

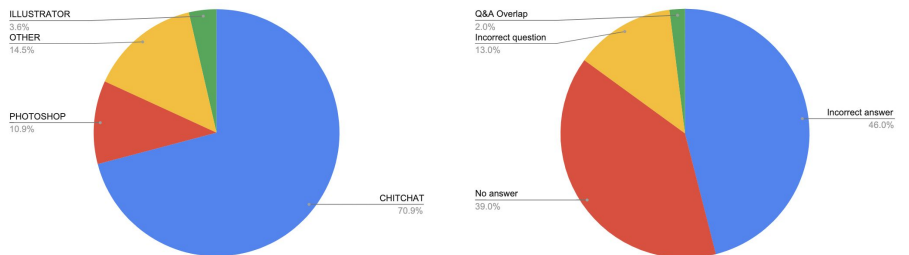
# BehanceQA: A New Dataset for Identifying Questions Pairs in Video Transcripts

Amir Pouran Ben Veyseh<sup>1</sup>, Viet Dac Lai<sup>1</sup>, Franck Dernoncourt<sup>2</sup>, Thien Huu Nguyen<sup>1</sup>

<sup>1</sup> University of Oregon, <sup>2</sup> Adobe Research

{apouranb,thien,viet}@cs.uoregon.edu, {franck.dernoncourt}@adobe.com

## Dataset and Analysis



## Statistics

Statistics	Train	Test	Dev	Total
# Sentences	220,520	23,273	14,104	257,897
# Questions	13,903	1,470	993	16,366
# Answers	4,371	527	303	5,272
Avg Question Length	7.78	6.13	8.88	7.79
Avg Answer Length	18.35	17.22	15.48	18.05

## Results

Model	Question			Answer			Overall		
	P	R	F1	P	R	F1	P	R	F1
RULE <sup>Q</sup> -BERT <sup>A</sup> (Pipeline)	43.83	51.72	47.45	36.72	2.80	5.22	41.47	37.99	39.65
RULE <sup>Q</sup> -BERT+CRF <sup>A</sup> (Pipeline)	43.83	51.72	47.45	25.98	2.38	4.37	38.52	37.88	38.19
BERT <sup>Q</sup> -BERT <sup>A</sup> (Pipeline)	81.31	93.37	86.92	42.25	4.70	8.47	69.95	68.49	69.21
BERT <sup>Q</sup> -BERT+CRF <sup>A</sup> (Pipeline)	81.31	93.37	86.92	41.13	4.16	7.56	69.64	68.34	68.98
Joint_BERT (Joint)	85.22	78.12	81.52	41.19	6.53	11.28	72.47	58.00	64.43
Joint_BERT+CRF (Joint)	84.16	76.88	80.36	40.39	5.79	10.13	71.50	56.91	63.37

## Conclusion and Future Work

- We propose a new dataset for identifying questions and answers in video transcripts
- The dataset is manually labeled and provides more than 257,000 sentences
- The dataset is annotated for the domain of Behance videos
- We conduct analysis by comparing pipeline and joint models and also rule-based and learning based techniques