Exploring a novel stand-up comedy dataset for quantifying humor.

...

Getting our feet wet with Live at the Apollo

Saurabh Gawande & Andreas Kaas Johansen

Why humor research?

- To improve our understanding of the complex human nature
  - What makes us laugh?
  - What makes us happy?
  - Why are some things funny when others are not? What does it say about people?

- In order to make more funny and relatable IT tools etc.
  - $20 and $60 billion were spent on humorous marketing in the US alone in 2008. [1]
  - ¾ of all super bowl ads were designed to be funny in 2008. [2]
  - In the future, making apps and personal assistants funny might improve interaction with products. [3]

Previous work:

- Machine learning algorithms trained on user-annotated funny comments [4]
- Attempt at creating a verbal humor taxonomy based on one-liners from blogs [5]
- Evaluation of features suspected to be correlated with humor on tagged tweets #humor #irony [6]

Research Question:

Will a new dataset, originating from standup comedy shows, containing audio signals and subtitles, allow humor researchers to gain more insights into what causes people to laugh?

Hypothesis:

Jokes of a negative sentiment are followed by significantly louder and/or longer laughter outbursts.

Example:

Negative joke: "*Her breath bloody stinks, but...*"
leads to "*HAHAHA!!!*"

Positive joke: "*Has anyone here actually been on a Jack The Ripper walk? Yes! You're like me. I love him...*"
leads to "*heh...*"

Methodology:

Dataset and Preprocessing:

- Live at the Apollo, season 11 episode 1 HDTV rip
- Live at the Apollo, season 11 episode 1 subtitles HDTV rip
    - “Clean” subtitles - so we have text and timestamps
    - “Concatenate” subtitles - so that a joke which takes place over several subtitles is concatenated into one “joke”.
    - Analyze the sentiment of each joke
    - Map sentiment of each joke to the laughter outburst which follows the joke
    - Note: A laughter outburst is defined by a start time, end time and the audio samples in-between
    - Note: Laughter outbursts are hand labeled

Tools:

R 3.3.0

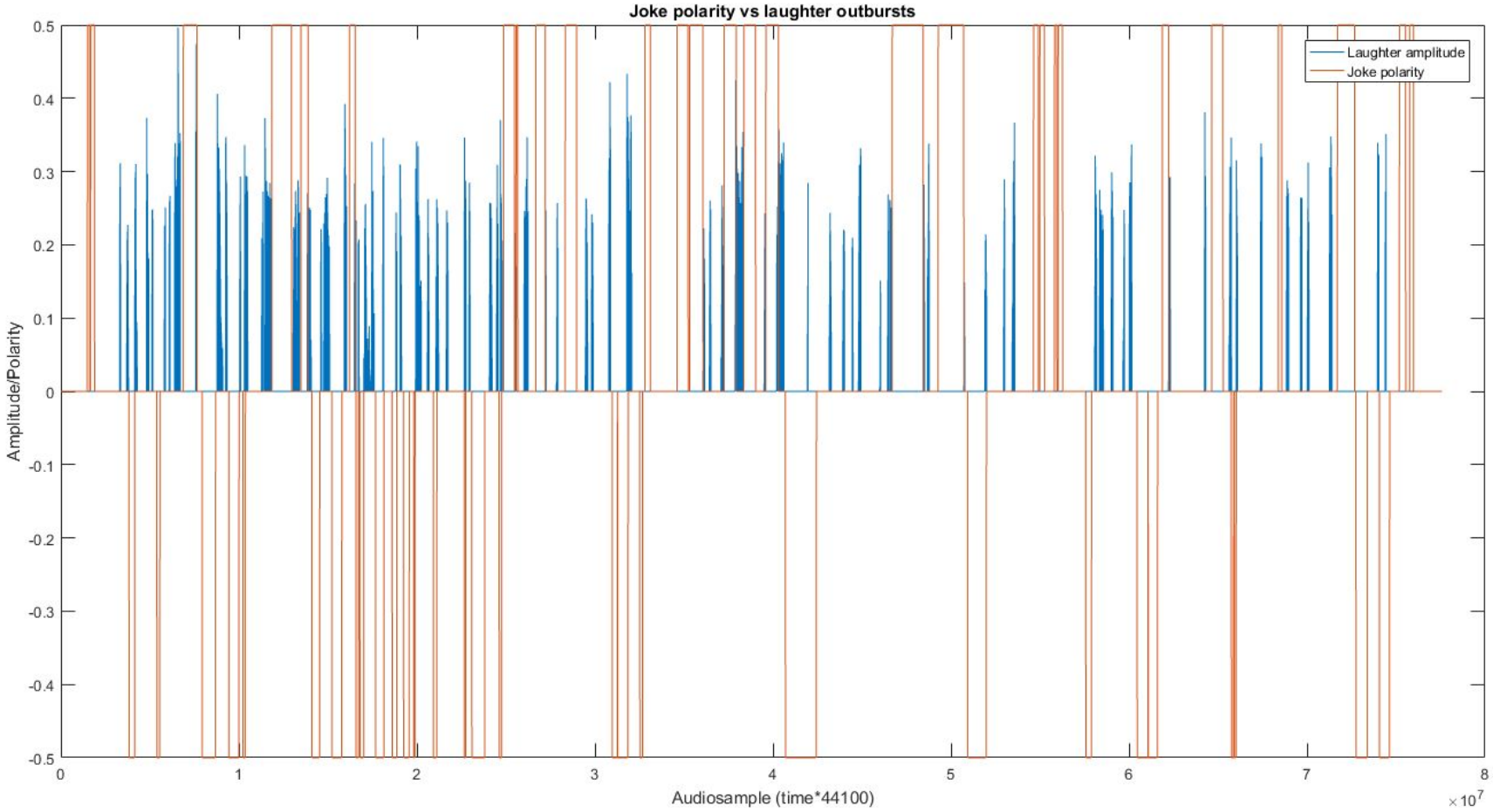
- Used for text preprocessing and sentiment analysis
  - Sentiment140 Natural Language Processing Service API by Stanford
    - Classifies text as either positive, neutral or negative.

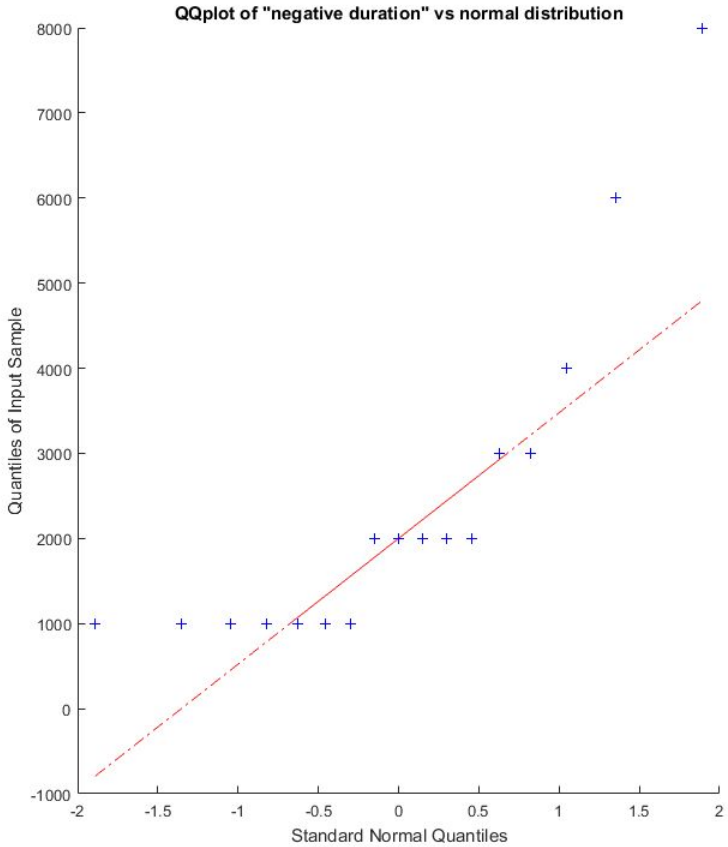
Matlab 2016a

- Used for audio preprocessing and performing statistical tests on audio samples

Data:

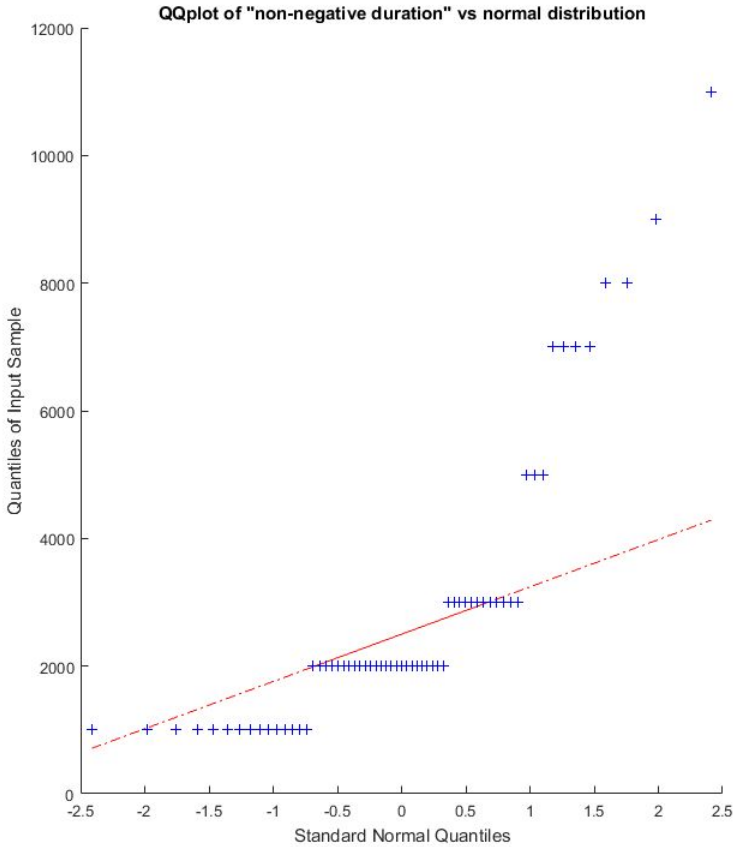
Pipeline	Total	Negative	Neutral	Positive
Subtitle mapped to sentiment	476	33	426	22
Concatenated subtitles mapped to sentiment	188	29	131	28
Laughter outbursts mapped to sentiment of joke it follows	80	17	52	11
Audio samples	77.6 million	?	?	?
Audio samples mapped to sentiment	9.8 million	1.8 million	8 million	
Mean(laughter outbursts)	80	17	63	


Joke polarity vs laughter outbursts
Laughter amplitude
Joke polarity
Amplitude/Polarity
Audiosample (time*44100)
×10^7
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
0
1
2
3
4
5
6
7
8



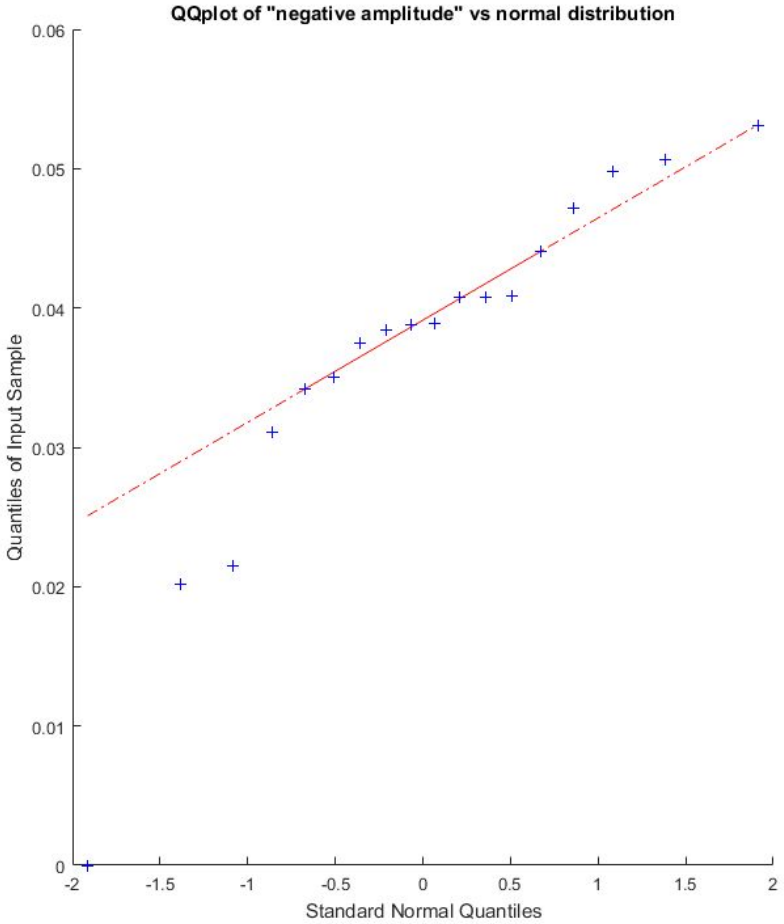
Shapiro wilk test for normality: p << 0.05

Normality rejected



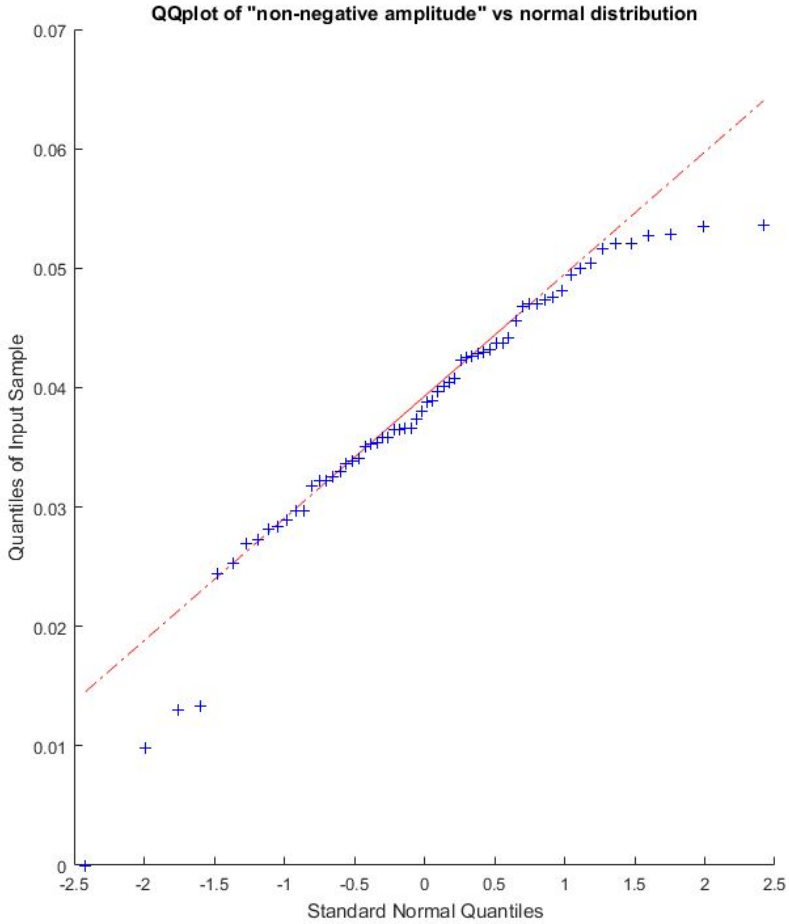
Shapiro wilk test for normality: p << 0.05

Normality rejected



Shapiro wilk test for normality: p=0.01884

Normality rejected



Shapiro wilk test for normality: p=0.001425

Normality rejected

Statistical Analysis (Results):

Laughter Duration
Two sided Wilcoxon Rank Sum test
mNeg = 2000ms , m!Neg = 2000ms
p-value 0.2919

Laughter Amplitude
Two sided Wilcoxon Rank Sum test
mNeg = 0.0389 , m!Neg = 0.0384
p-value 0.9911

H_o (Null Hypothesis) : $m_1 = m_2$

Discussion:

Will a new dataset, originating from standup comedy shows, containing audio signals and subtitles, allow humor researchers to gain more insights into what causes people to laugh?

Bias inherent caused by methods:

- Hand-labeled audio
- Inaccurate sentiment analysis - "*sounds like a good plan, can't possibly fail!*"
- Wrong approach in terms of methods

Discussion:

Will a new dataset, originating from standup comedy shows, containing audio signals and subtitles, allow humor researchers to gain more insights into what causes people to laugh?

Bias inherent caused by dataset:

- Audio remastering
- “Applaud sign” in the Apollo?
- Demographic bias in audience (Segmenting ‘clean’ audience who dont like swearing)
- Human nature, laughter is contagious
- Did not consider jokes might be related

Ethical and sustainability aspects:

- Key ethical issues:
- Data Collection - Permission to use BBC content
- We can’t promise that the data won’t end up in someone else's hands

- Sustainability aspects:
- Might make human beings less social. (see the movie “Her” 2013)
- Truly convincing AI’s might improve our daily habits, which might affect environmental and economical sustainability in the long run.

Questions ?

Backup - Definition of humor

Definition of Humor

- **Humor:** 3c: something that is or is designed to be comical or amusing

- **Comical:** 2: Causing laughter especially because of a startlingly or unexpectedly humorous impact <wearing a comical expression>
- **Amusing:** Definition for Language Learners: : causing laughter or enjoyment

-Source Miriam Webster Online Dictionary

Backup - Sources

[1] "Humor In The Advertising Business: Theory, Practice, And Wit". *Choice Reviews Online* 45.07 (2008): 45-3871-45-3871. Web.

[2] Gulas, Charles S., Kim K. McKeage, and Marc G. Weinberger. "It's Just A Joke". *Journal of Advertising* 39.4 (2010): 109-120. Web.

[3] Binsted, K. et al. "Computational Humor". *IEEE Intell. Syst.* 21.2 (2006): 59-69. Web.

[4] Reyes, Antonio et al. "Evaluating Humour Features On Web Comments". *LREC* (2010): n. pag. Web. 9 Sept. 2016.

[5] Reyes, A., P. Rosso, and D. Buscaldi. "Humor In The Blogosphere: First Clues For A Verbal Humor Taxonomy". *Journal of Intelligent Systems* 18.4 (2009): n. pag. Web.

[6] Reyes, Antonio, Paolo Rosso, and Davide Buscaldi. "From Humor Recognition To Irony Detection: The Figurative Language Of Social Media". *Data & Knowledge Engineering* 74 (2012): 1-12. Web.