

roeder_2022_data_driven_decision_making_in_credit_risk_management_the_information_value_of_analyst_reports

Year

2022

Author(s)

Jan Roeder and Matthias Palmer and Jan Muntermann

Title

Data-driven decision-making in credit risk management: The information value of analyst reports

Venue

Decision Support Systems

Topic labeling

Partially automated

Focus

Secondary

Type of contribution

Novel approach

Underlying technique

Manual labeling assisted by topic prevalence information (by company and year)

Topic labeling parameters

\

Label generation

“Topic label assignment is based on the detailed topic analysis plots”

Three key metrics are presented to improve our understanding of the characteristics of the topic.

- topic prevalence for a single company, illustrating the extent to which a topic focuses on a particular company.
- topic prevalence in relation to years.
- topic top words

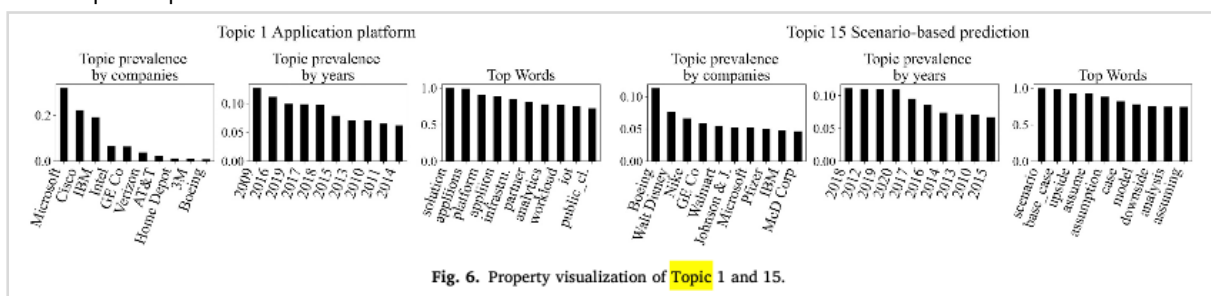


Fig. 6. Property visualization of Topic 1 and 15.

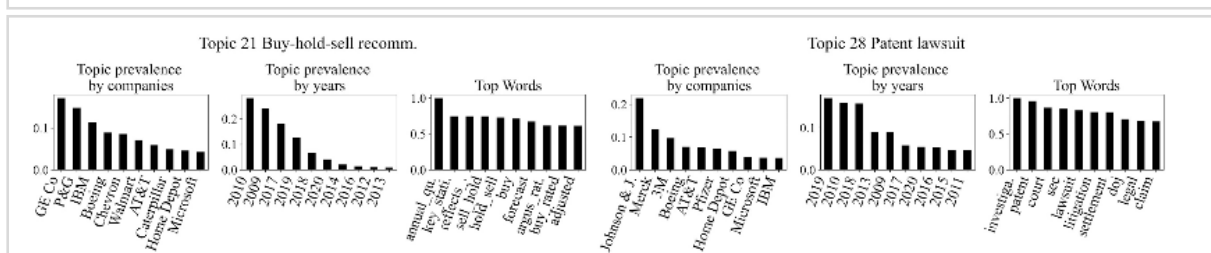


Fig. 7. Property visualization of Topic 21 and 28.

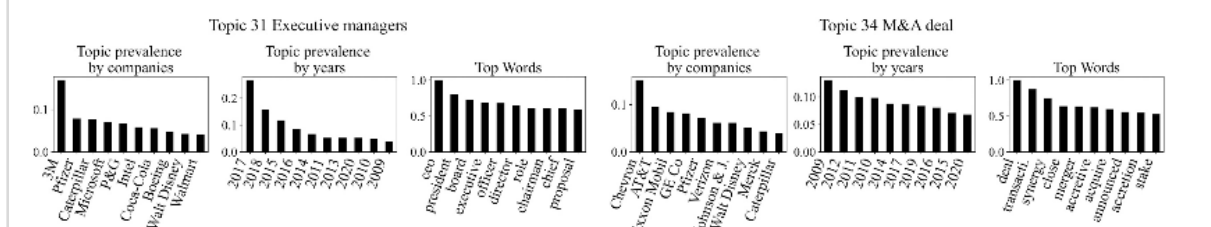


Fig. 8. Property visualization of Topic 31 and 34.

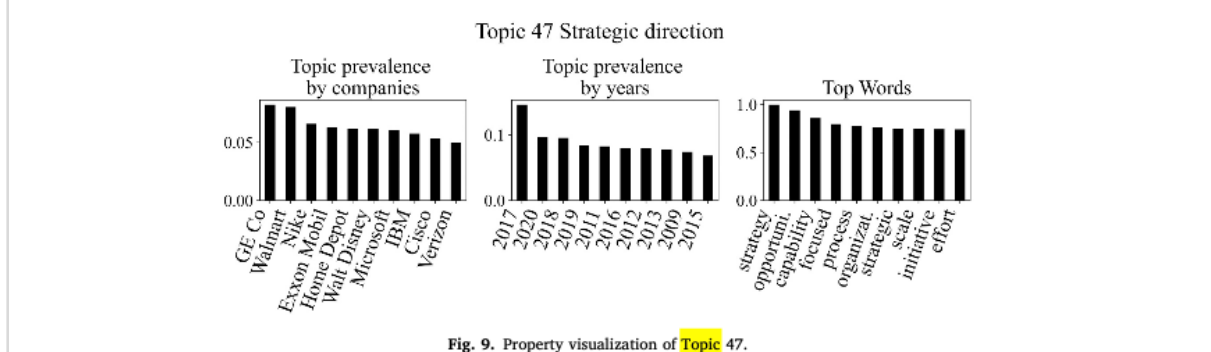


Fig. 9. Property visualization of Topic 47.

Table A.2

Top 10 words per topic estimated via LDA for analyst reports. Some topics are not part of the regression analysis (see Section 4.2). Topic label assignment is based on the detailed topic analysis plots as seen in Section 4.4.

Topic	Top 10 words with highest relevance per topic
1 Application platform	solution, applions, platform, applion, infrastructure, partner, analytics, workload, iot, public_cloud
2 Broker disclaimer	research, subject, compensation, specific_recommendation, received_compensation, view_expressed, analyst, reflect_personal, report_accurately, affiliate
3 Nike	dtc, adidas, basketball, demand_creation, digital, athletic, jordan, footwear_apparel, western_europe, woman
4 Smartphone	iphone, net_add, arpu, smartphone, churn, upgrade, prepaid, subsidy, post_paid, verse
5 Uptrend	upside, solid, momentum, improving, cycle, incremental, gain, healthy, near_term, strength
6 Walmart ecommerce	commerce, merchane, flipkart, ecommerce, food, online, sam, assortment, format, fuel
7 Financial statement analysis	charge, income, gaap, adjusted, related, item, loss, accounting, excluding, approximately
8 Sector trade	rank, sector, pair_trading, fed, best_rank, similarity_index, closest_competitor, attractiveness, quality, ecb
9 Earnings call	forward_looking, today, slide, statement, analyst, investor_relation, good_morning, qtr, grew, financial
10 Debt to cash flow	debt, free_cash, flow, repurchase, balance_sheet, shareholder, buyback, liquidity, fund, ratio
11 Orthopedic surgery	synthes, surgery, worldwide, ous, operational, spine, knee, hip, surgical, otc
12 Mobile service provider	spectrum, lte, carrier, sprint, verizons, fiber, unlimited, fcc, tower, telecom
13 Natural gas	perman, natural_gas, gas, lng, exploration, refining, bbl, liquid, xto, crude_oil
14 Caterpillar machinery	machinery, dealer_inventory, energy_transportation, erpillars, engine, machine, construction_equipment, oil_gas, mining_equipment, aftermarket
15 Scenario-based prediction	scenario, base_case, upside, assume, assumption, case, model, downside, analysis, assuming
16 Issues	issue, dont, number, fact, make, doe, big, long, problem, clear
17 Internet provider	support, solution, offer, ethernet, managed, access, feature, carrier, vpn, unified_communiions
18 Miscellaneous	thing, lot, kind, people, talk, great, little_bit, sort, good, question
19 Soft drinks	bottler, bottling, sparkling, refranchising, drink, pep, cce, concentrate, csd, sparkling_beverage
20 Regulatory disclaimer	research_analyst, exchange_regulated, regulation_authority, authority, security_plc, prudential_regulation, taiwan_security, registration_number, exchange, future_commission
21 Buy-hold-sell recomm.	annual_quarterly, key_statistic, reflects_previous, sell_hold, hold_sell, buy, forecast, argus_rating, buy_rated, adjusted
22 Intel chips	dcg, memory, cpu, ccg, altera, nand, amd, foundry, tsmc, mobileye
23 P&G body care	fabric_care, grooming, diaper, care, commodity, hair_care, family_care, gillette, developed, olay
24 Microsoft cloud & personal comp.	lidin, gaming, aws, intelligent_cloud, microsofts, personal_computing, github, window_oem, mpc, pbp

Table A.2 (continued)

Topic	Top 10 words with highest relevance per topic
25 Boeing airplanes	max, airline, airplane, airbus, plane, commercial_airplane, flight, aerospace, fleet, supplier
26 Intel mobile products	tablet, chip, notebook, atom, shipment, sandy_bridge, amd, microprocessor, arm, asp
27 Efficacy pharmaceuticals	dose, placebo, efficacy, fda, tofacitinib, phase_iii, clinical, dos, xarelto, therapy
28 Patent lawsuit	investigation, patent, court, sec, lawsuit, litigation, settlement, doj, legal, claim
29 Vaccines	vaccine, pfizers, prevnar, animal_health, xeljanz, wyeth, lipitor, generic, januvia, innovative
30 Survey data	survey, march, respondent, week, ubs_evidence, proprietary, basket, surveyed, monthly, indie
31 Executive managers	ceo, president, board, executive, officer, director, role, chairman, chief, proposal
32 Network infrastructure	ucs, routing, nexus, public_sector, switch, juniper, router, collaboration, recurring, switching_routing
33 Competitor analysis	firm, division, economy, moat, economic_moat, giant, competition, account, rise, growing
34 M&A deal	deal, transaction, synergy, close, merger, accretive, acquire, announced, accretion, stake
35 Healthcare input cost	rose, health_care, electronics_energy, safety_graphic, local_currency, raw_material, fell, acelity, lcd, non_recurring
36 Macroeconomic weakness	pressure, weakness, weak, macro, negative, lowering, near_term, spending, cut, environment
37 Johnson & Johnson products	zytiga, remicade, imbruvica, darzalex, invokana, ims, daratumumab, xarelto, olysio, stelara
38 COVID-19	covid, fy20, fy19, fy21, 4q19, important_closure, pandemic, 1q20, 2q20, 1q19
39 IBM product suite	strategic_imperative, signing, mainframe, constant_currency, watson, gts, red_hat, cognitive, outsourcing, analytics
40 Satellite TV	video, directv, dtv, time_warner, entertainment, ott, twx, warnermedia, bundle, verse
41 Option metrics	option, implied_volatility, volatility, considered_undervalued, spread, straddle, etf, strike, considered_undervalued, reuters_com
42 Revenue estimate	rev, street, guide, est., beat, mgmt, guided, model, miss, flat
43 Change prediction	grew, declined, sequentially, flat, slightly, prior, decreased, compared, expects, noted
44 Disney media and parks	star_war, box_office, cable_network, programming, theme_park, movie, park_resort, abc, cable, fox
45 Macroeconomic exposure	fiscal, basis_point, fourth_quarter, ago, compared, rose, period, foreign_currency, declined, totaled
46 Valuation	valuation, premium, relative, trade, count, historical, trading, yield, sector, dividend_yield
47 Strategic direction	strategy, opportunity, capability, focused, process, organization, strategic, scale, initiative, effort
48 Microsoft gaming and search	xbox, microsofts, search, nokia, window_phone, bing, server_tool, skype, user, surface
49 Cancer treatment	keytruda, nsccl, lung_cancer, cancer, tumor, keynote, chemotherapy, pfs, combo, melanoma
50 Home Depot	pro, hds, lowes, big_ticket, appliance, weather, online, lumber, supply_chain, diy
51 Meeting analysis	meeting, belief, noted, opportunity, highlighted, expects, update, analyst, reiterated, strategy
52 McDonald's revenue sources	breakfast, apmea, qsr, food, franchisees, sandwich, eotf, franchisee, chicken, burger
53 Sovereign energy demand	country, india, government, energy, plant, japan, chinese, local, world, facility
54 Energy production	alstom, oil_gas, gas_turbine, energy, wind, renewables, nbcu, finance, bhge, restructuring
55 Forecast evaluate	forecast, beat, slightly, previously, ahead, raised, raising, versus, upside, unchanged

Motivation

\

Topic modeling

LDA

Topic modeling parameters

Nr of topics (K): 20 to 100

Nr. of topics

55

Label

Manually assigned single or multi-word labels based on the three key topic metrics

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Finance

Dataset: Finance

Problem statement

Credit default swaps are commonly used financial instruments that provide information on a counterparty's creditworthiness. Although this metric can provide crucial insights, the underlying price dynamics often remain unknown and require further explanation. Data-driven decision-making is a key concept for identifying these reasons and supporting and justifying decisions.

In this paper, we provide such justifications by applying sentiment and topic analysis to company-related financial analyst reports.

While the contents of financial news have been analyzed in the past, analyst reports can offer additional insights, as seasoned analysts use them to disseminate in-depth research to experienced investors.

This analysis examines 3386 analyst reports covering constituents of the Dow Jones Industrial Average Index in the period from 2009 to 2020.

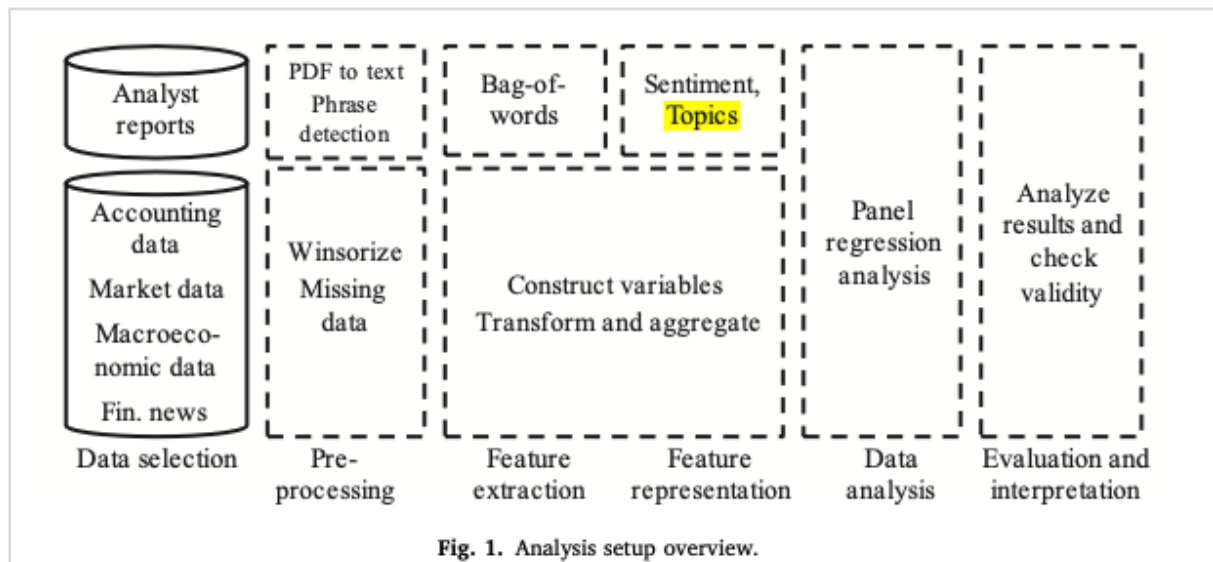
Corpus

Origin: Dow Jones Industrial Average (DJIA) index, Refinitiv EIKON

Nr. of documents: 3386

Details:

- The analyst reports stem from companies that are part of the Dow Jones Industrial Average (DJIA) index in the period from 2009 to 2020.
- CDS spreads from Refinitiv EIKON
- Constituents of the DJIA can be linked to CDS Spreads and analyst reports. For each earnings announcement date, the CDS spread is linked to analyst reports published up to ten days before the earnings date.



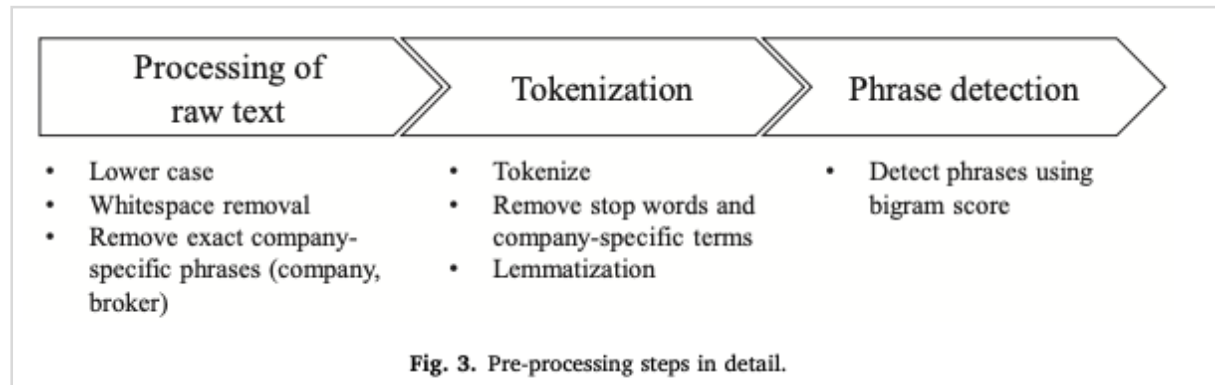
Document

An analyst report belonging to a company and a fiscal quarter.

Pre-processing

- Transform PDF files and into Excel files, where each paragraph is represented as an Excel cell.
- To remove duplicates, reports were compared on a paragraph basis. Those that exceeded a similarity threshold of 70% were removed.
- raw text is lowercased
- whitespaces are removed
- and company- and broker-specific phrases are removed.

- text is tokenized
- lemmatized
- numbers and stop words are removed.
- Commonly occurring two-word phrases (i.e., bigrams) are identified and concatenated with an underscore.



@article{roeder_2022_data_driven_decision_making_in_credit_risk_management_the_information_value_of_analyst_reports,

abstract = {Other than banks, non-financial companies also continuously monitor and analyze their credit risk exposure to avoid possible counterparty defaults. Credit default swaps are commonly used financial instruments that provide information on a counterparty's creditworthiness. Although this metric can provide crucial insights, the underlying price dynamics often remain unknown and require further explanation. Data-driven decision-making is a key concept for identifying these reasons and supporting and justifying decisions. In this paper, we provide such justifications by applying sentiment and topic analysis to company-related financial analyst reports. While the contents of financial news have been analyzed in the past, analyst reports can offer additional insights, as seasoned analysts use them to disseminate in-depth research to experienced investors. This analysis examines 3386 analyst reports covering constituents of the Dow Jones Industrial Average Index in the period from 2009 to 2020. The results suggest that even when established credit risk indicators and financial news are considered, the sentiment and a subset of topics are correlated with changes in the credit default swap spread, indicating a fundamental relationship between quantitative risk metric and analyst reports. We find that analyst reports contain information related to the change in credit default swap spreads, an insight that helps to improve our understanding of existing risk assessments. The outcome indicates that banks or

corporate risk managers can benefit from complementing established financial metrics and even financial news data with new insights derived from analyst reports.},

author = {Jan Roeder and Matthias Palmer and Jan Muntermann},

date-added = {2023-03-23 20:30:32 +0100},

date-modified = {2023-03-23 20:30:32 +0100},

doi = {https://doi.org/10.1016/j.dss.2022.113770},

issn = {0167-9236},

journal = {Decision Support Systems},

keywords = {Credit risk, Data-driven decision-making, Unstructured data, Text mining, Sentiment analysis, Topic mining},

pages = {113770},

title = {Data-driven decision-making in credit risk management: The information value of analyst reports},

url = {https://www.sciencedirect.com/science/article/pii/S0167923622000410},

volume = {158},

year = {2022}}

#Thesis/Papers/Initial