-Palestinian as figure 3 shows. We can conclude from performed outputs for our conducted experiments, that's Pro-Israeli bias has the majority by 5 outlets and 2 for Pro-Palestinian through western media coverage for 2014 Palestinian–Israel conflict.

V. CONCLUTION AND FUTURE WORK

In this paper, we investigate about bias at western media coverage about Palestinian–Israeli conflict specially at 2014 war. Our conducted experiments, which adopt text mining techniques to extract the sentiment form gathered articles from multiple western media agencies. We make comparative evaluation between three supervised machine learning algorithms under accuracy, precision, recall and F-measure. Sentiment analysis results shows that there exits majority bias for Israeli through western media coverage. As future work, we intend to use feature extraction to extract the most effective features at articles, to minimize the numbers of features and results be more accurate and interpretable.

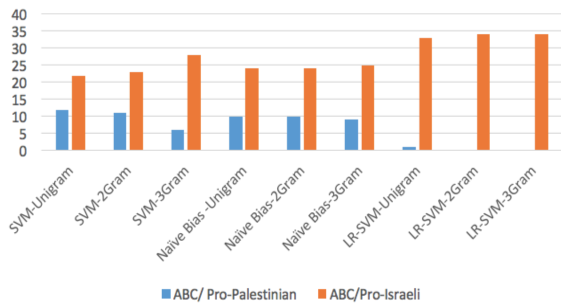

Figure 2: Sentiment analysis for ABC agency

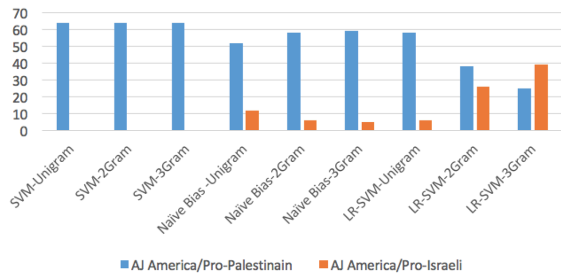

Figure 3: Sentiment analysis for Aljazeera America agency

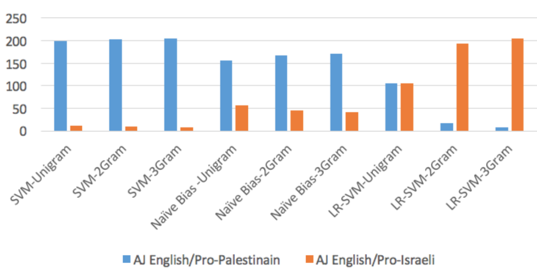

Figure 4: Sentiment analysis for Aljazeera English agency

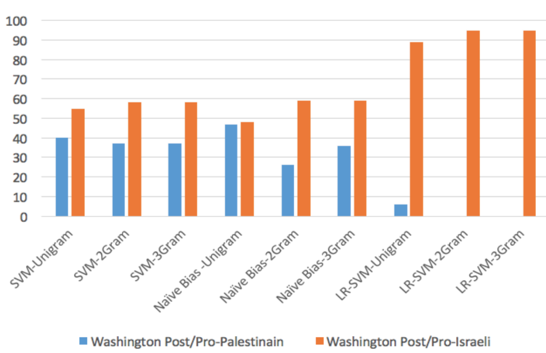

Figure 5: Sentiment analysis for Washington Post agency

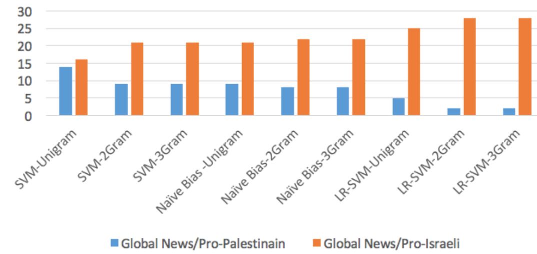

Figure6: Sentiment analysis for Global News agency

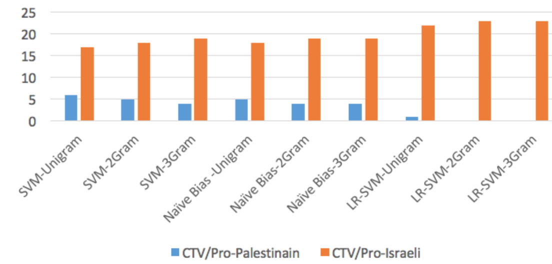

Figure7: Sentiment analysis for CTV agency

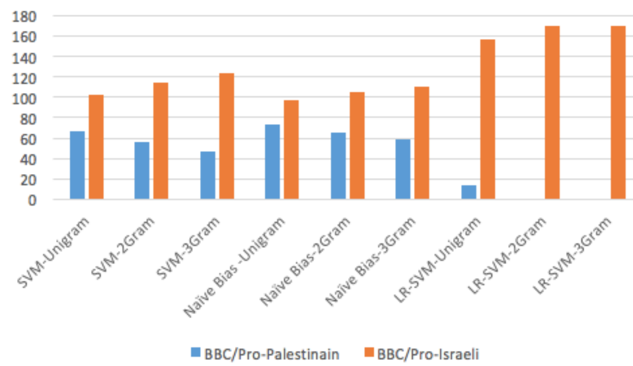

Figure 8: Sentiment analysis for BBC News agency

REFERENCES

[1] K. Lazaridou and R. Krestel, "Identifying Political Bias in News Articles," *Bulletin of the IEEE TCDL,* vol. 12, 2016.

[2] K. Lazaridou, R. Krestel, and F. Naumann, "Identifying Media Bias by Analyzing Reported Speech," in *Data Mining (ICDM), 2017 IEEE International Conference on*, 2017, pp. 943-948: IEEE.

[3] M. Kaya, G. Fidan, and I. H. Toroslu, "Sentiment analysis of turkish political news," in *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01*, 2012, pp. 174-180: IEEE Computer Society.

[4] S. M. Graber, "War of perception: a Habermasian discourse analysis of human shield newspaper reporting during the 2014 Gaza War," *Critical Studies in Media Communication,* vol. 34, no. 3, pp. 293-307, 2017.

[5] T. S. Zakzouk and H. I. Mathkour, "Comparing text classifiers for sports news," *Procedia Technology,* vol. 1, pp. 474-480, 2012.

[6] N. S. Holtzman, J. P. Schott, M. N. Jones, D. A. Balota, and T. Yarkoni, "Exploring media bias with semantic analysis tools: Validation of the Contrast Analysis of Semantic Similarity (CASS)," *Behavior Research Methods,* vol. 43, no. 1, pp. 193-200, 2011.

[7] S. Gupta, "Finding bias in political news and blog websites," ed, 2009.

[8] P. Tripathi, S. K. Vishwakarma, and A. Lala, "Sentiment analysis of english tweets using rapid miner," in *Computational Intelligence and Communication Networks (CICN), 2015 International Conference on*, 2015, pp. 668-672: IEEE.

[9] A. Balahur and R. Steinberger, "Rethinking Sentiment Analysis in the News: from Theory to Practice and back," *Proceeding of WOMSA,* vol. 9, 2009.

[10] L. Yan and Y. Zhang, "News Sentiment Analysis Based on Cross-Domain Sentiment Word Lists and Content Classifiers," in *International Conference on Advanced Data Mining and Applications*, 2012, pp. 577-588: Springer.

[11] P. F. Bakken, T. A. Bratlie, C. Marco, and J. A. Gulla, "Political News Sentiment Analysis for Under-resourced Languages," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, pp. 2989-2996.

[12] A. Dehghan, L. Montgomery, M. Arciniegas-Mendez, and M. Ferman-Guerra, "Predicting News Bias."

[13] A. Balahur *et al.*, "Sentiment analysis in the news," *arXiv preprint arXiv:1309.6202,* 2013.

[14] S. Vijayarani, M. J. Ilamathi, and M. Nithya, "Preprocessing techniques for text mining-an overview," *International Journal of Computer Science & Communication Networks,* vol. 5, no. 1, pp. 7-16, 2015.

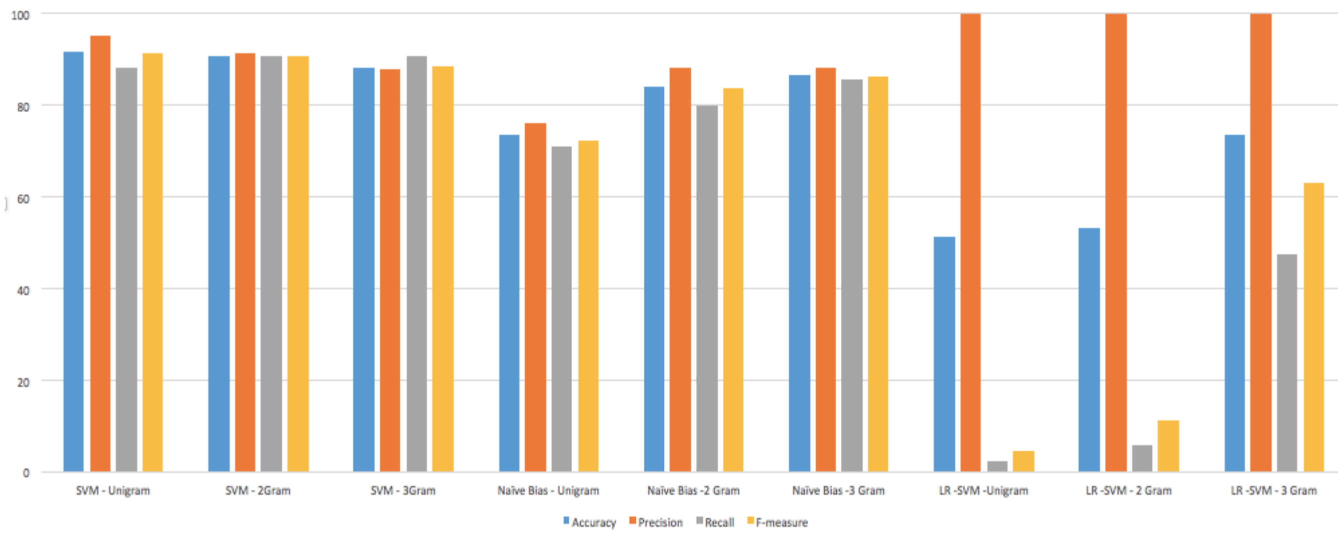

Figure 9: Supervised algorithms under evaluation metrics