

Automatic Thematic Content Analysis: Finding Frames in News

Daan Odijk¹, Bjorn Burscher², Rens Vliegenthart², and Maarten de Rijke¹

¹ Intelligent System Labs Amsterdam (ISLA), University of Amsterdam
d.odijk, derijke@uva.nl

² Department of Communication Science, University of Amsterdam
and Amsterdam School of Communication Research (ASCoR)
b.burscher, r.vliegenthart@uva.nl

Abstract. Framing in news is the way in which journalists depict an issue in terms of a ‘central organizing idea.’ Frames can be a perspective on an issue. We explore the automatic classification of four generic news frames: conflict, human interest, economic consequences, and morality. Complex characteristics of messages such as frames have been studied using thematic content analysis. Indicator questions are formulated, which are then manually coded by humans after reading a text and combined into a characterization of the message. We operationalize this as a classification task and, inspired by the way-of-working of media analysts, we propose a two-stage approach, where we first rate a news article using indicator questions for a frame and then use the outcomes to predict whether a frame is present. We approach human accuracy on almost all indicator questions and frames.

1 Introduction

There is a growing trend of applying computational thinking and linguistic approaches to social science research. In particular, language technology is proving to be a useful but underutilized approach that may be able to make significant contributions to research in a wide range of social science domains [2]. One particular domain in which this is happening is the study of news and its impact. Early examples focus mostly on analyzing factual aspects in news, such as [15], who analyzed the impact of news on corporate reputation by measuring the amount of news about specific issues. Increasingly, however, we are also seeing the use of language technology to analyze more subjective aspects of news for the purposes of social science research [12]. In this paper, we report on work aimed at analyzing the use of framing in news.

Framing in news is the way in which journalists depict an issue in terms of a ‘central organizing idea’ [6]. Frames can be regarded as a perspective on an issue. In the social sciences, mass communication (e.g., news) is often studied through a methodology called content analysis: “Who says what, to whom, why, to what extent and with what effect?” [11]. The aim of content analysis is to

systematically quantify specified characteristics of messages. When these characteristics are complex, thematic content analysis can be applied: first, texts are annotated for indicator questions (e.g., “Does the item refer to winners and losers?”) and the answers to such questions are subsequently aggregated to support a more complex judgment about the text (e.g., the presence of a conflict frame). Content analysis is a laborious process, and there is a clear need for a computational approach. This approach can improve the consistency, efficiency, reliability and replicability of the analyses, as larger volumes of news can be studied in a replicable manner, allowing the study of long-term trends.

For studying framing, some first computational approaches have been attempted, using dictionary-based methods. We approach the problem of frame detection in news as a two-stage classification task. We start by predicting the outcomes to indicator questions associated with a frame and then use the predicted outcomes to decide about the presence of the frame in a given text.

Our contribution in this paper consists in a two-stage approach to finding frames that allows us to answer the following research questions:

RQ1 Can we approach human performance on predicting answers to indicator questions?

RQ2 Can we approach human performance on predicting the presence of a frame?

The remainder of this article is organized as follows: in Section 2 we discuss media frame analysis and related work; Section 3 describes our proposed methods and Section 4 describes the experimental setup. We present and discuss our results in Section 5, after which we conclude in Section 6.

2 Media Frame Analysis

News coverage can be approached as an accumulation of “interpretative packages” in which journalists depict an issue in terms of a *frame*, i.e., a central organizing idea [6]. Frames are the dependent variable when studying the process of how frames emerge (*frame building*) and the independent variable when studying effects of frames on predispositions of the public (*frame setting*) [22]. When studying the adoption of frames in the news, content analysis of news media is the most dominant research technique.

Using questions as indicators of news frames in manual content analysis is the most widely used approach to manually detecting frames in text. Indicator questions are added to a codebook and answered by human coders while reading the text unit to be analyzed [26]. Each question is designed such that it captures the semantics of a given frame. Generally, several questions are combined as indicators for the same frame. This way of making inferences from texts is also referred to as thematic content analysis [20].

Automatic or semi-automatic frame detection is rare. The approaches that do exist follow a dictionary-based or rule-based approach. E.g., Ruigrok and Van Atteveldt [21] define search strings for the automatic extraction of a priori

defined concepts in newspaper articles, and then apply a probabilistic measure to indicate associations between such concepts. Similarly, Shah et al. [25] first define “idea categories,” then specify words that reveal those categories, and finally, program rules that combine the idea categories in order to give a more complex meaning as a frame.

In this paper we focus on four commonly used frames [24]. For convenience they are listed in Section 4, together with their indicator questions. The *conflict frame* highlights conflict between individuals, groups or institutions. Prior research has shown that the depiction of conflict is common in political news coverage [16], and that it has inherent news value [4, 27].

By emphasizing individual examples in the illustration of issues, the *human interest frame* adds a human face to news coverage. According to Iyengar [7], news coverage can be framed in a thematic manner, taking a macro perspective, or in an episodic manner, focusing on the role of the individual concerned by an issue. Such use of exemplars in news coverage is observed by several scholars [16, 24, 28] and connects to research on personalization of political news [7].

The *economic consequence frame* approaches an event in terms of its economic impact on individuals, groups, countries or institutions. Covering an event with respect to its consequences is argued to possess high news value and to increase the pertinence of the event among the audience [5].

The *morality frame* puts moral prescriptions or moral tenets central when discussing an issue or event. Morality as a news frame has been studied in various academic publications and is found to be applied in the context of various issues as, for example, gay rights [17] and biotechnology [1].

Over the past decade, language technology has witnessed a rapid broadening, along two dimensions. First, moving beyond an almost exclusive focus on working with news corpora, different genres of text are now being subjected to, e.g., semantic analysis [13, 14]. Second, from a strong focus on analyzing facts the field is broadening to also include more subjective aspects of language, such as opinions and sentiment [19], human values [3], argumentation [18] and user experiences from online forums [8].

In this paper, we contribute over and above the related work discussed, by presenting and evaluating an ensemble-based classification approach for frame detection in news. To the best of our knowledge, this is the first work in which statistical classification methods are applied to this central issue in studying media. Furthermore, we investigate whether explicitly modeling the thematic content analysis approach improves performance.

3 Frame Classification

We approach the task of frame detection in news as a classification task. The assumption underlying thematic content analysis is that frames manifest themselves in a news article in a manner that is measured using indicator questions. We follow this assumption and analyze the wording in a news article in order to make a decision about the presence of frames.

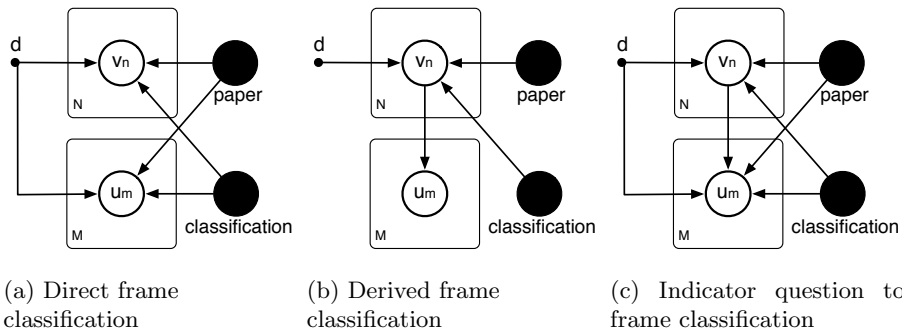


Fig. 1: Graphical models of the three classification approaches. The circles represent random variables, where the filled are observable. The rectangular plates indicate multiple of these variables.

Given a collection of documents D and a set of frames U for which a set of indicator questions V have been defined, we estimate the probability $P(u_m|d)$ that a frame $u_m \in U$ is present in document $d \in D$. In thematic content analysis this probability is deconstructed into $P(u_m|v_1, \dots, v_N)$ and a set of probabilities for each $v_n \in V$: $P(v_n|d)$. This formal definition of the task can be used for the automatic classification and for manual content analysis. In the latter case, the probability $P(v_n|d)$ is estimated using manually coding by humans after reading the document.

Document representation. We represent documents as a bag-of-words with TF.IDF scores for each word. We apply sublinear term frequency scaling, i.e., replace TF with $1 + \log(TF)$, use l2 normalization and smooth IDF weights by adding one to document frequencies. We have evaluated other representation (e.g., n-grams and topic models), but these did not improve classification performance and will not be reported here.

Besides the words represented in the document, we extend the document representation with information on the source of the document and with a classification for each document (i.e., a topic, such as finance, infrastructure, etc.). The extended bag-of-words document representation serves as the features for classification.

Frame and indicator question classification. We propose two baselines and three approaches for automatic indicator question and frame classification. These methods differ in how the coherence between indicator question and frame is modeled. The approaches are depicted in graphical models in Figure 1 and will be described below.

Stratified random classification baseline. Our first baseline approach is very naive and intended as a lower bound. It randomly chooses the answer to a indicator question or whether a frame is present or not, taking into account only the prevalence in training set. This naive baseline randomly assigns a classifica-

tion, without considering the document and its representation, with a probability based on the class distributions. This naive baseline will be more likely to randomly pick the majority class than the minority class.

Direct classification baseline. Our second baseline approach is to classify answers to indicator questions and the presence of each frame directly. More formally, we train a classifier to estimate $P(u_m|d)$ for each frame $u_m \in U$. This approach is the simplest approach and is depicted in Figure 1a. Note that for frames, we completely ignore the indicator questions in this baseline approach.

For classification we use Logistic Regression to optimize logistic loss using Pegasos-style regularization. For training we alternate between pairwise ROC-optimization and standard stochastic gradient steps on single examples [23]. This baseline approach aims to be flexible in dealing with issues such as class imbalance.

Ensemble-based direct classification. Our first approach is to improve binary classification decisions for indicator questions and for the presence of a frame by using an ensemble of binary-class linear classifiers (also depicted in Figure 1a). The predictions of all these classifiers are the features for a final classifier. The ensemble includes different linear support vector machines (SVMs), linear rank-based SVMs [9, 23], and Perceptron-based algorithms [10]. This ensemble-based approach aims to be flexible in dealing with the different complex characteristic of each of the classifications. We combine the classifiers in the ensemble using the same classifier as described above for the baseline approach.

Derived frame classification. Our second approach is to derive the presence or absence of a frame based on the classification for indicator questions. More formally, we train an ensemble-based classifier to estimate $P(\hat{v}_n|d)$ for each indicator question $v_n \in V$. We then derive the probability $P(u_m|d)$ for each frame $u_m \in U$ from $P(\hat{u}_m|\hat{v}_1, \dots, \hat{v}_N)$ for all indicator questions $v_m \in V$. This approach is depicted in Figure 1b and closely resembles the manual approach, where human coders make binary decisions for $P(v_n|d)$ for each $v_n \in V$ and $d \in D$.

Indicator question to frame classification. Our third approach is a cascade approach, where we first classify for the indicator question and then use the outcomes to classify the frames. More formally, we train an ensemble-based classifier to estimate $P(\hat{v}_n|d)$ for each indicator question $v_n \in V$. We then train an ensemble-based classifier to estimate the probability $P(u_m|d, \hat{v}_1, \dots, \hat{v}_N)$ for each frame $u_m \in U$. This approach is depicted in Figure 1c. Practically, we implement this by adding ensemble-based predictions for indicator questions as features for the frames classifiers.

4 Experimental Setup

To evaluate our methods we run a number of experiments. We describe the document collection used, outline how the four frames have been coded in the

manual content analysis that we use as training and test data, and explain how we evaluate the performance of our classification models.

Document collection. Our document collection consists of digital versions of front page news articles of three Dutch national daily newspapers (*De Volkskrant*, *NRC Handelsblad* and *De Telegraaf*) for the period between 1995 and 2011. These articles come from the Dutch Lexis-Nexis newspaper archive, and each article has a topical classification (based, e.g., on the location in the newspaper). The used sample is a stratified sample of 13% for each year.

Indicator Questions Annotations. For each year covered in our collection, a random sample of news articles was taken. This sample was filtered (based on manually assigned labels) to only contain articles that were political in nature. The resulting 5,875 documents have been manually coded for the presence of four generic news frames (described in Section 2). Indicator questions were used to code the news frames.

A total of thirteen yes-or-no-questions were used as indicators of the news frames. In previous research, these questions have been shown to be reliable indicators of the four frames [24]. The indicator questions for each frame are:

C *Conflict frame:*

C1 Does the item reflect disagreement between parties, individuals, groups or countries?

C2 Does the item refer to winners and losers?

C3 Does the item refer to two sides or more than two sides of the problem?

E *Economic consequence frame:*

E1 Is there a reference to the financial costs/degree of expense involved, or to financial losses or gains, now or in the future?

E2 Is there a reference to the non-financial costs/degree of expense involved, or to non-financial losses or gains, now or in the future?

E3 Is there a reference to economic consequences of pursuing or not pursuing a course of action?

H *Human interest frame:*

H1 Does the item provide a human example or human face on the issue?

H2 Does the item employ adjectives or personal vignettes that generate feelings of outrage, empathy caring?

H3 Does the item mention how individuals and groups are affected by the issue or problem?

H4 Does the item go into the private or personal lives of the actors?

M *Morality frame:*

M1 Does the item contain any moral message?

M2 Does the item make reference to morality, God or other religious tenets?

M3 Does the item offer specific social prescriptions about how to behave?

Manual coding was conducted by a total of 30 trained coders. All coders were communication science students and native speakers of the Dutch language. In order to assess inter-coder reliability, a random subset of 159 articles was coded by multiple coders. Measures of the percentage of inter-coder agreement range

from 70% to 94%. The inter-coder reliability is included in the results in Table 1 and Table 3, with the label ‘Human.’

Frame Annotations. Based on the annotations for indicator questions, a second annotation round gave rise to the construction of frame annotations, following the methodology described in [24]. To establish the coherence of the indicator questions and their relation to the frames a factor analysis is performed. We find a four factor solution for the answers to the indicator questions. In this solution each indicator question has a loading onto each factor (i.e., a weight).

In these factor loadings, we can identify the four frames, i.e., for each frame there is a factor with high loads for the corresponding indicator questions and low loadings for the others. For two indicator questions (C2 and E2) the factor load is below 0.5, and hence these were considered unreliable indicators (in line with [24]). This means that the remaining indicator questions can be considered reliable indicators of the four frames: a frame is considered present in a news document whenever any of the indicator questions corresponding to the frame is answered positively.

Evaluation metrics. We perform ten-fold cross-validation and compare the agreement between human annotators and our automatic approach in terms of agreement. Where possible, we evaluate both the answers to indicator questions and the frame annotations. Furthermore, we compare the approaches in receiver operating characteristics (ROC) space. We compare the ability to distinguish true positive classifications from false positives for different operating characteristics that will produce increasingly more positive results. In this ROC space, we can compute the area under the curve (AUC). The AUC metric for a classifier expresses the probability that the classifier will rank a positive document above a negative document.

5 Results and Discussion

Table 1 and Table 3 describe the agreement between our approaches and the human annotations for each of the eleven indicator questions and the four frames. For comparison, these tables also include the inter-annotator agreement for human coders. Table 2 and Table 4 describe the area under the curve (AUC) metric for our approaches.

Indicator Questions Classification Results. We can observe in Table 1 that our baseline single classifier direct approach (“Direct”) performs well on some of the indicator questions, but worse on others. The direct baseline is unable to consistently improve over the naive stratified random baseline (“Random”). Our ensemble-based approach (“Ensemble”) substantially improves over these baselines and achieves accuracy scores ranging from 65% accuracy upwards. While we observe that the accuracy varies among the four frames and the corresponding indicator questions, our ensemble-based approach is able to capture the complex characteristics for all questions and frames. The conflict indicator questions (C1 and C3) and human interest question H3 perform below average in the baselines, but perform substantially better in the ensemble-based approach.

Table 1: Agreement between automatic classification predictions and human annotations for each of the eleven indicator questions and the three approaches (two baselines and ensemble).

	C1	C3	E1	E3	H1	H2	H3	H4	M1	M2	M3
Random	.5214	.5980	.7093	.8419	.7963	.8346	.5144	.9122	.9348	.9397	.9535
Direct	.5709	.6140	.7093	.8419	.7963	.8346	.5750	.9122	.9348	.9397	.9535
Ensemble	.7064	.6945	.8511	.8650	.8007	.8393	.6489	.9137	.9345	.9460	.9535
<i>Coder biased ensemble run is included below for analysis.</i>											
Biased	.7200	.7413	.8553	.8819	.8213	.8494	.7045	.9185	.9346	.9501	.9525
<i>Human inter-coder agreement is included below for comparison. Note that this is evaluated on a small dataset.</i>											
Human	.7239	.6994	.8282	.8466	.7546	.7055	.6748	.8405	.9080	.9041	.9202

Table 2: Area under the curve (AUC) for ROC of automatic classification predictions compared to human annotations for each of the eleven indicator questions and the two direct approaches (baseline and ensemble).

	C1	C3	E1	E3	H1	H2	H3	H4	M1	M2	M3
Direct	.6235	.6601	.6973	.6885	.6283	.5802	.6027	.5960	.5572	.6591	.4903
Ensemble	.7744	.7672	.8966	.8432	.7483	.7419	.7051	.7990	.6917	.8884	.6509

Human interest question H4 and the morality questions (M1, M2 and M3) show high baseline performance, but do not show substantial improvements for the direct approaches, despite our pairwise optimization approach. This suggests that these questions are underrepresented and possibly less well represented using a bag-of-words approach than the other questions.

Looking at the AUC metric results in Table 2, we see the same substantial improvements of the ensemble-based approach over the direct classification baseline. We can also observe a substantial improvement for the aforementioned indicator questions H4, M1, M2 and M3. This suggest that while we are not better in terms of accuracy for these questions, we are indeed better at estimating the probability of a document belonging to a class.

Frame Classification Results. We can observe in Table 3 that accuracy scores on frames follow the same pattern as the indicator questions. The conflict and human interest frame prediction again performs worse than the others. Interestingly, we can observe a substantial improvement for the morality frame over the stratified random baseline. The ensemble-based approach is able to obtain substantial improvements over the baselines approaches. We can also observe that deriving the scores from the indicator questions does not perform well, directly predicting scores for frames using the ensemble-based approach performs substantially better. Interestingly, the two-stage indicator question to

Table 3: Agreement between automatic classification predictions and human annotations for each of the four frames and the five frame classification approaches.

	C	E	H	M
Random	.6403	.5755	.6231	.8679
Direct	.6654	.8134	.7779	.9668
Ensemble	.7241	.8506	.7949	.9668
Derived	.5709	.7093	.6158	.9348
IQ \rightarrow F	.7202	.8489	.8014	.9677
<i>Coder biased ensemble run is included below for analysis.</i>				
Biased	.7501	.8642	.8141	.9685
<i>Human agreement on small dataset included for comparison.</i>				
Human	.7730	.8160	.6442	.8528

Table 4: Area under the curve (AUC) for ROC of automatic classification predictions compared to human annotations for each of the four frames and four frame classification approaches.

	C	E	H	M
Direct	.6379	.6956	.6008	.5909
Ensemble	.7802	.8496	.7580	.7597
Derived	.5575	.5000	.5897	.5000
IQ \rightarrow F	.7677	.8436	.7748	.8025

frame classification approach does not perform better than the direct approach. The additional information we add by first classifying the indicator questions does not help in classifying the frames. The results for the AUC metric (described in Table 4) show a qualitatively similar pattern as the agreement.

Furthermore, we can observe from Table 1 and Table 3 that the morality frame and the corresponding questions perform strikingly well in all approaches in terms of agreement. A plausible explanation for this is that this frame is a lot less prevalent than the other three (present in 13% of the documents, compared to 64% for conflict, 58% for economic consequence and 62% for human interest). The AUC results in Table 2 and Table 4 provide some evidence that these classifiers still perform up to par.

To validate this, and to obtain more insight into the operation characteristics of the classifiers we take a more detailed look at the ROC curves. Figure 2 shows these ROC curves for the direct ensemble-based approach for the frames. We can observe a similar curve for each of the frames. From these graphs and the AUC results, we can conclude that while we can not perfectly classify the frame annotations, we are able to obtain a good rate of true positives if we allow some false positives.

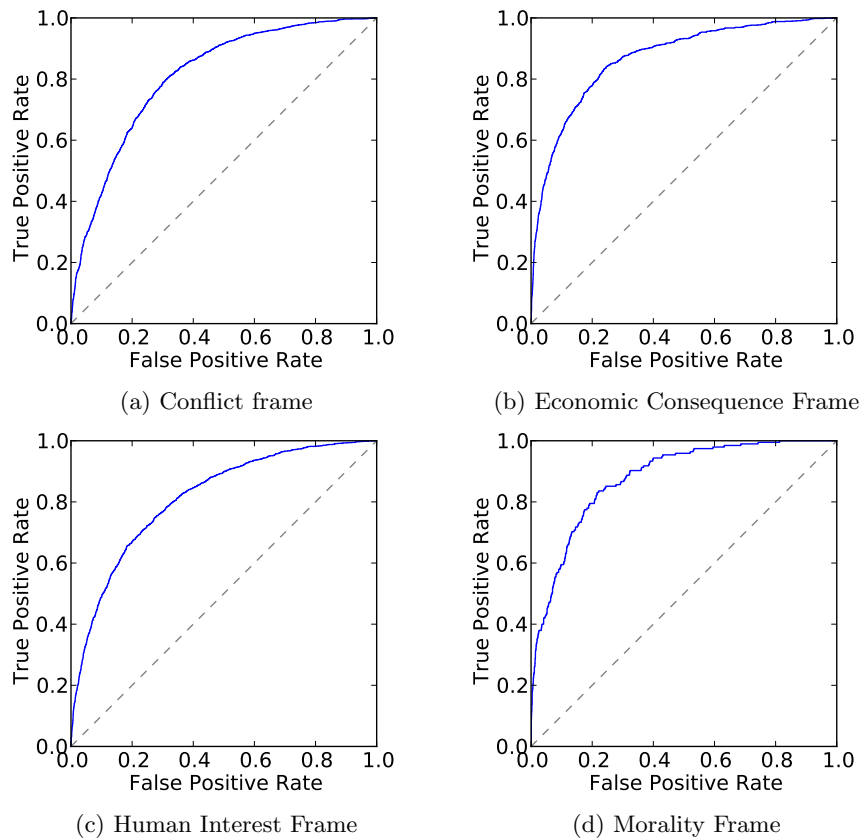


Fig. 2: ROC Curves for the ensemble-based direct approach for the four frames.

Human Inter-Coder Agreement. Compared to human inter-coder agreement, nearly all accuracy scores for the ensemble-based and two-stage approaches are at or above that level. Note, however, that human agreement is evaluated on a much smaller dataset. We observe a lower performance compared to human agreement for question H3, the conflict frame and corresponding questions C1 and C3. For the morality frame and the human interest and morality questions the human inter-coder agreement is even below the stratified random baseline.

To investigate the difficulty of each question and the quality of the human annotations, we look at whether the annotations for questions are stable across coders. We measure this by evaluating a new ensemble-based model where the document representation is extended with variables representing the coder. This creates the unrealistic but insightful scenario where we predict the answer of a specific coder to a specific question. This model allows us to compensate for a bias a coder might have, possibly resulting in higher performance compared to the regular ensemble-based approach.

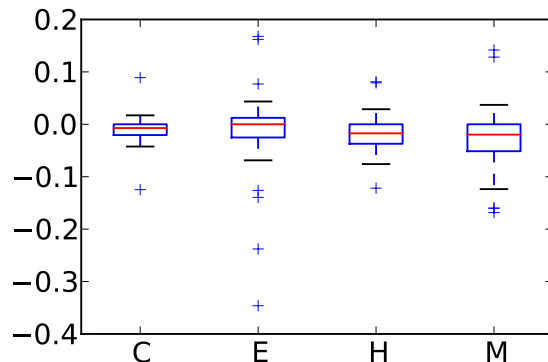


Fig. 3: Box plot of the weights on the binary coder variables for the four frames in one of the ensemble SVM classifiers.

Agreement for the biased model is included in Table 1 and Table 3. We observe increased performance for most questions, with C3, E3 and H3 standing out. For frames the performance is increased for the human interest frame, economic consequence frame and most substantially for the conflict frame.

To further investigate this, we look at the weights that the coder features get assigned in the biased model. If all coders would answer the indicator questions exactly the same, the coder features will have a weight very close to zero. A weight that differs from zero suggests a consistent difference in answers from one coder compared to the other coders. Figure 3 shows these weights for each frame in one of the classifiers in the direct frame ensemble classifier. We see that the weights do indeed deviate from zero, with a different range per frame. The economic consequence frame has the highest range, with a maximum of 0.5 bias per coder on a scale of -1 to 1 . These weights suggest consistently different interpretations of the indicator questions across coders.

6 Conclusion

We have proposed algorithmic approaches to finding frames in news that follow the manual thematic content analysis approach. Our results provide strong evidence that we are able to approach human performance on predicting both the answers to indicator questions as well as the presence of a frame.

Our ensemble-based approach to directly predicting the presence of a frame is the most effective and improves substantially over the baseline approach. The derived approach, which directly follows the manual approach, was the least effective. Surprisingly, the more informed indicator question to frame classification approach did not perform better than the ensemble-based direct classification approach. This suggests that for the task of frame classification, explicitly modeling the manual thematic content analysis does not improve performance. Our

ensemble-based direct classification approach is sufficient to capture the complex characteristics of frames that the indicator questions are aimed to represent.

The results of an analysis using a model that explicitly models coder bias and the relatively low inter-coder agreement suggest that coders have different interpretations of the indicator questions for the frames. Like the indicator questions that represent different aspects of complex characteristics of messages, it seems that human coders represent different views on these aspects and characteristics.

Finally, for the task of frame detection in news, we have shown that using an ensemble-based classification approach we are able to approach human performance in terms of accuracy on this task. A combined approach of human and automated frame detection seems to be the logical way forward.

Acknowledgments This research was supported by the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreements nr 258191 (PROMISE Network of Excellence) and 288024 (LiMoSINe project), the Netherlands Organisation for Scientific Research (NWO) under project nrs 640.004.802, 727.011.005, 612.001.116, HOR-11-10, 451-09-011, the Center for Creation, Content and Technology (CCCT), the QuaMerdes project funded by the CLARIN-nl program, the TROVe project funded by the CLARIAH program, the Dutch national program COMMIT, the ESF Research Network Program ELIAS, the Elite Network Shifts project funded by the Royal Dutch Academy of Sciences (KNAW), the Netherlands eScience Center under project number 027.012.105 and the Yahoo! Faculty Research and Engagement Program.

7 References

- [1] P. R. Brewer. Framing, value words, and citizens’ explanations of their issue opinions. *Political Communication*, 19(3):303–316, 2002.
- [2] A.-S. Cheng, K. R. Fleischmann, P. Wang, and D. W. Oard. Advancing social science research by applying computational linguistics. In *Proceedings of the Annual Conference of the American Society for Information Science and Technology*, 2008.
- [3] K. R. Fleischmann, D. W. Oard, A.-S. Cheng, P. Wang, and E. Ishita. Automatic classification of human values: Applying computational thinking to information ethics. *Proceedings of the American Society for Information Science and Technology*, 46(1):1–4, 2009.
- [4] J. Galtung and M. H. Ruge. The structure of foreign news. *Journal of Peace Research*, 2(1):64–90, 1965.
- [5] W. Gamson. *Talking Politics*. Cambridge University Press, 1992.
- [6] W. Gamson and A. Modigliani. Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, pages 1–37, 1989.
- [7] S. Iyengar. *Is anyone responsible?: How television frames political issues*. University of Chicago Press, 1991.

- [8] V. Jijkoun, M. de Rijke, W. Weerkamp, P. Ackermans, and G. Geleijnse. Mining user experiences from online forums: an exploration. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, pages 17–18. Association for Computational Linguistics, 2010.
- [9] T. Joachims. Optimizing search engines using clickthrough data. In *SIGKDD'02*, pages 133–142. ACM, 2002.
- [10] W. Krauth and M. Mézard. Learning algorithms with optimal stability in neural networks. *Journal of Physics A: Mathematical and General*, 20(11):L745, 1999.
- [11] H. D. Lasswell. The structure and function of communication in society. *The communication of ideas*, 37, 1948.
- [12] D. Lazer, A. S. Pentland, L. Adamic, S. Aral, A. L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al. Life in the network: the coming age of computational social science. *Science (New York, NY)*, 323(5915):721, 2009.
- [13] E. Meij, M. Bron, B. Huurnink, L. Hollink, and M. de Rijke. Learning semantic query suggestions. In *8th International Semantic Web Conference (ISWC 2009)*, October 2009.
- [14] E. Meij, W. Weerkamp, and M. de Rijke. Adding semantics to microblog posts. In *WSDM 2012: Fifth ACM International Conference on Web Search and Data Mining*, February 2012.
- [15] M. Meijer and J. Kleinnijenhuis. Issue news and corporate reputation: Applying the theories of agenda setting and issue ownership in the field of business communication. *Journal of Communication*, 56(3):543–559, 2006.
- [16] W. R. Neuman, M. R. Just, and A. N. Crigler. *Common knowledge: News and the construction of political meaning*. University of Chicago Press, 1992.
- [17] M. C. Nisbet and M. Huge. Attention cycles and frames in the plant biotechnology debate managing power and participation through the press/policy connection. *The Harvard International Journal of Press/Politics*, 11(2):3–40, 2006.
- [18] R. M. Palau and M.-F. Moens. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM, 2009.
- [19] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135, 2008.
- [20] C. Roberts. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*. Lawrence Erlbaum, New York, 1997.
- [21] N. Ruigrok and W. Van Atteveldt. Global angling with a local angle: How US, British, and Dutch newspapers frame global and local terrorist attacks. *The Harvard International Journal of Press/Politics*, 12(1):68–90, 2007.
- [22] D. A. Scheufele. Framing as a theory of media effects. *Journal of Communication*, 49(1):103–122, 1999.

- [23] D. Sculley. Combined regression and ranking. In *SIGKDD'10*, pages 979–988. ACM, 2010.
- [24] H. A. Semetko and P. M. Valkenburg. Framing european politics: A content analysis of press and television news. *Journal of Communication*, 50(2):93–109, 2000.
- [25] D. V. Shah, M. D. Watts, D. Domke, and D. P. Fan. News framing and cueing of issue regimes: Explaining clinton’s public approval in spite of scandal. *Public Opinion Quarterly*, 66(3):339–370, 2002.
- [26] A. Simon and M. Xenos. Media framing and effective public deliberation. *Political Communication*, 17(4):363–376, 2000.
- [27] R. Vliegthart, H. G. Boomgaarden, and J. W. Boumans. *Changes in political news coverage*. Palgrave Macmillan, 2011.
- [28] D. Zillmann and H. B. Brosius. *Exemplification in communication*. Hogrefe and Huber, 2000.