First Workshop on Computational Modeling of Human Multimodal Language
MOSEI Corpus Details

Workshop Organizers

1. Data Acquisition

Videos in the MOSEI dataset are all in the form of monologues - videos with only one person in front of the camera.

1.1. Crawl System

We developed a crawler that can crawl YouTube and filter videos with only one person in front of the camera. This filter is implemented by extracting a number of frames from each video, and then using OpenCV's [3] Haar cascades to estimate how many faces are in each video. The crawler is supplied a search term which it then forwards to the YouTube Data API. The search terms provide a rough estimate of topics in the datasets, since they are directly connected to meta-data provided by the uploader.

Figure 1 shows the distribution of the video topics used in MOSEI. The diversity of the video topics brings the following generalizability advantages: 1) the models trained on MOSEI will be generalizable across different topics and the notion of dataset domain is marginalized, 2) the diversity of topics bring variety of speakers, which allows the trained models to be generalizable across different speakers, and 3) the diversity in topics furthermore brings diversity in recording setups which allows the trained models to be generalizable across microphones and cameras with different intrinsic parameters. This diversity makes MOSEI a one-of-a-kind dataset for sentiment analysis and emotion recognition.

1.2. Transcripts

The crawled videos are limited to only videos with user-provided transcripts (which we rely on the YouTube Data API for). However to ensure that the user-provided transcript is reliable, we further post-process with the following filters: 1) punctuation – we use various heuristics about punctuation to ensure that the transcript is high quality, 2) alignment – we ensure that the forced alignment using P2FA [5] passes with high confidence. These filters allow us to filter out videos with bad transcripts.

Figure 1: The topics of videos in MOSEI, displayed as a Venn-style word cloud [1]. Larger words indicate more videos from that topic.

2. Dataset Splits

The MOSEI Mega Corpus facilitates both machine learning and behavioral studies. The dataset in full form can be used for machine learning research as it contains a rather balanced distribution across various sentiment scores. MOSEI Natural Split is a subset of the dataset that was crawled without any form of sentiment and emotion guidance, and reflects a more random sample of YouTube monologues. As a result, this subset contains fewer polarized videos. In the next subsection, we first discuss the Natural Split, and then discuss how we acquired video with more polarized sentiment and emotion.

2.1. MOSEI Natural Split

MOSEI Natural Split contains sentences randomly sampled from YouTube monologues. The distribution of annotated sentiment for this split is skewed towards neutral sentences. Due to a combination of factors such as video topic and gender in the selection of these videos, we believe this is the distribution of uttered sentences in YouTube monologues. From a machine learning perspective, this distribution is not ideal since there are many neutral sentences.

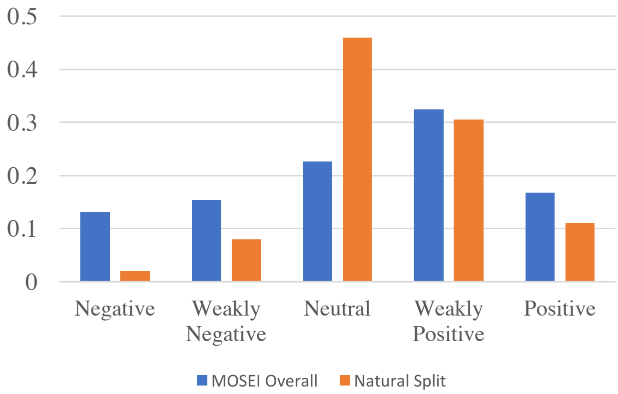

Figure 2: Distribution of sentiment labels for MOSEI Overall and Natural split. This figure does not represent magnitude (Overall 23,500 sentences but Natural has 7,500 sentences), only ratio.

This dataset contains a total of 7,500 sentences from more than 245 topics and 721 different speakers with almost an equal number of male/female speakers (57.2% male and 42.8% female). Figure 2 shows the distribution of sentiment in Natural Split compared to the overall dataset.

2.2. MOSEI Guided Crawl

To compensate for the lack of videos with polarized sentiment, we use a text-based sentiment analysis model based on the best performing text-based model in [6] trained on the CMU-MOSI dataset and sentences specifically annotated for this task. We use this model to detect the most polarized videos which we crawled, and sent these for annotation on Amazon Mechanical Turk. We also use 500 polarized videos of POM dataset [4] which have manual sentiment annotations and extract their sentences. A total of 16,000 sentences are extracted using guided crawl.

3. Crowdsourced Annotations

MOSEI is designed to capture the crowd's perception of a speaker's sentiment and emotions. We rely on minimal training for the annotations to limit the potential bias training may cause. Modeling the crowd's raw perception of sentiment and emotions is vital to creating real-world applications that model the thought processes of the general population. This is in contrast to datasets in the same domain of sentiment and emotion recognition which rely on experts' opinions, which may not agree with the general population's opinion. We prioritize the general population's perception over the psychological definitions of sentiment and emotions, which can only be inferred by experts.

All the monologue sentences in the dataset are annotated using Amazon Mechanical Turk (AMT). Each sentence is annotated thrice by different annotators. Only master annotators with an acceptance rate of over 98% were allowed to annotate the dataset. The following question is asked to the MTurk workers for annotations: "Watch the video clip and rate the sentiment and emotions of the speaker. Please note that you may or may not agree with what the speaker says. It is important that you only rate the sentiment and emotions of the speaker, not yourself."

	MOSEI Krippendorf alpha
Sentiment	0.53
Happiness	0.41
Anger	0.18
Sadness	0.12
Disgust	0.21
Fear	0.02
Surprise	0.09

Table 1: Agreement Krippendorf alpha values for annotations in the MOSEI dataset.

3.1. Annotator Agreement

Table 1 shows the agreement scores between annotators in terms of Krippendorf's Alpha. While MOSEI is annotated by crowdsourced workers in a fairly subjective manner – by asking their opinion about sentiment and emotion of the speaker with minimal training – the overall agreement scores are comparable with other datasets annotated by experts outlined in the submitted paper. Furthermore, a lower agreement would be expected from a dataset such as MOSEI due to its diversity of topics and speakers, and inherent variance in wild data. These factors impact the agreement since they increase the subjectivity of the task.

3.2. Sentiment and Emotions Definition

Due to the vast number of similar tasks on AMT, annotators are relatively familiar with broad definition of sentiment and emotions. However, through a five minute training video, we define *emotions* as the speaker's expression of state of mind and feeling while uttering the sentence. *Sentiment* is defined as the speakers attitude towards the topic of his/her discussion. The annotators were asked to annotate sentiment on a seven-step Likert scale of [-3: highly negative, -2: negative, -1: weakly negative, 0: neutral, 1: weakly positive, 2: positive, 3: highly positive]. The Emotions selected are the six basic Ekman emotions [2] of {happiness, sadness, anger, fear, disgust, surprise}. Each of the emotions is annotated at a four-step Likert scale for the presence of an emotion x: [0: no evidence of x, 1: weakly x, 2: x, 3: highly x]. The annotators are also asked to determine the speaker's gender.

3.3. Annotation User Interface

Figure 3 shows the sample annotation interface that AMT workers see when performing annotations. Each worker must finish watching the training video before starting their annotations. Furthermore, at any time during the annotation

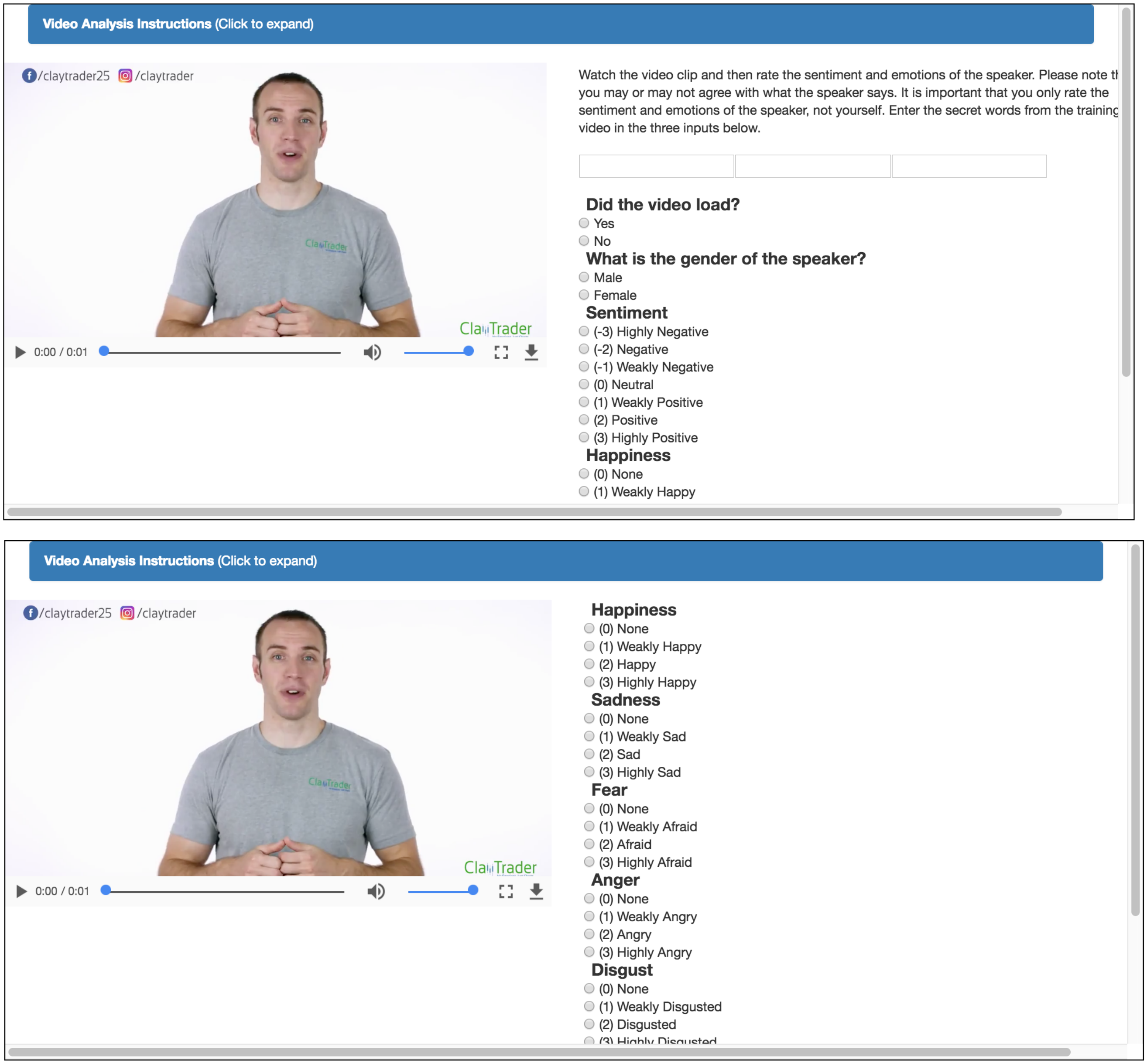

Figure 3: Annotation user interface for sentiment (top) and emotion (bottom) labeling.

session workers can rewatch the training video to refresh their memory for specific instructions.

4. General Statistics

4.1. Video Statistics

Table 2 shows some summary statistics of the MOSEI Mega Corpus such as the average length and duration of sentences and distribution of unigrams.

4.2. Video Topics

The crawl search terms included a total of 1962 terms. However, only around 250 resulted in acceptable videos. Since the search terms are related to the metadata of the videos provided by the uploader, they are related to the video's topic. These topics are shown in Figure 1 and are mentioned below with the total number of videos and percentage of all videos in the dataset:

reviews (524, 16.2%), debates (94, 2.91%), consulting (59, 1.83%), financial (59, 1.83%), speeches (51, 1.58%), statement (40, 1.24%), speech (40, 1.24%), advertising (39,

	MOSEI Statistics
Total number of sentences	23453
Total number of opinion sentences	18148
Total number of objective sentences	5305
Total number of videos	3228
Total number of distinct speakers	1000
Total number of distinct topics	250
Average number of sentences in a video	7.3
Average length of sentences	7.28 seconds
Average word count per sentence	19.2
Total number of words in sentences	447143
Total of unique words in sentences	23026
Total number of words appearing at least 10 times in the dataset	3413
Total number of words appearing at least 20 times in the dataset	1971
Total number of words appearing at least 50 times in the dataset	888

Table 2: MOSEI Mega Corpus summary statistics.

1.21%), consumers (37, 1.15%), faq (34, 1.05%), investment (32, 0.99%), consumer (30, 0.93%), political (26, 0.81%), loans (26, 0.81%), seeing (24, 0.74%), monologue (23, 0.71%), firms (22, 0.68%), product review (22, 0.68%), independent (22, 0.68%), analysis (21, 0.65%), testimony (21, 0.65%), business (21, 0.65%), marketing speech (21, 0.65%), speech marketing (20, 0.62%), investing (19, 0.59%), investors (19, 0.59%), comments (19, 0.59%), equity (18, 0.56%), summary (17, 0.53%), remarks (16, 0.50%), hearing (16, 0.50%), companies (16, 0.50%), politics (15, 0.46%), customers (15, 0.46%), financing (15, 0.46%), Response (15, 0.46%), description (14, 0.43%), convention (14, 0.43%), retail (14, 0.43%), marketing (14, 0.43%), review (14, 0.43%), advertisers (14, 0.43%), questions and answers (13, 0.40%), phd (13, 0.40%), retailers (13, 0.40%), banking (13, 0.40%), products (13, 0.40%), update (12, 0.37%), person (12, 0.37%), definition (12, 0.37%), lecture (12, 0.37%), application (11, 0.34%), marketplace (11, 0.34%), online courses (11, 0.34%), customer (11, 0.34%), dialogue (11, 0.34%), presentation (11, 0.34%), placement (10, 0.31%), sustainability (10, 0.31%), entrepreneurship (10, 0.31%), social (10, 0.31%), congress (10, 0.31%), economics (10, 0.31%), economic (10, 0.31%), committee (10, 0.31%), businesses (10, 0.31%), counter (9, 0.28%), corporate (9, 0.28%), speaker (9, 0.28%), seminar (9, 0.28%), hear (9, 0.28%), sociology (9, 0.28%), updates (9, 0.28%), religious (9, 0.28%), stocks (8, 0.25%), topic (8, 0.25%), instruction (8, 0.25%), conference (8, 0.25%), integrated (8, 0.25%), pricing (8, 0.25%), separate (7, 0.22%), meeting (7, 0.22%), evaluation (7, 0.22%), outsourcing (7, 0.22%), product marketing (7, 0.22%), buyers (7, 0.22%), narrative (6, 0.19%), summit (6, 0.19%), economies (6, 0.19%), My thoughts on (6, 0.19%), weekly update (6, 0.19%), distress (6, 0.19%), industry (6, 0.19%), conviction (6, 0.19%), listener (6, 0.19%), employers (6, 0.19%), eulogy (6, 0.19%), updated (6, 0.19%), business update (6, 0.19%), sector (6, 0.19%), talk (5, 0.15%), announcing (5, 0.15%), apology (5, 0.15%), due (5, 0.15%), storytelling (5, 0.15%), statements (5, 0.15%), presentations (5, 0.15%), journal (5, 0.15%), web (5, 0.15%), newsletter (5, 0.15%), home business (5, 0.15%), details (5, 0.15%), Discussion (5, 0.15%), automation (5, 0.15%), ads (5, 0.15%), announcement (5, 0.15%), investments (4, 0.12%), witness (4, 0.12%), textbook (4, 0.12%), responding (4, 0.12%), editing (4, 0.12%), explanation (4, 0.12%), arbitrator (4, 0.12%), corporations (4, 0.12%), inclusion (4, 0.12%), market (4, 0.12%), respond (4, 0.12%), underwriting (4, 0.12%), timing (4, 0.12%), handling (4, 0.12%), economy (4, 0.12%), How I feel about (3, 0.09%), stock (3, 0.09%), endorsement (3, 0.09%), accusations (3, 0.09%), announcements (3, 0.09%), speakers (3, 0.09%), daily dose of (3, 0.09%), entrepreneurial (3, 0.09%), textiles (3, 0.09%), commodity (3, 0.09%), sneezing (3, 0.09%), shareholders (3, 0.09%), lesson (3, 0.09%), sentencing (3, 0.09%), reaction (3, 0.09%), movie review (3, 0.09%), responses (3, 0.09%), keynote (3, 0.09%), dividends (3, 0.09%), witnesses (3, 0.09%), seminars (3, 0.09%), tutorial (3, 0.09%), branding (3, 0.09%), added (3, 0.09%), revision (3, 0.09%), upload (3, 0.09%), how to (3, 0.09%), investor (3, 0.09%), entity (3, 0.09%), exchanges (3, 0.09%), rhetoric (3, 0.09%), podcast (3, 0.09%), proposition (3, 0.09%), reviewed (3, 0.09%), refining (2, 0.06%), outline (2, 0.06%), chairperson (2, 0.06%), farms (2, 0.06%), manufacturing (2, 0.06%), prosecution (2, 0.06%), voiceover (2, 0.06%), informational (2, 0.06%), socialist (2, 0.06%), markets (2, 0.06%), pharmaceuticals (2, 0.06%), sermon (2, 0.06%), Things I like (2, 0.06%), overview (2, 0.06%), positioning (2, 0.06%), derivation (2, 0.06%), faqs (2, 0.06%), discussions (2, 0.06%), suppliers (2, 0.06%), stimulus (2, 0.06%), monologues (2, 0.06%), lenders (2, 0.06%), premium (2, 0.06%), movie reviews (2, 0.06%), quarterly (2, 0.06%), conferences (2, 0.06%), industries (2, 0.06%), debate (2, 0.06%), protest (2, 0.06%), unofficial (2, 0.06%), institutional (2, 0.06%), unified (2, 0.06%), QA (1, 0.03%), home speech (1, 0.03%), resignation (1, 0.03%), document (1, 0.03%), symposium (1, 0.03%), retailing (1, 0.03%), forestry (1, 0.03%), commercials (1, 0.03%), specialization (1, 0.03%), liberalism (1, 0.03%), reviewing (1, 0.03%), vol (1, 0.03%), futures (1, 0.03%), hearings (1, 0.03%), shops (1, 0.03%), independant (1, 0.03%), participatory (1, 0.03%), enterprises (1, 0.03%), bechtel (1, 0.03%), asset (1, 0.03%), addition (1, 0.03%), configuration (1, 0.03%), nyse (1, 0.03%), home marketing (1, 0.03%), narration (1, 0.03%), My response to (1, 0.03%), retailer (1, 0.03%), multiculturalism (1, 0.03%), industrial (1, 0.03%), comment (1, 0.03%), reply (1, 0.03%), colloquium (1, 0.03%), rewritten (1, 0.03%), autonomous (1, 0.03%), recitation (1, 0.03%), macroeconomics (1, 0.03%), ict (1, 0.03%), deadpan (1, 0.03%), monthly update (1, 0.03%), nasdaq (1, 0.03%), revised (1, 0.03%), crm (1, 0.03%), concise (1, 0.03%), discussing (1, 0.03%), securities (1, 0.03%), officially (1, 0.03%), enclave (1, 0.03%), citigroup (1, 0.03%), edit (1, 0.03%)

4.3. Channel IDs

There are a total number of 734 unique YouTube channels from which male and female videos are extracted. We use the channel id as a heuristic to approximate the number of speakers. Each channel at most gives two identities one male and one female. The following lists the top 250 channel IDs from which videos were obtained from, together with the gender of the speaker from that channel, the total number of videos obtained from that channel and percentage of all videos in the dataset:

UCQlGBspQdj17WOPBQMT1k9A (female, 132, 4.09%), UCnQznIHK_kZIEMUl2nT1mYQ (male, 117, 3.62%), UC-mJM2Qe4jcPa4Kd-4pGtXA (male, 105, 3.25%), UCnQznIHK_kZIEMUl2nT1mYQ (female, 91, 2.82%), UCQlGBspQdj17WOPBQMT1k9A (male, 90, 2.79%), UC3MPYO0ocVKu_FJfYem0Fig (female, 81, 2.51%), UC-mJM2Qe4jcPa4Kd-4pGtXA (female, 38, 1.18%), UC6ZhpmNnLxlOYipqh8wbM3A (male, 28, 0.87%), UCng0oDXtwDLqAteSqJi7tyg (male, 22, 0.68%), UCJGIMtVU-bIzi3L5Xm00HiA (male, 18, 0.56%), UCBVCi5JbYmfG3q5MEuoWdOw (male, 18, 0.56%), UCBVCi5JbYmfG3q5MEuoWdOw (female, 10, 0.31%), UC6ZhpmNnLxlOYipqh8wbM3A (female, 9, 0.28%), UC21bAFVlfgbEDy0QiNf-ygQ (female, 8, 0.25%), UC4XJnRPZjXhgvVMhXKNSJvQ (male, 8, 0.25%), UCeYuZB53ge-tdHWUDF87xQw (male, 7, 0.22%), UCgF4XB5sQdAKbbTLzZbQnnA (male, 7, 0.22%), UCQP-DKyFy2dMheEoJqTLzCw (female, 6, 0.19%), UCBFcDw3VWjhAJtAmAI7U6cA (male, 6, 0.19%), UC3MPYO0ocVKu_FJfYem0Fig (male, 6, 0.19%), UCxVAZswnxZz3gkjdARRa_-g (male, 6, 0.19%), UC15plGNCMRkMmREMNkbYJPA (male, 5, 0.15%), UC70MtJNetbXjbdgnOdbKkSA (female, 5, 0.15%), UCVWh8HXVtuaYeMHGvmMBLqQ (female, 5, 0.15%), UCxVAZswnxZz3gkjdARRa_-g (female, 5, 0.15%), UCHlEaKbepQ_S9iIoZPKVQew (male, 4, 0.12%), UCfp5MBK2IC90Bdm5IQUJvYw (male, 4, 0.12%), UCWJp3kRn3ZK2kOCqNHz0VBw (male, 4, 0.12%), UCS4e6OUhaT65zagSldHfA8A (female, 4, 0.12%), UCng0oDXtwDLqAteSqJi7tyg (female, 4, 0.12%), UCS7gcmlW0Bpm46AzbKn3img (male, 4, 0.12%), UCgF4XB5sQdAKbbTLzZbQnnA (female, 4, 0.12%), UCZmZhlG4G5DTg1FO2c34Phw (male, 4, 0.12%), UCJr29wlSJoQZUWmPjNDODFQ (male, 3, 0.09%), UCrlcu5KChYyHwXlIeD7oLUg (male, 3, 0.09%), UC0LNvguQUCKbDKosXo7xw4A (male, 3, 0.09%), UC1yk0FVuAQctI6yjRlqc1Eg (female, 3, 0.09%), UCpVPntWE2zYrafVWJRL7wlw (male, 3, 0.09%), UCVWh8HXVtuaYeMHGvmMBLqQ (male, 3, 0.09%), UCUnlt4u1kZR2iKEKbZdELCQ (male, 3, 0.09%), UCpFaxVbcP-2-vgKikuNmMew (male, 3, 0.09%), UChZ8zqDwmqvPn4FuEyA0pwg (male, 3, 0.09%), UC3wthuKCoOR9E-Y15sP_NlA (female, 3, 0.09%), UCpaOW3Vd6fl_U1FF0S1VRdg (male, 3, 0.09%), UCkIzis6MSbv5oR_spgHhP2w (female, 2, 0.06%), UCcxYfgjfspV1XQdVumVH5bg (male, 2, 0.06%), UCw-kH-Od73XDAt7qtH9uBYA (male, 2, 0.06%), UCsib3HIZtE7wKSKm08LRjLA (male, 2, 0.06%), UCCEgTszl0-4_-x4I7r4bDhA (male, 2, 0.06%), UCsEhUNwpPsQKeEdYWLjSg1A (male, 2, 0.06%), UC2sJZ3e3DYvAwCCsrr34S7w (male, 2, 0.06%), UCrlcu5KChYyHwXlIeD7oLUg (female, 2, 0.06%), UCufEhpxVSccRGT7FCHOa2Gw (female, 2, 0.06%), UCSpVHeDGr9UbREhRca0qwsA (male, 2, 0.06%), UC9bZmtwQYvpImKpBg5bDv5g (female, 2, 0.06%), UC0vuJJUpIguVOfmXvk1PQpA (male, 2, 0.06%), UCtL5WjEYhlHzmOan57lJgnw (male, 2, 0.06%), UCiANdkAuJAiiHCwDf_tJLYg (male, 2, 0.06%), UCRhnxMeE7PV_0H2J89FKxrg (male, 2, 0.06%), UCRhnxMeE7PV_0H2J89FKxrg (female, 2, 0.06%), UCVrYey5SZMid_VZk9D8tYmA (male, 2, 0.06%), UCDGknzyQfNiThyt4vg4MlTQ (male, 2, 0.06%), UCK_vX6MneWKCLtL0uJwYOxQ (male, 2, 0.06%), UCNDjJh8pTbLJphUOY-SrVDg (male, 2, 0.06%), UC8cXXCdLzcYoYGa_BqaPsgA (male, 2, 0.06%), UCQCTCDhr1KYCP0-NLz0IS0Q (female, 2, 0.06%), UCGPJn9Ciiwc3vT2539J6gJg (female, 2, 0.06%), UCFS0Ox4LDKIx6lJED9r51Cw (male, 2, 0.06%), UCFS0Ox4LDKIx6lJED9r51Cw (female, 2, 0.06%), UCGaVdbSav8xWuFWTadK6loA (male, 2, 0.06%), UCrKkwAsivzl0l7ooslc8klA (male, 2, 0.06%), UCeX20-rTnN46DUD6C7IplWg (male, 2, 0.06%), UC3pSSFaBz6fIvIU7kcxzN1Q (male, 2, 0.06%), UCpi0W79RZVWo3lvXeNR7HGQ (male, 2, 0.06%), UC67Vc0fkLYeUPBp1f02VY9Q (female, 2, 0.06%), UCeaNshO4Ydxnvw226674d0A (male, 2, 0.06%), UC67Vc0fkLYeUPBp1f02VY9Q (male, 2, 0.06%), UClEFD_Q-idenkbhOGwv_Jmw (female, 2, 0.06%), UCXmDODGuRorJE_rAU30vVvw (male, 2, 0.06%), UCRcMAMZ6CVVXSZtOT1AjpiQ (female, 2, 0.06%), UCDJFty377JHdWfCEkTXVS3Q (male, 2, 0.06%), UCESRu4djCKDPeCxprniB-yQ (male, 2, 0.06%), UCQqa08s5CzkCbs–Xf1bwig (male, 1, 0.03%), UC-Gz6qnBnBmT02Eq7InjATg (male, 1, 0.03%), UCGO-x25PyOx0IDm9A1dC5Ag (male, 1, 0.03%), UC23dbHbfH-kOntle-7eqVgw (female, 1, 0.03%), UCBkgSGQuGFsZZRLyyMYcjsQ (male, 1, 0.03%), UCr_CSpTPaedbMMC89sazTnA (male, 1, 0.03%), UCSkn1qgZCvy8GA165KwanrQ (male, 1, 0.03%), UCF52sgO4R26MOru41pfOOaA (female, 1, 0.03%), UCspyNrm1N_pzinlaegFkoKw (female, 1, 0.03%), UCx-1puI2WHFOutUxSJoce5w (female, 1, 0.03%), UC7cv_YhdEHiEOhVubLuFJWQ (male, 1, 0.03%), UCw-kH-Od73XDAt7qtH9uBYA (female, 1, 0.03%), UCXTSbRTLCr9r8Ne_7ElnZQQ (male, 1, 0.03%), UCPkvs3GgnM_pGRS42EKFw2g (male, 1, 0.03%), UCDiz3BBJPk4zbU5DhtIW3aA (male, 1, 0.03%), UCGwwcL370sFtL0bs1wIZjFw (female, 1, 0.03%), UCJTbtUFa6B7bKWUn1NR15QA (male, 1, 0.03%), UCbotZDKzpEhRhSB16sDEkGg (male, 1, 0.03%), UCaCtdex9SL4Pc7vQSf5ZGhA (female, 1, 0.03%), UCctRIfobmq8kh2hSS6DcQUg (male, 1, 0.03%), UChRfKCdgC_WijJt7nKuJJKw (male, 1, 0.03%), UCFJNcE0iHj7P6dhp5iCZRLg (male, 1, 0.03%), UCuv8z9NSyb5_CABggNpkKBQ (male, 1, 0.03%), UCoE-vWaFFFrJQFORbnh90UQ (female, 1, 0.03%), UCvWDr2xEBhIpc_Kb97cPdCQ (male, 1, 0.03%), UCvSal3J4UZPsDgZe5ZaKqaw (male, 1, 0.03%), UCK6NNQ5-fuxa5CWiyYkH97A (male, 1, 0.03%), UCJ2eJngkiPd2XjKuQlgn3pg (male, 1, 0.03%), UCLI_3vDvEjh3_ILTl1BkXhQ (female, 1, 0.03%), UC-cxqz-2fBOekclLlaERLJQ (male, 1, 0.03%), UChzR4YuJPLrYSIARuORmeQQ (female, 1, 0.03%), UCml9e27XuPFxHyl817JzyrA (male, 1, 0.03%), UCM2Wzxs3MTTUt0sL8ZnP85g (female, 1, 0.03%), UCqX4MqSEj02KPP2vFKwWucA (female, 1, 0.03%), UC6JXu_UyR87_SrOdpKT9hyA (male, 1, 0.03%), UCfn98TnOmGjHCFc9QrQuzEQ (male, 1, 0.03%), UC6suMjuyyKL8Hdq0Fo7h8rg (female, 1, 0.03%), UCM3XtsVwgt7NFV2KIFFVtlg (female, 1, 0.03%), UCyRLHNmtBPr64Mm0D9hYdgg (male, 1, 0.03%), UCWdJ3BxqWopIdJ7J_BNu1zQ (male, 1, 0.03%), UCVAgjqtPEcAShlzo6CFIyzA (male, 1, 0.03%), UCRh2wWe9wetYY8cfwiioAlw (male, 1, 0.03%), UCt2BwSAxGoMHvf-RBP_7YiQ (female, 1, 0.03%), UC-RlNcHZEucIevdNBeR-KPw (male, 1, 0.03%), UCjpf9rtBlQJaATIcDHpJU9Q (female, 1, 0.03%), UC6XgR7FMGjZ07vpxTqRwurw (male, 1, 0.03%), UCL1Ma8M4n-6olSIJTOhxhbQ (female, 1, 0.03%), UCsUAkN7fF3FPS-Y7bRnc2xA (male, 1, 0.03%), UCL9-bfld991n7mK2NAzuupA (male, 1, 0.03%), UC9P-Sn_WcR6Fx1n7UlW064g (male, 1, 0.03%), UCriLGWtM8JT-x-DkPx2G2AQ (male, 1, 0.03%), UCOtnu-KKoAbN47IuYMeDPOg (male, 1, 0.03%), UCXsH4hSV_kEdAOsupMMm4Qw (female, 1, 0.03%), UCmqYwsyr3BOVv0zoH4W48ww (female, 1, 0.03%), UCfFw0bNr4gKzoFwFs5oreBg (female, 1, 0.03%), UCXsH4hSV_kEdAOsupMMm4Qw (male, 1, 0.03%), UCufEhpxVSccRGT7FCHOa2Gw (male, 1, 0.03%), UCUpZdMB0kfhF-aZ9Kyni-KQ (male, 1, 0.03%), UCXF_lLs1IgXivMIynBimBOw (female, 1, 0.03%), UC0VvxW6DEAVxIEfWgBnuaMQ (male, 1, 0.03%), UCkRqFLQLXEnOkGkubCDaxNA (male, 1, 0.03%), UCXTAY9DlvBqMEXh1paOdmOg (male, 1, 0.03%), UC55XqZStTGZhzlHtF1yoCEA (male, 1, 0.03%), UCi9_4Z3yEh5tbEr6Y7gkU8w (female, 1, 0.03%), UCqFunGBhuL-l7xiNJsSjvEQ (female, 1, 0.03%), UCv5967rhojn0IaPxJusxy1A (female, 1, 0.03%), UCU1XozSx35hUCqhpIbuzX7g (male, 1, 0.03%), UCANEhA9r4Uq60V0y3WsG1CQ (male, 1, 0.03%), UCfbPPGeagedB-wUwFmFP5CA (female, 1, 0.03%), UCJdrLE4VFwF691fDw8lTAiQ (male, 1, 0.03%), UCnyQyZu_fkxVa_CbT6hS1qg (male, 1, 0.03%), UCCtVTbqOnc_JTTD2DQSthSQ (male, 1, 0.03%), UCsFm77UH05WDV0q5QshI8TA (female, 1, 0.03%), UCZGpy6QwTt2NHJLGPsZ-s6Q (male, 1, 0.03%), UCK7-wD91qr5Ugah_XrqvWgg (female, 1, 0.03%), UCV7WzZeBxwlvUFe9Tx0INMg (male, 1, 0.03%), UCalOGPnHpKr-0tjdF-ab_mQ (male, 1, 0.03%), UCiXBo8LN1E3eevHtYC37uGQ (male, 1, 0.03%), UCATgN3Em0CwZoMr9IQ6FCwg (male, 1, 0.03%), UCUv8dqCeYKczXIhnL4-UyQA (female, 1, 0.03%), UCWOkEnBl5TO4SCLfSlosjgg (male, 1, 0.03%), UCPm0GHnxLBl4Qy_-XI-zAzg (female, 1, 0.03%), UCaZ-1zEs-YWKbi-L-2p_RFQ (male, 1, 0.03%), UCiMg06DjcUk5FRiM3g5sqoQ (female, 1, 0.03%), UC4tIVHTJBZeMFHHfGb2IqLQ (female, 1, 0.03%), UC9PmsKBVV_kaOd8p4wuWkuA (male, 1, 0.03%), UCqTZ4mzzR6qy4MPHyuxicXw (female, 1, 0.03%), UC5uCqa7W8d4dZvMU2tAcR-Q (female, 1, 0.03%), UChDay1MZzkP-X3cgITy9GCA (male, 1, 0.03%), UCau23h1uU0VjL7uGR3We2Dg (male, 1, 0.03%), UCBRwxMdCy3UbhLfANgz0_Rg (male, 1, 0.03%), UCovtFObhY9NypXcyHxAS7-Q (male, 1, 0.03%), UCfp5MBK2IC90Bdm5IQUJvYw (female, 1, 0.03%), UCzRipjQYhKGNFdTHNMzN-fIQ (male, 1, 0.03%), UCbt8SycZej80mabDtc-wuEQ (male, 1, 0.03%), UClASEfEoOpXL1_eyzJ-F0IQ (male, 1, 0.03%), UCMH7hOaE_V-aKV8c7NjZloA (male, 1, 0.03%), UC11RRI571HrJkBUJWNGYsxg (male, 1, 0.03%), UCCk3Mt0fwYlZmTDgYz2J6WA (female, 1, 0.03%), UCvn_XCl_mgQmt3sD753zdJA (male, 1, 0.03%), UC11RRI571HrJkBUJWNGYsxg (female, 1, 0.03%), UCAkmVsd_pi2eLlfaPJ0iSCA (male, 1, 0.03%), UCqJkAAmi4QKCPCF62r_-BhQ (female, 1, 0.03%), UCzq2rnoZ4iC-30wYZkqRrGw (female, 1, 0.03%), UC1yk0FVuAQctI6yjRlqc1Eg (male, 1, 0.03%), UC6fUahKiPDn1-3476TU-ovA (male, 1, 0.03%), UC15plGNCMRkMmREMNkbYJPA (female, 1, 0.03%), UCCLrseuBNA1tArAny_e9bRA (male, 1, 0.03%), UCd3hTZw4b4foQwnDB8NWz6Q (male, 1, 0.03%), UCuXnKeesX_QALxTBfg17lSQ (female, 1, 0.03%), UCewFEPHPaIs4cx54tyDMEIA (male, 1, 0.03%), UCFadJmP9WNNHfaGVIPMj61A (male, 1, 0.03%), UCtdvIYsWH5ExPtRiCR0dejQ (female, 1, 0.03%), UCfOdDDLoWlP3aXvNi146CxA (male, 1, 0.03%), UCH1BOfbaZEndPaQAJ1sstHg (female, 1, 0.03%), UC0ZPKM4RA-m1thhxjmers5g (female, 1, 0.03%), UC-Zg0Q4lE-Uk9XYAHX-Nj6A (male, 1, 0.03%), UCK0unFYixOTVMCmiU6NY9Nw (male, 1, 0.03%), UCpE5S-iuV221aM4vxaHbU6w (male, 1, 0.03%), UC2SDCQ_wTaR4jOGA4LVivuA (female, 1, 0.03%), UCWJp3kRn3ZK2kOCqNHz0VBw (female, 1, 0.03%), UCMyM7x2ZX0w2qbnrRFX91Nw (female, 1, 0.03%), UCS8806_2El-sadP7rwkI5Lg (male, 1, 0.03%), UCdzO_9VoCiO35EHOLjaH6MA (male, 1, 0.03%), UCKPKN15xKoE9uKBVjZ6sagw (male, 1, 0.03%), UCGZuCKmrPQ6vlTVxL_DBmIQ (female, 1, 0.03%), UCuTcp5DjqOvAQY136jYGXqg (female, 1, 0.03%), UCIvZRoVl9ADtQ22CQOs24_w (male, 1, 0.03%), UC4NkS_w8o50U6jw2oksEMxQ (female, 1, 0.03%), UCXrfD2R9cxNJ40udLMXu-OA (female, 1, 0.03%), UC0KHKeRR9X946Ott3C8tKpw (female, 1, 0.03%), UC0i3XXC3tiZyPN81yrVsOdw (female, 1, 0.03%), UC_aP7p621ATY_yAa8jMqUVA (male, 1, 0.03%), UCC_XHoVUfIMhx_7AG1NQQjw (male, 1, 0.03%), UC_aP7p621ATY_yAa8jMqUVA (female, 1, 0.03%), UCIFdw_tmUSk-Ziiw0X-taCw (male, 1, 0.03%), UCpWLE7oAq7ngU4NOg5rJdKg (male, 1, 0.03%), UClXVcgLnEjOozczj6oqJW5w (male, 1, 0.03%), UCe24MwwMf93_45ZmHkGrNXw (male, 1, 0.03%), UCZ85UhK5tgU2OzXWin81GWA (male, 1, 0.03%), UCfYcdHJeCgBTCCFEio4LkXA (female, 1, 0.03%), UCAu4ni9sAuwXrbXb9tAjFVQ (female, 1, 0.03%), UCE3IWJtce6NoHkAoaMlfZyQ (male, 1, 0.03%), UCsgP8iNXTPtbNhDIcQOz3dg (female, 1, 0.03%), UClYdSKazSIYR9VLmJf9P8_w (female, 1, 0.03%), UCiwkNpVfiF8ElHFf03cSu4A (male, 1, 0.03%), UCe6rQbQU8Dc23k6LQ-owNJA (female, 1, 0.03%), UCko_tGt6GkTIkVHbztNgC4w (female, 1, 0.03%), UClPkNCC1Gd_cyI7QTe7lJwA (male, 1, 0.03%), UCMPfTh0U3CMFWCHUuXvDLQw (male, 1, 0.03%), UCXfywPIGeHlataAi1hsB1wg (female, 1, 0.03%), UCWkcASmjGP_FhM08UHtNbsQ (male, 1, 0.03%), UC_x5XG1OV2P6uZZ5FSM9Ttw (male, 1, 0.03%), UCS1Z1tKNGwMAa0Cow8fRSVw (female, 1, 0.03%), UCW5t7tx4ke9elfwBtaC6zRg (female, 1, 0.03%), UCSWHwmbKPQnoYKwGVvVWYog (male, 1, 0.03%), UCl6vWwMCjufI8OPtOInHf0g (female, 1, 0.03%), UCPb5b5NUdThpyewy2YKlPwA (female, 1, 0.03%), UCFTFP0bv7o59qW_1XJAuqoA (male, 1, 0.03%), UCFv5pvZwJ2ZORd8QaY3SnMw (female, 1, 0.03%), UC_TLNbBZBPFAjXecf7NTKfg (female, 1, 0.03%), UC6fp3DKaK5INfZ3gfpVcM0A (male, 1, 0.03%), UCfSAVdVthSxiVhnu3fP5ijg (male, 1, 0.03%), UC0JO-FDmcF2lFP1l-rsmBWg (male, 1, 0.03%), UC7VegB4HHzMLEwOE3E7XJ6w (male, 1, 0.03%), UCeBE2qQgwNpm7RvQ31BkGxw (male, 1, 0.03%), UC6ZfX0Yq82WTROPFHqOz9pw (male, 1, 0.03%), UCRvCxhTMns6Bx_y1s9h80Pw (female, 1, 0.03%)

4.4. Video Transcript Statistics

The most common words in video transcripts are listed below:

the (19546, 4.37%), and (14824, 3.32%), to (14205, 3.18%), a (10891, 2.44%), of (10173, 2.28%), you (8374, 1.87%), that (8259, 1.85%), i (7791, 1.74%), in (6883, 1.54%), is (6700, 1.50%), it (6177, 1.38%), this (4621, 1.03%), for (4131, 0.92%), so (3435, 0.77%), have (3055, 0.68%), on (2979, 0.67%), are (2938, 0.66%), with (2826, 0.63%), we (2819, 0.63%), it's (2781, 0.62%), be (2765, 0.62%), but (2659, 0.59%), your (2622, 0.59%), was (2558, 0.57%), not (2517, 0.56%), they (2470, 0.55%), like (2354, 0.53%), movie (2269, 0.51%), as (2210, 0.49%), just (2200, 0.49%), or (2199, 0.49%), about (2119, 0.47%), if (2076, 0.46%), really (1976, 0.44%), can (1899, 0.42%), know (1867, 0.42%), what (1862, 0.42%), at (1808, 0.40%), my (1767, 0.40%), all (1749, 0.39%), one (1609, 0.36%), do (1595, 0.36%), going (1452, 0.32%), out (1425, 0.32%), from (1385, 0.31%), will (1366, 0.31%), an (1348, 0.30%), people (1332, 0.30%), because (1278, 0.29%), get (1227, 0.27%), i'm (1210, 0.27%), there (1206, 0.27%), more (1202, 0.27%), very (1200, 0.27%), he (1199, 0.27%), how (1157, 0.26%), don't (1148, 0.26%), our (1146, 0.26%), think (1137, 0.25%), some (1117, 0.25%), when (1086, 0.24%), who (1069, 0.24%), has (1055, 0.24%), by (1053, 0.24%), good (1044, 0.23%), their (1041, 0.23%), up (1039, 0.23%), me (1017, 0.23%), want (992, 0.22%), would (968, 0.22%), you're (950, 0.21%), them (917, 0.21%), see (912, 0.20%), that's (870, 0.19%), which (858, 0.19%), go (852, 0.19%), now (849, 0.19%), make (830, 0.19%), also (828, 0.19%), time (818, 0.18%), other (750, 0.17%), lot (748, 0.17%), then (739, 0.17%), these (720, 0.16%), into (710, 0.16%), way (678, 0.15%), kind (675, 0.15%), things (673, 0.15%), great (664, 0.15%), here (663, 0.15%), been (657, 0.15%), well (645, 0.14%), those (645, 0.14%), two (631, 0.14%), much (629, 0.14%), even (628, 0.14%), first (620, 0.14%), had (599, 0.13%), where (593, 0.13%), were (588, 0.13%), actually (586, 0.13%), his (585, 0.13%), new (578, 0.13%), there's (571, 0.13%), no (568, 0.13%), something (561, 0.13%), need (549, 0.12%), any (538, 0.12%), say (521, 0.12%), little (512, 0.11%), over (510, 0.11%), work (503, 0.11%), got (496, 0.11%), most (489, 0.11%), could (475, 0.11%), take (474, 0.11%), than (471, 0.11%), right (465, 0.10%), thing (463, 0.10%), give (455, 0.10%), back (448, 0.10%), they're (447, 0.10%), only (445, 0.10%), today (443, 0.10%), didn't (437, 0.10%), look (431, 0.10%), many (424, 0.09%), through (413, 0.09%), help (410, 0.09%), she (409, 0.09%), business (408, 0.09%), find (396, 0.09%), should (395, 0.09%), we're (393, 0.09%), pretty (393, 0.09%), different (392, 0.09%), her (391, 0.09%), mean (389, 0.09%), doing (387, 0.09%), being (379, 0.08%), use (373, 0.08%), why (372, 0.08%), important (366, 0.08%), world (366, 0.08%), three (362, 0.08%), video (357, 0.08%), did (356, 0.08%), movies (353, 0.08%), best (351, 0.08%), story (350, 0.08%), sure (348, 0.08%), come (340, 0.08%), money (334, 0.07%), probably (333, 0.07%), life (331, 0.07%), better (330, 0.07%), i've (329, 0.07%), years (329, 0.07%), looking (328, 0.07%), he's (326, 0.07%), too (325, 0.07%), every (325, 0.07%), him (322, 0.07%), five (322, 0.07%), talk (321, 0.07%), day (318, 0.07%), might (316, 0.07%), down (315, 0.07%), am (313, 0.07%), before (313, 0.07%), may (310, 0.07%), around (309, 0.07%), bad (307, 0.07%), hi (306, 0.07%), basically (303, 0.07%), again (302, 0.07%), does (297, 0.07%), part (292, 0.07%), after (290, 0.06%), tell (290, 0.06%), maybe (284, 0.06%), same (279, 0.06%), course (271, 0.06%), name (271, 0.06%), its (270, 0.06%), another (270, 0.06%), own (269, 0.06%), put (267, 0.06%), next (264, 0.06%), always (264, 0.06%), made (262, 0.06%), gonna (262, 0.06%), love (261, 0.06%), big (258, 0.06%), said (258, 0.06%), definitely (258, 0.06%), information (257, 0.06%), long (257, 0.06%), end (256, 0.06%), start (256, 0.06%), off (256, 0.06%), able (254, 0.06%), us (254, 0.06%), year (252, 0.06%), everything (251, 0.06%), job (251, 0.06%), review (250, 0.06%), bit (250, 0.06%), school (249, 0.06%), called (248, 0.06%), never (245, 0.05%), try (245, 0.05%), whole (244, 0.05%), still (242, 0.05%), watch (241, 0.05%), thought (240, 0.05%), recommend (239, 0.05%), person (236, 0.05%), making (236, 0.05%), trying (234, 0.05%), funny (232, 0.05%), real (231, 0.05%), last (231, 0.05%), done (229, 0.05%), working (227, 0.05%), doesn't (226, 0.05%), seen (225, 0.05%), together (225, 0.05%), feel (224, 0.05%), company (222, 0.05%), getting (219, 0.05%), show (217, 0.05%), stuff (211, 0.05%), film (209, 0.05%), anything (209, 0.05%), you'll (207, 0.05%), let (206, 0.05%), can't (206, 0.05%), wasn't (205, 0.05%), talking (203, 0.05%), believe (199, 0.04%), few (198, 0.04%), you've (197, 0.04%), students (194, 0.04%), guys (193, 0.04%), keep (193, 0.04%), ever (192, 0.04%), understand (192, 0.04%), guy (191, 0.04%), learn (188, 0.04%)

References

[1] G. Coppersmith and E. Kelly. Dynamic wordclouds and vennclouds for exploratory data analysis. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pages 22–29, Baltimore, Maryland, USA, June 2014. Association for Computational Linguistics.

[2] P. Ekman, W. V. Freisen, and S. Ancoli. Facial signs of emotional experience. *Journal of personality and social psychology*, 39(6):1125, 1980.

[3] Itseez. Open source computer vision library. `https://github.com/itseez/opencv`, 2015.

[4] S. Park, H. S. Shim, M. Chatterjee, K. Sagae, and L.-P. Morency. Computational analysis of persuasiveness in social multimedia: A novel dataset and multimodal prediction approach. In *Proceedings of the 16th International Conference on Multimodal Interaction*, ICMI '14, pages 50–57, New York, NY, USA, 2014. ACM.

[5] J. Yuan and M. Liberman. Speaker identification on the scotus corpus. *Journal of the Acoustical Society of America*, 123(5):3878, 2008.

[6] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency. Tensor fusion network for multimodal sentiment analysis. In *Empirical Methods in Natural Language Processing, EMNLP*, 2017.