Q1 Please enter your group id (please use exactly the same string you used in your submission file(s)).

Answered: 20   Skipped: 0

#	RESPONSES		DATE
1	COMPLETE	iam	8/10/2018 3:32 PM
2	COMPLETE	uhhLT	8/7/2018 7:49 PM
3	COMPLETE	fkieITF	8/7/2018 7:18 PM
4	COMPLETE	HIIwiStJS	8/7/2018 7:11 PM
5	COMPLETE	upInf	8/7/2018 4:15 PM
6	COMPLETE	dfkilt	8/7/2018 4:00 PM
7	COMPLETE	cluzh	8/7/2018 3:55 PM
8	COMPLETE	FoSIL	8/7/2018 1:27 PM
9	COMPLETE	spMMMP	8/7/2018 10:37 AM
10	COMPLETE	TUWienKBS	8/7/2018 9:26 AM
11	COMPLETE	Potsdam	8/7/2018 12:55 AM
12	COMPLETE	SaarOffDe	8/6/2018 6:52 PM
13	COMPLETE	InriaFBK	8/6/2018 2:46 PM
14	COMPLETE	RuG	8/6/2018 11:07 AM
15	COMPLETE	hshl	8/6/2018 7:42 AM
16	COMPLETE	HaUA	8/2/2018 9:47 PM
17	COMPLETE	hpiTM	8/1/2018 3:33 PM
18	COMPLETE	hda	7/31/2018 5:55 PM
19	COMPLETE	UdSW	7/31/2018 2:08 PM
20	COMPLETE	KAUSTmine	7/27/2018 8:22 PM

Q2 Please enter your affiliation.

Answered: 20   Skipped: 0

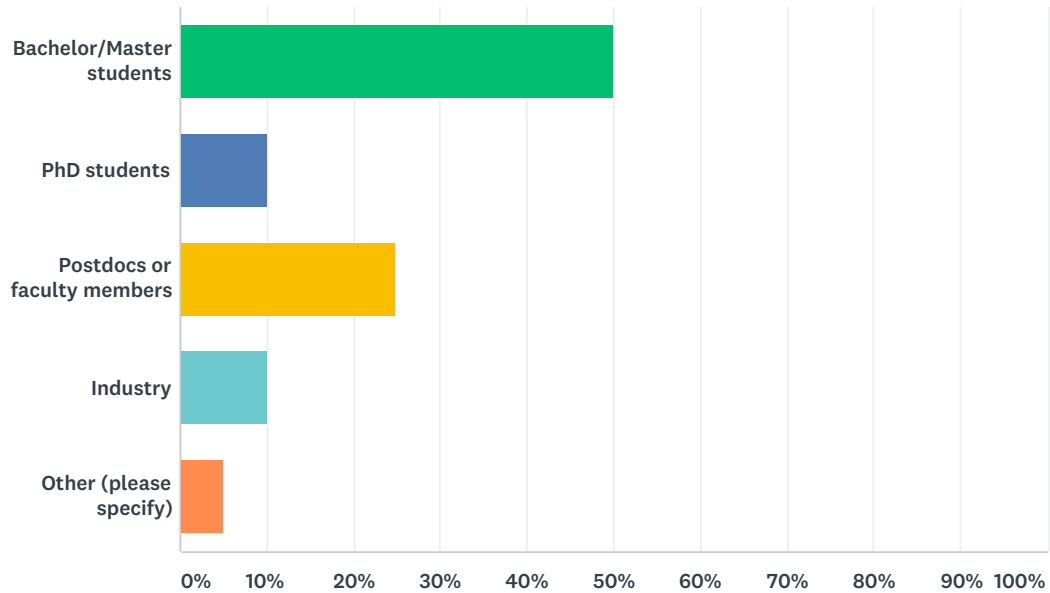| #	RESPONSES	DATE
1	IIIT-Hyderabad, i.am+ LLC	8/10/2018 3:32 PM
2	Hamburg University	8/7/2018 7:49 PM
3	Fraunhofer Institute	8/7/2018 7:18 PM
4	Institut für Informationswissenschaft & Sprachtechnologie, Universität Hildesheim	8/7/2018 7:11 PM
5	Department of Computer Science and Mathematics, University of Passau	8/7/2018 4:15 PM
6	DFKI GmbH	8/7/2018 4:00 PM
7	Computational Linguistics, University of Zurich	8/7/2018 3:55 PM
8	Forensic Science Investigation Lab. University of Applied Sciences Mittweida	8/7/2018 1:27 PM
9	ZHAW & SpinningBytes	8/7/2018 10:37 AM
10	TU Wien, Institut für Logic and Computation	8/7/2018 9:26 AM
11	University of Potsdam, Germany	8/7/2018 12:55 AM
12	University des Saarlandes	8/6/2018 6:52 PM
13	Fondazione Bruno Kessler	8/6/2018 2:46 PM
14	University of Groningen	8/6/2018 11:07 AM
15	Hochschule Hamm-Lippstadt	8/6/2018 7:42 AM
16	Hildesheim/Antwerp	8/2/2018 9:47 PM
17	Hasso Plattner Institute, University of Potsdam	8/1/2018 3:33 PM
18	Darmstadt University of Applied Sciences	7/31/2018 5:55 PM
19	Saarland University	7/31/2018 2:08 PM
20	Visiting Student	7/27/2018 8:22 PM

Q3 How many people are in your group?

Answered: 20   Skipped: 0

#	RESPONSES	DATE
1	3	8/10/2018 3:32 PM
2	4	8/7/2018 7:49 PM
3	1	8/7/2018 7:18 PM
4	1	8/7/2018 7:11 PM
5	5	8/7/2018 4:15 PM
6	5	8/7/2018 4:00 PM
7	1	8/7/2018 3:55 PM
8	4	8/7/2018 1:27 PM
9	4	8/7/2018 10:37 AM
10	2	8/7/2018 9:26 AM
11	4	8/7/2018 12:55 AM
12	3	8/6/2018 6:52 PM
13	7	8/6/2018 2:46 PM
14	5	8/6/2018 11:07 AM
15	2	8/6/2018 7:42 AM
16	2	8/2/2018 9:47 PM
17	4	8/1/2018 3:33 PM
18	8	7/31/2018 5:55 PM
19	8	7/31/2018 2:08 PM
20	3	7/27/2018 8:22 PM

Q4 Which position does the majority of your group members have? Please mark one of the following options.
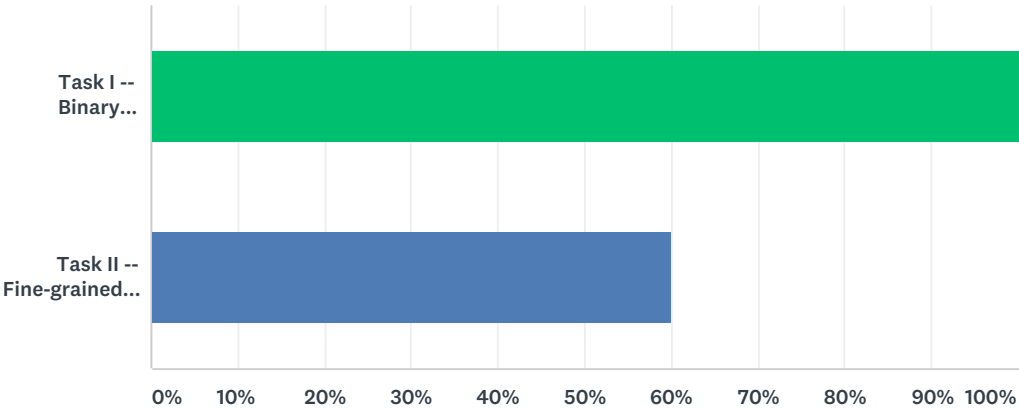
Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
Bachelor/Master students	50.00%	10
PhD students	10.00%	2
Postdocs or faculty members	25.00%	5
Industry	10.00%	2
Other (please specify)	5.00%	1
TOTAL		20

#	OTHER (PLEASE SPECIFY)	DATE
1	One is Master Student, one is Postdoc, so the majority is hard to determine.	8/7/2018 9:26 AM

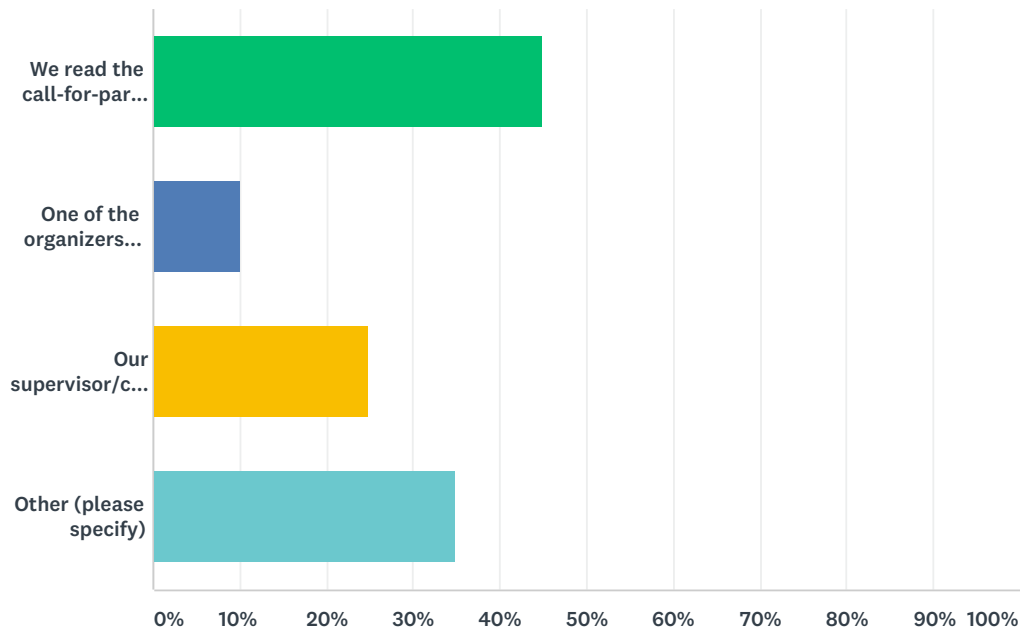# Q5 Which task did you participate in? Please mark all answers that apply.

Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
Task I -- Binary classification	100.00%	20
Task II -- Fine-grained classification	60.00%	12
Total Respondents: 20		

Q6 How did you come across this shared task? Please mark all answers that apply.

Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
We read the call-for-participation on some mailing list (e.g. corpora-list).	45.00%	9
One of the organizers informed us about this shared task.	10.00%	2
Our supervisor/course instructor encouraged us to participate in this shared task.	25.00%	5
Other (please specify)	35.00%	7
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	working on German Sentiment Analysis	8/10/2018 3:32 PM
2	Konvens 2018 website	8/7/2018 7:11 PM
3	Iggsa member	8/7/2018 3:55 PM
4	I worked on NER task and found NER shared task 2014	8/7/2018 1:27 PM
5	Found the Germeval 2017 proceedings during a literature review, checked the website and saw that 2018 was on the way :)	8/6/2018 7:42 AM
6	member of the group is task organizer	7/31/2018 2:08 PM
7	Internet research	7/27/2018 8:22 PM

Q7 Did you have some experience with the automatic detection of offensive language before? Please mark all answers that apply.

Answered: 20   Skipped: 0
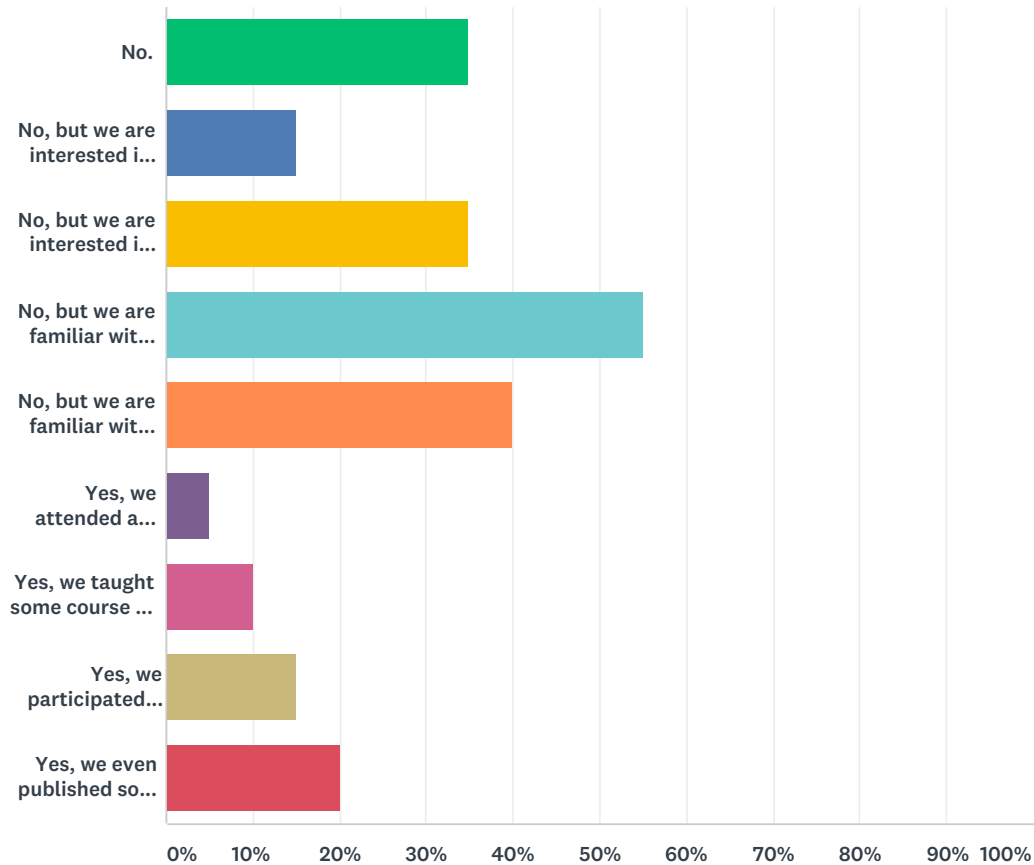
ANSWER CHOICES	RESPONSES	
No.	35.00%	7
No, but we are interested in social linguistics.	15.00%	3
No, but we are interested in NLP for social media.	35.00%	7
No, but we are familiar with text classification tasks.	55.00%	11
No, but we are familiar with sentiment analysis.	40.00%	8
Yes, we attended a course on this topic in the past.	5.00%	1
Yes, we taught some course on this topic in the past.	10.00%	2
Yes, we participated in a very similar shared task which was run on another language.	15.00%	3
Yes, we even published some paper/article on this task.	20.00%	4
Total Respondents: 20		

Q8 What type of classification approach did you pursue? Please mark all answers that apply.

Answered: 20   Skipped: 0
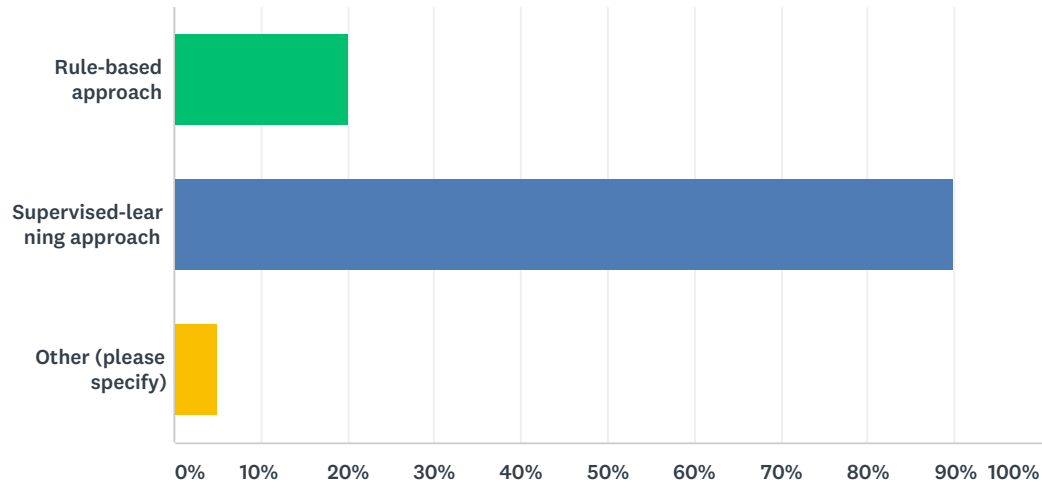
ANSWER CHOICES	RESPONSES	
Rule-based approach	20.00%	4
Supervised-learning approach	90.00%	18
Other (please specify)	5.00%	1
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	Semi-supervised approach that combines unsupervised language models with a supervised classifier	8/6/2018 7:59 AM

Q9 What training data did you use for a supervised-learning approach? Please mark all answers that apply.

Answered: 20   Skipped: 0
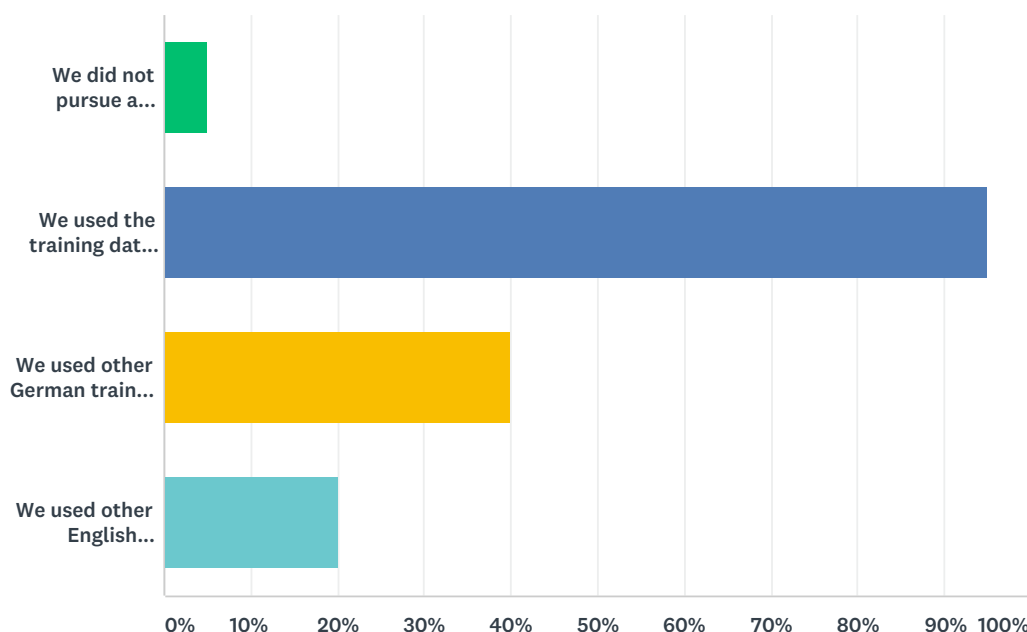
ANSWER CHOICES	RESPONSES	
We did not pursue a supervised-learning approach.	5.00%	1
We used the training data provided by the shared task.	95.00%	19
We used other German training data.	40.00%	8
We used other English training data.	20.00%	4
Total Respondents: 20		

#	IN CASE YOU USED OTHER GERMAN/ENGLISH TRAINING DATA, PLEASE SPECIFY THEM.	DATE
1	https://github.com/upInf/germeval2018; https://github.com/t-davidson/hate-speech-and-offensive-language	8/7/2018 4:24 PM
2	https://github.com/lukovnikov/hatespeech	8/7/2018 10:51 AM
3	pretrained German embeddings from Ruppenhofer http://www.cl.uni-heidelberg.de/english/research/downloads/resource_pages/GermanTwitterEmbeddings/GermanTwitterEmbeddings_data.shtml	8/7/2018 9:34 AM
4	Political Speech Project, https://rania.shinyapps.io/PoliticalSpeechProject/	8/6/2018 11:17 AM
5	For the unsupervised language models we used a dump of the entire German wikipedia and about 300k German tweets that we collected with a custom script	8/6/2018 7:59 AM
6	Twitter Embeddings and part of our Polly Corpus	8/2/2018 10:11 PM
7	Kaggle Toxic Comment Classification Dataset	8/1/2018 3:38 PM
8	dataset by Waseem et al. (2016); dataset by Wulczyn et al. (2017)	7/31/2018 2:18 PM

Q10 What types of classifier(s) did you use for supervised learning? Please mark all answers that apply.

Answered: 20   Skipped: 0
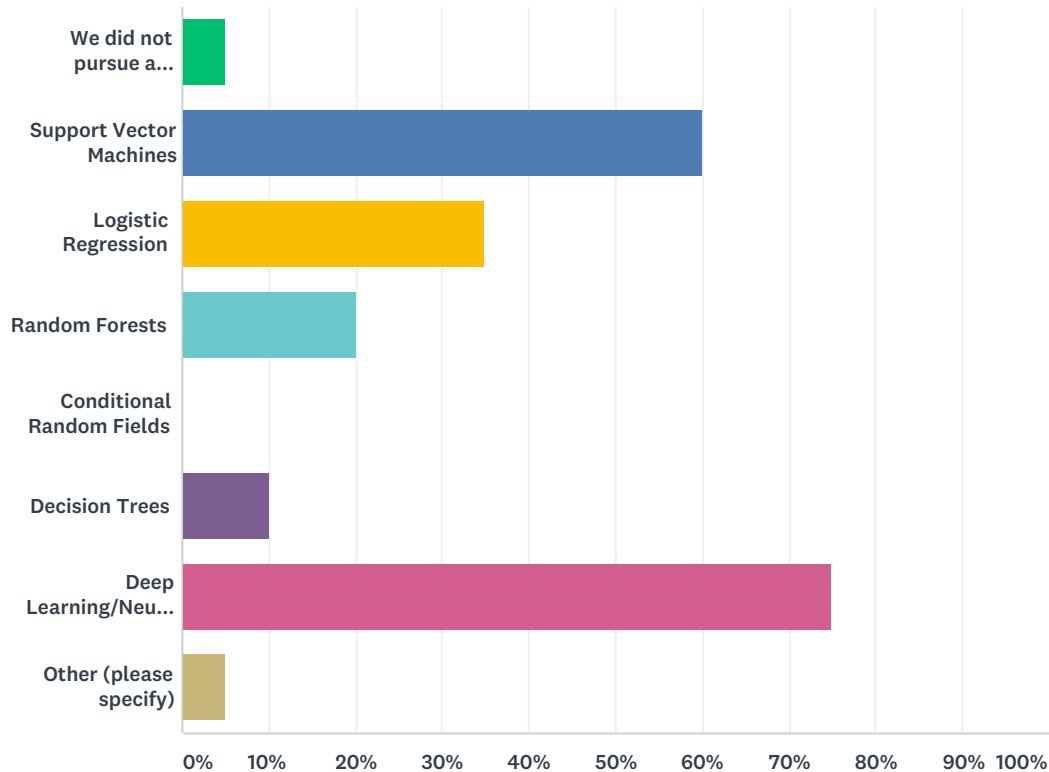
ANSWER CHOICES	RESPONSES	
We did not pursue a supervised-learning approach.	5.00%	1
Support Vector Machines	60.00%	12
Logistic Regression	35.00%	7
Random Forests	20.00%	4
Conditional Random Fields	0.00%	0
Decision Trees	10.00%	2
Deep Learning/Neural Networks	75.00%	15
Other (please specify)	5.00%	1
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	Naive Bayes, MLP	8/7/2018 1:34 PM

Q11 Did you use deep learning/neural networks? If so, what type of classifier did you use? Please mark all answers that apply.

Answered: 20   Skipped: 0
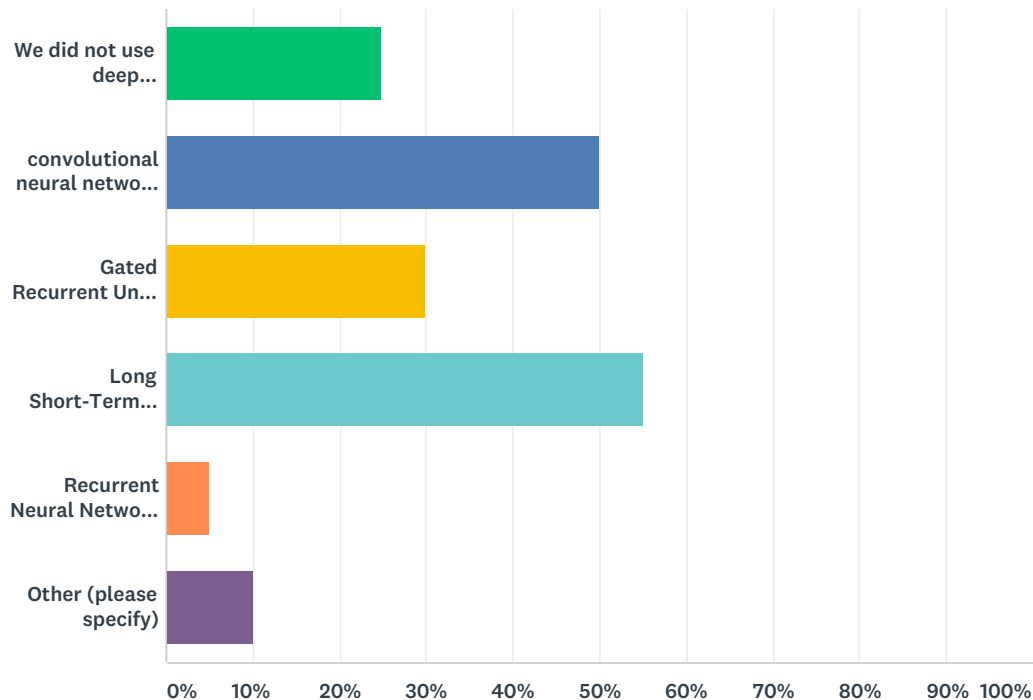
ANSWER CHOICES	RESPONSES	
We did not use deep learning/neural networks.	25.00%	5
convolutional neural networks (CNN)	50.00%	10
Gated Recurrent Unit neural networds (GRU)	30.00%	6
Long Short-Term Memory neural networks (LSTM)	55.00%	11
Recurrent Neural Network (RNN) -- if not GRU or LSTM	5.00%	1
Other (please specify)	10.00%	2
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	we experimented with deep NNs and describe experiments in the paper but did not use them for the submission	8/7/2018 9:34 AM
2	Specifically we used a variation of LSTM with dropout (AWD-LSTM)	8/6/2018 7:59 AM

Q12 On what aspect did you focus when you built your system?

Answered: 20   Skipped: 0
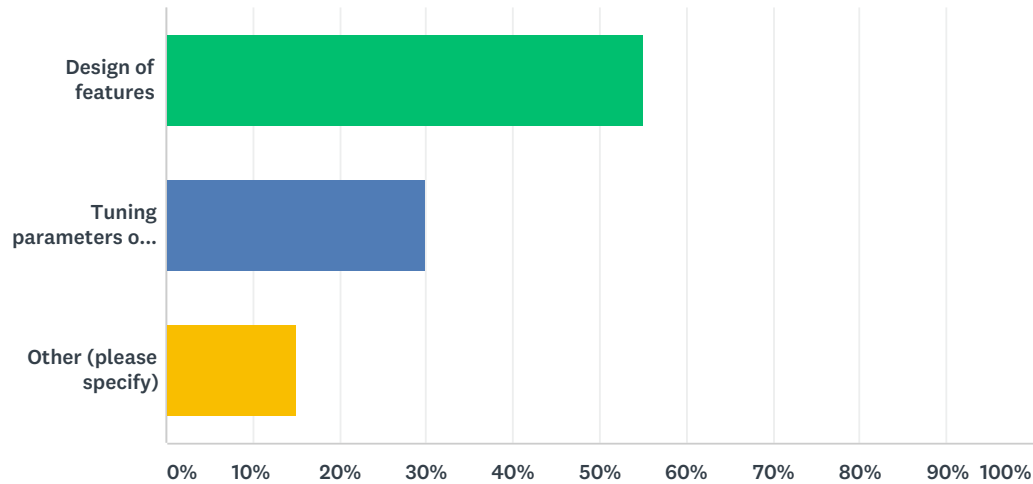
ANSWER CHOICES	RESPONSES	
Design of features	55.00%	11
Tuning parameters of a (supervised) classifier	30.00%	6
Other (please specify)	15.00%	3
TOTAL		20

#	OTHER (PLEASE SPECIFY)	DATE
1	Transfer learning	8/7/2018 7:54 PM
2	understanding implicit offensive language	8/7/2018 4:01 PM
3	tuning parameters of the underlying German language model	8/6/2018 7:59 AM

Q13 Did you use any word embeddings in your approach? If so, what type of embeddings did you employ? Please mark all answers that apply.

Answered: 20   Skipped: 0
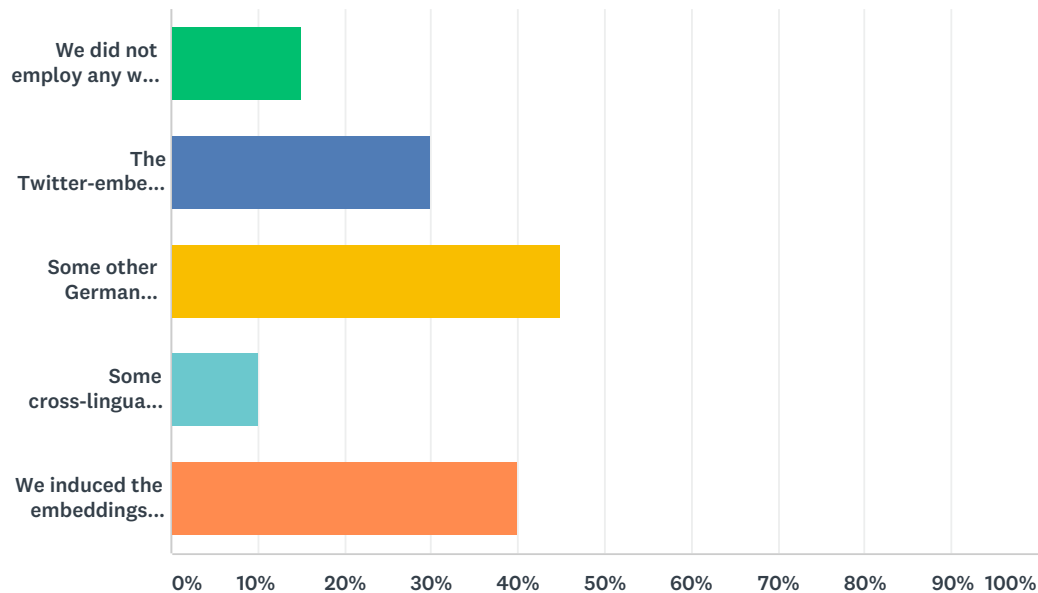
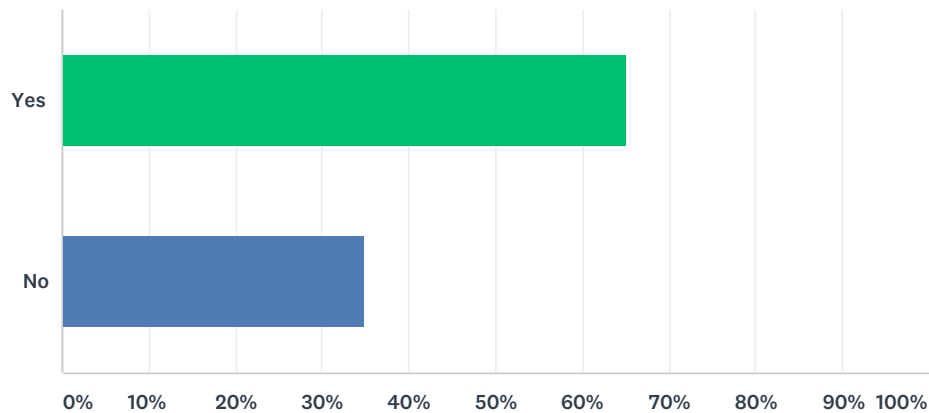| ANSWER CHOICES	RESPONSES	
We did not employ any word embeddings.	15.00%	3
The Twitter-embeddings provided by Heidelberg University.	30.00%	6
Some other German pre-trained embeddings that are publicly available.	45.00%	9
Some cross-lingual embeddings.	10.00%	2
We induced the embeddings ourselves.	40.00%	8
Total Respondents: 20		

#	IN CASE YOU USED EMBEDDINGS OTHER THAN THE GERMAN TWITTER-EMBEDDINGS FROM HEIDELBERG UNIVERSITY, PLEASE SPECIFY THE RESOURCE. (IN CASE YOU INDUCED THE EMBEDDINGS YOURSELF, PLEASE SPECIFY THE CORPUS ON WHICH YOU INDUCED EMBEDDINGS.)	DATE
1	Twitter Embeddings by Spinningbytes	8/10/2018 3:44 PM
2	https://github.com/upInf/germeval2018	8/7/2018 4:24 PM
3	We got twitter data from a co-worker.	8/7/2018 4:04 PM
4	own corpus of newspaper texts (Swiss newspapers)	8/7/2018 4:01 PM
5	Word Embedding based on German Wikipedia	8/7/2018 1:34 PM
6	http://spinningbytes.com	8/7/2018 10:51 AM
7	we also tried other German pretrained embeddings but the results were worse so we did not use them for the competition	8/7/2018 9:34 AM
8	- German fastText embeddings - Emoji embeddings (which we aligned with German ones) from Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. 2016. "What does this Emoji Mean? A Vector Space Skip-Gram Model for Twitter Emojis". In Proceedings of LREC.	8/6/2018 9:53 PM

9	We used as a corpus Facebook comments from hate-oriented public pages. Moreover we induced additional embeddings from the training data, initialized with the Facebook embeddings.	8/6/2018 11:17 AM
10	german FastText embeddings (obtained from common crawl corpus)	8/1/2018 3:38 PM
11	embeddings induced on COW16 (Schäfer & Bildhauer, 2012; Schäfer 2015); embeddings induced on deWaC (Baroni, 2009); crosslingual embeddings induced with VecMap (AreTxte et al., 2017) with corpora bing deWaC and Bing Liu's sentiment corpus (Jindal & Liu 2008)	7/31/2018 2:18 PM

Q14 Did you consider any subword analysis (e.g. character n-grams, morphological decomposition) in your approach (for these purposes, lemmatization or stemming does not count as subword analysis)?

Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
Yes	65.00%	13
No	35.00%	7
TOTAL		20

#	IN CASE YOU CARRIED OUT SUBWORD ANALYSIS, PLEASE SPECIFY WHAT YOU DID:	DATE
1	character n-gram TF-IDF, learning word representation through character embeddings	8/10/2018 3:44 PM
2	fastText embeddings	8/7/2018 7:54 PM
3	character n-grams	8/7/2018 7:19 PM
4	character n-grams	8/7/2018 4:04 PM
5	generic character n-grams (length 3 to 7) + TFIDF	8/7/2018 9:34 AM
6	character ngrams	8/7/2018 1:01 AM
7	character n-grams	8/6/2018 11:17 AM
8	character n-grams	8/2/2018 10:11 PM
9	character n-grams	8/1/2018 3:38 PM
10	character n-grams; lexicon matching function matched prefixes rather than full (lemmatized) tokens	7/31/2018 2:18 PM

Q15 Did you employ any task-specific lexicon, i.e. some lexicon of offensive words? (A simple list of offensive words is also regarded as a task-specific lexicon.) If so, describe that lexicon.

Answered: 20   Skipped: 0
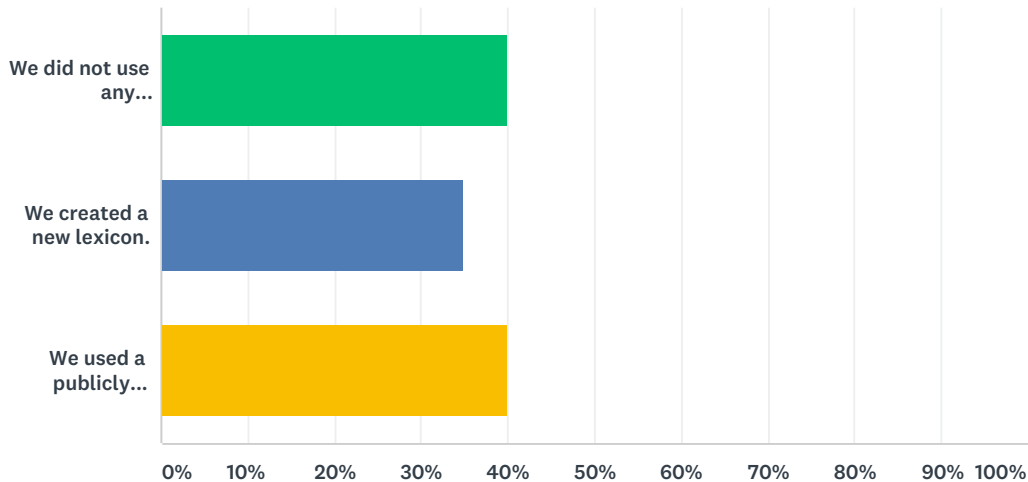
ANSWER CHOICES	RESPONSES	
We did not use any task-specific lexicon.	40.00%	8
We created a new lexicon.	35.00%	7
We used a publicly available lexicon.	40.00%	8
Total Respondents: 20		

#	IN CASE YOU USED A PUBLICLY AVAILABLE LEXICON, PLEASE SPECIFY IT:	DATE
1	German Slur Dictionary	8/10/2018 3:44 PM
2	German Sentiment Lexicon from University of Zurich; SentiWS; Schimpfword-wiki	8/7/2018 7:19 PM
3	from the website of germeval: swearwords	8/7/2018 4:01 PM
4	Multi-Domain Sentiment Lexicon for German, http://hyperhero.com/de/insults.htm, https://www.hatebase.org/search_results/keywords%3D%7Cfilter_about_ethnicity%3D%7Cfilter_about_nationality%3D%7Cfilter_about_religion%3D%7Cfilter_about_gender%3D%7Cfilter_about_sexual_orientation%3D%7Cfilter_about_disability%3D%7Cfilter_about_class%3D%7Cfilter_archaic%3D%7Cinclude_meaning%3D%7Cfilter_language%3Ddeu	8/7/2018 10:51 AM
5	LIWC German categories 18 (anger) and 66 (swear words); very few manually added words (e.g. "Lügenpresse")	8/7/2018 1:01 AM
6	https://www.schimpfwoerter.de	8/6/2018 6:58 PM
7	http://www.hyperhero.com/de/insults.htm	8/6/2018 11:17 AM
8	abusive lexicon by Wiegand et al. (2018) translated into German	7/31/2018 2:18 PM

Q16 Did you create a new task-specific lexicon? If so, how? Please mark only one option.

Answered: 20   Skipped: 0
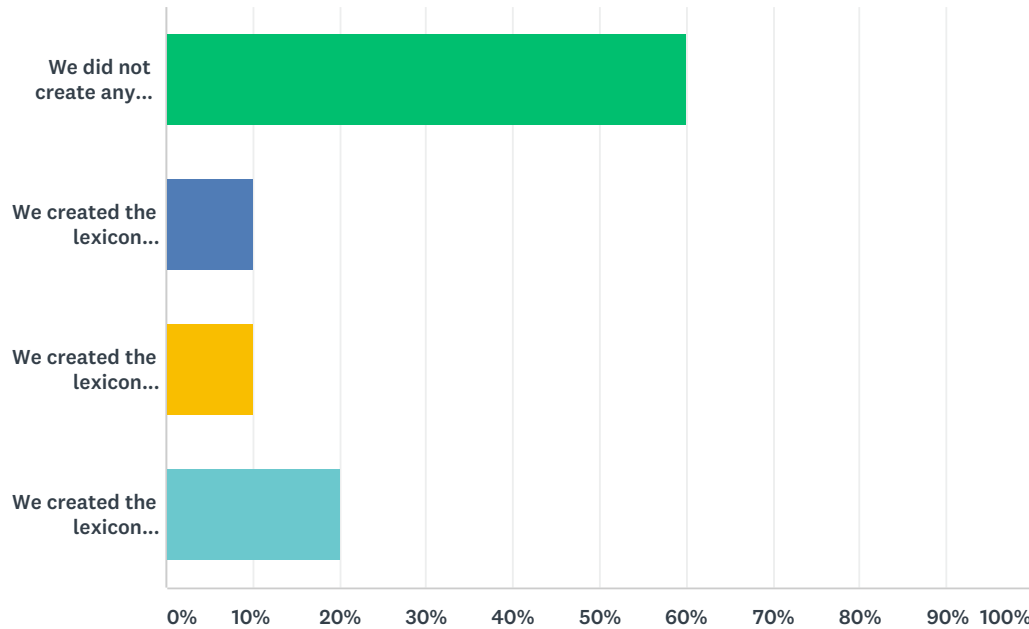
ANSWER CHOICES	RESPONSES	
We did not create any task-specific lexicon.	60.00%	12
We created the lexicon manually.	10.00%	2
We created the lexicon automatically.	10.00%	2
We created the lexicon semi-automatically.	20.00%	4
TOTAL		20

Q17 If you used a task-specific lexicon, please specify its size (number of word entries).

Answered: 20   Skipped: 0
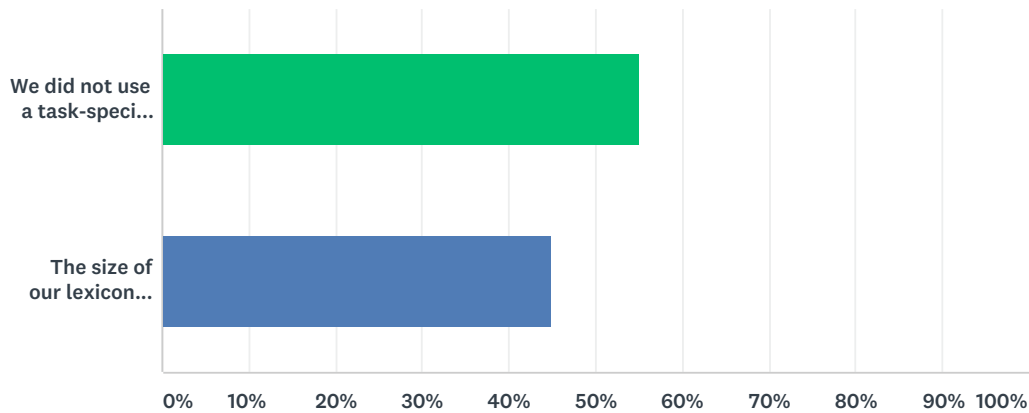
ANSWER CHOICES	RESPONSES	
We did not use a task-specific lexicon.	55.00%	11
The size of our lexicon was:	45.00%	9
TOTAL		20

#	THE SIZE OF OUR LEXICON WAS:	DATE
1	3000	8/7/2018 7:54 PM
2	67	8/7/2018 7:19 PM
3	about 600 words	8/7/2018 4:01 PM
4	2917, 11300, 48	8/7/2018 10:51 AM
5	432	8/7/2018 1:01 AM
6	1,750	8/2/2018 10:11 PM
7	approx 10.000 entries	8/1/2018 3:38 PM
8	3545	7/31/2018 5:57 PM
9	1566 (unigram) entries	7/31/2018 2:18 PM

Q18 Contents of the task-specific lexicon? Please mark all that apply.

Answered: 20   Skipped: 0
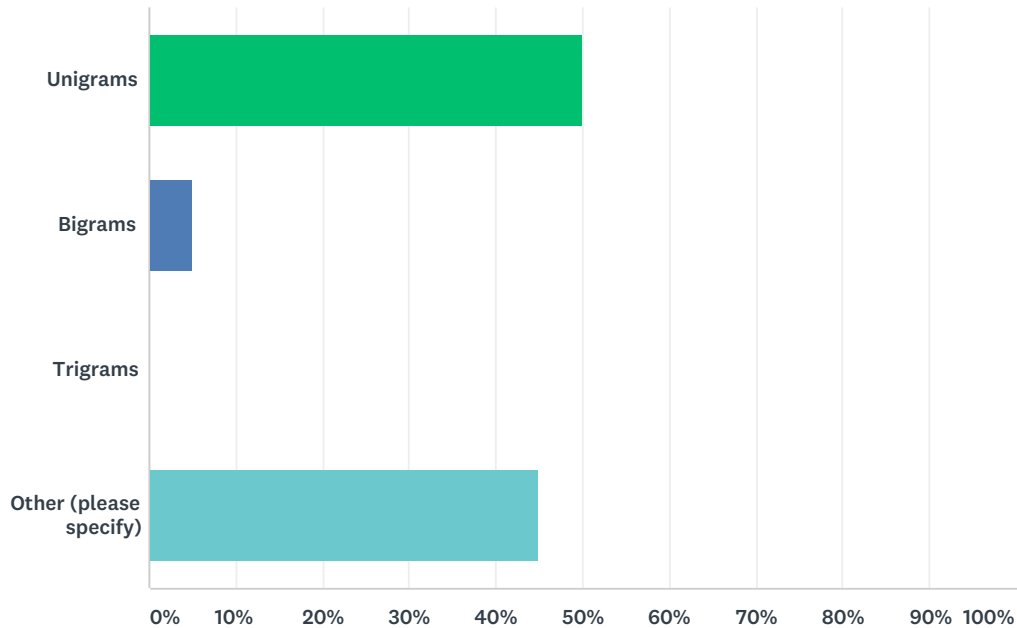
ANSWER CHOICES	RESPONSES	
Unigrams	50.00%	10
Bigrams	5.00%	1
Trigrams	0.00%	0
Other (please specify)	45.00%	9
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	None	8/7/2018 4:24 PM
2	No task specific lexicon.	8/7/2018 4:04 PM
3	We did not task-specific lexicon	8/7/2018 1:34 PM
4	n/a	8/7/2018 9:34 AM
5	No task-specific lexicon	8/6/2018 9:53 PM
6	did not use a task-specific lexicon	8/6/2018 6:58 PM
7	no task-specific lexicon	8/6/2018 11:17 AM
8	Did not use a TSL.	8/6/2018 7:59 AM
9	none	7/27/2018 8:24 PM

Q19 What type of general-purpose lexicons (e.g. GermaNet) did you use for your system? Please mark all answers that apply.

Answered: 20   Skipped: 0
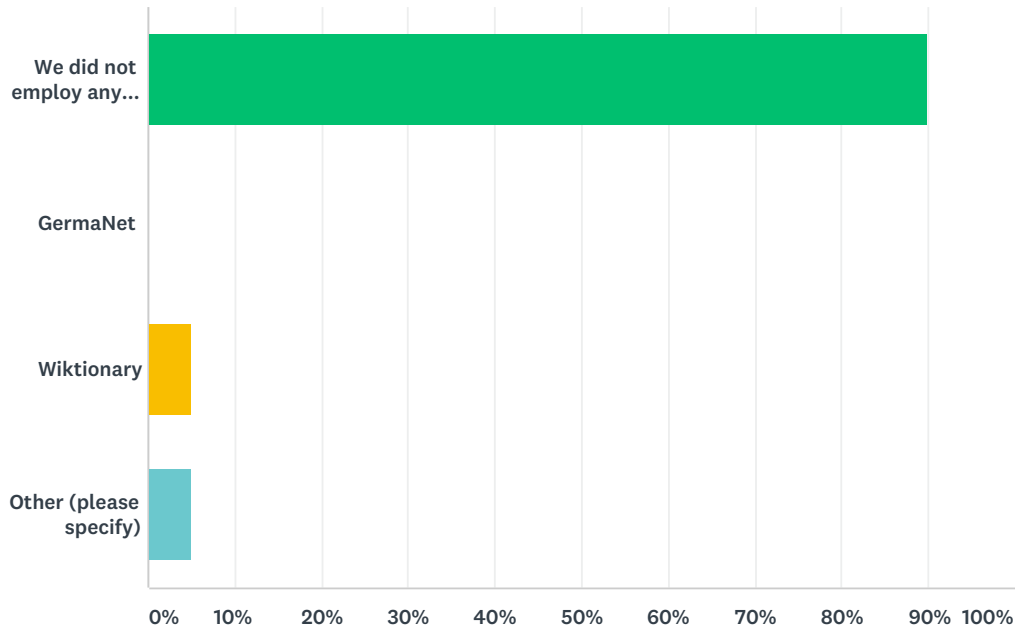
ANSWER CHOICES	RESPONSES	
We did not employ any general-purpose lexicon.	90.00%	18
GermaNet	0.00%	0
Wiktionary	5.00%	1
Other (please specify)	5.00%	1
Total Respondents: 20		

#	OTHER (PLEASE SPECIFY)	DATE
1	GermanPolarityClues	8/10/2018 3:44 PM

Q20 Information used from general-purpose lexicons.

Answered: 20   Skipped: 0
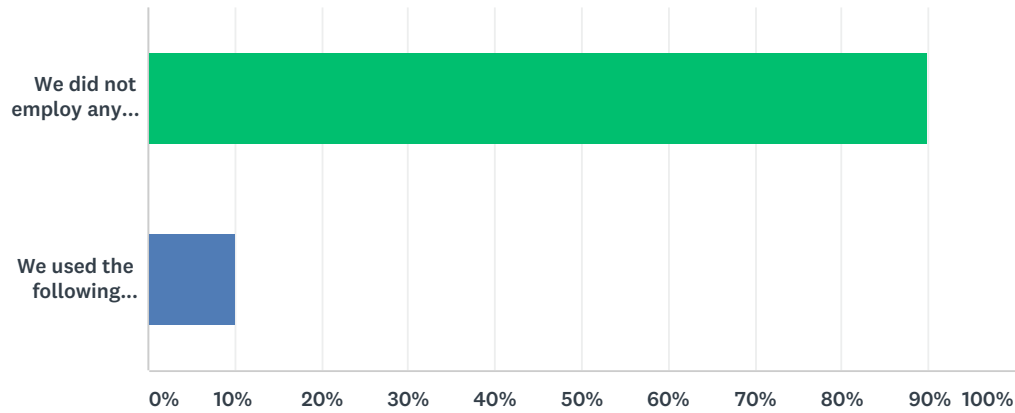
ANSWER CHOICES	RESPONSES	
We did not employ any general-purpose lexicon.	90.00%	18
We used the following information provided by the resource (e.g. GermaNet synsets or hyperonymy relations):	10.00%	2
TOTAL		20

#	WE USED THE FOLLOWING INFORMATION PROVIDED BY THE RESOURCE (E.G. GERMANET SYNSETS OR HYPERONYMY RELATIONS):	DATE
1	words from slur dictionary	8/10/2018 3:44 PM
2	word-usage information; synonymy relations	7/31/2018 2:18 PM

Q21 Does your system incorporate any sentiment information (e.g. polarity classification, subjectivity detection) as some auxiliary task?

Answered: 20   Skipped: 0
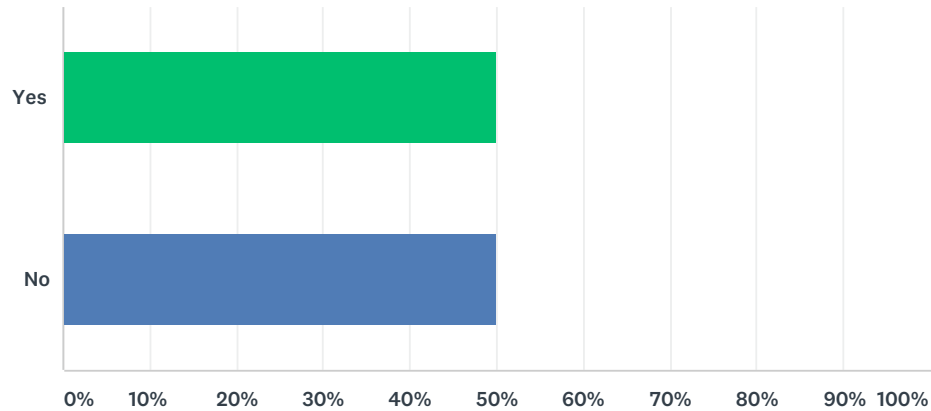
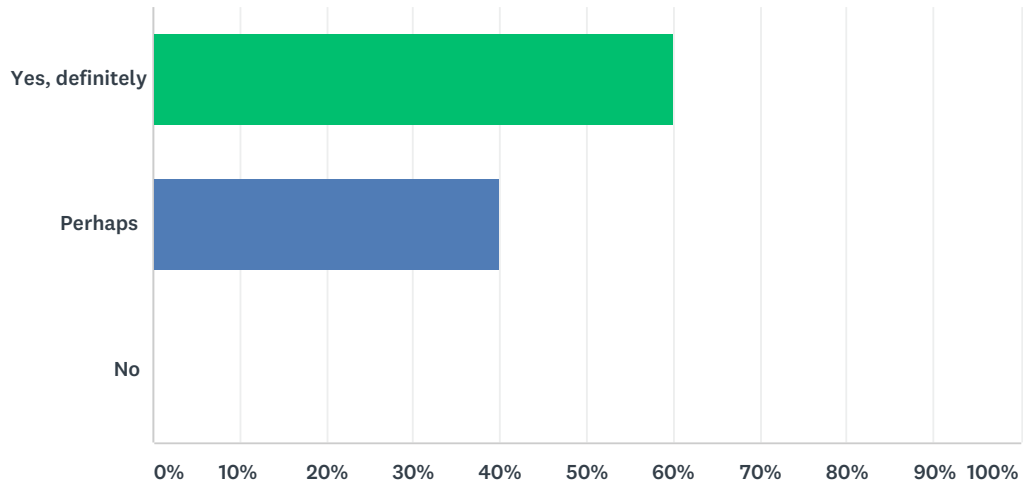| ANSWER CHOICES	RESPONSES	
Yes	50.00%	10
No	50.00%	10
TOTAL		20

Q22 What, in your view, is the most important/effective component (e.g. feature, rule, classifier etc.) of your system?

Answered: 20   Skipped: 0

#	RESPONSES	DATE
1	character n-grams	8/10/2018 3:44 PM
2	neural model pretrained on tweets assigned to topic clusters	8/7/2018 7:54 PM
3	character n-grams	8/7/2018 7:19 PM
4	feature combination	8/7/2018 7:19 PM
5	word embeddings	8/7/2018 4:24 PM
6	Combination of feature and classifier	8/7/2018 4:04 PM
7	offensive language lexicon, verb lexicon for stance analysis	8/7/2018 4:01 PM
8	contextual information in texts is important for classification	8/7/2018 1:34 PM
9	Ensemble	8/7/2018 10:51 AM
10	the per-class most important (according to TFIDF) features that are used to compute additional features and help with the small PROFANITY class to increase the score after macro-averaging	8/7/2018 9:34 AM
11	character ngrams (4- and 5grams) + SVM	8/7/2018 1:01 AM
12	classifier	8/6/2018 9:53 PM
13	Word embeddings	8/6/2018 6:58 PM
14	word embeddings + ensemble classifier	8/6/2018 11:17 AM
15	The German Language Model (which is an AWD-LSTM)	8/6/2018 7:59 AM
16	character n-grams + lexicon of offensive words + feature selection + oversampling	8/2/2018 10:11 PM
17	task-specific lexicons	8/1/2018 3:38 PM
18	lexicon	7/31/2018 5:57 PM
19	a robust means of word generalization (e.g. word embeddings)	7/31/2018 2:18 PM
20	Quality of the word embeddings that we used	7/27/2018 8:24 PM

Q23 Would your group be participating in another edition of this shared task in the future?
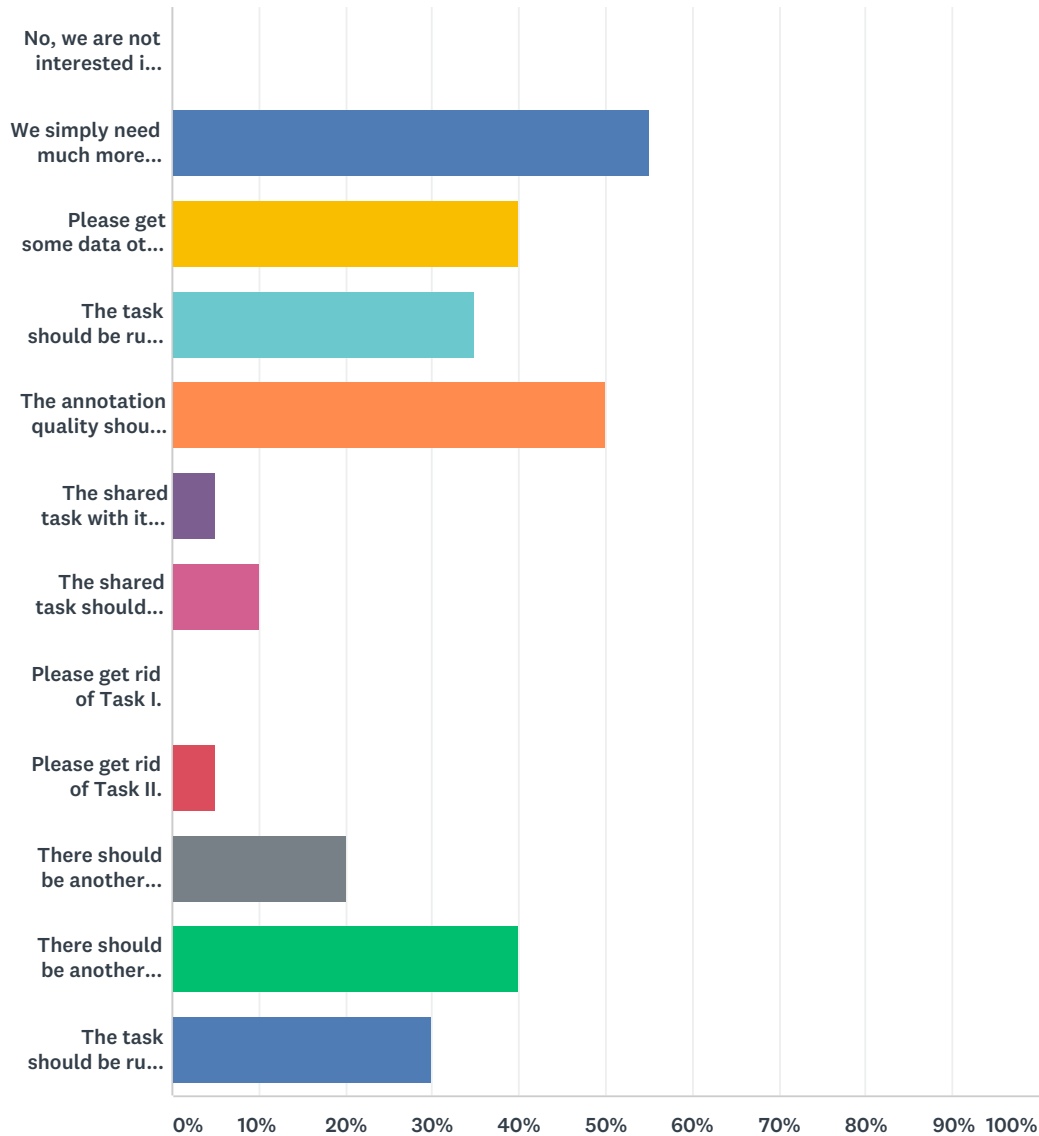
Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
Yes, definitely	60.00%	12
Perhaps	40.00%	8
No	0.00%	0
TOTAL		20

Q24 If you were interested in another edition, what should be changed about the setting? Please mark all answers that apply.

Answered: 20   Skipped: 0

ANSWER CHOICES	RESPONSES	
No, we are not interested in another edition of the shared task.	0.00%	0
We simply need much more training data.	55.00%	11
Please get some data other than Twitter.	40.00%	8
The task should be run on several domains.	35.00%	7
The annotation quality should be improved, particularly, the quality of Task II.	50.00%	10
The shared task with its two present tasks should be the same as before.	5.00%	1
The shared task should maintains its 2 (sub)tasks, but there should be at least some further (optional) (sub)task.	10.00%	2

Please get rid of Task I.	0.00%	0
Please get rid of Task II.	5.00%	1
There should be another (sub)task on distinguishing between explicit and implicit offence. (We define explicit offence as offence in which at least offensive word, e.g. idiot, slut, nigger, occurs.)	20.00%	4
There should be another (sub)task in which the offended target groups are to be determined (e.g. personal insult vs. misogyny vs. homophobia vs. Islamophobia sv. anti-semitism etc.).	40.00%	8
The task should be run on several languages, not just German.	30.00%	6
Total Respondents: 20		

#	OTHER SUGGESTIONS:	DATE
1	The training data set should have been more balanced. PROFANITY type has very few samples.	8/10/2018 3:51 PM
2	I guess it was a conscious design decision to only provide the text of the tweets (and not userId for example). For more reliable approaches to this task, being able to include other (identity based) information would make a lot of sense (insofar as that is legally and practically/technically possible of course).	8/7/2018 4:07 PM
3	Label distribution of Task II was bad. The deadline was in the middle of the summer, so it was difficult to be active in the last weeks.	8/7/2018 10:56 AM
4	I find the restriction on only textual data very limiting. In my opinion it would be important to include metadata such as the author (profile), discourse context (e.g., previous comments or answers), time, etc. The annotation was also sometimes hard to follow. I.e., any mention of the word "scheiß" lead to an "OFFENSE" even though many might not find this offensive (e.g, if it's about the weather). However, offensive words that induce a certain denigrating frame did often not lead to classifications (e.g., "Asylanten"). I guess the way to address this subjectivity would be by including more ratings per item.	8/7/2018 1:19 AM
5	If a class is included in the classification task, please provide enough training data.	8/6/2018 9:54 PM
6	Make annotation guidelines explicit and available to task participants	8/6/2018 11:21 AM
7	a public leaderboard during the task	8/1/2018 3:40 PM

Q25 What suggestions do you have for improving the administration of the shared task?

Answered: 20   Skipped: 0

#	RESPONSES	DATE
1	more subtasks could have been included	8/10/2018 3:51 PM
2	provide a central dev set for all participants	8/7/2018 7:56 PM
3	It would be much easier to write the paper if the results on the test set would be made available before writing. Perhaps it would be possible to stagger the deadlines and hand in the predicted tags a few days before the paper.	8/7/2018 7:22 PM
4	none	8/7/2018 7:20 PM
5	keep up the good work!	8/7/2018 4:26 PM
6	Nothing, the administration was rather good, everybody was quite nice answering emails etc.	8/7/2018 4:07 PM
7	none	8/7/2018 4:02 PM
8	The organization of this shared task is good. Clear schedule, good mail-list service	8/7/2018 1:38 PM
9	codalab for the submission and evaluation (see VarDial)	8/7/2018 10:56 AM
10	the submission format could be more clear, but this is a very minor issue. maybe macro-averaging is not such a good evaluation criterion in an imbalanced dataset? (but then maybe it encourages solutions that can generalize better)	8/7/2018 9:37 AM
11	What would be really cool is a leaderboard (run on a dev set, not the real test set!), where one could see how things are going. Also some kind of baseline.	8/7/2018 1:19 AM
12	None	8/6/2018 9:54 PM
13	None	8/6/2018 7:01 PM
14	System results on test data should be available when submitting the system description paper.	8/6/2018 11:21 AM
15	Keep up the good work :)	8/6/2018 8:00 AM
16	It would be great if the workshop was a little longer (a day maybe, including a common dinner for brainstorming and networking).	8/2/2018 10:18 PM
17	no suggestions	8/1/2018 3:40 PM
18	--	7/31/2018 5:58 PM
19	N/A	7/31/2018 2:19 PM
20	None, perfectly fine!	7/27/2018 8:25 PM