

PogChamp or WutFace?

Twitch Chat Sentiment Analysis

Armin Bernstetter

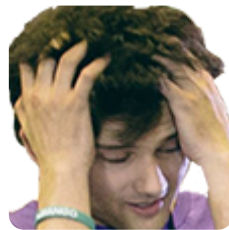
2019-06-19

Inhalt

1. Twitch
2. Sentiment Analysis
3. Vorarbeit
4. Twitch Chat Sentiment Analysis
5. Ergebnisse/Case Studies

Paper Details

- Eingereicht bei ACM Transactions on Social Computing
- Special Issue on Emoji Understanding and Applications in Social Media
- Eingereicht November 2018 (!!!)
- Erste Reviews Anfang Juni 2019 (!!!)



Aktuell: Arbeit an Minor Revision

Teil I: Twitch

Details

- Live streaming Plattform
- Fokus auf Gaming
- Gegründet 2011 als Teil der Plattform Justin.tv
- 2014 aufgekauft von Amazon

Content

- Ursprünglich gedacht für Esport Broadcasts
- Aktuell:
 - Let's Plays
 - IRL streams
 - Esport
 - Events (Ankündigungen, PKs etc)

Stream

Interaktion

- Wichtig für Twitch: Interaktion der Viewer
- Esport/Events: “Public viewing”, gemeinsame Reaktion auf Stream
- Streamer: Interaktion mit Person vor der Kamera
- Chat, Donations, Subscriptions

Chat

Emotes

- Für Nutzer oft wichtigster Aspekt des Chats
- Twitch-Emotes sind Memes
- Deutlich aussagekräftiger als normale Emojis
- Besitzen oft eigene Hintergrundgeschichten

Arten von Emotes

- Globale Twitch Emotes
- Subscriber Emotes
- “Externe” Emotes (via Browser Extensions BetterTwitchTV, FrankerFaceZ)

Beispiele

Emote	Name	Bedeutung/Benutzung
	Kappa	Ironie/Sarkasmus
	WutFace	Ekel, negative Überraschung
	PogChamp	Begeisterung
	ANELE	Häufig rassistisch z.B. Dota Hero Techies
	cmonBruh	Ebenfalls häufig rassistisch. z.B. Streamer benutzt das Wort "black"

Häufigste Wörter im Corpus

Rank	Twitch	Twitch (no emotes)	English	Twitter
1	LUL 	like	new	tinyurl.com
2	Kappa 	get	home	new
3	<3 	lol	us	like
4	PogChamp 	u	page	good
5	like	good	search	get

Fazit

- Emotes besitzen großen Anteil am Chat
- Emotes sind sehr aussagekräftig
- Emotes können eindeutig positiv oder negativ sein

⇒ Können Emotes benutzt werden um Sentiment Analysis auf Twitch Chat Text auszuführen?

Teil II: Sentiment Analysis Allgemein

Basics

- Forschungsbereich im Text Mining/NLP
- Extraktion von Emotionen, Meinungen etc aus Text
- Beispiele: Produkt-/Filmreviews, Social Media

Ansätze

- Lexikon
 - Wörterbuch mit meist numerischen Werten zu jedem Wort
 - Einfachster Ansatz: Durchschnitt über Textabschnitt (z.B. Satz)
 - Problem: Unbekannte Wörter
- Machine Learning
 - Gelabelter Datensatz (meist Satz + Sentiment label) als Trainingsdaten

Beispiel: VADER

- Valence Aware Dictionary and sEntiment Reasoner
- Teil der Python Library NLTK
- Angepasst an Social Media

Beispiel aus dem Wörterbuch:

Token	Mean-Sentiment-Rating	Standard Deviation	Raw-Human-Sentiment-Ratings
wtf	-2.8	0.74833	[-4, -3, -2, -3, -2, -2, -2, -4, -3, -3]

Beispiel: VADER (cont.)

Text	Result
VADER is smart, handsome, and funny.	{‘pos’: 0.746, ‘compound’: 0.8316, ‘neu’: 0.254, ‘neg’: 0.0}
VADER is smart, handsome, and funny!	{‘pos’: 0.752, ‘compound’: 0.8439, ‘neu’: 0.248, ‘neg’: 0.0}

Problem

Existierende Lexika beinhalten maximal Emoticons wie “:.)”, “:(“, “:D” etc

⇒ Ein spezielles Twitch-Sentiment Lexikon wird benötigt

Teil III: Vorarbeit

Twitch Datensatz

- Ungelabelt
- Ca. 3 Milliarden Nachrichten
- Daten aus April, Mai und Juni 2018
- Crawling über die Twitch API

Gelabelter Testdatensatz

- 2000 Nachrichten
- Handgelabelt durch Crowd Sourcing (Figure Eight)
- Zur Evaluation der Sentiment Analysis Methoden

Sentiment Lexikon

Drei Lexika:

- VADER Lexikon
- Emoji Lexikon (unicode emojis)
- Neu erstelltes Twitch Emote Lexikon

Emote Sentiment Umfrage

- Top 100 Emotes (Häufigkeit)
- Twitch Emotes sowie BTTV und FFZ Emotes
- Google forms Umfrage, verbreitet über Twitter und Reddit
- 108 Antworten
 - Ca. 80% davon “fairly familiar” oder “extremely familiar” mit Twitch emotes
 - Ca. 60% benutzen browser extensions wie BTTV

Umfrageergebnisse (Ausschnitt)

Emote	Negative	Neutral	Positive	Unknown/NA
FeelsBadMan 	71	17	19	1
FeelsGoodMan 	1	7	98	2
LUL 	11	23	72	2
OMEGALUL 	17	26	62	3
⋮	⋮	⋮	⋮	⋮
jinnythype 	7	31	17	53

Teil IV: Twitch Chat Sentiment Analysis

Disclaimer

Kurz gehaltener Abschnitt.

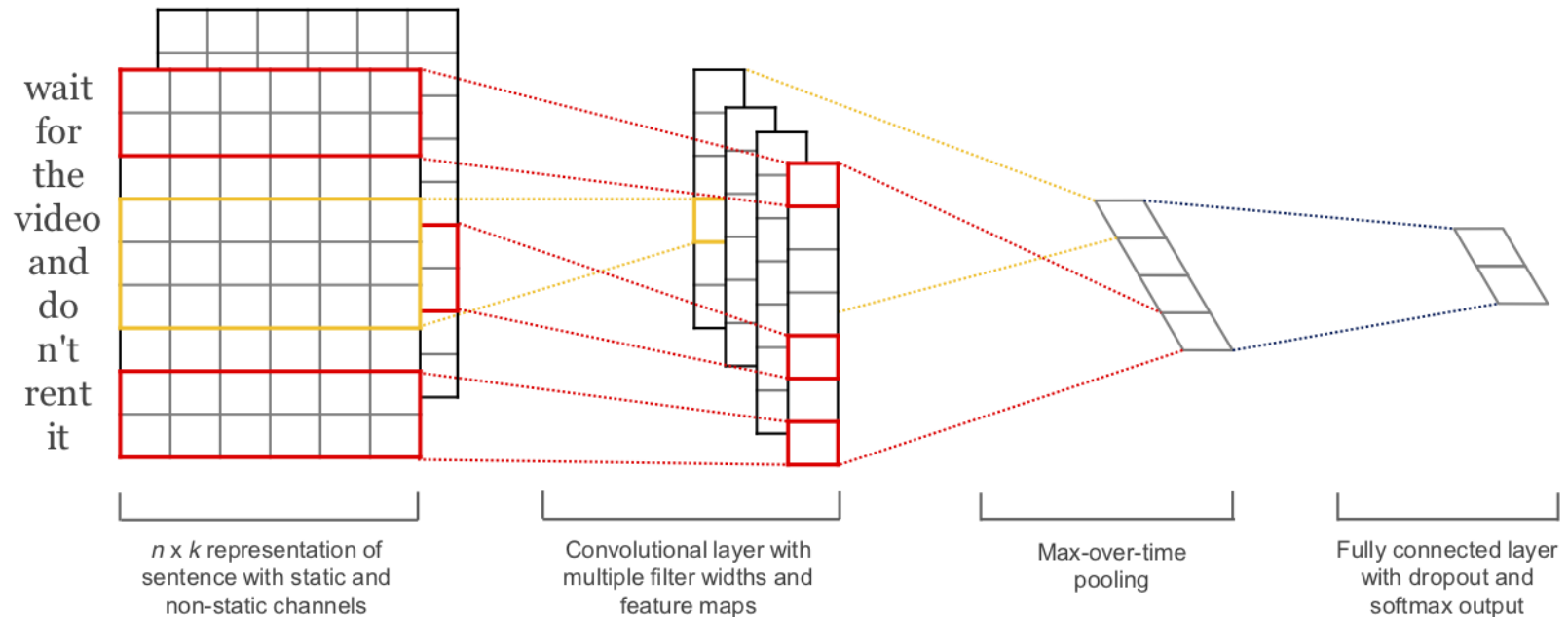
Empfehlung: Im Paper nachlesen.

Lexikon

- Kombination der drei Lexika
- Ansätze:
 - Average based (Tokens sowie gelabelte Nachrichten haben nur einen numerischen Wert zwischen -1 und 1)
 - Distribution based (Tokens sowie gelabelte Nachrichten haben je Werte für neg/neu/pos)

Sentence CNN

- Convolutional Neural Network
- Trainingsdatensatz gelabelt mittels Lexikonansatz
- Hoffnung: CNN lernt, mit unbekannten Tokens umzugehen



Teil V: Ergebnisse

Ergebnisse auf dem Evaluationsdatensatz

Method	accuracy	macro recall	macro F1 score
Random Baseline	33.3%	33.3%	32.7%
Random Baseline (sampling from target distribution)	35.6%	33.3%	33.3%
Majority Baseline	40.1%	33.3%	19.1%
NLTK Baseline	43.0%	39.3%	34.0%
Average Based Lexicon Approach	61.8%	58.9%	60.5%
Distribution Based Lexicon Approach	62.8%	60.5%	61.7%
Sentence CNN	63.8%	61.4%	62.6%

Auswirkung der Twitch Emotes

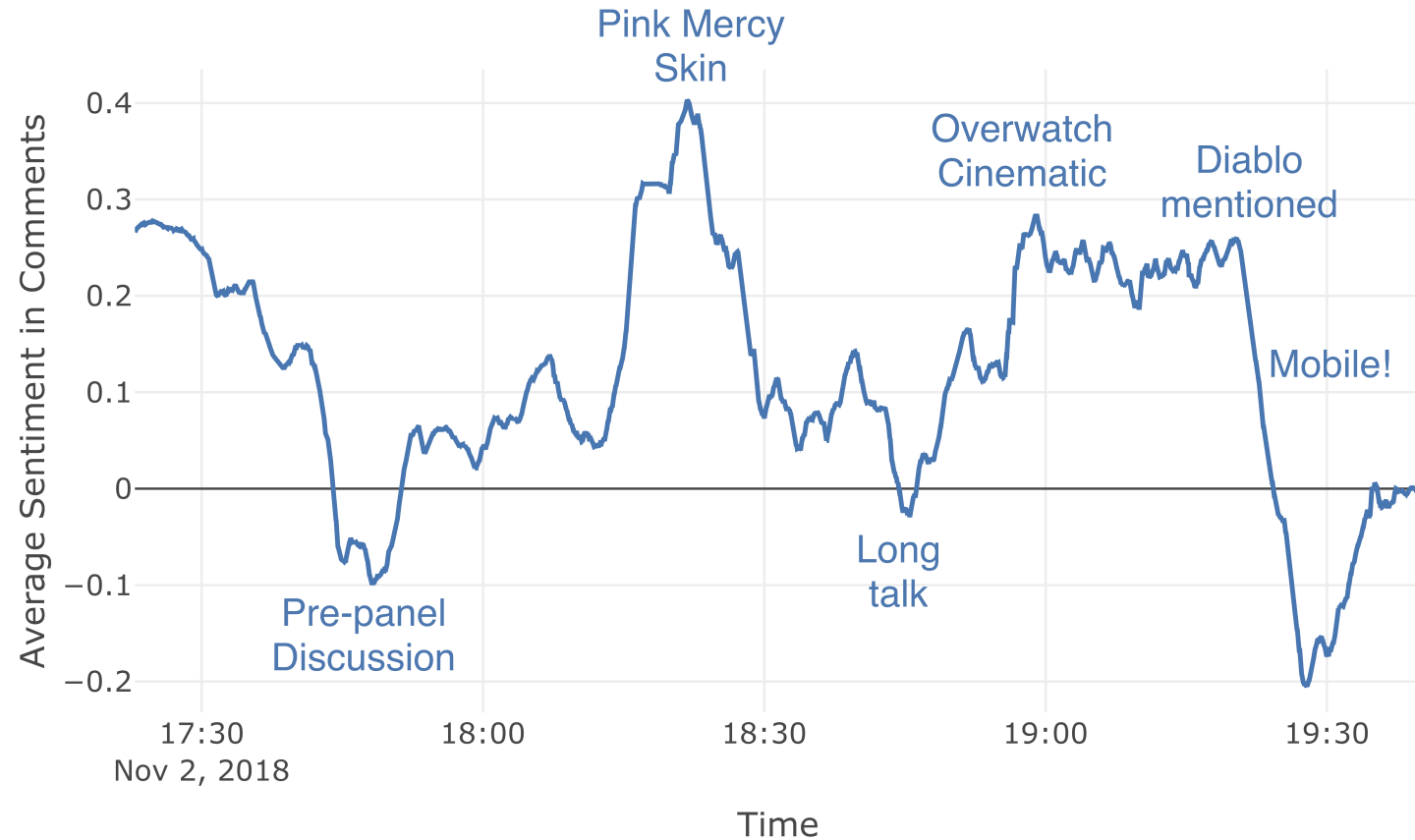
	accuracy Average/Distribution	macro recall Average/Distribution	macro F1 score Average/Distribution	% of comments containing signal tokens
Emoji	39.5 % / 39.8 %	34.0 % / 34.3 %	21.1 % / 21.4 %	3.8 %
Vader	48.3 % / 45.9 %	45.1 % / 43.1 %	42.1 % / 38.1 %	26.6 %
Emoji + Vader	48.5 % / 46.4 %	45.3 % / 43.6 %	42.7 % / 39.2 %	29.1 %
Emote	58.9 % / 59.7 %	54.3 % / 55.2 %	55.1 % / 56.0 %	48.0 %
Emote + Emoji	58.9 % / 60.3 %	54.4 % / 55.8 %	55.3 % / 56.7 %	50.6 %
Emote + Vader	61.8 % / 62.4 %	58.8 % / 60.2 %	60.4 % / 61.3 %	63.4 %
Emote + Emoji + Vader	61.8 % / 62.8 %	58.9 % / 60.5 %	60.5 % / 61.7 %	65.2 %

Word Embeddings

- Einige interessante Ergebnisse
- z.B. Steigerung von Emotes wie LUL -> OMEGALUL
- Nachzulesen im Paper

Blizzcon: Diablo Immortal Announcement

Blizzcon Sentiment



Ausblick

- Paper wird im Moment leicht überarbeitet
- Anfang Juli endgültige Abgabe
- Veröffentlichung irgendwann?



