

BehanceCC: A ChitChat Detection Dataset For Livestreaming Video Transcripts

¹ Viet Dac Lai, ¹ Amir Poursan Ben Veyseh, ² Franck Deroncourt, and ¹ Thien Huu Nguyen
¹Department of Computer and Information Science, University of Oregon, Eugene, OR, USA
²Adobe Research, San Jose, CA, USA
{vietl, apouran, thien}@cs.uoregon.edu
franck.deroncourt@adobe.com

1. Abstract

Livestreaming videos have become an effective broadcasting method for both video sharing and educational purposes. However, livestreaming videos contain a considerable amount of off-topic content (i.e., up to 50%) which introduces significant noises and data load to downstream applications. This paper presents BehanceCC, a new human-annotated benchmark dataset for off-topic detection (also called chitchat detection) in livestreaming video transcripts. In addition to describing the challenges of the dataset, our extensive experiments of various baselines reveal the complexity of chitchat detection for livestreaming videos and suggest potential future research directions for this task. The dataset will be made publicly available to foster research in this area.

The dataset is freely accessible at <https://github.com/nlp-uoregon/behancecc>.

2. Chitchat Detection

Non-chitchat
Brainstorming, commenting, discussing of designing idea
Mentioning a tool
Mentioning an artistic object
Mentioning an action
Mentioning an action with computer hardware such as keyboard, mouse, and drawing tablet
Planning/Introduction of the work in the video
Mentioning color, shape, size, pattern, direction
Mentioning a graphical user interface of the graphical design tool
Chitchat
Welcoming
Small talk with audience about unrelated topics
Verbal pauses
Transitional sentences
Filling sentences, confirmation sentences
Talking about the streamer interests
Talking about tips not related to the purpose of the video
Talking about traveling/careers/politics/breaking news

Figure 1. Examples of chitchat/non-chitchat sentences in livestreaming videos in BehanceCC dataset.

2. Examples

Sure.
Alright, OK.
Get the head in here.
Really rough.
We may be right here.
Caller.
He appreciated jerk.
The same icon, just clicking it a second time has a new label.
Oh, really, so it actually does say like this.

Figure 1. Chitchat utterances in livestreaming transcript (Chitchat sentences are highlighted in orange)

3. Challenges

Close topics
I was going to say something along the lines of like when Blizzard announced classic servers, I thought that we'd never see the day.
And if we did, it would be near the end of World of Warcraft, like for the Warcraft Lifespan.
I'm not really getting that feeling anymore.
I think that Classic in retail can easily coexist.
Right?
Did I did I want to make that head bigger.
Thanks alright.
See the ear.
Space here.
Holler That thing this comes way down lower.
A little higher.

Figure 2. An example of game-related chitchat texts in the transcript of a livestreaming videos about graphical design. (Chitchat sentences are highlighted in orange)

Word errors
Have you purchased the **PS five?**
I probably won't push it.
Purchase it until a year after.
Um, or whenever their second generation of **PS five** come out.
...
Versions of the **PSS**, right?

Figure 3. Examples of inconsistent and noisy texts in transcripts of livestreaming videos. "PS" may refer to the Sony Play Station game console or the Adobe Photoshop software. (Chitchat sentences are highlighted in orange, word errors are bold and underlined, correct word are highlighted in green.)

4. Data Statistics

#	Train	Dev	Test
Document	2,514	198	199
Sentence	154,897	11,175	12,216
Token	1,466,035	99,947	105,128
Max sentence/document	105	88	88
Chitchat Sentence	75,980	6,031	5,566

Table 1. Statistics of the BehanceCC dataset.

5. Evaluation

	Model	Test		
		P	R	F
BERT	MLP	66.7	91.5	77.2
	CRF	63.8	94.3	76.1
	BiLSTM	75.2	91.8	82.7
	BiLSTM+CRF	71.6	93.3	81.0
RoBERTa	MLP	68.0	90.4	77.6
	CRF	66.8	91.9	77.4
	BiLSTM	70.5	94.7	80.8
	BiLSTM+CRF	69.3	95.1	80.2

Table 2. Performances of the examined models on the BehanceCC dataset.

6. References

Cieri, Christopher, et al. Fisher English training speech part 1 transcripts. *Philadelphia: Linguistic Data Consortium* (2004).

Konigari, Rachna, et al. Topic Shift Detection for Mixed Initiative Response. *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 2021.