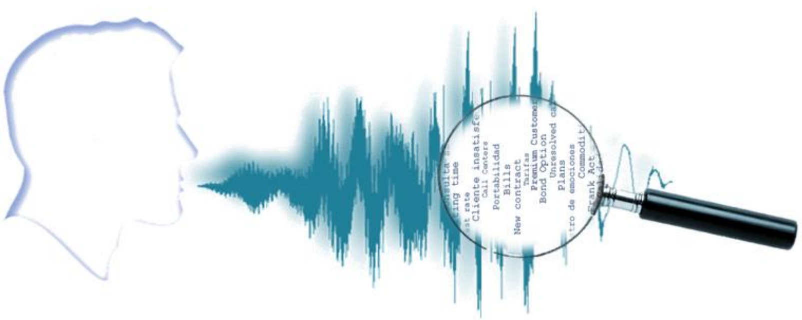
Using Unstructured Data for Business Analytics

Associate Professor
HUANG, Ke-Wei



Agenda

1. Image Mining for Business Analytics
    1. History of Computer Vision by DNN
    2. Facial images
    3. Satellite images
    4. Manufacturing images
    5. Medical images for prescription (there are many studies but are omitted in this module.)

2. Audio Mining for Business Analytics

3. Video Mining for Business Analytics

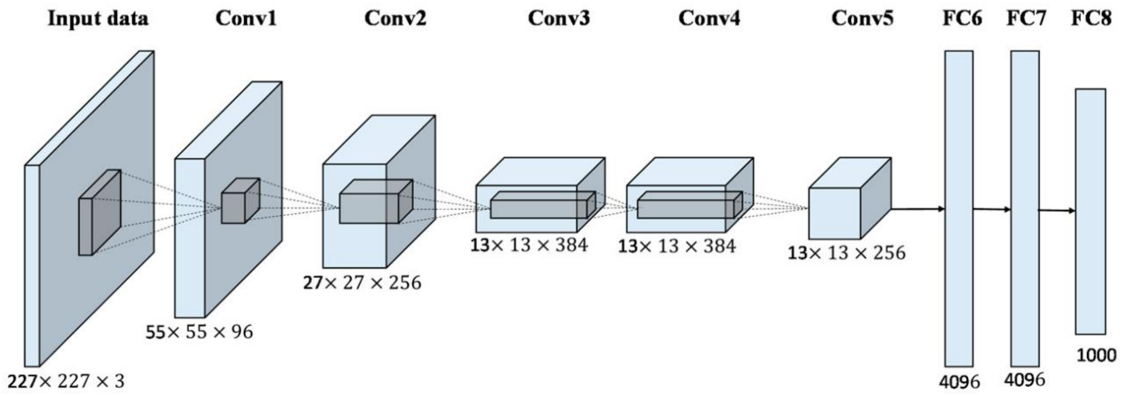
1. Image Mining

- Deep learning is most successfully implemented for analyzing image files and the huge success make computer vision applications become a fastest-growing research area.
- Computer vision started around 1970s (surely without neural network, no deep learning).
- At the time, computer vision is concerned with the theory behind artificial systems that extract information from images.
- The first modern deep learning model appeared after 30 years. It is called LeNet, published in 1998 by Pofessor Yann LeCun, who is also recognized as the father of Convolutional Neural Network (CNN).

Evolution of DNN for Computer Vision

- LeNet-5, the pioneering 7-level convolutional network by LeCun et al. in 1998 classifies digits, was applied by several banks to recognise hand-written numbers on checks (cheques) digitized in 32x32 pixel images (no color, black and white).
- The ability to process higher resolution images requires larger and more layers of convolutional neural networks, so this technique is constrained by the availability of computing resources at that time.
- After 1998, computing hardware kept being improved and GPU computing also improved. The same idea of LeNet-5 became very powerful after we have much better hardware.

Evolution of DNN for Computer Vision

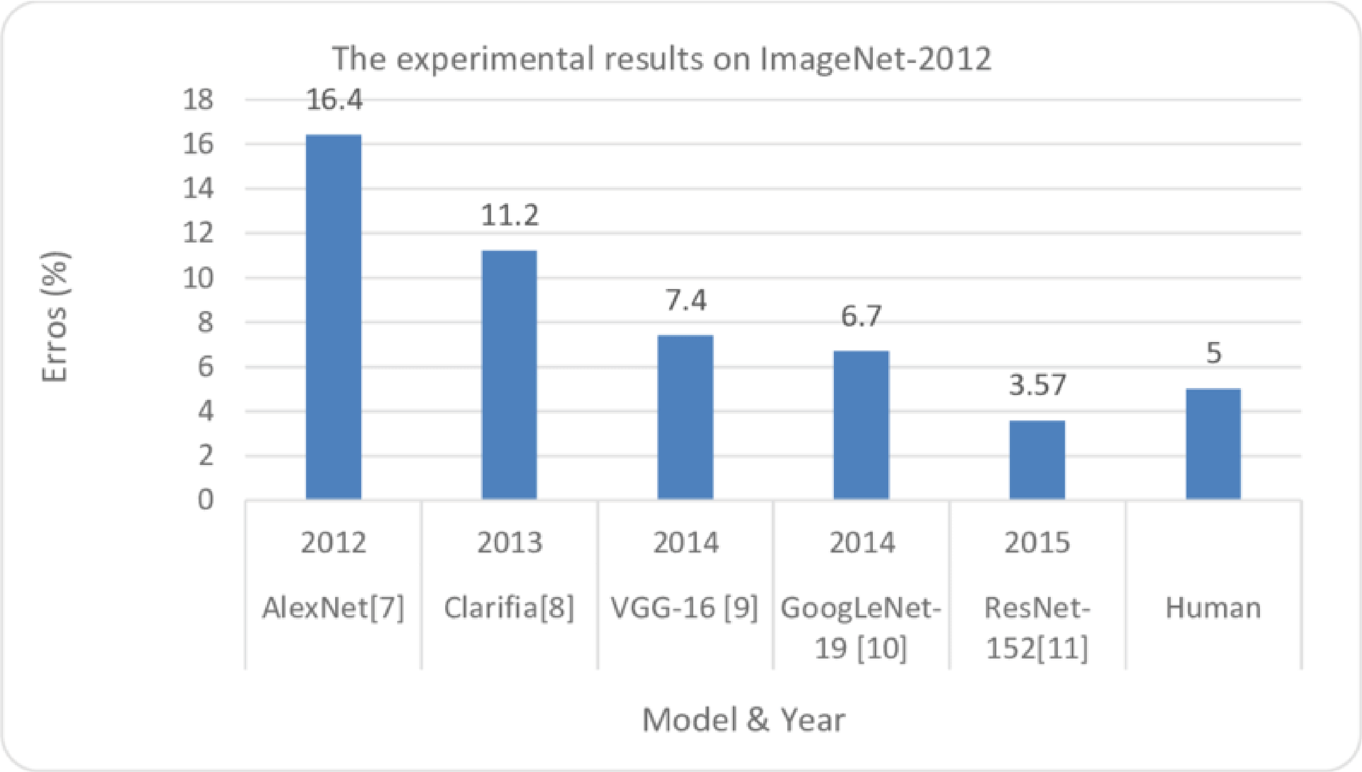


- After almost 15 years, there is another significant breakthrough, called AlexNet published in 2012.
- AlexNet won the most important image data competition in 2012, beating the #2 by 10% error, achieving error rate at 15.3%.
- AlexNet is an 8-layer neural network based on the core idea of LeNet, improvement with ReLu activation function.
- AlexNet has huge impact on the CV field.

ImageNet Large Scale Visual Recognition Challenge

- ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is the largest annual data competition.
- ImageNet (started in 2009-10) is a huge image database for training and testing modern algorithms.
- Now ImageNet has about 14 million images with human coded categories. There are 20,000 categories indicating what is shown in the image.
- ILSVRC competes by classifying images into 1000 selected categories. Before AlexNet, the state-of-art can achieve about 25% Top-5 error rate (by the true category is not in the top-5 predicted category).

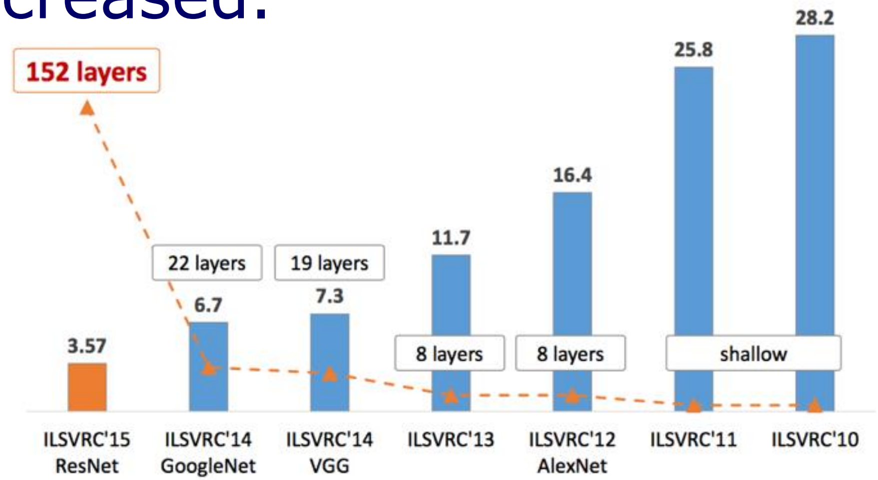
ImageNet Large Scale Visual Recognition Challenge



- 2015 is another important year in which ResNet delivered another breakthrough performance 3.57% much better than state-of-art at that time. 3.57% is important because it is better than human being for this specific task (but human can recognize more categories and also recognize context better).

ImageNet Large Scale Visual Recognition Challenge

- The other reason why ResNet is a breakthrough because the number of layers sharply increased.

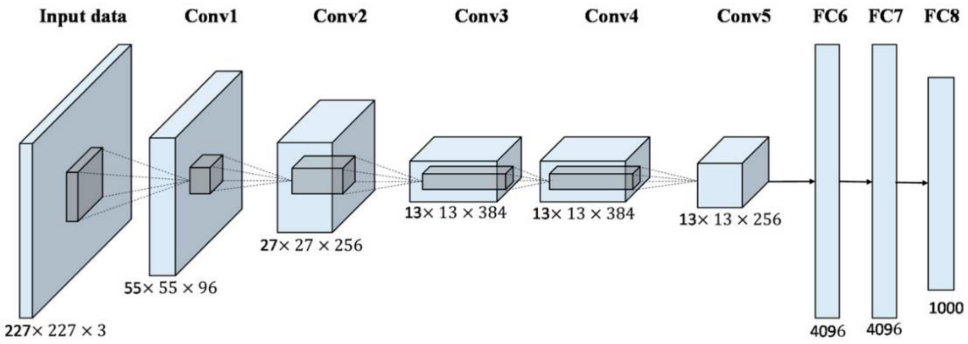


- 2016 winning performance is around 3% but lower than 3%.
- 2017 winning performance is 2.251%.
- 2017 is the last competition and the organizer decided to change the topic in 2018.

Pre-Trained Models?

- LeNet for simple images.
- All names mentioned in the previous slides are potential pre-trained models that you can try. VGG 2014 and ResNet 2015 are the two famous ones researchers mostly use.
- This link is an excellent reference
- One VGG pre-trained model online is at https://www.cs.toronto.edu/~frossard/post/vgg16/
- A very comprehensive list https://github.com/tensorflow/models/tree/master/research/slim#pre-trained-models

Pre-Trained Models?



- Conceptually, using pretrained model means we remove the output layer, which is a soft-max layer for classifying images into 1000 types.
- In this AlexNet example, we can use the output of FC6 or FC7 as the output of the pre-trained model. You can think of this as we send image into the first 8 layers and AlexNet turn each image into a numerical vector, which is like the image characteristics.
- Then you build your own supervised learning model by using the output of FC7 as input features and your coded image label to train by any supervised learning method.

Examples of Papers Covered in this Part

- Psychology => It seems there are some facial traits correlate with useful characteristics.

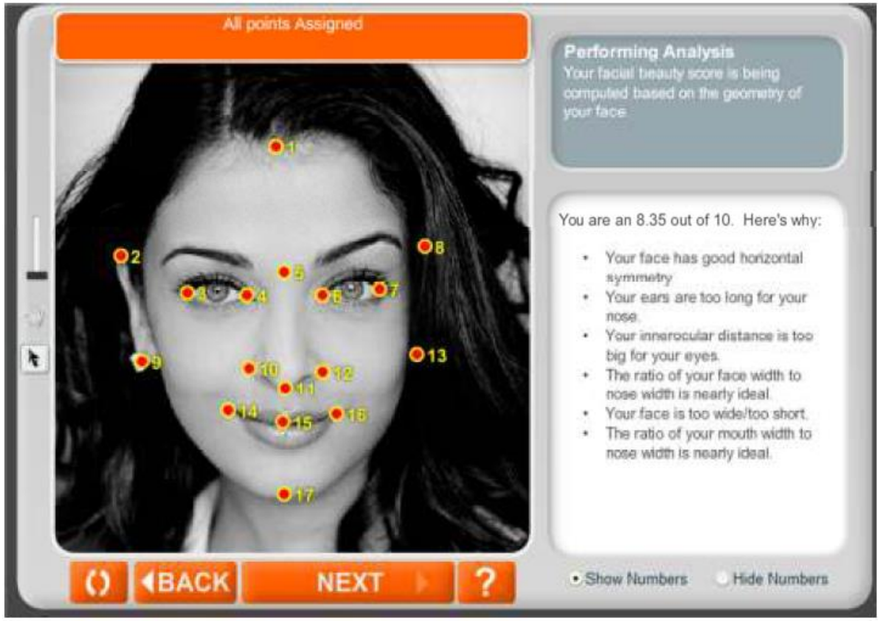
1. 2014. "Beauty Is Wealth: CEO Appearance and Shareholder Value". (unpublished paper)
   - There are new papers using pre-trained VGG for classifying beauty or not.
2. 2017. "The Role of Facial Appearance on CEO Selection after Firm Misconduct". (Psychology paper, not a top business school paper)
3. 2017. "Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images". (Psychology paper, but using VGG pre-trained model.)

Study 1: Image of CEO

- Research Objective: this paper examines whether and the appearance of CEOs affects shareholder value.
- Sample firms: facial Attractiveness Index of 677 CEOs from the S&P 500 firms based on their facial geometry.
- In this paper, Facial Attractive Index of CEOs is calculated from *anaface.com*, which uses similar techniques to those used by Schmid, Marx, and Samal (2008).
- *Anaface.com's* specific algorithm is proprietary, but it takes into account many factors from neoclassical beauty, modern research papers, and their own scientific studies/statistical analysis. Examples include things such as comparing innerocular distance to mouth width and nose width to face height.
- There very recent papers use DNN for similar studies.

Study 1: Findings

- CEOs with a higher Facial Attractiveness Index are associated with better stock returns around their first days on the job, higher acquirer returns upon acquisition announcements, and higher total compensation to the CEO.
- CEO appearance matters for shareholder value and provide an explanation why more attractive CEOs receive “beauty premiums” in their compensation.



Study 2: Image of CEO

1. Can we measure trustworthiness of CEOs by photos?
2. How does this quantified trustworthiness correlate with financial/accounting variables?
3. Data:
    1. Firms with financial misconduct have been studied thoroughly in accounting and is well-defined.
    2. Sample firms are firms with accounting restatements (only downward, negative revisions) from 2003 to 2006
    3. Matched each restating firm with a nonrestating firm as a benchmarking group
    4. Using photos available from company web sites and Google Images.
    5. 394 restating firms in the end.

Study 2: Computational Approach

Successor CEO facial integrity is the key variable.

1. It is measured by using the facial width-to-height ratio (fWHR), which is the ratio of bizygomatic width (i.e., width between the two cheekbones) to upper facial height (i.e., distance between upper lip and mid eyebrow).
2. This objective facial structure measure has been associated with leadership characteristics, such as general trustworthiness in “Psychology” (Dr. Huang: it depends on how much you trust psychology research).
3. For instance, men with higher fWHRs are more dominant and aggressive and have been linked to leading more financially successful firms than men with lower fWHRs.
4. Focus only on male CEO because this metric works better for male.

Psychology Study Findings

- Past fWHR research has found that men with wider faces (relative to the height of their face) are viewed as less trustworthy than are men with narrower faces, who are viewed as more trustworthy
- Men with narrower faces are less likely to cheat or exploit others' trust. Importantly, men with wider faces not only behave in untrustworthy ways, but also are perceived as being untrustworthy and self-interested.
- In summary, results across a number of studies have demonstrated a positive relationship between fWHR and unethical behavior, exploitation and cheating, which suggests a strong link to perceptions of untrustworthiness generally, and likely lack of integrity in particular.

Study 2: Findings

1. Firms with financial misconduct (financial restatements leading to lower earnings) are more likely to select a new CEO who looks like having more integrity. Analysts also react more positively to CEOs who looks like trustworthy.
2. But it is the opposite for normal firms (?) according to this paper.

Table 2

Effects of Restatement on Successor CEO FWHR[a,b]

Variables	DV = Successor CEO fwhr Model 1	Model 2
Firm age	.03 (.03)	.01 (.02)
Firm size	−.01 (.01)	−.01 (.01)
Year, 2004	−.01 (.06)	.01 (.06)
Year, 2005	−.01 (.05)	−.00 (.03)
Year, 2006	.11* (.06)	.14* (.06)
ROA	1.08 (1.74)	1.82 (1.23)
Incumbent CEO fwhr	.17*** (.04)	.22*** (.06)
Incumbent CEO smile	−.01 (.01)	−.01* (.00)
Successor CEO smile	−.03 (.04)	.01 (.03)
Incumbent difficulty	−.01 (.01)	−.01* (.01)
Successor difficulty	−.02* (.01)	−.01 (.01)
Successor CEO experience	−.08*** (.02)	−.09** (.03)
Successor CEO elite education	−.03 (.03)	−.02 (.03)
Successor CEO from Fin./Acctg.	.03 (.04)	.03 (.04)
Restatement		−.06* (.03)
Constant	.16*** (.00)	.05*** (.00)
N total	501	501
N uncensored	42	42
Log Lik	−61.94	−59.92
LR test (χ^2)		4.04*
Lambda	−.15*	−.15*
athrho	−14.52 (56.79)	−15.74 (102.33)
lnsigma	−1.91*** (.12)	−1.89*** (.12)

[a] Regression model is from second stage Heckman; standard errors are in parentheses. [b] For conciseness, the output from the selection equation (first stage) is not shown.

* $p < .05$. ** $p < .01$. *** $p < .001$, two-sided tests.

Study 2: Findings

Table 3

Effects of Restatement and Successor CEO FWHR on External Reactions[a]

percentage of negative words in news media

Variables	DV = Change in analysts forecasts			DV = Negative coverage		
	Model 1	Model 2	Model 3	Model 7	Model 8	Model 9
Firm age	.34^{+} (.19)	.47* (.21)	.47*** (.10)	.53* (.24)	.40** (.13)	.44 (.27)
Firm size	−.02 (.07)	.02 (.06)	.08 (.06)	−.03 (.19)	.12 (.08)	.07 (.11)
Year, 2004	−1.64*** (.45)	−1.83** (.57)	−1.79*** (.23)	1.26 (.77)	1.36*** (.41)	1.22** (.45)
Year, 2005	−.07 (.39)	−.02 (.37)	.43 (.34)	1.08** (.38)	.93* (.46)	.64^{+} (.38)
Year, 2006	.79* (.34)	.96*** (.25)	1.93*** (.15)	.63 (.53)	.97^{+} (.59)	.44 (.68)
ROA	10.72 (16.18)	13.66 (13.29)	18.70 (13.24)	−14.33 (13.13)	−36.37*** (10.60)	−39.28^{+} (21.11)
Incumbent CEO fwhr	−5.95** (2.19)	−6.36^{+} (3.58)	−4.47** (1.70)	−7.32*** (2.14)	−6.59*** (1.42)	−8.42*** (1.28)
Incumbent CEO smile	−.35** (.14)	−.45^{+} (.27)	−.74*** (.06)	−.56 (1.07)	−.79*** (.22)	−.48 (.40)
Successor CEO smile	1.10** (.43)	1.04** (.39)	1.67*** (.14)	−1.82*** (.51)	−1.65*** (.39)	−1.80*** (.45)
Incumbent difficulty	.33*** (.08)	.33*** (.10)	.32*** (.06)	.19*** (.05)	.11 (.09)	.08 (.08)
Successor difficulty	−.08 (.16)	−.17^{+} (.09)	−.16 (.10)	.17 (.22)	.06 (.10)	.07 (.22)
Successor CEO experience	.75*** (.23)	.77* (.32)	.47*** (.11)	−.15 (.49)	−.10 (.15)	−.45 (.37)
Successor CEO elite education	−1.09*** (.12)	−.96*** (.22)	−.92*** (.11)	−.05 (.27)	−.38* (.18)	−.46^{+} (.27)
Successor CEO from Fin./Acctg.	−.31 (.39)	−.14 (.31)	−.33 (.30)	−1.73*** (.38)	−1.49*** (.31)	−1.45*** (.35)
Restatement	−.05 (.15)	.02 (.22)	12.15 (.00)	−.81 (.76)	−1.05** (.34)	−8.96 (.00)
Change in number of forecast	−.31*** (.09)	−.25** (.08)	−.28** (.09)			
Duration from last forecast	.01 (.01)	.02 (.03)	.07^{+} (.04)			
Successor CEO fwhr		−.91* (.41)	4.22*** (.24)		1.45*** (.31)	−1.71** (.56)
Restatement × Successor CEO fWHR			−6.56*** (.13)			4.31*** (.24)
Constant	−1.99*** (.00)	−.18*** (.00)	−11.97*** (1.85)	3.02*** (.00)	1.42*** (.00)	7.52*** (1.90)
N total	484	483	483	495	494	494
N uncensored	26	25	25	36	35	35
Log Lik	−63.40	−58.71	−52.32	−134.9	−127.8	−126.8
LR test (χ^2)		9.38**	22.16***		14.2***	16.2***
Lambda	.38***	.33***	.25***	−1.50**	−1.36*	−1.36^{+}
athrho	13.73*** (.17)	13.45*** (.17)	11.94*** (.18)	−15.58*** (.14)	−17.43*** (.14)	−13.09*** (.14)
lnsigma	−.97*** (.16)	−1.09*** (.17)	−1.38*** (.27)	.41** (.15)	.31* (.15)	.30 (.19)

[a] Regression model is from second stage Heckman; standard errors are in parenthes

$^{+}$ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$, two-sided tests.

Study 3: Sexual Orientation

1. Research objective: can we use AI to predict sexual orientation by face images?

- There are existing literature about identifying sexual orientation by photos, but not by deep learning or AI.
- Facial images from public profiles posted on a U.S. dating website. This paper recorded 130,741 images of 36,630 men and 170,360 images of 38,593 women between the ages of 18 and 40, who reported their location as the United States.
- 35,326 photos after data cleaning.
- Gay and heterosexual people were represented in equal numbers (on dating site, users self-reported that (man looking for man…etc.)).

Method

- Facial features were extracted from the images using a widely employed DNN called VGG-Face.
- VGG-Face translates a facial image into 4,096 scores subsuming its core features.
- Unfortunately, unlike psychometric scores, VGG-Face scores are not easily interpretable. A single score might subsume differences in multiple facial features typically considered to be distinct by humans (e.g., nose shape, skin tone, or eye color).
- After dimension reduction by singular value decomposition SVD, 500 selected features were entered into a logistic regression aimed at classifying sexual orientation.

Two Advantages of VGG-Face

- VGG-Face aims at representing a given face as a vector of scores that are as unaffected as possible by facial expression, background, lighting, head orientation, image properties such as brightness or contrast, and other factors that can vary across different images of the same person.
- Second, employing a DNN trained on a different sample and for a different purpose reduces the risk of overfitting (i.e., discovering differences between gay and heterosexual faces that are specific to our sample rather than universal).
  - This paper uses a trained model (VGG-Face) built from millions of photos.
  - Authors do not train a new model.

Findings

- These features were entered into a logistic regression aimed at classifying sexual orientation.

- Given a single facial image, a classifier could correctly distinguish between gay and heterosexual men in 81% of cases (AUC), and in 71% of cases for women (AUC).
- The precision of the algorithm increased to 91% (AUC) and 83% (AUC), respectively, given five facial images per person.
- Human judges achieved much lower precision: 61% for men and 54% for women.

Managerial implications?

1. We can design customized display ads or marketing messages to these special groups of customers just by their FB/LinkedIn photos.
2. Companies with membership program with photos can provided customized services.
3. This area is very new and there are potential to apply deep learning on photos to identify other personalities. Findings will be useful for marketing or Human Resources management purpose. For example, customers are more loyal or not is very critical both to marketing (less likely to switch) and HRM (less likely to switch), lazy => good for marketing, (wont check bills, wont complain, wont switch), bad for HRM.

Agenda

1. Image Mining for Business Analytics
    1. Face images
    2. Satellite images
    3. Manufacturing images
    4. Medical images for prescription (there are many applications but are omitted in this module.)

2. Audio Mining for Business Analytics

3. Video Mining for Business Analytics

Satellite Image Mining

1. 2016. “Combining satellite imagery and machine learning to predict poverty” (Science, very top journal)
2. 2017. “Data Mining Techniques on Satellite Images for Discovery of Risk Areas” (2-3rd tiered IS journal)
3. 2017. "Using deep learning and Google Street View to estimate the demographic makeup of neighbourhoods across the United States.“ (top journal)

- Hedge funds already started using this kind of images to predict the agricultural products’ yield for trading commodity futures.

Study 1

- Survey and satellite data from five African countries — Nigeria, Tanzania, Uganda, Malawi, and Rwanda.
- This paper shows that convolutional neural network can be trained to identify image features that can explain up to 75% of the variation in local-level economic outcomes.
- Publicly available (image) data only.
- This can alleviate the economic data collection efforts and costs in these poor countries.
- This paper uses a semi-supervised learning approach because economic survey results are available only in few areas.

Study 1

- Both day-time (roof material quality) and night-time images (night activities) are important for this project’s purpose.
- The trained model is then used to estimate either average household expenditures or average household wealth at the “cluster” level (roughly equivalent to villages in rural areas or wards in urban areas).
- This paper demonstrates that existing high resolution daytime satellite imagery can be used to make fairly accurate predictions about the spatial distribution of economic well-being across five African countries.

Study 2

- This paper focuses on remote sensing satellite data processing using data mining methods to discover risk areas of the epidemic disease by connecting the environment, climate and health.
- Supervised classification algorithm
- This paper establishes the links between the environment and the epidemic.
- This will allow national governments, local authorities and the public health officials to effective management according to risk areas.

Study 3: Google Street View

- A similar but quite different source of image is Google Street View. If you don’t know this, this is the 360 degree photos that captured by Google cars from almost all streets in the world.
- Google applied deep learning to enrich their map datasets.
- For example,
  - addresses are extracted from photos.
  - Merchant stores are extracted from banners shown in the photos
  - It can help identify the types of building and be used with satellite image to produce more information about that location.

Study 3

- Using deep learning and Google Street View to estimate the demographic makeup of neighbourhoods across the United States http://www.pnas.org/content/114/50/13108

- This paper shows that socioeconomic attributes such as income, race, education, and voting patterns can be inferred from cars detected in Google Street View images using deep learning.

Agenda

1. Image Mining for Business Analytics
    1. Face images
    2. Satellite images
    3. Manufacturing images
    4. Medical images for prescription (there are many applications but are omitted in this module.)

2. Audio Mining for Business Analytics

3. Video Mining for Business Analytics

Agenda

1. Image Mining for Business Analytics
    1. Face images
    2. Satellite images
    3. Manufacturing images
    4. Medical images for prescription (there are many applications but are omitted in this module.)
2. Audio Mining for Business Analytics
3. Video Mining for Business Analytics

Manufacturing Image Mining

- 2010. "An Effective and Efficient Defect Inspection System for Tft-Lcd Polarised Films Using Adaptive Thresholds and Shape-Based Image Analyses“ (2nd tier journal in OM/IEOR/Decision Science)
- This paper proposes a defect inspection system for TFT-LCD film images that detects film defects and classifies them based on the type of a defect.

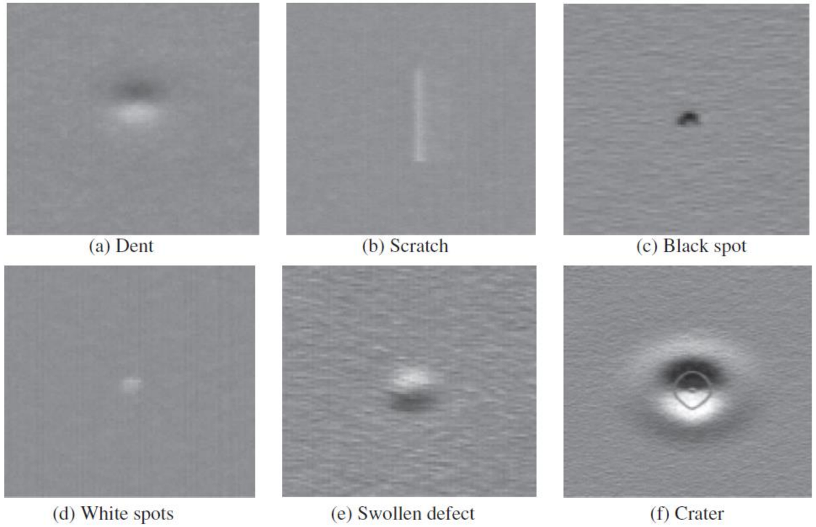


Figure 1. Defect images captured by line-scan cameras and LED illumination sources.

Agenda

1. Image Mining for Business Analytics
    1. Face images
    2. Satellite images
    3. Manufacturing images
    4. Medical images for prescription (there are many applications but are omitted in this module.)

2. Audio Mining for Business Analytics

3. Video Mining for Business Analytics

Audio Analytics

- There are two types of audio analytics in academics and in practice: automatic speech recognition (using the content from voice) and the pitch/tone/speed of speech, we focus on the 2nd one.
- There are limited papers published in top business journals.
- Psychologists showed that “hesitate=uncertain” can be detected from speech.
- Another large category is “emotions” detection, but this is still not well-accepted among psychologists, especially lie detection by voice.
- 4 papers covered in this part (see footnotes).

R Packages

- R packages like tuneR make it possible to record, import, modify and export audio objects.
- tuneR and seewave packages can be used to analyze sound. Many complimentary functions are offered by these packages and they can be used to extract and compare relevant amplitude, temporal, phase and frequency parameters. The package seewave is now widely used in bioacoustics.
- New packages like playitbyr and audiolyzR can be used for sonification. Playitbyr is used to map data onto sonic parameters like pitch, tempo and rhythm.

Application 1: Call Center Monitoring

- Monitoring can be done by processing speech associated with the call and by identifying indicators of profanity, inappropriate language or the use of word “supervisor”.
- If any of these indicators are found in the processed speech, then monitoring profile is searched to determine whether the speech is normal for the customer or the call should be directed to supervisor.
- Pitch and intensity of the call is also measured by detecting the energy levels of the call.
- Semantic categorization system, parsing system, and Automatic Speech Recognition (ASR) system have been proposed.

Other Audio Analytics Applications

1. Military applications: please see the paper
2. Aviation applications: please see the paper
3. Financial applications: 3 examples later
4. Healthcare applications:

- It can be used to detect Parkinson‘s disease in a patient.
- It is also used to detect depression and other mental issues.
- One NUS CS professor has been working on how music can help patients recover better.

Financial Study 1: Analyzing Speech to Detect Financial Misreporting

Whether vocal markers of cognitive dissonance are useful for detecting financial misreporting?

Data

- Streamed audio from earnings conference calls available on ThomsonReuters StreetEvents was encoded in mono directly onto computer hard disk, using Total Recorder 7.1 software, at 11.025 kHz sampling rate and saved as uncompressed .wav files.
- This speech corpus comprises 1,647 quarterly earnings conference calls spanning the period Jan 1 2007 to December 31 2007.

Method: LVA dissonance measure

- Processed audio files (.wav files) were then analyzed using a commercial version (Ex-Sense Pro R) of the LVA software that uses LVA technology developed by Nemesysco Ltd.
- LVA is based on a set of proprietary signal–processing algorithms purported to identify different types of stress, cognitive processes, and emotional reactions.
- The algorithms measure features of the speech waveform to create a foundation for identifying the speak
- The software produces four “fundamental” measures, emotional stress level, cognition level, general stress level, and thinking level.

Method: LVA dissonance measure

- Pertinent to our study is Cognition Level, which is purported to measure cognitive dissonance.
- The software also produces other measures deemed "conclusion" variables (e.g., Lie Stress), which are proprietary combinations of the fundamental LVA measures and are meant to indicate when a speech segment may represent untruthful statements.
- From Dr. Huang: some psychology researchers questioned the validity of this kind of software and the voice-based lie detector. For your future works, you need to double-check whether the software you use can effectively identify some emotions.

Findings

TABLE 4—*Continued*

Panel B: Results after including control variables							
Variable[a]	Predicted Sign	(A) *IRREG_RES*	(B) *IRREG_RES*	(C) *IRREG_RES*	(D) *IRREG_RES*	(E) *IRREG_RES*	(F) *IRREG_RES*
Intercept	(?)	−5.238***	−5.934***	−6.329***	−1.239	−2.225	−3.029
		(0.641)	(0.983)	(0.984)	(2.572)	(2.606)	(2.643)
Vocal Dissonance Marker							
COGDIS	(+)	5.241**	4.729**	4.959**	4.510**		4.677**
		(2.492)	(2.432)	(2.461)	(2.597)		(2.767)
Financial Statement–Based Predictor							
FSCORE	(+)	0.436**		0.324**		0.410**	0.434**
		(0.175)		(0.174)		(0.196)	(0.196)
ACCT_RISK	(+)		0.026***	0.024***		0.028***	0.027***
			(0.010)	(0.010)		(0.009)	(0.009)
Non-misreporting Dissonance Drivers							
RET	(?)				0.807*	0.958**	0.998**
					(0.430)	(0.436)	(0.430)
lnMVE	(?)				−0.435***	−0.568***	−0.563***
					(0.120)	(0.127)	(0.129)
VOL	(?)				−1.110	−1.198	−4.226
					(18.689)	(18.899)	(19.979)
ROA	(?)				2.099	4.237	3.936
					(3.864)	(3.588)	(3.796)
UE	(?)				−10.988*	−8.803	−6.392
					(5.872)	(6.484)	(6.654)
BM	(?)				0.041	−0.336	−0.284
					(1.022)	(0.996)	(1.029)

(*Continued*)

The DV is irregular restatement

Financial Study 2

- "The Power of Voice: Managerial Affective States and Future Firm Performance," The Journal of Finance.
- Main findings:
  - When managers are scrutinized by analysts during conference calls, positive and negative affects displayed by managers are informative about the firm's financial future.
  - Analysts do not incorporate this information when forecasting near-term earnings. When making stock recommendation changes, however, analysts incorporate positive but not negative affect.

Data

- This paper uses CEO and CFO voice recordings from earnings conference calls to develop measures of managers' emotive states when communicating information to analysts and investors.
- All conference calls held between January 1 and December 31, 2007 available on the Thomson Reuters StreetEvents database.
- For each conference call, this paper separately measures positive and negative affects for the CEO and CFO because each individual speaker has a different vocal profile that requires separate calibration.

Method

- This paper constructs measures of affective states with the help of a computer software program that uses LVA technology. LVA was invented in 1997 by Nemesysco Ltd. in Israel.
- This paper uses the LVA-based Ex-Sense Pro-R (version 4.3.9) Digital Emotion Analyzer application.
- Emotion-free benchmarking voice is collected during the beginning of a conversation.
- 4 outputs: Emotion Level, Cognition Level, Global Stress, and Thinking Level. See footnote for more explanations about these 4 outputs.
- This paper uses Emotion Level & Cognition Level

Table III

Estimation of the Association between Affect and Contemporaneous Stock Returns

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and the contemporaneous stock market reaction (*CAR*(0,1)). Robust standard errors are presented in parentheses below the coefficient estimates. *** and **: significant at 0.01 and 0.05 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	(1)	(2)
Intercept	?	−0.0108	−0.0034
		(0.0205)	(0.0200)
PAFF	+	0.1647**	
		(0.0776)	
NAFF	−	−0.0290	
		(0.0270)	
$PAFF^{HS}$	+		0.1263*
			(0.0961)
$NAFF^{HS}$	−		−0.1522***
			(0.0440)
$PAFF^{LS}$	+		0.1507**
			(0.0817)
$NAFF^{LS}$	−		0.0432
			(0.0316)
UE_t	+	0.8204***	0.2576
		(0.2494)	(0.2681)
LNMVE	?	0.0003	−0.0008
		(0.0015)	(0.0015)
MOM	?	0.0040	0.0011
		(0.0107)	(0.0104)
BM	?	−0.0037	−0.0006
		(0.0073)	(0.0071)
VOL	?	−0.1926	−0.2094
		(0.3432)	(0.3310)
POSWORDS	+	0.0290***	0.0236***
		(0.0072)	(0.0071)
NEGWORDS	−	−0.0453***	−0.0399***
		(0.0086)	(0.0085)
N		1,647	1,647
Adjusted R^2		7.64%	10.65%

Table IV

Estimation of the Association between Affect and Analyst One-Quarter-Ahead Forecast Revisions

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and analyst earnings forecast revisions (*FREV*) and recommendation revisions (*RECREV*). See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	*FREV* (1)	*RECREV* (2)
Intercept	?	0.0022*	−0.2723***
		(0.0013)	(0.0568)
$PAFF^{HS}$	+	−0.0047	0.4200**
		(0.0076)	(0.2266)
$NAFF^{HS}$	−	0.0019	−0.0121
		(0.0036)	(0.1025)
$PAFF^{LS}$	+	0.0014	0.1056
		(0.0051)	(0.1861)
$NAFF^{LS}$	−	−0.0002	0.0862
		(0.0018)	(0.0732)
UE_t	+	0.1273***	0.9383**
		(0.0256)	(0.4967)
CAR(0,1)	+	0.0111***	0.5224***
		(0.0021)	(0.0852)
LNMVE	?	−0.0001	−0.0016
		(0.0000)	(0.0035)
MOM	?	0.0031***	0.0715***
		(0.0007)	(0.0223)
BM	?	−0.0032***	−0.0014
		(0.0007)	(0.0212)
VOL	?	−0.0834***	0.2570
		(0.0247)	(0.6596)
LAGREC	−		−0.1073***
			(0.0126)
POSWORDS	+	0.0006	−0.0148
		(0.0005)	(0.0218)
NEGWORDS	−	−0.0005	−0.0194
		(0.0007)	(0.0271)
N		1,647	1,647
Adjusted R^2		25.72%	12.06%

Estimations of the Association between Affect and Future Earnings News

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and future earnings surprises (*UE*). Superscripts *HS* and *LS* represent high- and low-scrutiny partitions. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	UE_{t+1} (1)	UE_{t+2} (2)	$UE_{t+1,t+2}$ (3)
Intercept	?	0.0081	0.0157**	0.0224**
		(0.0045)	(0.0068)	(0.0106)
$PAFF^{HS}$	+	0.0234	0.0690**	0.0753*
		(0.0231)	(0.0338)	(0.0510)
$NAFF^{HS}$	−	−0.0027	−0.0307**	−0.0431*
		(0.0129)	(0.0185)	(0.0294)
$PAFF^{LS}$	+	0.0107	0.0207	0.0215
		(0.0151)	(0.0223)	(0.0291)
$NAFF^{LS}$	−	−0.0056	−0.0121	−0.0192*
		(0.0053)	(0.0114)	(0.0118)
UE_t	+	0.4767***	0.4424***	0.7472***
		(0.1104)	(0.1531)	(0.2429)
FREV	+	0.4298**	0.2856	0.5811*
		(0.1878)	(0.2663)	(0.3921)
FDISP	−	−0.0357**	−0.0447*	−0.0824*
		(0.0163)	(0.0295)	(0.0521)
LNMVE	?	−0.0003	−0.0006	−0.0007
		(0.0003)	(0.0005)	(0.0008)
MOM	?	0.0078***	0.0092***	0.0123***
		(0.0028)	(0.0031)	(0.0047)
BM	?	−0.0074***	−0.0144***	−0.0211***
		(0.0023)	(0.0039)	(0.0061)
VOL	?	−0.1355	−0.3020**	−0.3334*
		(0.0977)	(0.1465)	(0.1949)
POSWORDS	+	−0.0020	−0.0022	−0.0028
		(0.0015)	(0.0018)	(0.0026)
NEGWORDS	−	0.0013	−0.0030	−0.0022
		(0.0021)	(0.0038)	(0.0053)
N		1,647	1,146	1,146
Adjusted R^2		28.80%	17.12%	20.25%

Table VI

OLS Estimations of the Association between Affect Variables and Future Stock Returns

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and future stock returns (*CAR*(2,180)). Superscripts *HS* and *LS* represent high- and low-scrutiny partitions. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	
Intercept	?	0.1900
		(0.1123)
$PAFF^{HS}$	+	0.5280
		(0.5649)
$NAFF^{HS}$	−	−0.6463***
		(0.2580)
$PAFF^{LS}$	+	−0.2258
		(0.4790)
$NAFF^{LS}$	−	0.0104
		(0.1697)
UE_t	+	2.2959
		(1.8075)
LNMVE	?	−0.0127
		(0.0085)
MOM	?	0.2404***
		(0.0553)
BM	?	−0.0974**
		(0.0487)
VOL	?	−3.7099*
		(1.9100)
POSWORDS	+	0.0226
		(0.0397)
NEGWORDS	−	0.0161
		(0.0476)
N		1,647
Adjusted R^2		5.39%

Variables Definitions

Appendix: Variable Definitions

Variable Name	Definition
PAFF	Positive affect measured as the percentage of spoken audio by management during the conference call with Emotion Level scores above the critical value of 110 as measured by LVA.
NAFF	Negative affect measured as the percentage of spoken audio by management during the conference call with Cognition Level scores above the critical value of 120 as measured by LVA.

$PAFF^{HS}$ ($NAFF^{HS}$)	Represents positive (negative) affect under a high-scrutiny partition. That is, $PAFF^{HS}$ ($NAFF^{HS}$) is set to *PAFF* (*NAFF*) when UE_t is less than zero, and zero otherwise.
$PAFF^{LS}$ ($NAFF^{LS}$)	Represents positive (negative) affect under a low-scrutiny partition. That is, $PAFF^{LS}$ ($NAFF^{LS}$) is set to *PAFF* (*NAFF*) when UE_t is greater than or equal to zero, and zero otherwise.

Financial Study 3: Voice Pitch & the CEO Labor Market Success

- A deep voice is evolutionarily advantageous for males, but does it confer benefit in competition for leadership positions?
- Data:

1. A list of male CEOs from the Standard & Poor's 1500 stock index
2. CEO speech corpus from earnings conference calls archived in the Thomson Reuters StreetEvents database.
3. 792 unique CEOs, with 216, 551 and 25 of the cross sectional observations having fiscal years ending in 2006, 2007 and 2008, respectively.
4. Male S&P 1500 CEOs exhibit little ethnic heterogeneity, as over 90% of the CEOs are Caucasian.

Voice Pitch

- Vocal depth is captured by measuring each CEO’s vocal fundamental frequency (F0), which listeners perceive as voice pitch.
- To estimate each CEO’s fundamental frequency, we stream approximately the first 20s of CEO speech and encode in mono directly onto a computer hard disk.
- The opening conference call presentation remarks are less prone to purposeful changes in voice pitch
- Each .wav file is then digitally analyzed using PRAAT acoustics software version 5.2.05

Findings

Regression analysis of the association between voice pitch and CEO labor market success (N = 792).

Dependent Variable:	Firm Size (I)	Firm Size (II)	Total Compensation (III)	Total Compensation (IV)	Total Compensation (V)	Tenure (VI)	Tenure (VII)
Intercept	14.603***	9.972***	10.772***	5.432***	4.468***	9.941***	0.046
	(2.180)	(2.961)	(1.152)	(0.885)	(1.269)	(1.127)	(1.584)
Voice Pitch	−1.367***	−1.199***	−0.523**	−0.023[a]	0.019	−0.520***	−0.457***
	(0.450)	(0.450)	(0.238)	(0.177)	(0.178)	(0.233)	(0.228)
Firm Size				0.366***	0.362***		−0.070***
				(0.015)	(0.015)		(0.020)
Age[b]		0.993***			0.164		2.545***
		(0.467)			(0.215)		(0.275)
Education		0.196*			0.173***		−0.056
		(0.117)			(0.049)		(0.064)
Formant Position		−0.428**			0.056		0.005
		(0.217)			(0.083)		(0.118)
Adjusted R^2	0.0113	0.0214	0.0048	0.4378	0.4456	0.0043	0.1060
Overall F	9.21***	5.67***	4.85**	302.18***	128.05***	4.97**	19.31***

All dependent variables are log-transformed, as are all continuous independent variables. The log transformation for P_f is equal to $\ln(2 + P_f)$. We add a constant of 2 to P_f because the minimum value is −1.6411, which cannot be log transformed. This table displays unstandardized regression coefficients. Heteroscedasticity robust standard errors are shown in parenthesis. Significantly different from zero at the *10% level, **5% level, and ***1% level, two-tailed.

Coefficients are statistically significant! (Economic significance could be small)

Agenda

1. Image Mining for Business Analytics
    1. Face images
    2. Satellite images
    3. Manufacturing images
    4. Medical images for prescription (there are many applications but are omitted in this module.)

2. Audio Mining for Business Analytics

3. Video Mining for Business Analytics

Video Analytics

1. 2014. "The Power of Fear: Facial Emotion Analysis of CEOs to Forecast Firm Performance“ (2nd tier conference paper)
2. 2017. "Firm Performance in the Face of Fear: How CEO Moods Affect Firm Performance“ (2nd tier journal paper)
3. 2017. "Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows“ (top 3 accounting journal paper)

4. MIT Video Enhancement
5. Amazon Unmanned Shop

Study 1

1. 2014. "The Power of Fear: Facial Emotion Analysis of CEOs to Forecast Firm Performance"
2. One caveat is this is not from a top CS conference but seems interesting.

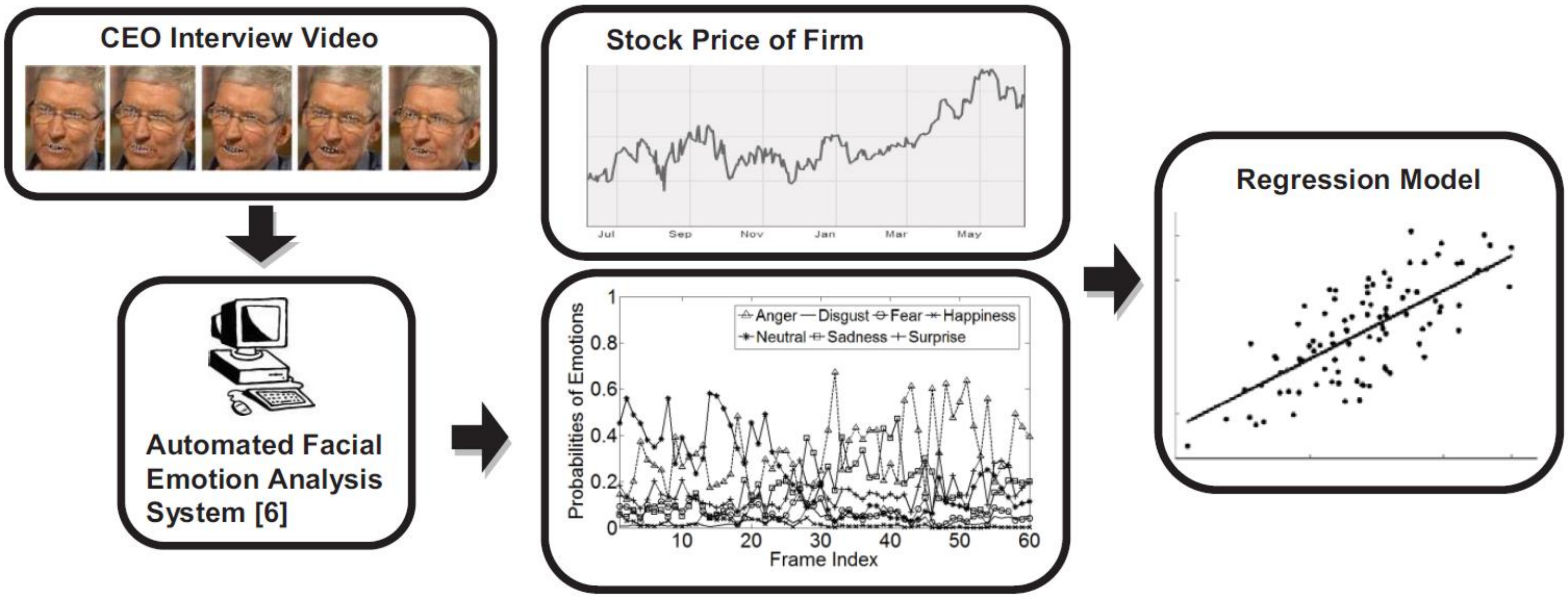


Figure 1. Framework to forecast the market value of firm based on CEO's measured facial emotions.

Study 1

- The interview videos of CEOs of Fortune 500 companies from 2006 through 2012 are obtained from YouTube and CNBC.
- 531 videos are selected for the experiment.
  1. Videos of CEO long enough, front face, not including other people.
  2. Companies that can be matched to Eventus for Event Study
- Baum-Welch algorithm is used in this study.
- The widely referenced Cohn-Kanade database is used to recognize the seven basic facial emotions including Neutral. The average accuracy is 86.87%.

Study 1

- 6 Facial Emotions are identified: Anger, Disgust, Fear, Happiness, Sadness, and Surprise.

Table 2. Library of Emotion Types and States of Face Regions [6].

Emotion Type	**States of Face Regions**
Anger	*Eyebrows fall, eyes close, mouth close, and lips corners pucker*
Disgust	*Eyebrows fall, eyes close, mouth close and lips corners pull*
Sadness	*Eyebrows raise, eyes close, mouth close and lips corners down*
Happiness	*Eyebrows neutral, eyes neutral, mouth open, and lips corners up*
Surprise	*Eyebrows raise, eyes open, mouth open, and lips corners neutral*
Fear	*Eyebrows raise, eyes open, mouth open, and lips corners pull*

Study 1

$$\sum_{i=-1}^{n} ar_i = \alpha + \beta_1 Anger + \beta_2 Disgust + \beta_3 Fear \qquad (4)$$
$$+ \beta_4 Happiness + \beta_5 Sadness + \beta_6 Surprise + \varepsilon$$

Table 3. The effect of CEO's emotion types on the firm's CARs before noise suppression.

	(-1,1)	(-1,5)	(-1,10)	(-1,30)	(-1,60)	(-1,90)
A	0.0043	0.0041	0.007	0.0104	0.0118	0.0232
D	0.0015	0.0049	0.0047	0.0153	0.0249	0.0214
F	0.01**	0.01*	0.01**	0.02**	0.03**	0.0256
H	-0.001	-0.006	-0.013	-0.020	-0.032	-0.037
SA	-0.001	-0.003	0.0020	0.0101	0.0099	0.0183
SU	0.0014	0.0020	0.0050	0.0031	0.0024	0.0009

Only fear has significant and "**positive**" impact? Why?

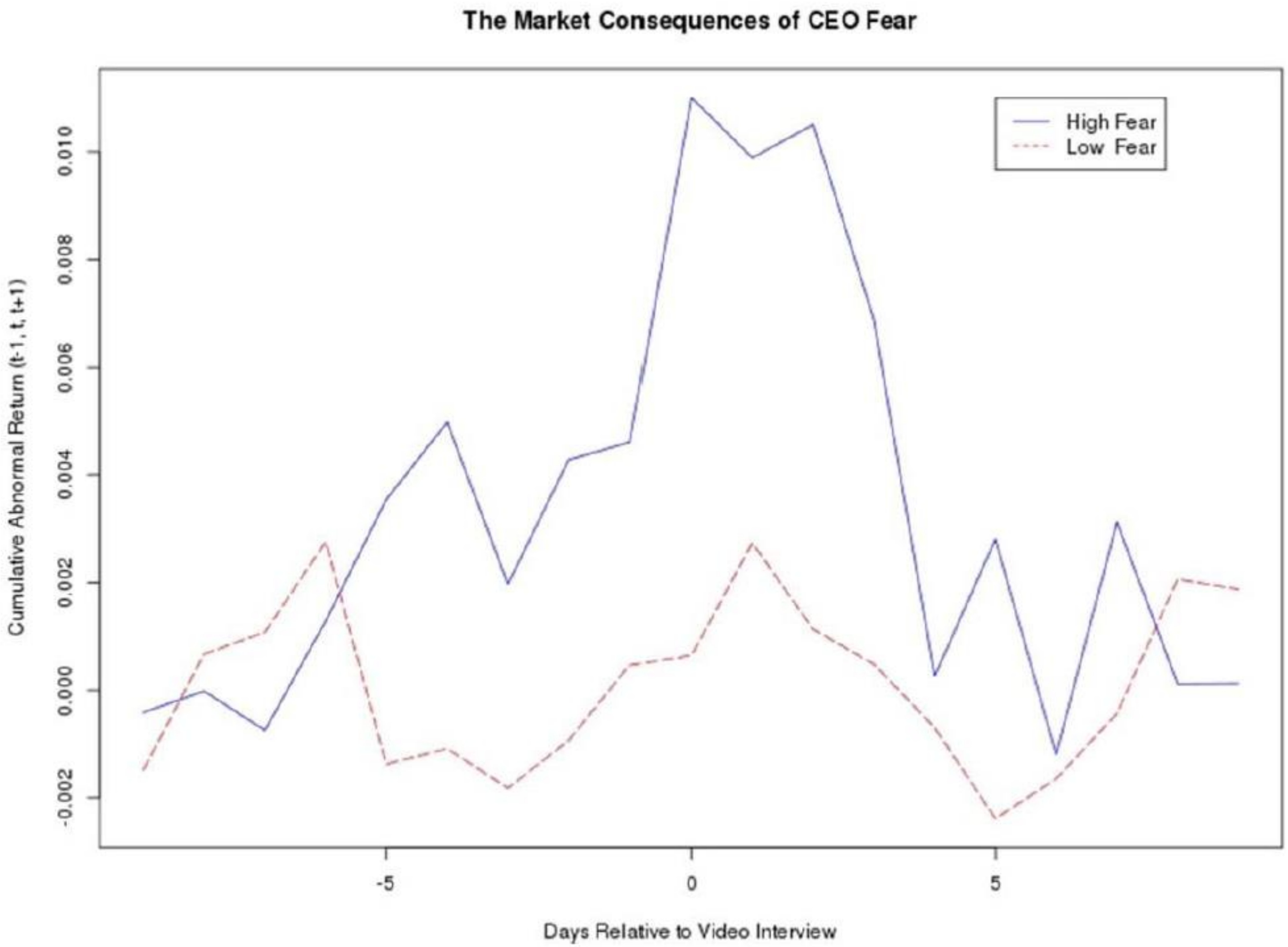
Video Analytics Study 2

- 2017. "Firm Performance in the Face of Fear: How CEO Moods Affect Firm Performance‟ (not a top journal)
- Some authors are the same as Study 1.

Main findings:

- Anger or disgust motivates a CEO to work harder to improve his or her situation; thus firm profitability improves in the subsequent quarter.
- Happy CEOs are less likely to work on hard or unpleasant tasks; thus profitability decreases in the subsequent quarter.
- In the short term, fear explains the firm’s announcement period market performance. However, fear is transient and performance improvement is short term.

Study 2



The Market Consequences of CEO Fear
High Fear
Low Fear
Cumulative Abnormal Return (t-1, t, t+1)
-0.002
0.000
0.002
0.004
0.006
0.008
0.010
-5
0
5
Days Relative to Video Interview

Study 2

Table 7. Profitability and CEO emotive states.

	PM	ROE	Sales growth (%)	ROA
Emotive state				
Anger	−0.004 (0.006)	0.011 (0.012)	0.011 (0.003)***	0.000 (0.000)
Disgust	0.093 (0.037)**	0.066 (0.076)	0.003 (0.022)	0.006 (0.003)**
Fear	−0.016 (0.027)	0.002 (0.056)	−0.010 (0.016)	−0.003 (0.002)
Happiness	−0.122 (0.048)**	−0.244 (0.094)**	−0.026 (0.029)	−0.012 (0.003)***
Sadness	0.028 (0.028)	−0.083 (0.057)	0.003 (0.017)	0.001 (0.002)
Surprise	−0.002 (0.014)	−0.005 (0.028)	0.003 (0.008)	0.000 (0.001)
Information perspective				
Good news	−0.037 (0.054)	0.020 (0.107)	−0.001 (0.033)	0.000 (0.004)
Bad news	−0.038 (0.076)	−0.044 (0.151)	0.018 (0.046)	−0.005 (0.005)
Length of interview	−0.014 (0.032)	0.113 (0.065)*	−0.016 (0.019)	−0.002 (0.002)
Interview topic				
Economy	0.037 (0.111)	0.081 (0.225)	−0.057 (0.067)	0.008 (0.008)
Industry and/or competition	0.037 (0.102)	−0.166 (0.208)	−0.061 (0.061)	0.006 (0.007)
Performance, growth and/or expansion	−0.019 (0.101)	0.088 (0.206)	−0.086 (0.061)	0.005 (0.007)
Company guidance	0.037 (0.098)	0.021 (0.199)	−0.093 (0.059)	0.004 (0.007)
Mergers, acquisitions, and/or divestitures	0.050 (0.119)	0.068 (0.243)	−0.076 (0.072)	0.002 (0.009)
Firm characteristic				
Market to book ratio	−0.002 (0.004)	0.056 (0.008)***	−0.000 (0.002)	0.000 (0.000)
Market capitalization (natural log)	0.043 (0.020)**	0.029 (0.038)	−0.002 (0.012)	0.005 (0.001)***
SD of returns	−4.636 (2.222)**	0.362 (3.838)	−0.316 (1.336)	−0.368 (0.135)***
Industry fixed effects	yes	yes	yes	yes
R^2	10.37%	17.33%	7.66%	18.73%
F stat	1.46	2.81	1.04	3.09

This table reports the relationship between the emotive states and firm profitability. We measure firm profitability 4 ways: profit margin (PM), return on equity (ROE), sales growth, and return on assets (ROA). Profitability is sampled in the first quarter after the quarter in which the interview occurs. We tabulate the results of the following multivariate model:

Video Analytics Study 3

- 2017. "Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows‟
- This paper is published in top 3 accounting journal.
- However, this paper does not use AI to code videos. It is uses human being to code videos.
- Results are interesting.
- 30-second content-filtered video clips of initial public offering (IPO) roadshow presentations.
- Viewers' overall perceptions of a CEO, is positively associated with pricing at all stages of the IPO (proposed price, offer price, and end of first day of trading).

Study 3: Data

- This paper uses video capture software to obtain IPO roadshows from RetailRoadshow.com, a website that posts roadshow presentation videos for public offerings.
- All U.S. firms that completed an original IPO on NASDAQ or NYSE in the United States from March 24, 2011 to December 31, 2013.
- 224 IPO cases after standard data pre-processing.
- Only 30 seconds of presentation for the raters to rate 3 aspects from 1 to 7: competence, trustworthiness, and attractiveness.

Study 3: Findings

Panel B: Perception Components and Firm Value

VARIABLES	Prediction	L(MVE_Proposed) (1)	(2)	(3)	L(MVE_Offer) (4)	(5)	(6)	L(MVE_Final) (7)	(8)	(9)
Competent	+	0.2457**			0.2461**			0.3174**		
		(2.37)			(2.21)			(2.42)		
Trustworthy	+		0.1305*			0.1252			0.1706	
			(1.84)			(1.44)			(1.45)	
Attractive	+			0.1657***			0.2257***			0.2764***
				(3.22)			(3.65)			(4.51)
Remaining Controls		Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry Fixed Effects		Included	Included	Included	Included	Included	Included	Included	Included	Included
Time Fixed Effects		Included	Included	Included	Included	Included	Included	Included	Included	Included
Observations		224	224	224	224	224	224	224	224	224
Adjusted R-squared		0.624	0.619	0.625	0.617	0.613	0.624	0.575	0.569	0.583

Table 9. Perception and Post-IPO performance

Panel A: Perception and Subsequent Stock Returns

VARIABLES	$BHAR_{2Y}$ (1)	(2)	$BHAR_{Max}$ (3)	(4)
Perception	0.1556	0.0701	0.0311	0.1614
	(1.12)	(0.32)	(0.16)	(0.64)
Remaining Controls	Excluded	Included	Excluded	Included
Industry Fixed Effects	Included	Included	Included	Included
Time Fixed Effects	Included	Included	Included	Included
Observations	224	224	224	224
R-squared	0.033	0.104	0.090	0.170

But perceptions of 3 aspects are not correlate with the buy and hold stock return

Study 3: Findings

Table 8. *Perception*: Measurement using Videos or Pictures

Panel A: Multiple Measures of Aggregate Perception and IPO Outcomes

	L(MVE_Proposed)	*L(MVE_Offer)*	*L(MVE_Final)*	*Underwriter*
VARIABLES	(1)	(2)	(3)	(4)
Perception	0.1843*	0.2477**	0.3046**	0.3046**
	(1.77)	(2.11)	(2.37)	(2.37)
Perception_Pic	0.2166*	0.1749	0.2291	0.2291
	(1.79)	(1.34)	(1.53)	(1.53)
Remaining Controls	Included	Included	Included	Included
Industry Fixed Effects	Included	Included	Included	Included
Time Fixed Effects	Included	Included	Included	Included
Observations	224	224	224	224
R-squared	0.630	0.623	0.584	0.340

Notes: Table 8, Panel A presents the results from an OLS regression of firms' IPO outcomes on various CEO, firm, and offering characteristics. *L(MVE_Proposed)* is the natural log of the firm's market value of common equity calculated using the proposed offer price. *L(MVE_Offer)* is the natural log of the firm's market value of common equity calculated using the final offer price. *L(MVE_Final)* is the natural log of the firm's market value of common equity calculated at the end of its first day trading on the secondary market. *Underwriter* is the average Carter-Manaster IPO ranking for the firm's lead underwriters. *Perception* is the average of *Competent*, *Trustworthy*, and *Attractive* (based on surveys using videos of the CEO). *Perception_Pic* is the CEO-specific average of Competent, Trustworthy, and Attractive ratings obtained from a MTurk survey that showed survey participants a picture of the CEO. See Appendix B for all other variable definitions. Standard errors are double-clustered by Fama-French 48 industry and year-week. T-statistics are provided in parentheses below the coefficients. *** designates two-tailed statistical significance at 1%, ** at 5%, and * at 10%.

Study 3: Findings

Table 9. *Perception* and Post-IPO performance, continued

Panel C: Perception and Subsequent Operating Performance

	ROA_{2Y}		ROA_{Max}	
VARIABLES	(1)	(2)	(3)	(4)
Perception	-0.0014	-0.0019	-0.0009	0.0013
	(-0.32)	(-0.25)	(-0.18)	(0.15)
Remaining Controls	Excluded	Included	Excluded	Included
Industry Fixed Effects	Included	Included	Included	Included
Time Fixed Effects	Included	Included	Included	Included
Observations	224	224	224	224
R-squared	0.368	0.486	0.316	0.463

Notes: Table 9, Panel C presents the results from an OLS regression of post-IPO operating performance on various CEO, firm, and offering characteristics. ROA_{2Y} is the firm's average quarterly net income divided by average quarterly total assets subsequent to its IPO. ROA_{Max} is defined similarly but includes data for all quarters subsequent to each firm's IPO through December 31, 2015. *Perception* is the average of *Competent*, *Trustworthy*, and *Attractive*. See Appendix B for all other variable definitions. Standard errors are double-clustered by Fama-French 48 industry and year-week. T-statistics are provided in parentheses below the coefficients. *** designates two-tailed statistical significance at 1%, ** at 5%, and * at 10%.

Perceptions also do not correlate with post-IPO firm performance measured by ROA.

Video Analytics Study 4

- MIT Video Magnification Technology
- http://people.csail.mit.edu/mrub/vidmag/
- Not much impact so far, probably not that accurate but is interesting to know and maybe one day it will become a disruptive tech.
- This kind of software can magnify small/tiny motions, especially pulse, leading to better identifying the emotion of a person in a video.
- Music affecting leaves to "shake". Software can recover music by seeing those leaves shaking… see the following video.
- https://www.ted.com/talks/abe_davis_new_video_technology_that_reveals_an_object_s_hidden_properties?language=en

Video Analytics Example 5

- Unmanned Shop seems like a very promising future of retailers. It is also important for SG because we have limited space and manpower.
- What is Amazon Go? https://www.amazon.com/b?ie=UTF8&node=16008589011
- "**Inside Amazon Go, a Store of the Future**" --- NY Times
- "**China is both ahead of and behind Amazon in cashier-less stores**" https://qz.com/1185081/amazon-go-china-is-both-ahead-of-and-behind-amazon-in-cashier-less-stores/

Other Video Analytics Examples

- China's CCTV surveillance network took just 7 minutes to capture BBC reporter https://techcrunch.com/2017/12/13/china-cctv-bbc-reporter/
- Retailers use various technologies to track the path and traffic flow of customers within a store or within a mall (indoor location-based services). http://www.behavioranalyticsretail.com/7-technologies-to-track-people/
- IBM Intelligent Video Analytics is another new product.