

Close Listening: Vocal Cues for Predicting Future Yields

SPEAKER



DORON REICHMANN
Ruhr-University Bochum

MODERATOR



PATRICK KISS
IR Club

PRÄSENTIERT VON  **cometis** 

Wir danken unseren Hauptsponsoren



Verbale und Nonverbale Signale in Finanzmärkten

IR Club

September 2023

Der Markt hört ganz genau hin

Wort für Wort...

TECHNOLOGY

T-Mobile CEO John Legere Unleashed

Count the F-bombs. T-Mobile chief executive **John Legere** has a way with words and occasionally profane ones, but never boring. The CEO of the nation's fourth-largest wireless carrier Monday at the inaugural **Code/Mobile** conference reported record subscriber gains but a surprising and before a conference call with analysts on T

MarketWatch "You are a disaster": Cleveland-Cliffs CEO berates analyst on earnings call

"You are an embarrassment to your parents," CEO says
Shares of Cleveland-Cliffs Inc. fell more than 6% Friday after Chief Executive Lourenco Goncalves slammed analysts during the mining company's earnings call earlier in the day, accusing them of failing to understand numbers and targeting a Goldman Sachs analyst for most of his angry remarks.

© CBS NEWS

Worst corporate euphemism ever? GM's "unallocated" factories a contender

up with a new way to describe being fired: *You're unallocated.*

base production at five plants in the U.S. and Canada and sack up to 14,000 factories and facilities "will

in straight, GM said in a news release.

invocative, combative



¹Footnotes

Wie die Wortwahl Märkte beeinflusst

Auswahl sprachlicher Dimensionen und Kapitalmarkteffekte

Ton/Sentiment

Beschreibt wie positive oder negative eine Nachricht ist.

Analysemethode:

- Dictionary-based approach
- Machine learning

Kapitalmarkteffekte

- Der Ton qualitativer Berichterstattung wird inkrementell zu quantitativer Offenlegung eingepreist.
- Weiche Informationen sind schwieriger zu verifizieren und können Investoren fehlleiten (Tone Management).

Komplexität

Beschreibt wie komplex/schwer verständlich Texte sind

Analysemethode:

- Gunning-Fog-Index
- Kombination aus (i) durchschnittlicher Satzlänge und (ii) Anzahl mehrsilbiger Worte im Text

- Übermäßig komplizierte Berichterstattung wird unvollständig vom Kapitalmarkt aufgenommen.
- Manager neigen dazu strategisch Sprache zu verkomplizieren, um negative Neuigkeiten zu verschleiern.

Scripting

Approximiert inwieweit Antworten in Q&As geskriptet sind

Analysemethode:

- Analyse von Function-Words (*there, just, very, not, no,...*)
- Approximierung der Ähnlichkeit von Funktion-Words in der Q&A und der geskripteten Präsentation

- Kapitalmärkte reagieren negativ auf mangelnde Spontanität von Managern in Q&A sessions.
- Scripting ist prominenter wenn Manager schlechte Ergebnisse erzielen.

Mit Machine Learning Sprache verstehen

Wie können Maschinen Sprache verstehen?

Herausforderung

Text repräsentiert natürliche Sprache, Algorithmen können jedoch nur mit numerischen Werten arbeiten.

Texte müssen also zunächst in numerische Werte umgewandelt werden

S1: "We **accomplished** our **goal** to gain profits"

S2: "We **achieved** our **objective** to gain sales"

S3: "My **dog** really gained some **weight**!"

Guess the odd one out!

Bag-Of-Words Approach

$T(S1, S2)$

= [accomplish, goal, gain, profit, achieve, objective, sales]

$TF(S1)=[1, 1, 1, 1, 0, 0, 0]$; $TF(S2)=[0, 0, 1, 0, 1, 1, 1]$

$$\frac{TF(S1) \cdot TF(S2)}{\|TF(S1)\| \times \|TF(S2)\|} = \frac{(1*0+1*0+1*1+1*0+0*1+0*1+0*1)}{(\sqrt{1^2+1^2+1^2+1^2+0^2+0^2+0^2}) \times (\sqrt{0^2+0^2+1^2+0^2+1^2+1^2+1^2})} = \frac{1}{4 \times 4} = \mathbf{0.062}$$

Demnach sind S1 und S2 so unterschiedlich wie S1 und S3.

Lösungsansatz

Moderne Textanalysemethoden wandeln Text in numerische Repräsentationen um, die sich aus dem Kontext des Textes ergeben.

Die Idee basiert auf einer alten Regel der Linguistik (Harris, 1954): *Worte mit ähnlichen Wortnachbarn haben eine ähnliche Bedeutung.*

Numerische Repräsentationen werden anhand von Wortnachbarn erlernt.

Machine Learning Approach

Die Modelle *lernen* numerische Repräsentationen für Worte, Sätze oder gesamte Dokumente.

Word	d-dimensional embeddings
accomplish	(0.213, 0.321, 0.032, 0.123 [...])
goal	(0.543, 0.124, 0.543, 0.123 [...])
gain	(0.976, 0.324, 0.844, 0.245 [...])
profit	(0.213, 0.787, 0.987, 0.314 [...])
Average	(0.486, 0.389, 0.602, 0.201 [...])

Word Embeddings aus MD&A Files

Ergebnis aus über 40.000 MD&A Sections aus 10-Ks

Code Example

```
import os
import logging # Setting up the Loggings to monitor gensim
import multiprocessing
from gensim.models import Word2Vec

# set path to data
path_to_files = r'D:\Datensammlung\SEC 10-K
files\lm_full_databse_10Ks\phrased_files'

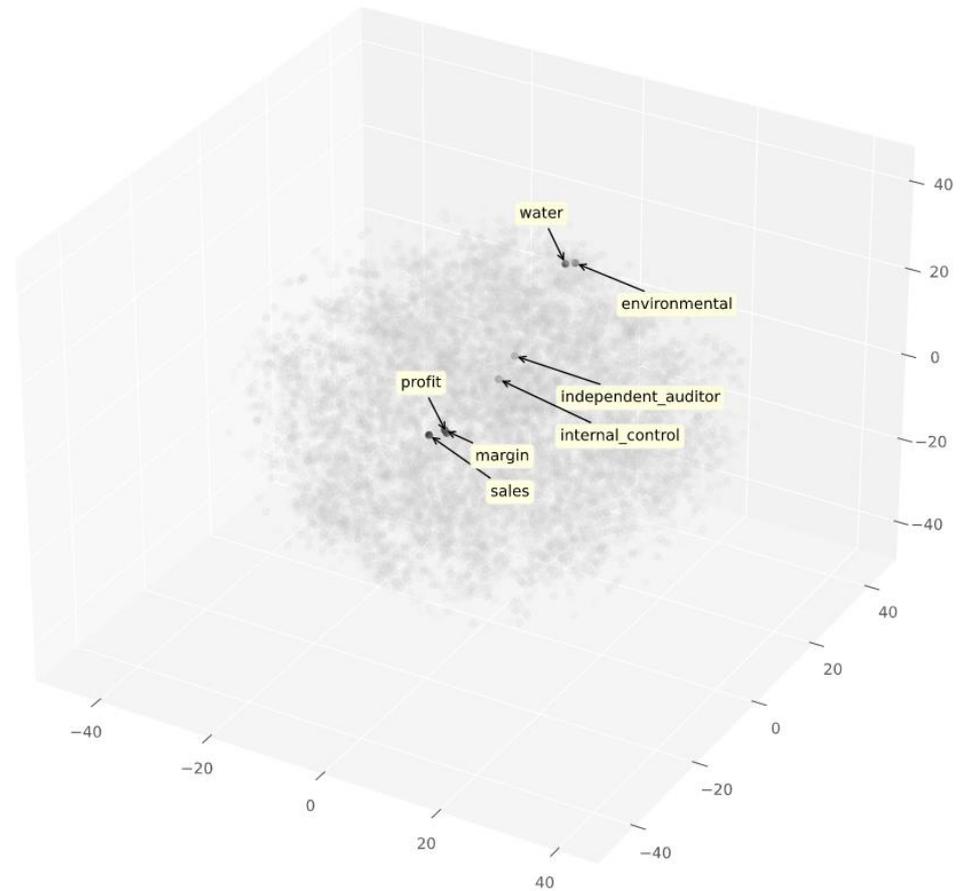
# use available cores
cores = multiprocessing.cpu_count() # Count the number of cores in a computer

# set logging to track progress
logging.basicConfig(format="%(levelname)s - %(asctime)s: %(message)s",
                    datefmt='%H:%M:%S', level=logging.INFO)

# collect data is one huge list
sec_files = []
for fname in os.listdir(path_to_files):
    sec_files.append(
        open(os.path.join(path_to_files, fname), encoding='utf-8').read()
    )

# set embedding dimension
embedding_dim = 300

# train model
w2v_model = Word2Vec(window=5,
                      size=embedding_dim,
                      min_count=50,
                      sg=0,
                      negative=5,
                      iter=20,
                      workers=cores-1)
```



Textanalyse und Machine Learning

Ein Beispiel

Studie

Out-of-Sample Predictability of Firm-Specific Stock Price Crashes: A Machine Learning Approach

(Kaya, Reichmann, Reichmann, 2023), *R&R at the Journal of Business Finance and Accounting*

Forschungsfrage:

Wir beobachten Regelmäßig firmen-spezifische, abrupte Aktiencrash die nicht durch allgemeine Marktbewegungen getrieben werden. Die Literatur besagt, dass Aktiencrashes durch das „horten“ negativer Informationen entstehen. In dieser Studie untersuchen wir, ob Informationen in qualitativer Berichterstattung helfen können, Aktiencrashes besser vorherzusagen.

Vorgehen



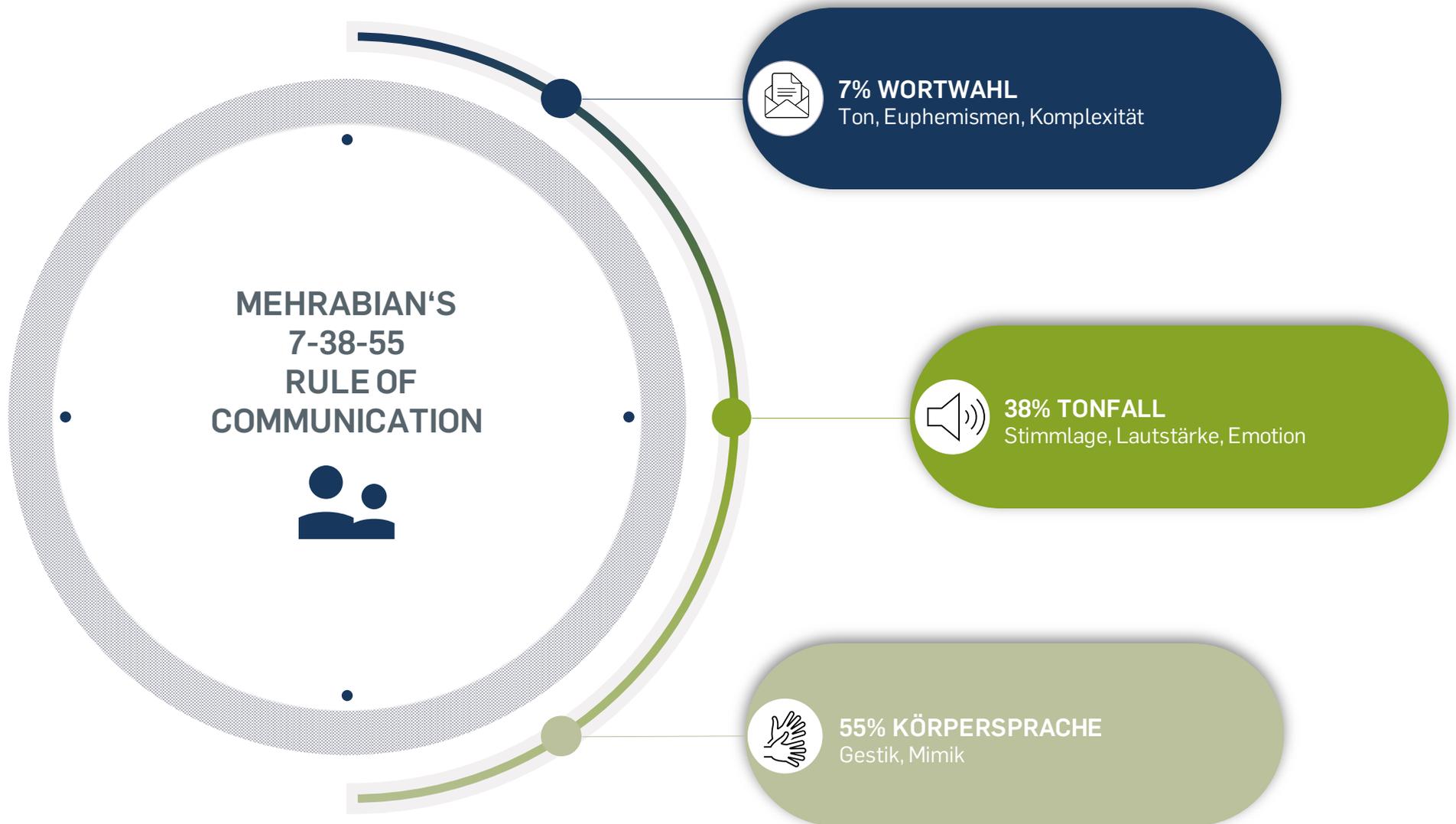
firm	year	data	crash	model_predict
AAPL	2016	item -cardinal- management_discussion_and_ana...	0	(train)
ALV.DE	2016	item -cardinal- management_discussion_and_ana...	1	(train)
AAPL	2017	item -cardinal- management_discussion_and_ana...	0	(train)
ALV.DE	2017	tem -cardinal- management_discussion_and_ana...	0	(train)
AAPL	2018	item -cardinal- management_discussion_and_ana...	0	(train)
ALV.DE	2018	item -cardinal- management_discussion_and_ana...	1	(train)
AAPL	2019	item -cardinal- management_discussion_and_ana...	?	0.1523
ALV.DE	2019	item -cardinal- management_discussion_and_ana...	?	0.3017

Ergebnis

- Moderne Machine Learning Verfahren können helfen Firmen-spezifische Aktiencrashes besser vorherzusagen
- Textdaten performen überraschend gut und haben einen vergleichbaren Informationsgehalt wie quantitative Bilanzdaten.
- Die Kombination aus Text und Finanzdaten führt zu den besten Prognoseergebnissen.

Die 7-38-55 Regel der Kommunikation

Nonverbale Kommunikation macht einen Großteil der Botschaft aus



Informationsquelle Tonfall – Evidenz aus der Psychologie

Was die Stimme verrät

Evolutionäres Kommunikations-Tool

Wenn immer wir sprechen, vermitteln wir nicht nur kontextbezogene Informationen, sondern auch Emotionen die in der Klangstruktur unserer Stimme enthalten sind.

Emotionale Sprache ist tief im menschlichen Sozialleben verwurzelt.

Forschung auf dem Gebiet der Psychologie legt nahe, dass wir über 24 Emotionen kommunizieren können, sogar ohne Worte zu formen.

Unkontrollierbar

Gemessen an der Anzahl an beweglichen Härchen und Muskeln gilt der Sprachprozess als eine der komplexesten Mechanismen des menschlichen Körpers.

Es ist weitgehend anerkannt, dass nonverbale Signale und insbesondere stimmliche Signale sehr schwierig zu kontrollieren sind.

Dementsprechend gibt unsere Stimme Informationen preis, die wir nicht preisgeben wollen.

Situationsbezogen

Unsere Emotionen sind sehr situationsabhängig und werden meist durch einen externen Auslöser determiniert.

Forschung aus der Psychologie zufolge werden die erleben wir die meisten Emotionen in zwischenmenschlichen Kommunikationen.

Dementsprechend ist es wichtig zu beachten in welchem Rahmen Emotionen analysiert werden.

Messung von Emotionen

Es gibt bis heute keinen allgemeinen Konsens wie allgemein anerkannte phonetische Merkmale (z.B., Pitch, Lautstärke) kombiniert werden um spezifische Emotionen abzubilden.

Es bleibt zudem unklar, wie verschiedene Emotionen wir über unsere Stimmen kommunizieren können.

Vereinzelte Software Anbieter bieten Möglichkeiten spezifische Emotionen zu kategorisieren.

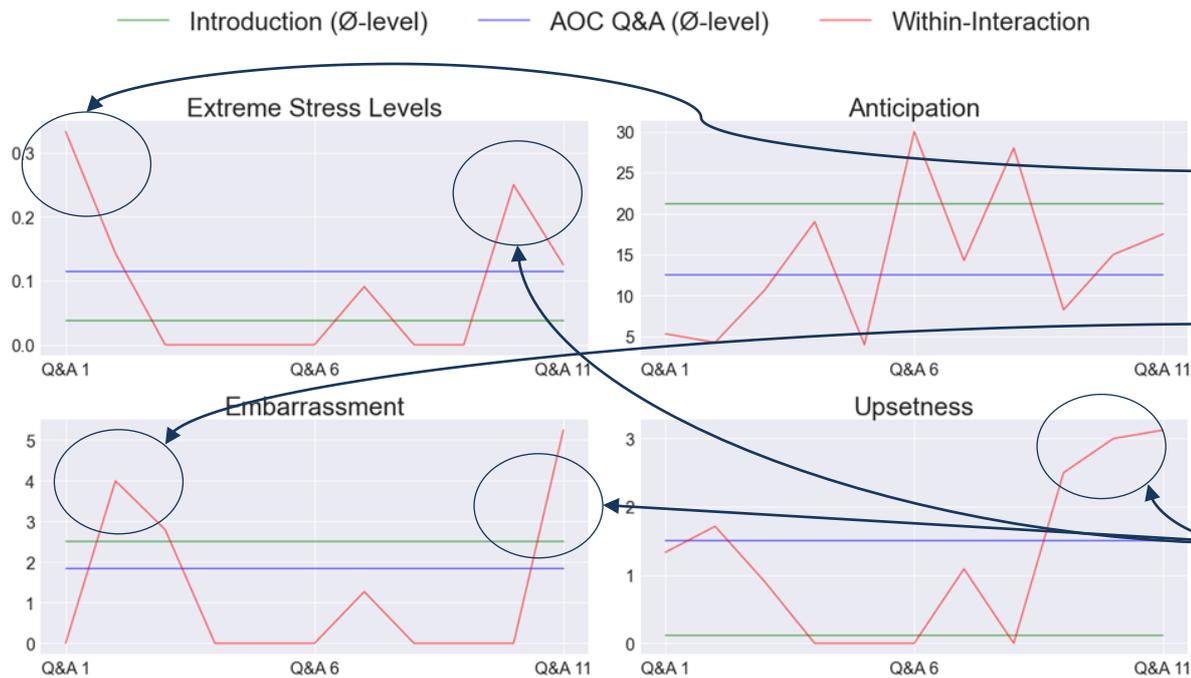
Emotional Leakage in Q&A Settings

US House Financial Services Committee on 23rd October 2019

The Guardian

Ocasio-Cortez stumps Zuckerberg with questions on far right and Cambridge Analytica

Democratic lawmaker challenges Facebook CEO during hearing over Libra cryptocurrency



Zuckerbergs emotionale Reaktion auf folgende Fragen von Alexandria Ocasio-Cortez:

What year and month did you first become aware of Cambridge Analytica?

Could I advertise predominantly black audience the wrong election date?

One more question: in your ongoing dinner parties with far-right figures, some of whom advanced the theory that white supremacy is a hoax, did you discuss social media bias against conservatives, and do you think there is a bias?

¹ Die Stimmanalyse Software wurde von Nemesysco Ltd, einer israelischen High-Tech Firma bereitgestellt.

Forschungsprojekt Voice Analysis

Listen Closely

Studie

Listen Closely: Using Vocal Cues to Predict Future Earnings

(Ewertz, Knickrehm, Nienhaus, Reichmann, 2023), Working Paper

Forschungsfrage:

Der prognostizierte Unternehmensgewinn ist der Ausgangspunkt für Unternehmensbewertungsmodelle. Dennoch bleibt die Prognose von Unternehmensgewinnen eine große Herausforderung in der Wirtschaftsforschung und Praxis. Kann eine Stimmanalyse von CEOs in Earnings Conference Calls helfen Unternehmensgewinne vorherzusagen?

Daten

- Audio und Transkripte von Earnings Conference Calls (Refinitiv Eikon).
- 8436 Q4 Earnings Calls von über 2000 US Unternehmen.
- Periode: 2015-2020 (um den Unternehmensgewinn in 2016-2021 zu prognostizieren).
- Unternehmensgewinn wird mittels EPS gemessen (IBES).
- Alle Observationen werden mit Finanzdaten aus Compustat gematcht.

Vorbereitung

- Fokus auf Q&A Session der Earnings Calls.
- Entwicklung von Algorithmen, um Sprachaufnahmen des CEOs zu extrahieren.
- Prognose: Performt das Unternehmen im nächsten Jahr besser oder schlechter?
- $\Delta EPS_{it+1} = EPS_{it+1} - EPS_{it} - drift_{it}$
- Darstellung als binärer Klassifikation (1 besser / 0 schlechter).

Herausforderungen der Stimmanalyse

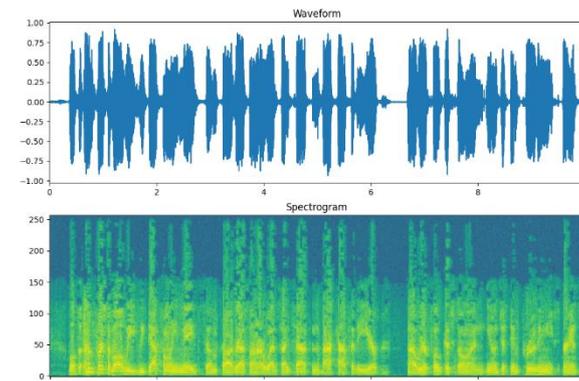
Listen Closely

Herausforderung

- Die menschliche Stimme ist äußerst Vielschichtig und die Bearbeitung von Audiodaten ist herausfordernd.
- Sprache ist
 - *granular*: selbst kleinste sprachliche Muster auf dem Silben-Level könnten Hinweise auf Emotionen liefern.
 - *sequentiell*: es ist nicht der einzelne Datenpunkt relevant sondern viel eher wie einzelne, granulare Datenpunkten zusammenhängen und somit eine Sprachmelodie bilden.
- Darüber hinaus bietet die Literatur keinerlei Hinweise, wie Stimmuster im unternehmerischen Kontext analysiert werden können.
- Es ist daher unklar welche Merkmale überhaupt relevant sind, um Emotionen zu identifizieren.
- Eine Stimmanalyse sollte daher möglichst umfassend sein und sowohl die Granularität als auch die sequentielle Natur von Stimmustern berücksichtigen.

Lösungsansatz

- Wir wandeln Audiodaten in Spektrogramme um, die das gesamte Frequenzspektrum abbilden, was Menschen üblicherweise hören.
- Die Spektrogramme dienen als Input in ein Convolutional Neural Network (CNN).
- CNNs spezialisiert in gitterartigen Strukturen, wie etwa Bildern, Muster zu erkennen.
- Da CNNs auf Pixel-Ebene arbeiten und sequenzielle Strukturen Berücksichtigen erfüllt das Modell die Anforderung an eine
 - *granulare* und
 - *sequentielle* Analyse der Stimmstrukturen

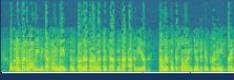
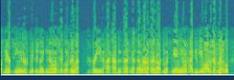
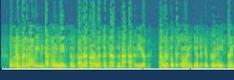
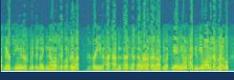
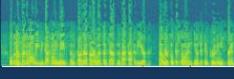
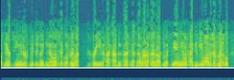
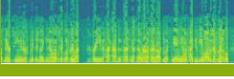


Herausforderungen der Stimmanalyse

Listen Closely

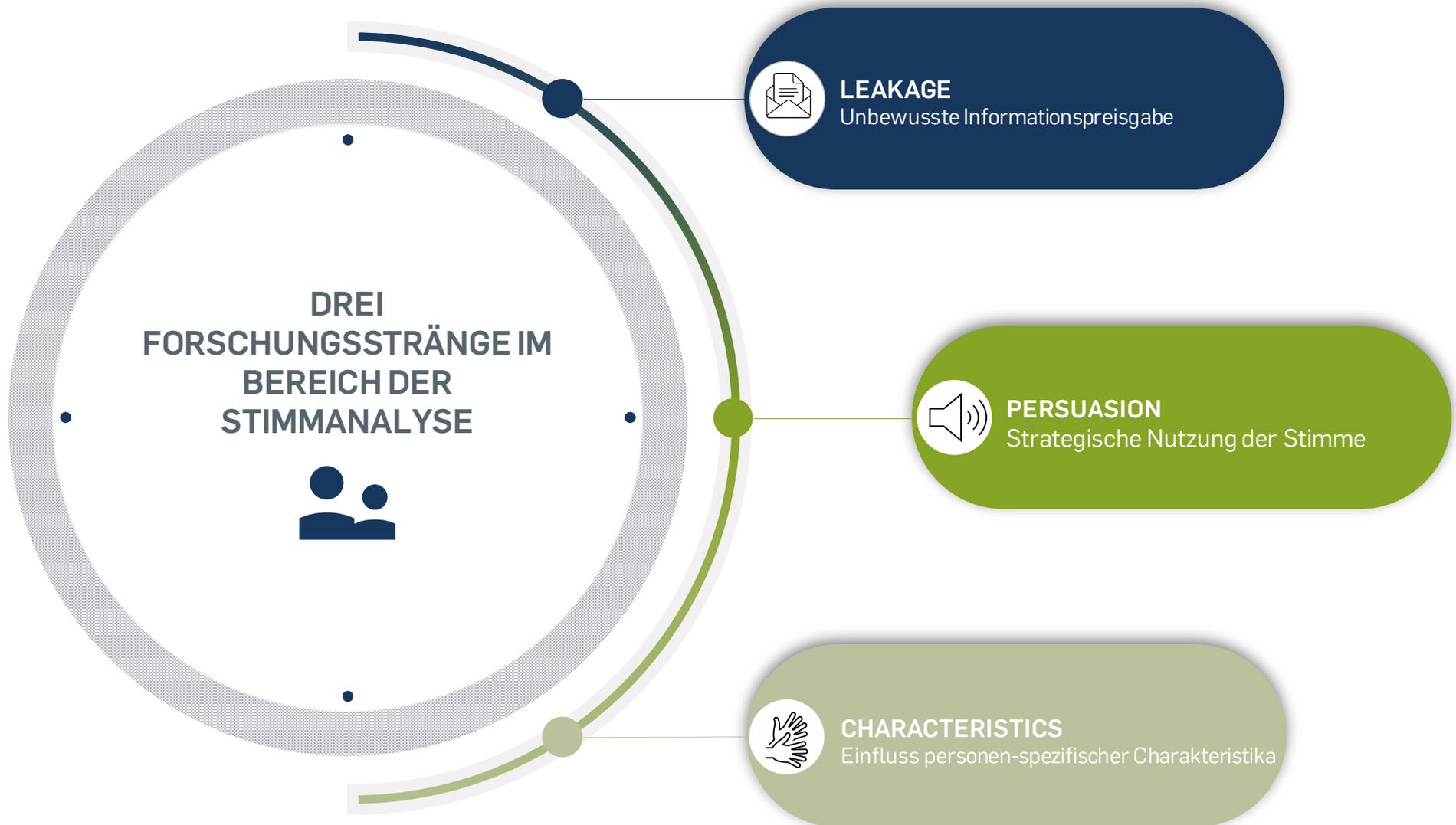
Ergebnisse

- Wir trainieren das CNN auf historischen Daten, um auf Basis der Spektrogramme vorherzusagen, ob sich das Ergebnis im Folgejahr verbessert oder verschlechtert.
- Das „Vocal Cue“ Modell outpermt
 - herkömmliche Prognosemodelle auf Basis detaillierter Finanzdaten
 - und Modelle die das gesprochene Wort in Earnings Calls verarbeitet
- Zudem können die Prognosen des Modells genutzt werden
 - um profitable Trading-Strategien zu implementieren (ca. 5% excess return p.a.)
 - und Analysten-Forecasts zu verbessern (relative 40% Verbesserung)
- Zusatzanalysen zeigen, dass die Prognosen
 - genauer sind, wenn Manager negative News kommunizieren
 - genauer sind, wenn Manager dazu neigen Antworten in der Q&A zu verweigern
 - ungenauer für sehr volatile Firmen sind,
- Unsere Ergebnisse zeigen, dass Stimmanalysen wertvolle Signale offenlegen, die bisher vom Kapitalmarkt überhört werden.

firm	year	data	earn_chg	model_predict
AAPL	2015		1	(train)
ALV.DE	2015		0	(train)
AAPL	2016		0	(train)
ALV.DE	2016		1	(train)
AAPL	2017		0	(train)
ALV.DE	2017		0	(train)
AAPL	2018		0	(train)
ALV.DE	2018		1	(train)
AAPL	2019		?	0.2678
ALV.DE	2019		?	0.9127

Ausblick

Forschung im Bereich der Stimmanalyse



Vielen Dank! Ich freue mich auf die Diskussion!

Kontakt: doron.reichmann@rub.de



autumn
EDITION **ir** 23

PRÄSENTIERT VON:  **cometis**  **irclub**