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

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	nlp-augsburg-04	7/8/2025 10:26 AM
2	Flauschgummi	7/8/2025 10:24 AM
3	TUM NLP Group	7/4/2025 1:47 PM
4	NLPSuedwestfalen	7/4/2025 9:43 AM
5	Die SuperGLEBer	7/3/2025 3:20 PM
6	AIxcellent Vibes	7/1/2025 11:04 AM
7	HHUflauschig	6/30/2025 1:58 PM
8	Georg Hofmann	6/30/2025 11:50 AM
9	Quabynar	6/30/2025 10:28 AM

Q2 Please provide the contact email address of your team.

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	evrenatasin@gmail.com / malika.abitova@uni-a.de	7/8/2025 10:26 AM
2	Bjoern.Kiesswetter@stud.uni-regensburg.de	7/8/2025 10:24 AM
3	faeze.ghorbanpour@tum.de	7/4/2025 1:47 PM
4	anna-hoff@hotmail.com	7/4/2025 9:43 AM
5	supergleber@informatik.uni-wuerzburg.de	7/3/2025 3:20 PM
6	bialonski@fh-aachen.de	7/1/2025 11:04 AM
7	petersew@hhu.de	6/30/2025 1:58 PM
8	georg.hofmann@uni-a.de	6/30/2025 11:50 AM
9	kwabena-odame.akomeah@ur.de	6/30/2025 10:28 AM

Q3 Please enter all participants' names of your team and their affiliations (one line per person).

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	Evren Ataş - University of Augsburg Malika Abitova - University of Augsburg	7/8/2025 10:26 AM
2	Kießwetter Björn, Lukas Christoph, Sieber Tim, Jain Parag	7/8/2025 10:24 AM
3	Sophie Opheiden TU Munich, Elisei Shushpanov TU Minich, Faeze Ghorbanpour TU Minich	7/4/2025 1:47 PM
4	Emir Bradaric - Fachhochschule Suedwestfalen Manuel Alexander Falk - Fachhochschule Suedwestfalen Anna Hoff - Fachhochschule Suedwestfalen	7/4/2025 9:43 AM
5	Jan Pfister, CAIDAS, Uni Würzburg Julia Wunderle, CAIDAS, Uni Würzburg Andreas Hotho, CAIDAS, Uni Würzburg	7/3/2025 3:20 PM
6	Christian Rene Thelen, Academic and Research Department Engineering Hydrology, RWTH Aachen University, Aachen, Germany Patrick Gustav Blaneck, IT Center, RWTH Aachen University, Aachen, Germany Tobias Bornheim, ORDIX AG Niklas Grieger, Institute for Data-Driven Technologies, Aachen University of Applied Sciences & Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands Stephan Bialonski, Institute for Data-Driven Technologies, Aachen University of Applied Sciences, Aachen, Germany	7/1/2025 11:04 AM
7	Wiebke Petersen, Heinrich-Heine-Universität Düsseldorf Lara Eulenpesch, Heinrich-Heine-Universität Düsseldorf	6/30/2025 1:58 PM
8	Georg Hofmann, Student, Faculty of Applied Computer Science, University of Augsburg	6/30/2025 11:50 AM
9	Kwabena ODame Akomeah University of Regensburg, Germany Udo Kruschwitz University of Regensburg, Germany Bernd Ludwig University of Regensbug, Germany Kwame Boateng Akomeah, Lewis Universty, IL, USA	6/30/2025 10:28 AM

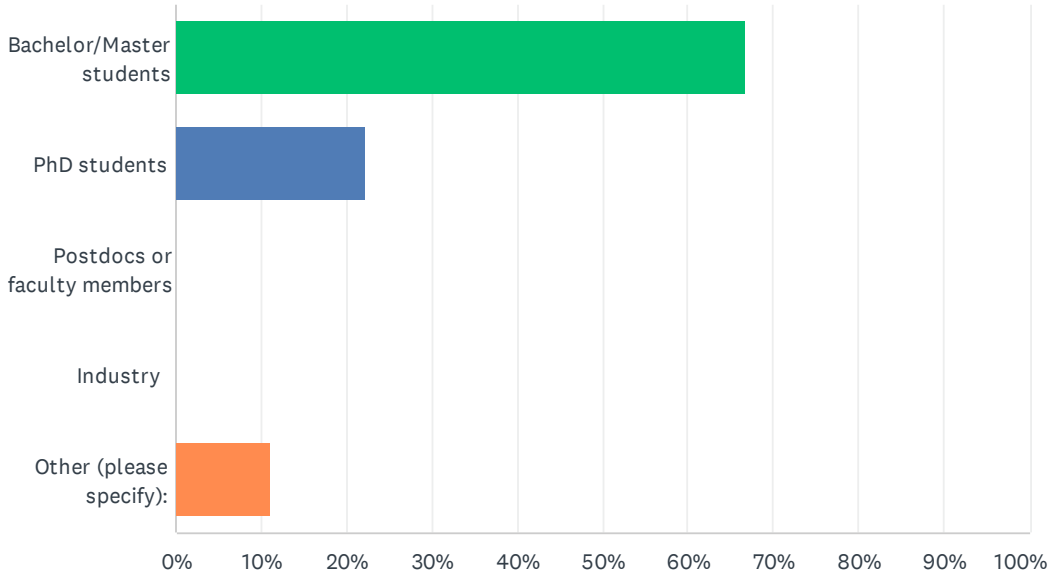
Q4 How many people are in your team?

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	2	7/8/2025 10:26 AM
2	4	7/8/2025 10:24 AM
3	3	7/4/2025 1:47 PM
4	3	7/4/2025 9:43 AM
5	3	7/3/2025 3:20 PM
6	5	7/1/2025 11:04 AM
7	2	6/30/2025 1:58 PM
8	1	6/30/2025 11:50 AM
9	4	6/30/2025 10:28 AM

Q5 Which position do(es) your team member(s) have who did most of the work? Please mark one of the following options.

Answered: 9 Skipped: 0

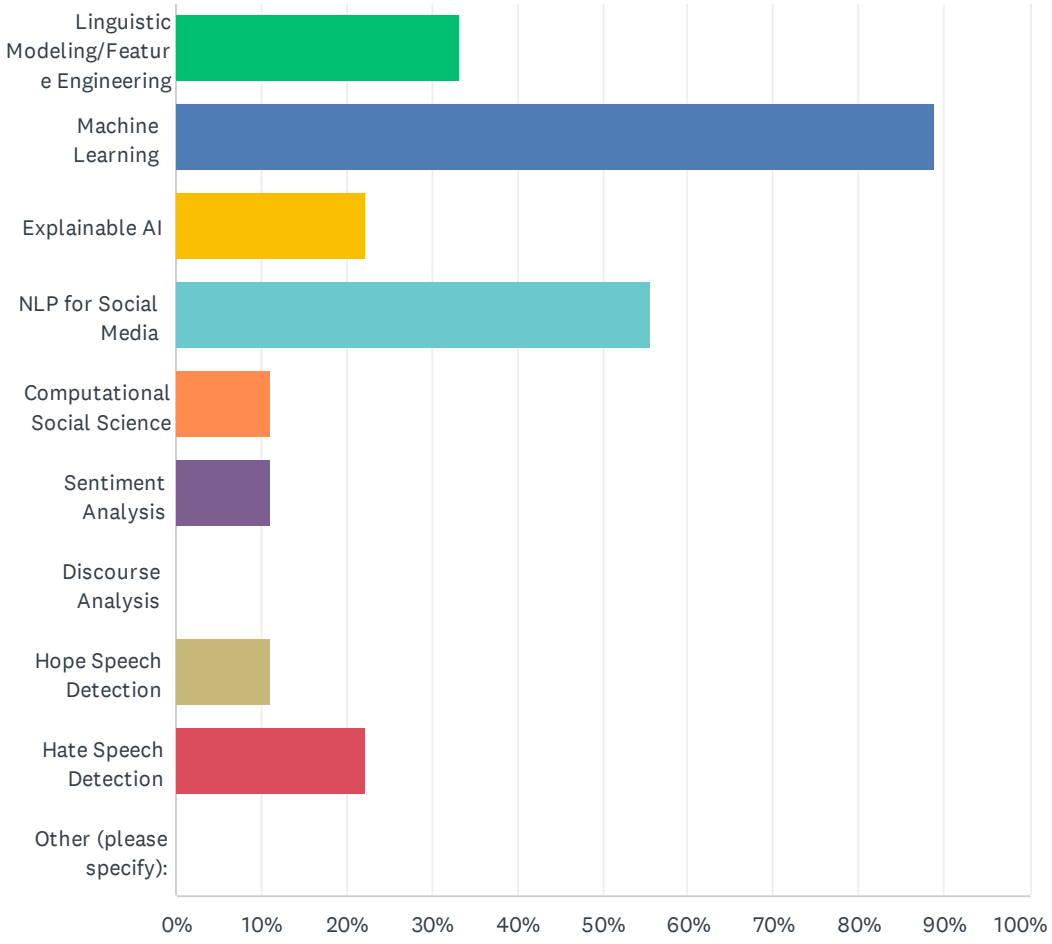


ANSWER CHOICES	RESPONSES	
Bachelor/Master students	66.67%	6
PhD students	22.22%	2
Postdocs or faculty members	0.00%	0
Industry	0.00%	0
Other (please specify):	11.11%	1
TOTAL		9

#	OTHER (PLEASE SPECIFY):	DATE
1	equal: faculty member and BA student	6/30/2025 1:58 PM

Q6 Where do your main interest(s) lie? Please mark all answers that apply.

Answered: 9 Skipped: 0

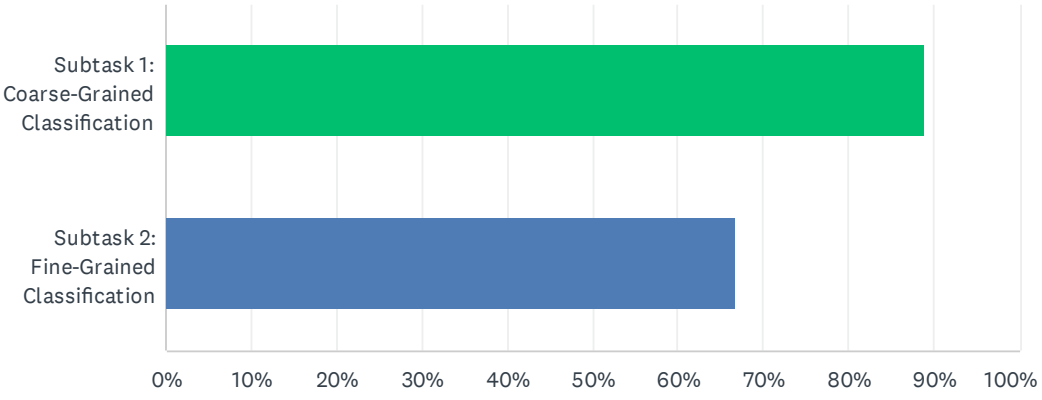


ANSWER CHOICES	RESPONSES	
Linguistic Modeling/Feature Engineering	33.33%	3
Machine Learning	88.89%	8
Explainable AI	22.22%	2
NLP for Social Media	55.56%	5
Computational Social Science	11.11%	1
Sentiment Analysis	11.11%	1
Discourse Analysis	0.00%	0
Hope Speech Detection	11.11%	1
Hate Speech Detection	22.22%	2
Other (please specify):	0.00%	0
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
	There are no responses.	

Q7 Which subtask(s) did you participate in? Please mark all answers that apply.

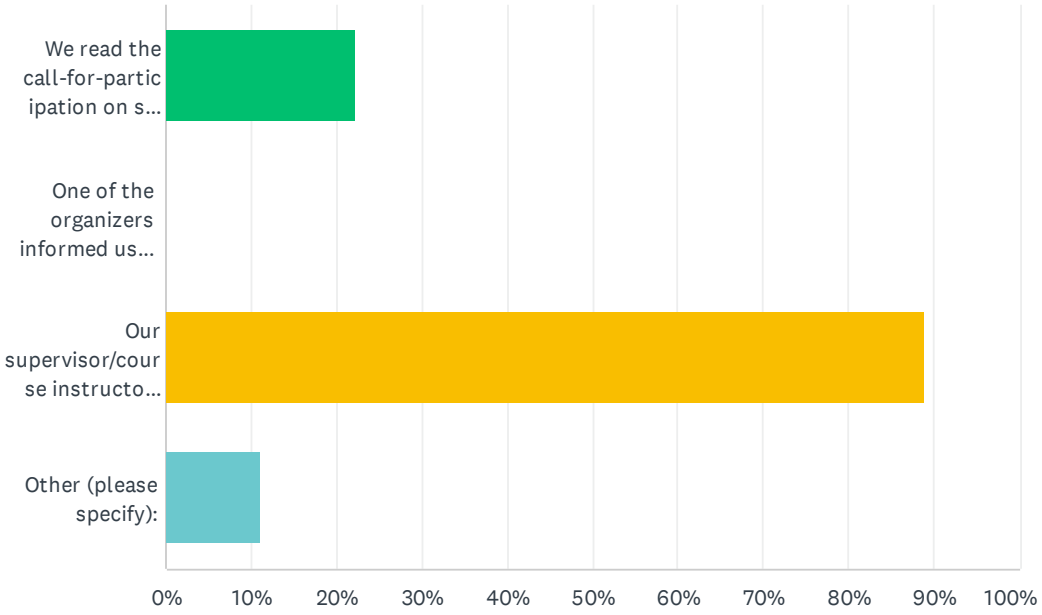
Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Subtask 1: Coarse-Grained Classification	88.89%	8
Subtask 2: Fine-Grained Classification	66.67%	6
Total Respondents: 9		

Q8 How did you come across this shared task? Please mark all answers that apply.

Answered: 9 Skipped: 0

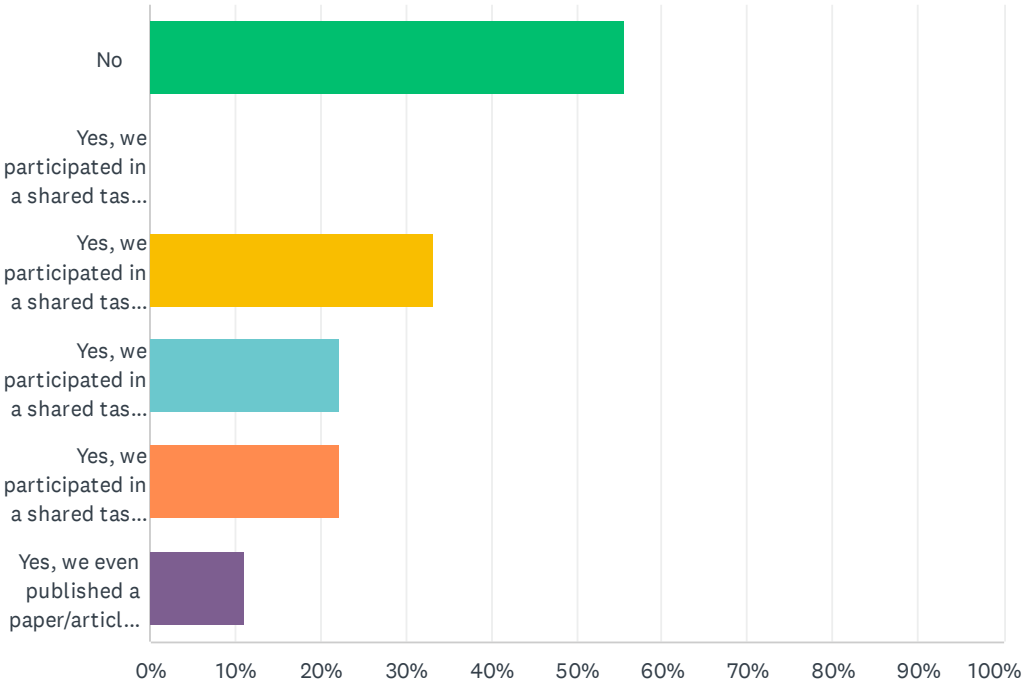


ANSWER CHOICES	RESPONSES	
We read the call-for-participation on some mailing list (e.g. corpora-list).	22.22%	2
One of the organizers informed us about this shared task.	0.00%	0
Our supervisor/course instructor encouraged us to participate in this shared task.	88.89%	8
Other (please specify):	11.11%	1
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
1	we already knew about GermEval	7/3/2025 3:20 PM

Q9 Did you or some team members have some experience with related tasks before? Please mark all answers that apply.

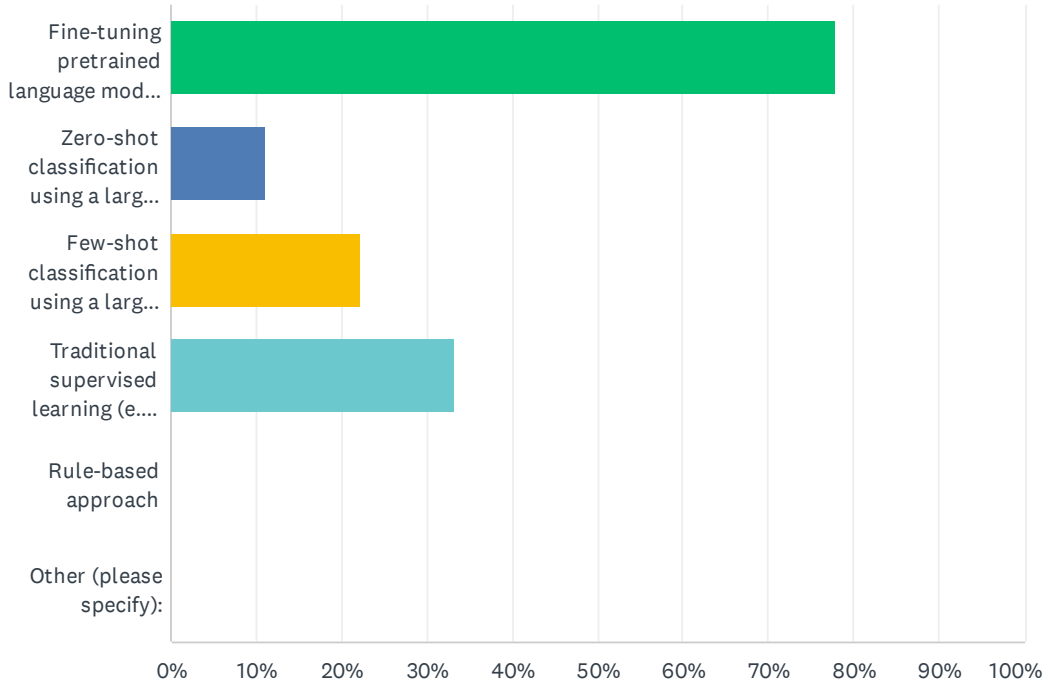
Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
No	55.56%	5
Yes, we participated in a shared task on hope speech detection before.	0.00%	0
Yes, we participated in a shared task on hate speech detection.	33.33%	3
Yes, we participated in a shared task on sentiment analysis.	22.22%	2
Yes, we participated in a shared task on German NLP.	22.22%	2
Yes, we even published a paper/article on this task.	11.11%	1
Total Respondents: 9		

Q10 What type of classification approach did you pursue? Please mark all answers that apply.

Answered: 9 Skipped: 0

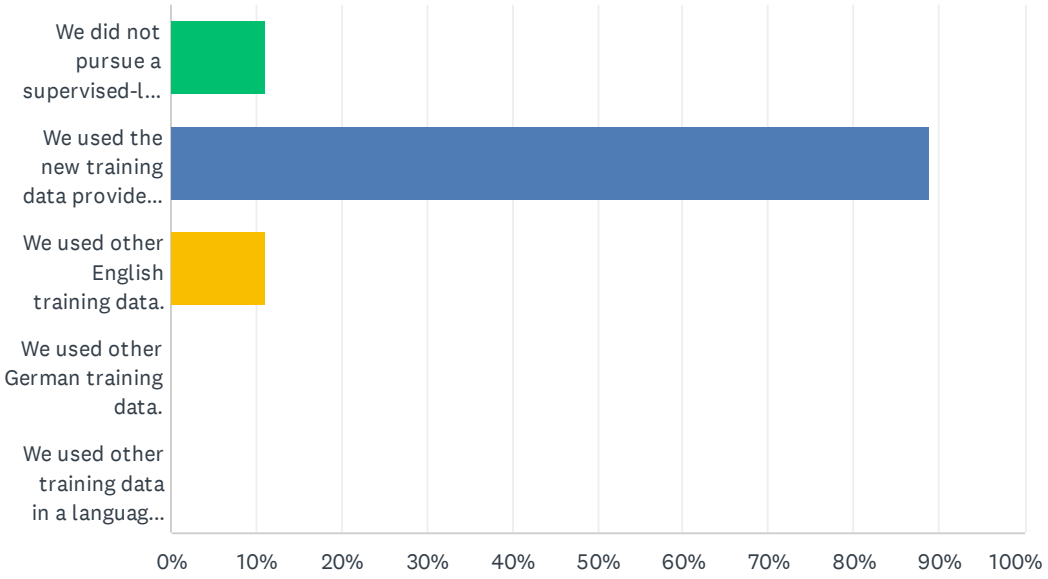


ANSWER CHOICES	RESPONSES	
Fine-tuning pretrained language models (e.g. BERT, RoBERTa, DeBERTa)	77.78%	7
Zero-shot classification using a large language model (e.g. GPT, LLaMA, Gemini)	11.11%	1
Few-shot classification using a large language model (e.g. GPT, LLaMA, Gemini)	22.22%	2
Traditional supervised learning (e.g. Logistic Regression, SVM)	33.33%	3
Rule-based approach	0.00%	0
Other (please specify):	0.00%	0
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
	There are no responses.	

Q11 What training data did you use for your supervised-learning approach? Please mark all answers that apply.

Answered: 9 Skipped: 0

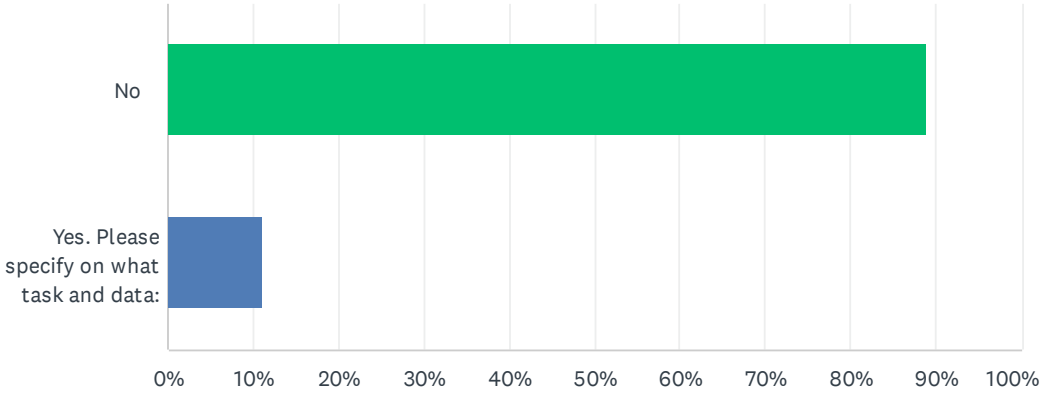


ANSWER CHOICES	RESPONSES	
We did not pursue a supervised-learning approach.	11.11%	1
We used the new training data provided by this year's GermEval shared task on Candy Speech Detection.	88.89%	8
We used other English training data.	11.11%	1
We used other German training data.	0.00%	0
We used other training data in a language other than English or German.	0.00%	0
Total Respondents: 9		

#	IF YOU USED ANY TRAINING DATA OTHER THAN THAT PROVIDED BY GERMEVAL 2025, PLEASE SPECIFY IT:	DATE
1	Go-Emotion dataset via hugging face	7/8/2025 10:24 AM

Q12 Did you pre-train a language model yourself?

Answered: 9 Skipped: 0

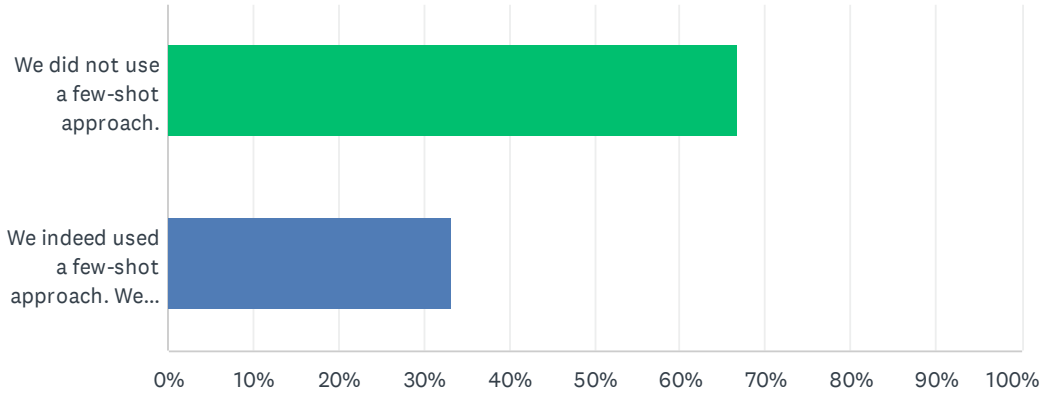


ANSWER CHOICES	RESPONSES	
No	88.89%	8
Yes. Please specify on what task and data:	11.11%	1
TOTAL		9

#	YES. PLEASE SPECIFY ON WHAT TASK AND DATA:	DATE
1	mDeBERTa for fine-tuning, Falcon for augmentation	7/4/2025 1:47 PM

Q13 If you used a few-shot approach, how many examples did you include in your prompt?

Answered: 9 Skipped: 0

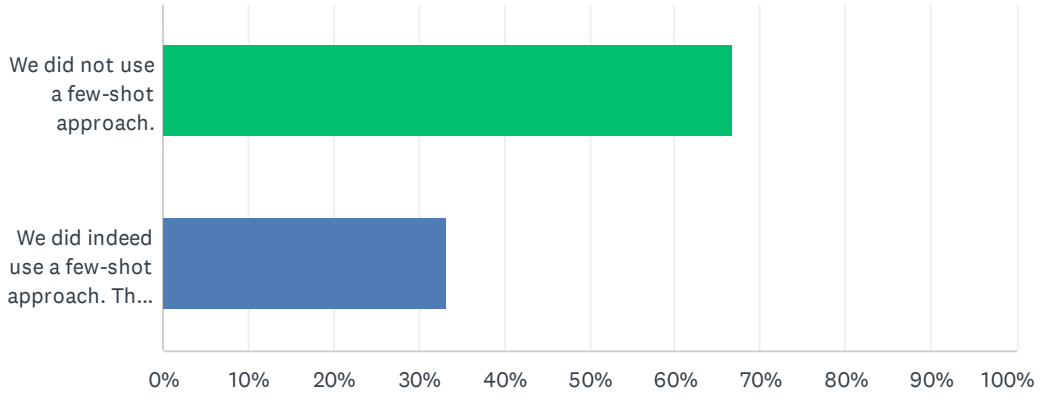


ANSWER CHOICES	RESPONSES	
We did not use a few-shot approach.	66.67%	6
We indeed used a few-shot approach. We selected the following number of examples:	33.33%	3
TOTAL		9

#	WE INDEED USED A FEW-SHOT APPROACH. WE SELECTED THE FOLLOWING NUMBER OF EXAMPLES:	DATE
1	20	7/4/2025 1:47 PM
2	50	6/30/2025 11:50 AM
3	5	6/30/2025 10:28 AM

Q14 If you used a few-shot approach, how did you choose the examples included in your prompt?

Answered: 9 Skipped: 0

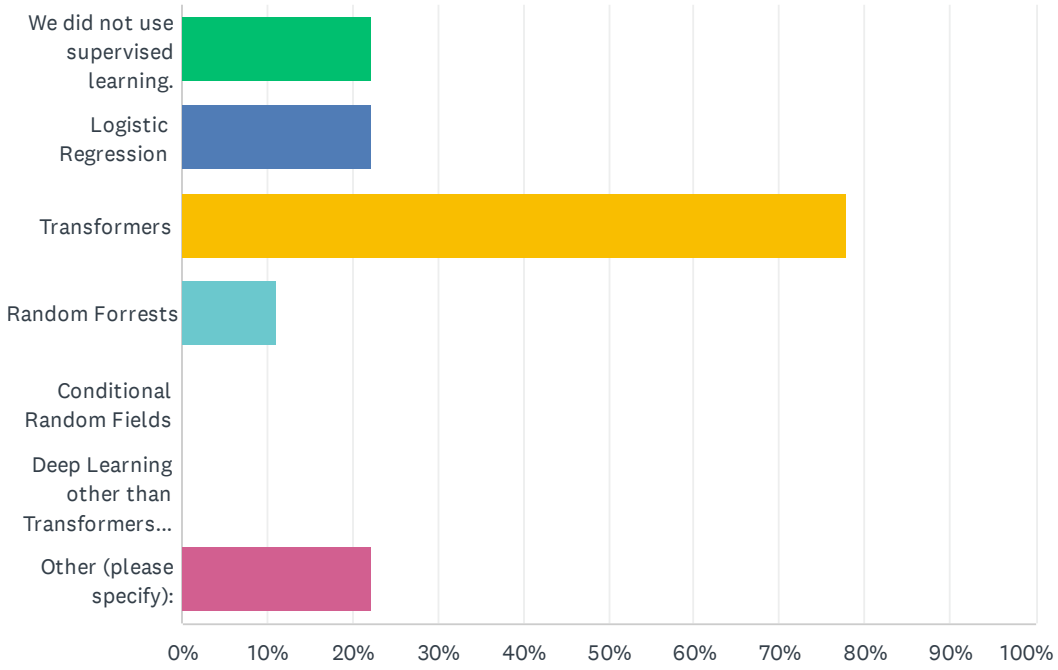


ANSWER CHOICES	RESPONSES	
We did not use a few-shot approach.	66.67%	6
We did indeed use a few-shot approach. The examples were selected as follows:	33.33%	3
TOTAL		9

#	WE DID INDEED USE A FEW-SHOT APPROACH. THE EXAMPLES WERE SELECTED AS FOLLOWS:	DATE
1	randomly	7/4/2025 1:47 PM
2	We used a Retrieval-Augmented Generation approach. Instead of manually picking examples, we automatically retrieved similar texts from a set of demonstrations using cosine similarity on text embeddings.	6/30/2025 11:50 AM
3	3 to 2 on both labels	6/30/2025 10:28 AM

Q15 What type(s) of classifier did you use for supervised learning (including fine-tuning)? Please select all that apply.

Answered: 9 Skipped: 0

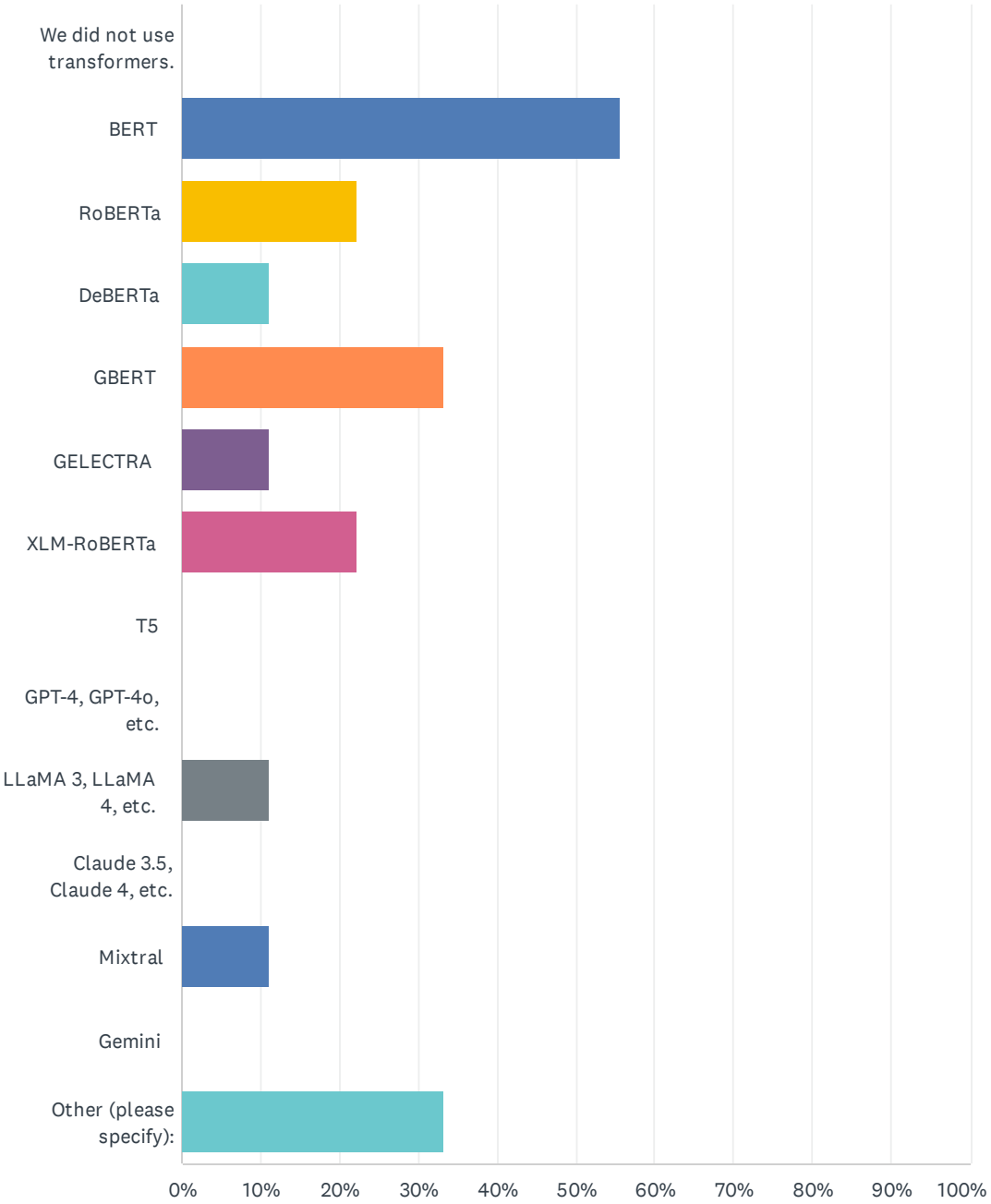


ANSWER CHOICES	RESPONSES	
We did not use supervised learning.	22.22%	2
Logistic Regression	22.22%	2
Transformers	77.78%	7
Random Forrests	11.11%	1
Conditional Random Fields	0.00%	0
Deep Learning other than Transformers (e.g. LSTM, GRU, RNN)	0.00%	0
Other (please specify):	22.22%	2
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
1	SVM, kNN, MLP	7/8/2025 10:24 AM
2	SVM	7/1/2025 11:04 AM

Q16 What type of transformer(s) did you use? Please mark all answers that apply.

Answered: 9 Skipped: 0

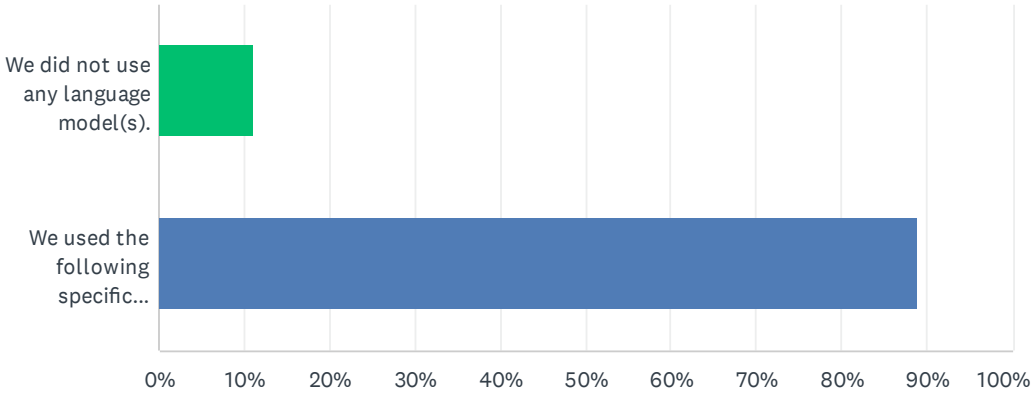


ANSWER CHOICES	RESPONSES	
We did not use transformers.	0.00%	0
BERT	55.56%	5
RoBERTa	22.22%	2
DeBERTa	11.11%	1
GBERT	33.33%	3
GELECTRA	11.11%	1
XLM-RoBERTa	22.22%	2
T5	0.00%	0
GPT-4, GPT-4o, etc.	0.00%	0
LLaMA 3, LLaMA 4, etc.	11.11%	1
Claude 3.5, Claude 4, etc.	0.00%	0
Mixtral	11.11%	1
Gemini	0.00%	0
Other (please specify):	33.33%	3
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
1	pretty much most German models, most of them can be found here: https://lsx-uniwue.github.io/SuperGLEBer-site/leaderboard_v1	7/3/2025 3:20 PM
2	Qwen3-Embedding	7/1/2025 11:04 AM
3	Qwen2.5-72B-Instruct	6/30/2025 11:50 AM

Q17 In case you used some language model(s), please list the exact model(s) that you used (e.g. bert-base-cased, deberta-large, llama-3-8B).

Answered: 9 Skipped: 0

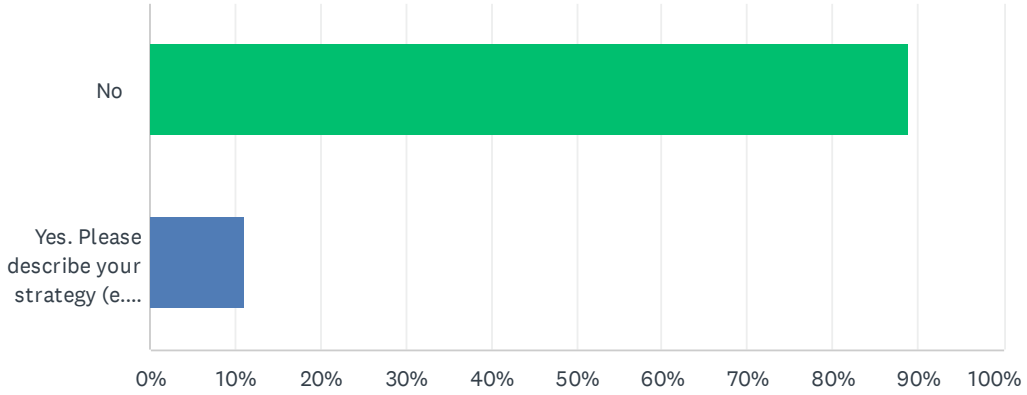


ANSWER CHOICES	RESPONSES	
We did not use any language model(s).	11.11%	1
We used the following specific models:	88.89%	8
TOTAL		9

#	WE USED THE FOLLOWING SPECIFIC MODELS:	DATE
1	bert-base-german-cased	7/8/2025 10:26 AM
2	oliverguhr/german-sentiment-bert	7/8/2025 10:24 AM
3	mDeBERTa for fine-tuning, Falcon for augmentation	7/4/2025 1:47 PM
4	"LSX-UniWue/LLaMmlein_120M", "LSX-UniWue/LLaMmlein_1B", "LSX-UniWue/LLaMmlein_7B", "DiscoResearch/Llama3-German-8B", "EuroBERT/EuroBERT-2.1B", "EuroBERT/EuroBERT-210m", "EuroBERT/EuroBERT-610m", "FacebookAI/xlm-roberta-base", "GeistBERT/GeistBERT_base", "LSX-UniWue/ModernGBERT_134M", "LSX-UniWue/ModernGBERT_1B", "LeoLM/leo-hessianai-7b", "Qwen/Qwen2.5-0.5B", "Qwen/Qwen2.5-7B", "Qwen/Qwen3-0.6B", "Qwen/Qwen3-1.7B", "Qwen/Qwen3-4B", "TUM/GottBERT_large", "benjamin/gerpt2", "benjamin/gerpt2-large", "bert-base-german-cased", "bigscience/bloomz-560m", "dbmdz/german-gpt2", "deepset/gbert-base", "deepset/gbert-large", "deepset/gelectra-base", "deepset/gelectra-large", "facebook/mbart-large-50", "flair/bueble-lm-2b", "gerturax/gerturax-1", "gerturax/gerturax-2", "gerturax/gerturax-3", "malteos/bloom-6b4-clp-german", "maxidl/DOSMo-7B-v0.2", "meta-llama/Llama-3.2-1B", "meta-llama/Llama-3.2-3B", "meta-llama/Meta-Llama-3.1-8B", "uklfr/gottbert-base", "utter-project/EuroLLM-1.7B"	7/3/2025 3:20 PM
5	deepset/gbert-large, FacebookAI/xlm-roberta-large, Qwen/Qwen3-Embedding-8B	7/1/2025 11:04 AM
6	gbert-large, bert-base-german-cased, roberta-large	6/30/2025 1:58 PM
7	all-mpnet-base-v2, Qwen2.5-72B-Instruct, Llama-3.3-70B-Instruct	6/30/2025 11:50 AM
8	Mistral-nemo	6/30/2025 10:28 AM

Q18 Did you employ a specific strategy for prompt engineering?

Answered: 9 Skipped: 0

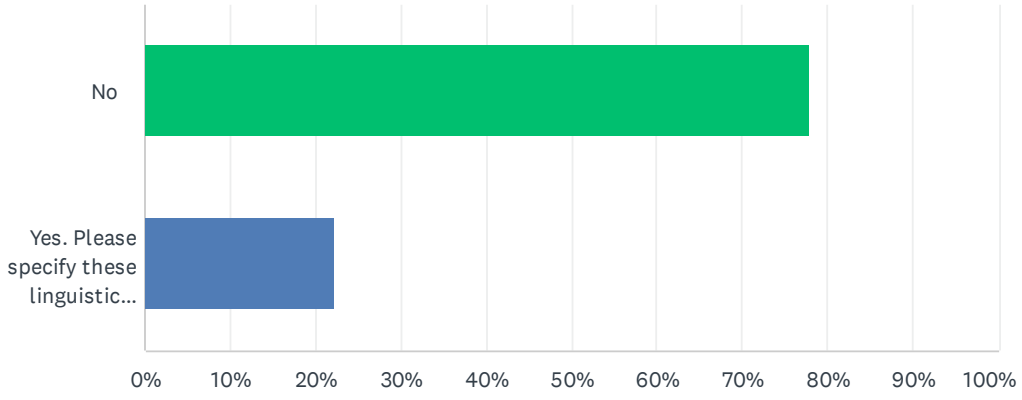


ANSWER CHOICES	RESPONSES	
No	88.89%	8
Yes. Please describe your strategy (e.g. chain-of-thought prompting):	11.11%	1
TOTAL		9

#	YES. PLEASE DESCRIBE YOUR STRATEGY (E.G. CHAIN-OF-THOUGHT PROMPTING):	DATE
1	but only to generate a list of candy speech words (using chatGPT)	6/30/2025 1:58 PM

Q19 Did you employ some linguistic heuristics or features (e.g. as part of prompts for an LLM)?

Answered: 9 Skipped: 0

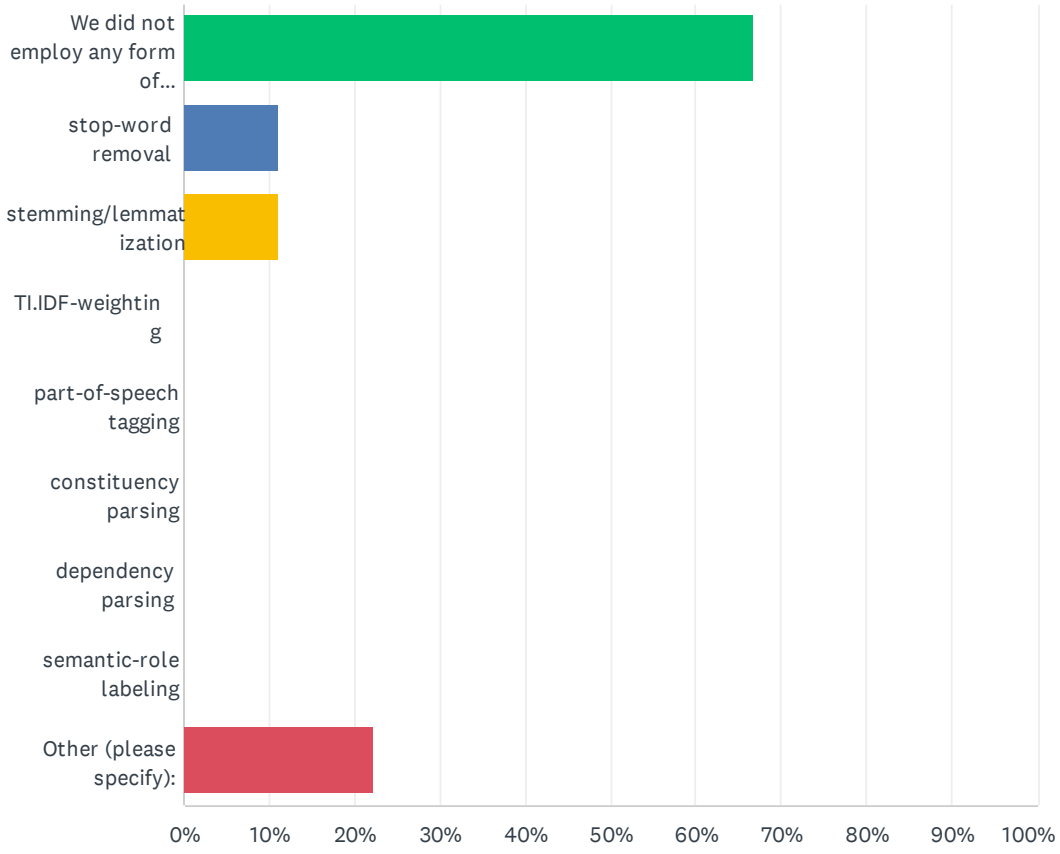


ANSWER CHOICES	RESPONSES	
No	77.78%	7
Yes. Please specify these linguistic heuristics or features:	22.22%	2
TOTAL		9

#	YES. PLEASE SPECIFY THESE LINGUISTIC HEURISTICS OR FEATURES:	DATE
1	diverse linguistic features for classification by logistic regression (e.g., postive word/token count, emoji count, all caps, ...) for span detection: spacy dependency trees	6/30/2025 1:58 PM
2	The use of certain words to restrict LLM from answering beyond what we expected(halucination). it worked in subtask1 but not 2.	6/30/2025 10:28 AM

Q20 Did you employ some form of (linguistic) pre-processing? (Standard tokenization as part of transformers does not count as such pre-processing.) Please mark all answers that apply.

Answered: 9 Skipped: 0

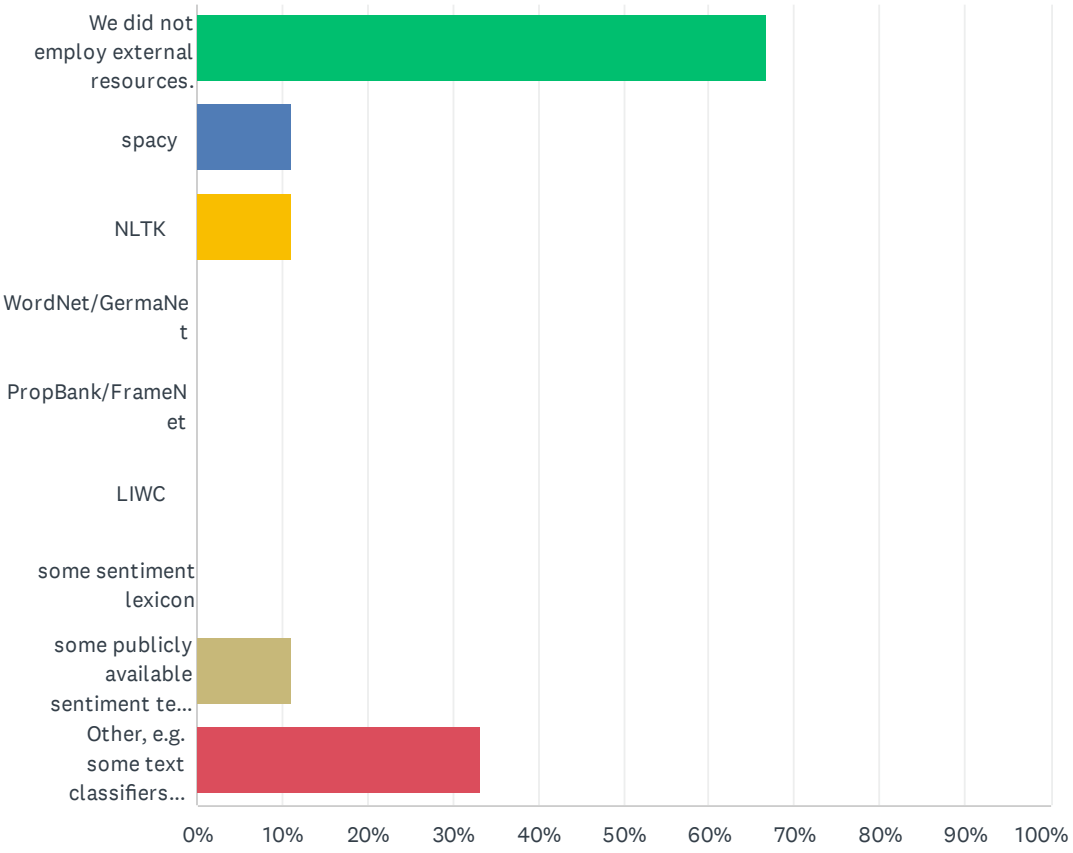


ANSWER CHOICES	RESPONSES	
We did not employ any form of pre-processing.	66.67%	6
stop-word removal	11.11%	1
stemming/lemmatization	11.11%	1
TI.IDF-weighting	0.00%	0
part-of-speech tagging	0.00%	0
constituency parsing	0.00%	0
dependency parsing	0.00%	0
semantic-role labeling	0.00%	0
Other (please specify):	22.22%	2
Total Respondents: 9		

#	OTHER (PLEASE SPECIFY):	DATE
1	First we converted each comment from latin1 to the utf-8 format. After that we cleaned each line as follows: \\ Remove the @ (e.g., @user), each link, each smiley, special character, underscore, single letter word, words with numbers (e.g., h4x0r), too many dots (e.g., ... $\rightarrow$.) and multiple question marks. \\ We also normalized the colloquial language, replaced slang terms with standard language (e.g., jz $\rightarrow$ jetzt), shortened excessively long letter repetitions, and stripped all punctuation marks.	7/8/2025 10:24 AM
2	spelling correction and translation	6/30/2025 1:58 PM

Q21 What type of external (linguistic) resources (e.g. tools, ontologies, knowledge bases, dictionaries) did you employ for your approach? (Statistical models, such as word embeddings/language models, should not be listed here.) Please mark all answers that apply.

Answered: 9 Skipped: 0

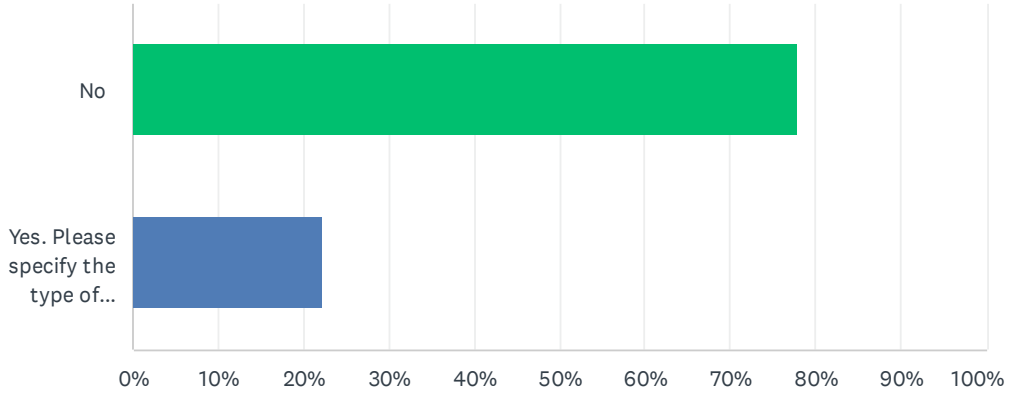


ANSWER CHOICES	RESPONSES	
We did not employ external resources.	66.67%	6
spacy	11.11%	1
NLTK	11.11%	1
WordNet/GermaNet	0.00%	0
PropBank/FrameNet	0.00%	0
LIWC	0.00%	0
some sentiment lexicon	0.00%	0
some publicly available sentiment text classifier	11.11%	1
Other, e.g. some text classifiers from huggingface (please specify):	33.33%	3
Total Respondents: 9		

#	OTHER, E.G. SOME TEXT CLASSIFIERS FROM HUGGINGFACE (PLEASE SPECIFY):	DATE
1	Doc2Vec (gensim)	7/8/2025 10:24 AM
2	We did try different resources and approaches but not in the final model.	7/4/2025 1:47 PM
3	spelling correction with /oliverguhr/spelling-correction-german-base and emotion detection with emotion-english-distilroberta-base and machine translation with Helsinki-NLP/opus-mt-de-en and TextBlob for sentiment polarity	6/30/2025 1:58 PM

Q22 Did you use some form of ensemble method in your approach?

Answered: 9 Skipped: 0

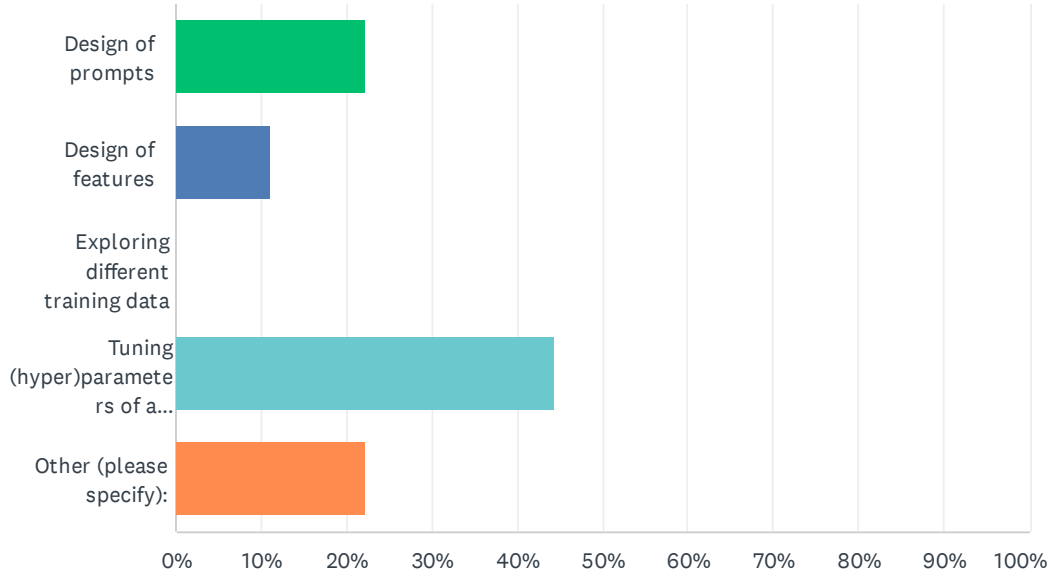


ANSWER CHOICES	RESPONSES	
No	77.78%	7
Yes. Please specify the type of ensemble:	22.22%	2
TOTAL		9

#	YES. PLEASE SPECIFY THE TYPE OF ENSEMBLE:	DATE
1	from sklearn.ensemble import RandomForestClassifier	7/8/2025 10:24 AM
2	stacking: output of BERT classifier as input for meta classifier (logistic regression) which also uses linguistic features. for task 2: stacking two BERT models: one for span detection and one for span classification	6/30/2025 1:58 PM

Q23 On what aspect did you focus when you built your system?

Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Design of prompts	22.22%	2
Design of features	11.11%	1
Exploring different training data	0.00%	0
Tuning (hyper)parameters of a (supervised) classifier	44.44%	4
Other (please specify):	22.22%	2
TOTAL		9

#	OTHER (PLEASE SPECIFY):	DATE
1	Comparing different model versions across subtasks.	7/8/2025 10:26 AM
2	systematic evaluation of available models against available German tasks	7/3/2025 3:20 PM

Q24 What, in your view, is the most important component (e.g. feature, rule, classifier, language model etc.) of your system?

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	The classifier	7/8/2025 10:26 AM
2	Main-MLP	7/8/2025 10:24 AM
3	classifier, language model	7/4/2025 1:47 PM
4	classifier	7/4/2025 9:43 AM
5	language model	7/3/2025 3:20 PM
6	language model	7/1/2025 11:04 AM
7	improving transformer classifications with linguistic features	6/30/2025 1:58 PM
8	The LLM (Qwen2.5-72B-Instruct)	6/30/2025 11:50 AM
9	Language model	6/30/2025 10:28 AM

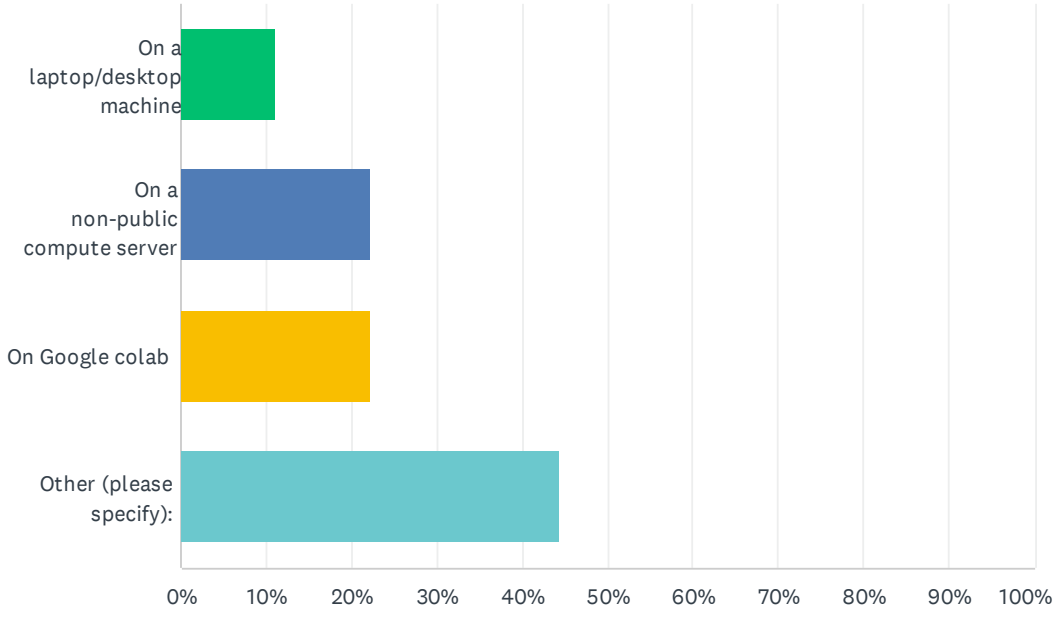
Q25 What, in your view, is the most effective component (e.g. feature, rule, classifier, language model etc.) of your system?

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	The BIO tagging setup	7/8/2025 10:26 AM
2	Main-MLP, maybe data cleaning	7/8/2025 10:24 AM
3	augmentation and Lora fine-tuning	7/4/2025 1:47 PM
4	language model	7/4/2025 9:43 AM
5	language model	7/3/2025 3:20 PM
6	language model	7/1/2025 11:04 AM
7	BERT models	6/30/2025 1:58 PM
8	The LLM (Qwen2.5-72B-Instruct)	6/30/2025 11:50 AM
9	Classifier	6/30/2025 10:28 AM

Q26 Where did you train/run your system?

Answered: 9 Skipped: 0

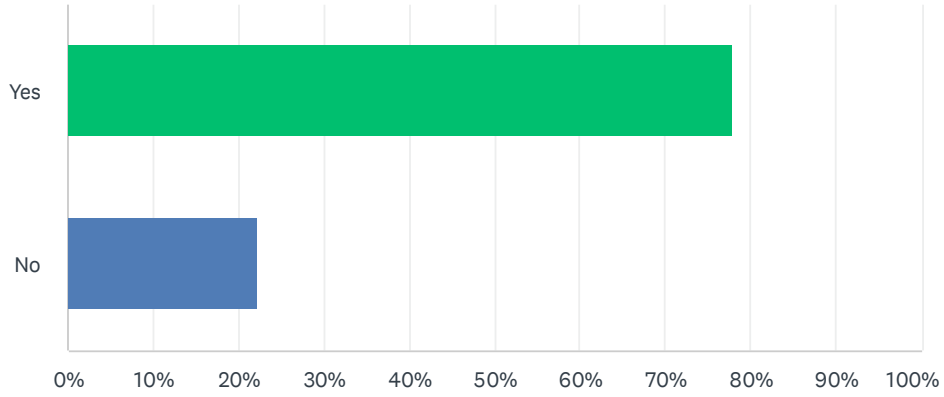


ANSWER CHOICES	RESPONSES	
On a laptop/desktop machine	11.11%	1
On a non-public compute server	22.22%	2
On Google colab	22.22%	2
Other (please specify):	44.44%	4
TOTAL		9

#	OTHER (PLEASE SPECIFY):	DATE
1	GPU Cluster of University	7/8/2025 10:26 AM
2	laptop/desktop machine and google collab	7/8/2025 10:24 AM
3	FH-Cluster	7/4/2025 9:43 AM
4	NHR@FAU cluster	7/3/2025 3:20 PM

Q27 Did you use GPUs to run your experiments? Note: Using an LLM such as GPT-4 via an API-based web service does not count as using GPUs.

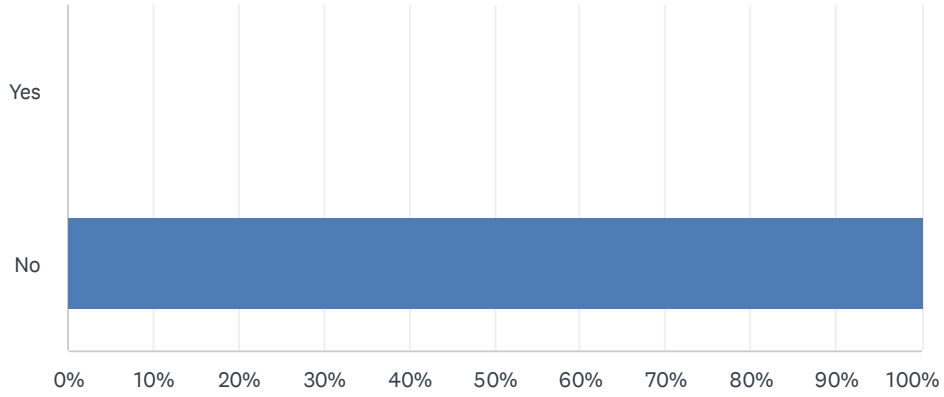
Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes	77.78%	7
No	22.22%	2
TOTAL		9

Q28 Did your approach rely on a subscription-based platform that provides access to an LLM via an API-based web service, e.g. using OpenAI’s API to obtain responses from GPT-4o?

Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes	0.00%	0
No	100.00%	9
TOTAL		9

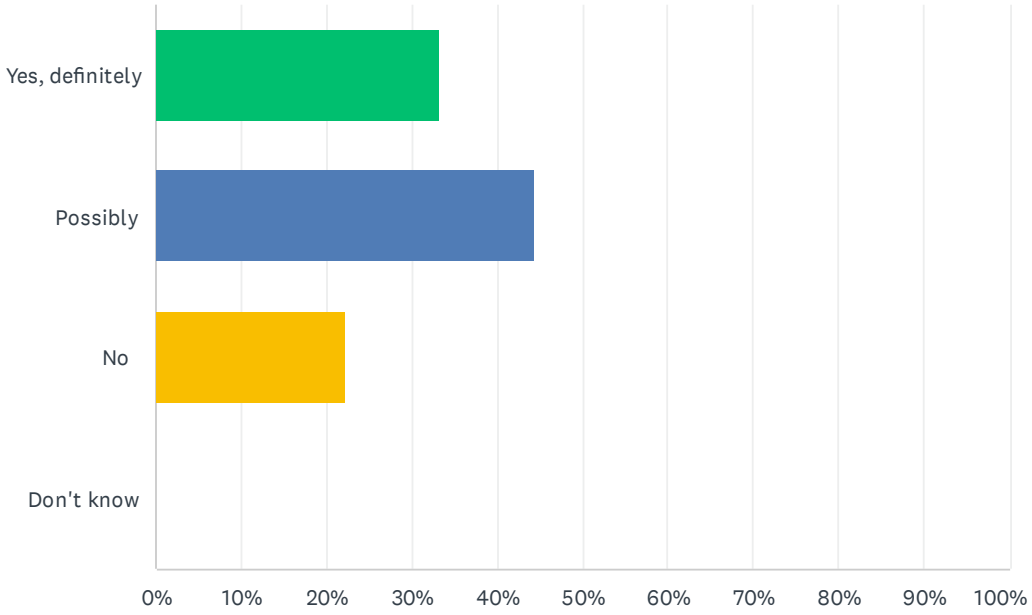
Q29 If possible, specify the essential components of your IT infrastructure (e.g. storage, GPUs etc.).

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	https://www.uni-augsburg.de/de/fakultaet/med/profs/exposure-science/aktuelles/linux-compute-cluster-augsburg/	7/8/2025 10:26 AM
2	NVIDIA RTX 2060 SUPER, AMD RYZEN (?) Prozessor, NVIDIA RTX 4070 SUPER, INTEL I7 14700 KF	7/8/2025 10:24 AM
3	We used Google Colab	7/4/2025 1:47 PM
4	We have not the details of the FH Cluster only that https://www.ki.fh-swf.de/	7/4/2025 9:43 AM
5	is this a question for quantities or type, or..? we didn't really need storage, or cpus, or... just GPUs, of which we had access to the H100 type	7/3/2025 3:20 PM
6	AMD Ryzen Threadripper PRO 7995WX with two NVIDIA RTX 6000 Ada Generation GPUs + Intel(R) Xeon(R) Platinum 8168 CPU @ 2.70GHz with a single NVIDIA RTX A6000 GPU	7/1/2025 11:04 AM
7	Google Colab T4 for fine-tuning BERT models	6/30/2025 1:58 PM
8	We got four NVIDIA H100 NVL	6/30/2025 11:50 AM
9	maximized effort to run on low level resources, no particular gpu or staorage needed	6/30/2025 10:28 AM

Q30 Do you think that the quality of your approach could be substantially improved if you simply had access to more powerful computing resources (e.g. stronger GPUs)?

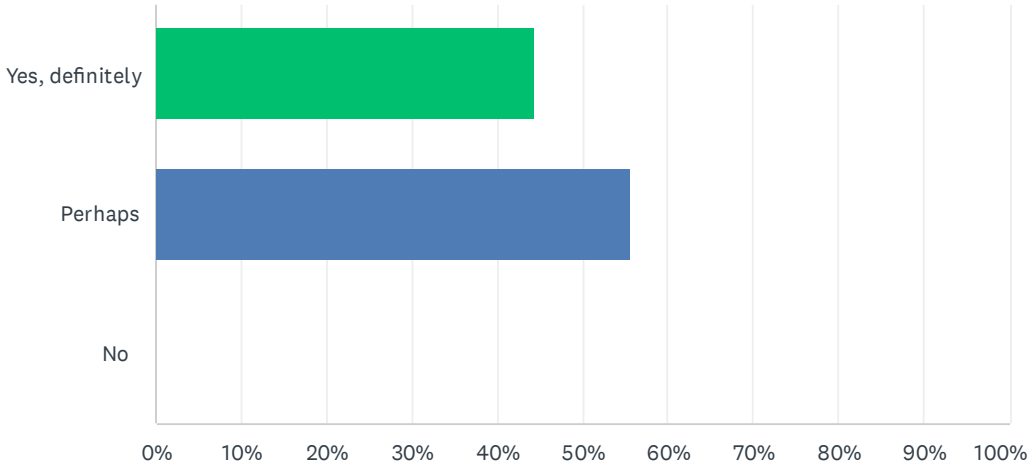
Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes, definitely	33.33%	3
Possibly	44.44%	4
No	22.22%	2
Don't know	0.00%	0
TOTAL		9

Q31 Would your team be interested in participating in another edition of this task in the future?

Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes, definitely	44.44%	4
Perhaps	55.56%	5
No	0.00%	0
TOTAL		9

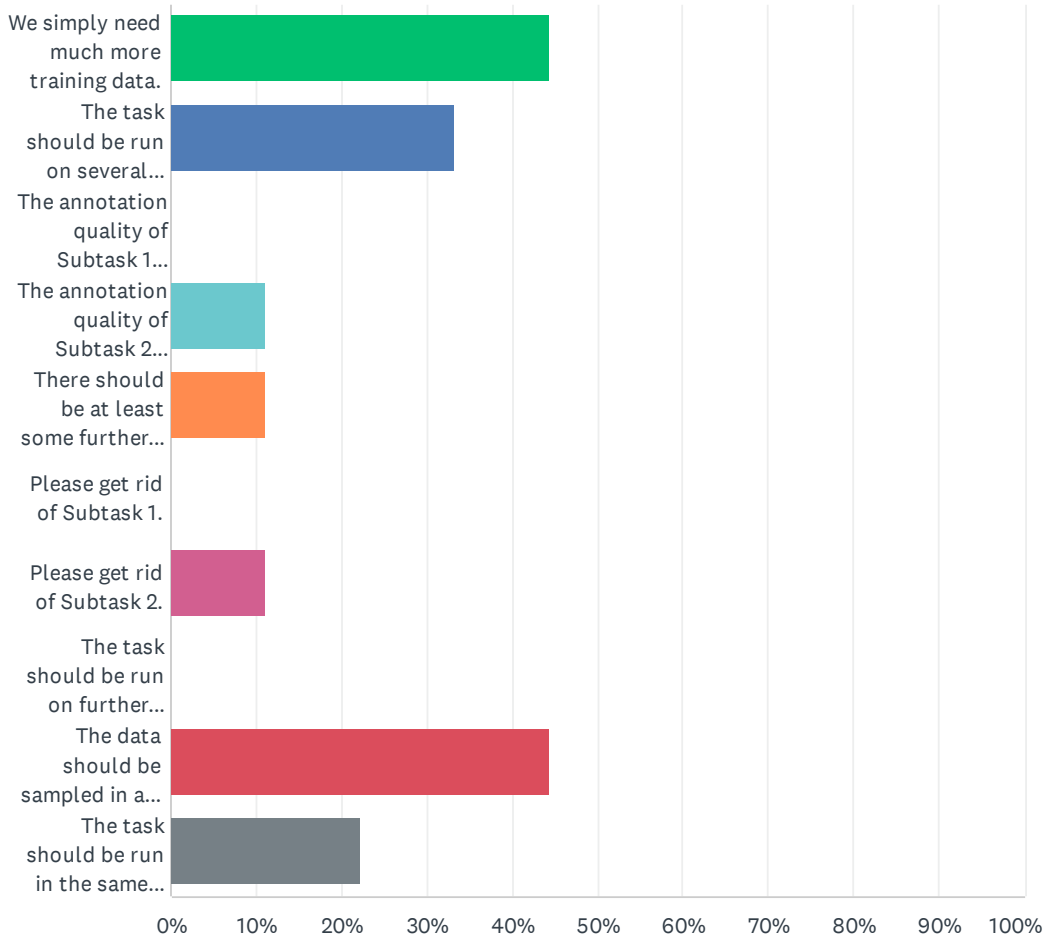
Q32 What are your main learnings or key takeaways (1–3 keywords) from participating in this shared task?

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1	linguistic techniques might help substantially model’s understanding sarcasm and labeling it correctly is hard	7/8/2025 10:26 AM
2	Consistency, patience, GPU	7/8/2025 10:24 AM
3	classification models and augmentation methods	7/4/2025 1:47 PM
4	Coding in a team and limitaitons for students for coding sametime.	7/4/2025 9:43 AM
5	tagging had a huge variance in results across models, but classification didn't. also the evaluation metric for tagging didn't really seem to correlate to our BIO-label-based F1 metric	7/3/2025 3:20 PM
6	solving more difficult tasks can result in better performance on easier downstream tasks	7/1/2025 11:04 AM
7	2step stacking better than direct model	6/30/2025 1:58 PM
8	When using Llama-3.3-70B-Instruct to add XML tags representing spans in a text, the model removes all whitespaces before punctuation marks. This causes issues when trying to extract the exact spans accurately.	6/30/2025 11:50 AM
9	It was a worthy challenge. Further studies will continue	6/30/2025 10:28 AM

Q33 What should be changed about the setting of a future edition of this task? Please mark all answers that apply.

Answered: 9 Skipped: 0



ANSWER CHOICES	RESPONSES	
We simply need much more training data.	44.44%	4
The task should be run on several domains.	33.33%	3
The annotation quality of Subtask 1 should be improved.	0.00%	0
The annotation quality of Subtask 2 should be improved.	11.11%	1
There should be at least some further (optional) subtask.	11.11%	1
Please get rid of Subtask 1.	0.00%	0
Please get rid of Subtask 2.	11.11%	1
The task should be run on further languages.	0.00%	0
The data should be sampled in a more unbiased manner.	44.44%	4
The task should be run in the same format as before.	22.22%	2
Total Respondents: 9		

#	OTHER SUGGESTIONS (INCLUDING SUGGESTIONS FOR ADDITIONAL SUBTASKS):	DATE
1	as far as we noticed, there were block comments from limited number of youtubers and fans of a specific youtuber might have their own language, and this might decrease model’s ability to generalize	7/8/2025 10:26 AM
2	Maybe also providing a GPU remotley.	7/4/2025 9:43 AM
3	the datapoints looked a bit "uniform" at times	7/3/2025 3:20 PM
4	the heterogeneity of the commenters' age and demographics could be increased. the commenters seemed to be quite young overall. some data points included sensitive personal information (i.e., addresses)	7/1/2025 11:04 AM
5	it is always nice to have a development phase in which more models can be submitted that are ranked.	6/30/2025 1:58 PM

Q34 What suggestions do you have for improving the administration of this task?

Answered: 9 Skipped: 0

#	RESPONSES	DATE
1		7/8/2025 10:26 AM
2	Nothing :)	7/8/2025 10:24 AM
3	Please do not force attendance for the shared task to participate in the conference, similar to SemEval shared tasks. We would like to have our paper published by you, but we may not be able to attend the conference.	7/4/2025 1:47 PM
4	You are great!	7/4/2025 9:43 AM
5	none	7/3/2025 3:20 PM
6	being able to see the other teams' scores on a seperate validation set (additional trial phase?) to increase the competitive spirit. more than three submissions would be nice	7/1/2025 11:04 AM
7	nothing, was perfect	6/30/2025 1:58 PM
8	The administration was great. Keep up the good work! Thank you all!	6/30/2025 11:50 AM
9	There should be a live leaderboard so we see how we are doing. Doing this was blind and not knowing what direction you are going	6/30/2025 10:28 AM