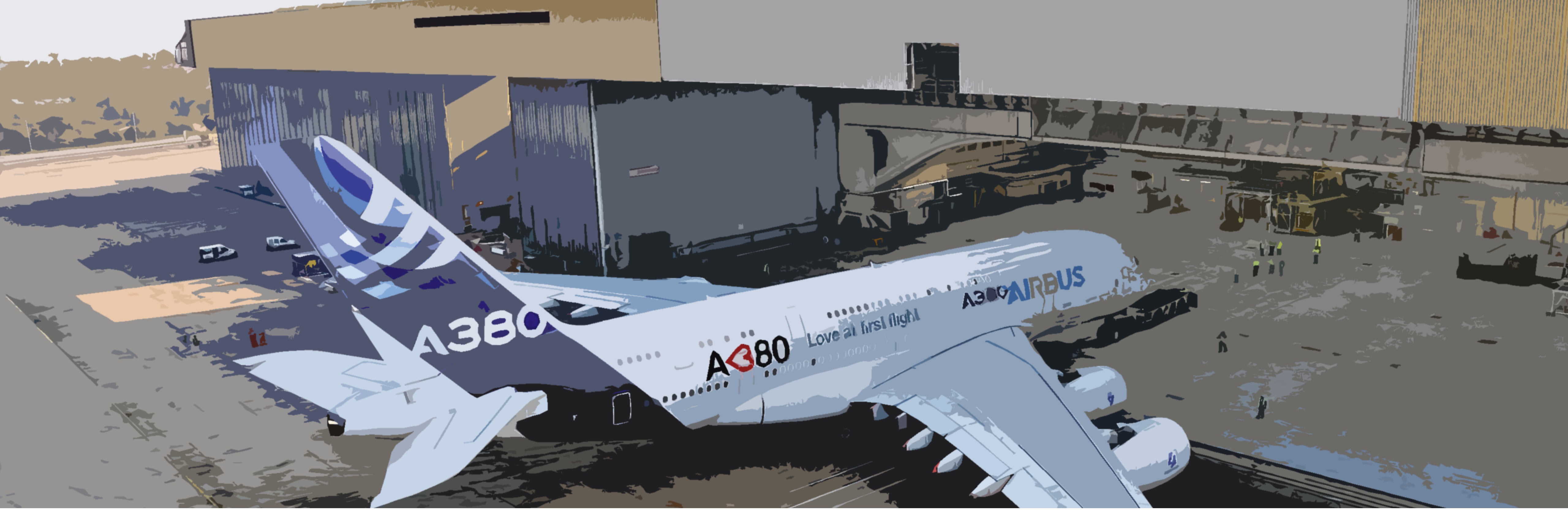


Determinants of Credit Ratings of State-Owned-Enterprises

8,276 - Foundations in Data Science and Machine Learning

Alessandro Fosselard, Antoine Ameye
May 2025

Universität St.Gallen

Very Short Introduction: “What is the current weight of S-O-Es in today’s economy?”

Global Presence	In 2023, SOEs accounted for approximately **12% of global market capitalization**.
Revenue and Assets	Between 2000 and 2023, the number of SOEs among the largest 500 enterprises by revenue worldwide **increased from 34 to 126**.
Regional Variations	In the **OECD area**, the market capitalization of listed firms with more than 25% public sector ownership is **just 2%**. In contrast, this figure is **16% in Latin America** and over **40% in some markets**.
Developing Countries	In developing countries, revenues from businesses with at least 10% state ownership are equivalent to **17% of GDP on average**.
China's SOEs	China's SOEs are particularly prominent, accounting for over 60% of the country's market capitalization and generating **about 23-28% of its GDP**.

Sources: OECD, World Bank

Agenda: we want to check whether state-ownership influences the rating, depending on the localization of the state or the industry.

Data Management

Modeling (ML)

Descriptive Analysis

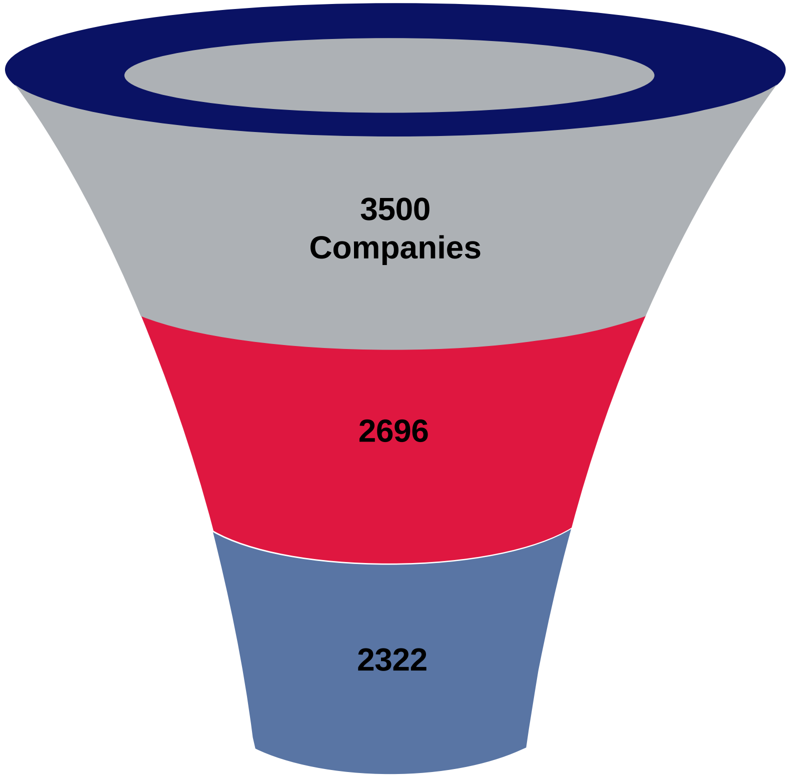
Getting the data: large DataFrames from Bloomberg® containing 35 columns of data for the 3500 rated firms across the 3 main credit rating agencies.

Bloomberg

MOODY'S RATINGS

S&P Global Ratings

FitchRatings



01 We download all the rated companies by Moody's, S&P and Fitch Ratings

02 We drop the duplicates companies, those that are rated by two or three rating agencies simultaneously (we keep S&P when this is the case) and the "NA"

03 We drop some financial values that had not enough data (with many blanks) to avoid losing too many companies on which to train the model

The resulting cleaned DataFrame is composed of:

- a (2322x26) X-features matrix
- a (2322x1) Y-target vector

Main assumption behind the data clustering across the 3 agencies: the rating correlating between the 3 is high and they all have the same number of notches (22).

In 90% of the cases where Moody’s, S&P, and Fitch all rated the same bond, **Fitch gave the same letter rating as at least one of the other agencies**.

Mean rating differences: **Fitch’s ratings were, on average, 0.3 notches higher** than Moody’s and S&P in the 3-agency sample; in larger samples, the **difference can reach 0.74 notches vs Moody’s and 0.56 vs S&P**.

Jewell, J. J., & Livingston, M. (1999). A Comparison of Bond Ratings from Moody's, S&P, and Fitch IBCA.

A Comparison of Bond Ratings from Moody’s S&P and Fitch IBCA

BY JEFF JEWELL AND MILES LIVINGSTON

Previous research has found that the bond market values the ratings of Moody’s and Standard & Poor’s. This paper extends earlier research by comparing the ratings of Moody’s, Standard and Poor’s, and Fitch IBCA. The authors examine a very large database with monthly observations of bonds and bond ratings over a five-year time period. The analysis focuses on comparing rating levels, rating changes, and the impact of ratings on bond yields. The results show that firms with publicly available Fitch IBCA ratings have higher ratings from Moody’s and S&P than firms without Fitch IBCA ratings. The typical firm releasing a Fitch IBCA rating has a lower yield (controlling for Moody’s and S&P rating), a more stable rating, and is more likely to receive an upgrade. For split-rated bonds (Moody’s vs. S&P), Fitch IBCA serves as a tiebreaker. This evidence is consistent with the bond market valuing the ratings of all three raters—Moody’s, Standard & Poor’s, and Fitch IBCA.

12–14% of firms had ratings **one notch higher** from S&P.

1–2% had **two or more** notches higher

Caridad, J. M., Arencibia, O., & Seda, P. (2020). Do Moody's and S&P Firm's Ratings Differ?

173
Caridad, L., Núñez-Tabales, J., Seda, P., & Arencibia, O.
ISSN 2071-789X
INTERDISCIPLINARY APPROACH TO ECONOMICS AND SOCIOLOGY
ECONOMICS & Sociology
Caridad, L., Núñez-Tabales, J., Seda, P., & Arencibia, O. (2020). Do Moody's and S&P firm's ratings differ?. *Economics and Sociology, 13*(4), 173-186. doi:10.14254/2071-789X.2020/13-4/11

DO MOODY'S AND S&P FIRM'S RATINGS DIFFER?

Lorena Caridad
University of Córdoba, Córdoba, Spain
E-mail: d92calol@uco.es
ORCID 0000-0002-3406-9917

Julia Núñez-Tabales
University of Córdoba, Córdoba, Spain
E-mail: es2nutaj@uco.es
ORCID 0000-0001-6597-6029

Petr Seda
VSB Technical University of Ostrava, Czech Republic
E-mail: petr.seda@vsb.cz
ORCID 0000-0002-2205-6377

Orlando Arencibia
VSB Technical University of Ostrava, Czech Republic
E-mail: orlando.arencibia@vsb.cz
ORCID 0000-0001-8217-2399

Received: December, 2019
1st Revision: September, 2020
Accepted: December, 2020

DOI: 10.14254/2071-789X.2020/13-4/11

ABSTRACT. Credit rating agencies produce public statements about the financial health of companies, institutions, geographical entities and financial assets. The main available information about firms, useful for analyzing their long-term creditworthiness is their public accounts about the activities and results, besides the audit reports and their credit ratings. The agencies’ results are mainly based on these data, but they claim to use additional qualitative information, with a methodology only partially disclosed. When different agencies produce long-term ratings about a particular firm, it should be expected that they were coincident, or at least similar, so that investors could use any of them to assess the potential financial risk. This is not the case, as the same companies can be rated differently by different agencies. This is the case with Standard and Poor's and Moody's: although their rating methods are not coincident, but their aim is to measure a similar latent variable – the firm's credit risk. These divergences could be caused, at least in part, by possible conflicts of interest or by a phenomenon called 'rating inflation'. A difference index is proposed to measure the differences in ratings when comparing several agencies' evaluations. The situation with the two main agencies is examined, using two large samples in a five-year period: clear discrepancies are observed, in some economic sectors, and similarities in others, with some evidence about getting higher ratings depending on a chosen agency. Also, a convergence of ratings during the period of 2014-2018 is observed, more prominent in some sectors, suggesting that additional regulation is needed to increase the market transparency.

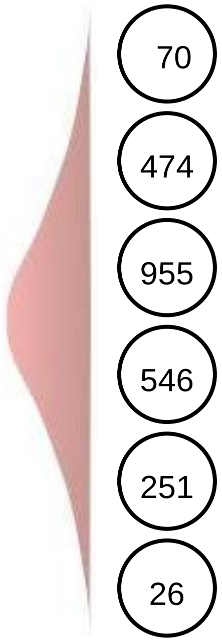
JEL Classification: G24, G17, F65, D53

Keywords: long-term ratings, conflict of interest, rating inflation, credit rating agencies, ratings evolution.

Sources: Caridad, J. M., Arencibia, O., & Seda, P. (2020). Do Moody's and S&P Firm's Ratings Differ?. Economics & Sociology, 13(4), 167-183. Jewell, J. J., & Livingston, M. (1999). A Comparison of Bond Ratings from Moody's, S&P, and Fitch IBCA. Financial Markets, Institutions & Instruments, 8(4), 1-45.

In addition, we solve the problem by regrouping the ratings in 6 categories such that the likelihood of companies being in different groups among credit raters falls drastically.

Distribution of Ratings



S&P	Fitch	Moody's	Grade
AAA	AAA	Aaa	Investment grade: *Prime*
AA+	AA+	Aa1	
AA	AA	Aa2	*High Grade*
AA-	AA-	Aa3	
A+	A+	A1	
A	A	A2	*Upper Medium Grade*
A-	A-	A3	
BBB+	BBB+	Baa1	
BBB	BBB	Baa2	
BBB- (India)*	BBB- (India)*	Baa3 (India)*	*Lower Medium Grade*
BB+	BB+	Ba1	
BB	BB	Ba2	Non-Investment Grade: *Speculative*
BB-	BB-	Ba3	
B+	B+	B1	
B	B	B2	*Highly Speculative*
B-	B-	B3	
CCC+	CCC+	Caa1	
CCC	CCC	Caa2	*Substantial Risk*
CCC-	CCC-	Caa3	
CC	CC	Ca	*Extremely speculative*
C	C	C	

Rating Categories

"a High Grade"

"b Upper Medium Grade"

"c Lower Medium Grade"

"d Non-Investment Grade Speculative"

"e Highly Speculative"

"f Substantial Risk"

Observation: we use letter at the beginning of the category title to have an alphabetical order (will be useful for confusion matrix later) and to better define them later on.

Data Management | Modeling (ML) | Descriptive Analysis

Sources: *S&P, Moody's and Fitch*

Extracting the SOE from the large sample using 47 batches of 50 companies sent to ChatGPT client through API key. Cost: ± $2, we lose 5 companies ► 2317

Using OpenAI

Classifying "Yes" or "No", splitting in 47 batches

- Results: 10.16% of the companies are identified as SOE
- Critic: Quite precise, also catches companies with state minority stakes

Extracting the 236 SOEs from the rest of the companies

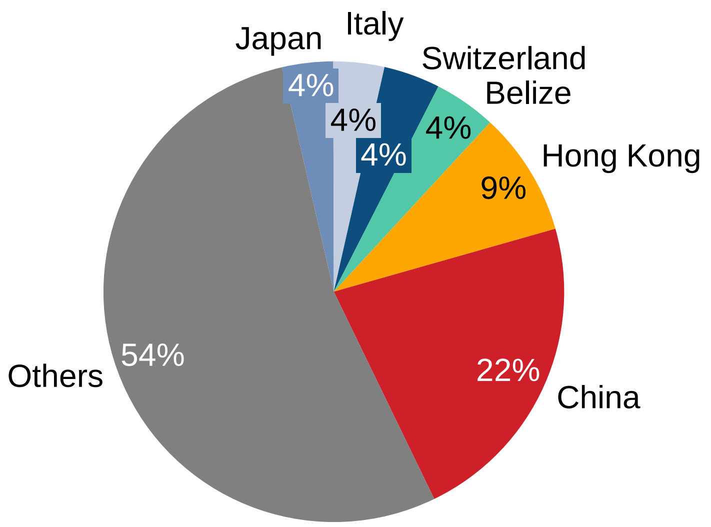
Observation: *many SOE's* are located in China, and the majority of them are in the financial sector, this could result in a model calibrated on the non-SOE firms (which are less skewed towards financials) and **reduce the precision** because the financial ratios of banks are highly influenced by regulations

Using Keywords

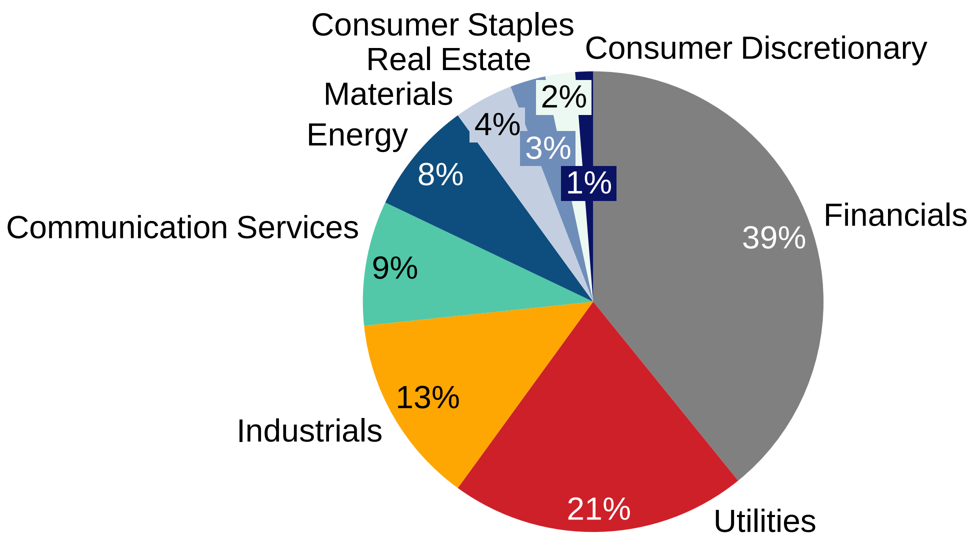
"holding", "state", "government" …

- Results: 14.99% of the companies are identified as SOE
- Critic: Not precise, also many private companies are "holdings"

Where are the SOE's located?



What are the SOE's sectors?



Data Management | **Modeling (ML)** | **Descriptive Analysis**

How do we split the data after this?

Non-SOE DataFrame with remaining 1850*13 random drawn Non-SOE
2322 – (2*236) = 1850

Non-SOE 236*13 random drawn from the Non-SOE DataFrame

236*13 **SOE DataFrame**, built using OpenAPI

Remove all the non-numerical columns for X

Set Category as Y

X_Train (1850*12), Y_Train (1850*1)

X_Test_1 (236*12), Y_Test_1 (236*1)

X_Test_2 (236*12), Y_Test_2 (236*1)

Model: Logistic Regression with k-Fold Cross-Validation, why did we choose this model?

Model Presentation

Logistic regression is a method that helps us predict categories (like credit ratings) based on input data. It looks at patterns in past data and estimates the probability that something belongs to each possible category.

Linear regression predicts a continuous number (like income or price), while Logistic regression predicts a category (like credit grade A, B, or C) by estimating probabilities for each option.

Objective: Classify **credit ratings** of **non-SOE (state-independent)** firms using **Logistic Regression.**

Code Explanation

We used LogisticRegressionCV, a tool that automatically tests the model’s performance using internal checks (cross-validation).

Specifically, we applied 5-fold cross-validation, which means that the training data is split into 5 parts. The model is trained on 4 parts and tested on the 5th — and this process is repeated 5 times.

The solver we used, "newton-cg", supports: Multi-class predictions (since we have multiple rating categories).

Comparing two different calibrations of the model

Results with no Penalty

C_value_logit = 1e20

We tested it on a random "virgin" sample of nongovernment stake

AUC: 79.4939 %
Acc: 52.1186 %

Then we tested it on companies with government stake

AUC: 64.7098 %
Acc: 38.5593 %

Results with penalty

C_value_logit = [0.01, 0.1, 1, 10, 100]

We tested it on a random "virgin" sample of nongovernment stake

AUC: 80.5986 %
Acc: 49.1525 %

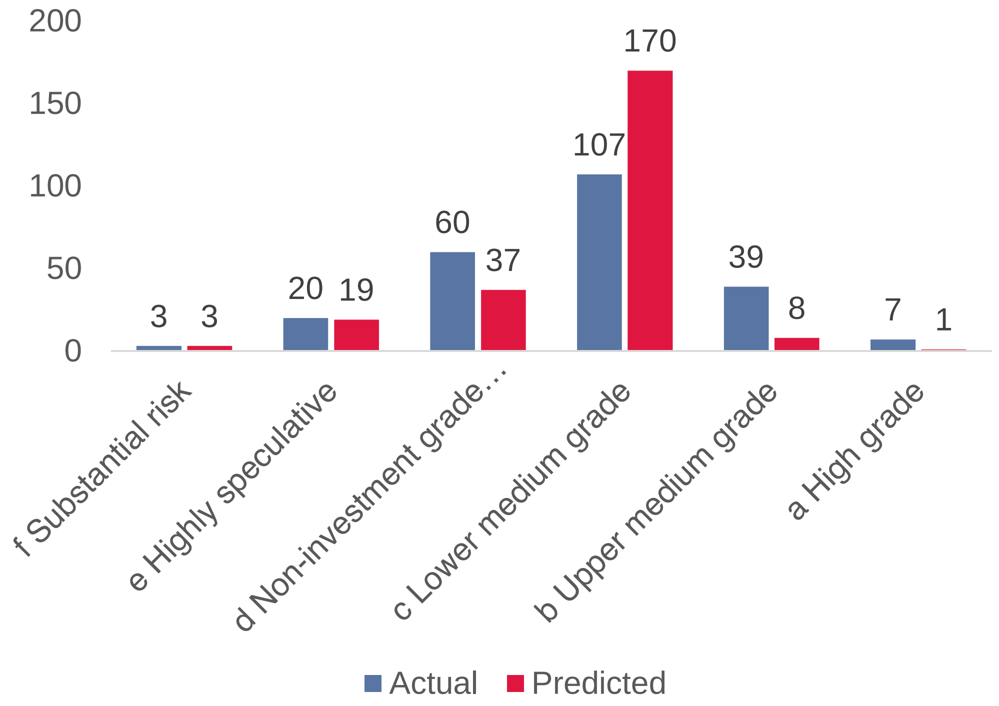
Then we tested it on companies with government stake

AUC: 60.8330 %
Acc: 38.1356 %

Logistic Regression - X_TEST_1 (Non-SOE): Centralizing effect, many companies are categorized in the Lower Medium Grade

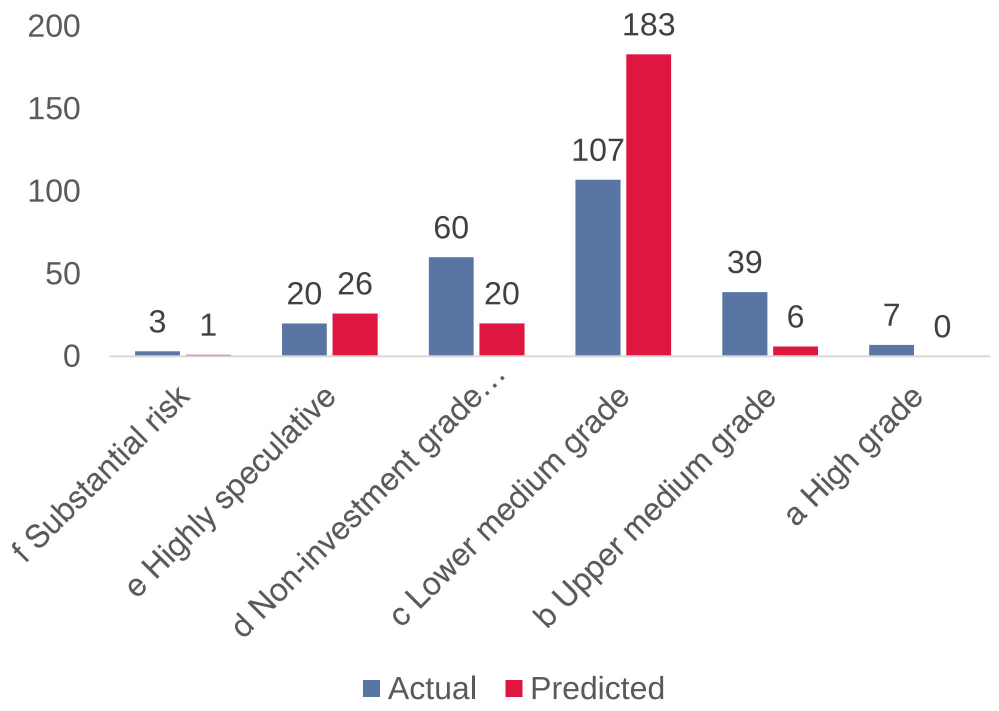
No penalty

The model predicts **less** D and B than the actual number. The model predicts a lot of C this might be due to the fact that the model was trained on a majority of C.



With penalty

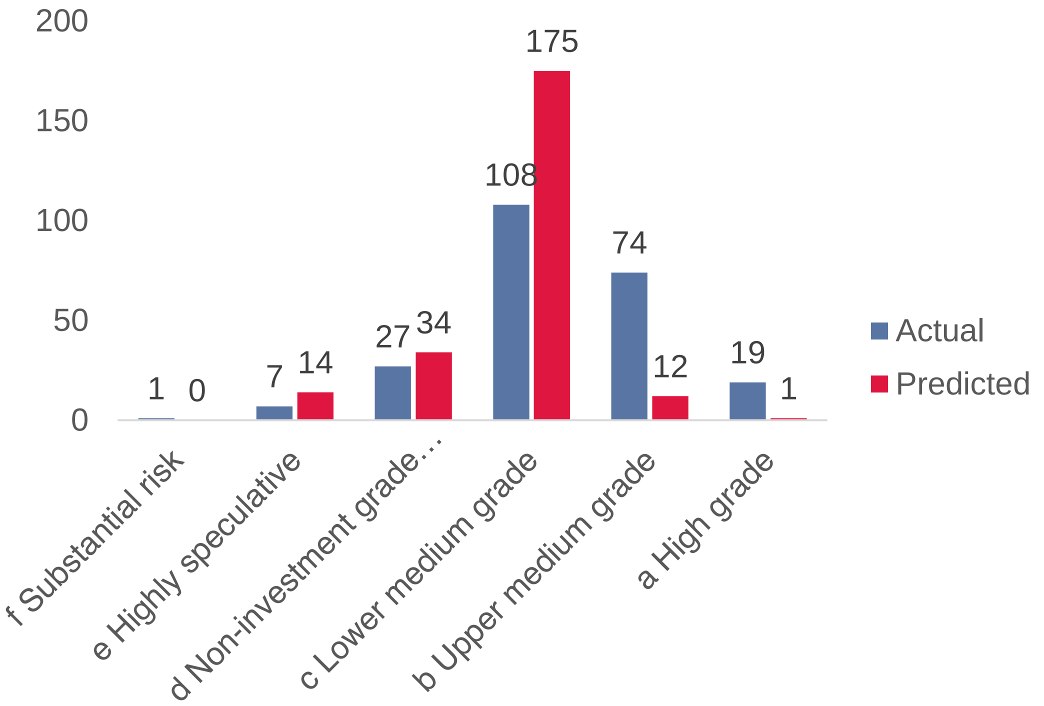
Same as with no penalty but now the model is predicting **more** E.



Logistic Regression - X_TEST_2 (SOE): The model is too negative, which confirms our query, state support seems positive

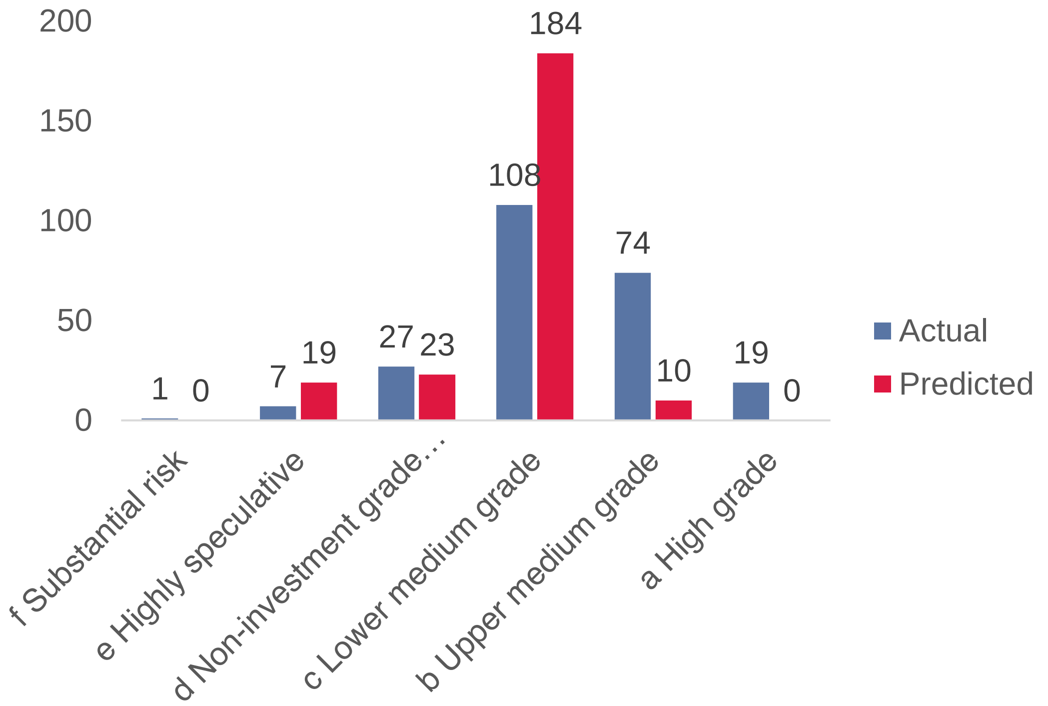
No penalty

This time the model predicts **more** D and E than there actually are. However, it still predicts **less** A and B, but now this is consistent with a potential state benefit.



With penalty

The model predicts **less** D but **more** E than the actual number. This is also in line with a potential state benefit. It still predicts **less** A and B than the actual number.



Confusion Matrix Logistic Regression: in the SOE a lot of mistakes were from C to B. Without Penalty

Non-SOE

Confusion Matrix - Non-SOE Companies

ACTUAL \ PREDICETD	a High grade	b Upper medium grade	c Lower medium grade	d Non-investment grade speculative	e Highly speculative	f Substantial risk
a High grade	0	1	6	0	0	0
b Upper medium grade	0	2	36	0	0	1
c Lower medium grade	0	3	93	10	1	0
d Non-investment grade speculative	1	2	29	20	7	1
e Highly speculative	0	0	6	7	7	0
f Substantial risk	0	0	0	0	2	1

SOE

Confusion Matrix - SOE Companies

ACTUAL \ PREDICETD	a High grade	b Upper medium grade	c Lower medium grade	d Non-investment grade speculative	e Highly speculative	f Substantial risk
a High grade	0	2	16	1	0	0
b Upper medium grade	0	6	63	1	4	0
c Lower medium grade	0	2	77	22	7	0
d Non-investment grade speculative	1	1	18	6	1	0
e Highly speculative	0	1	1	3	2	0
f Substantial risk	0	0	0	1	0	0

ACTUAL

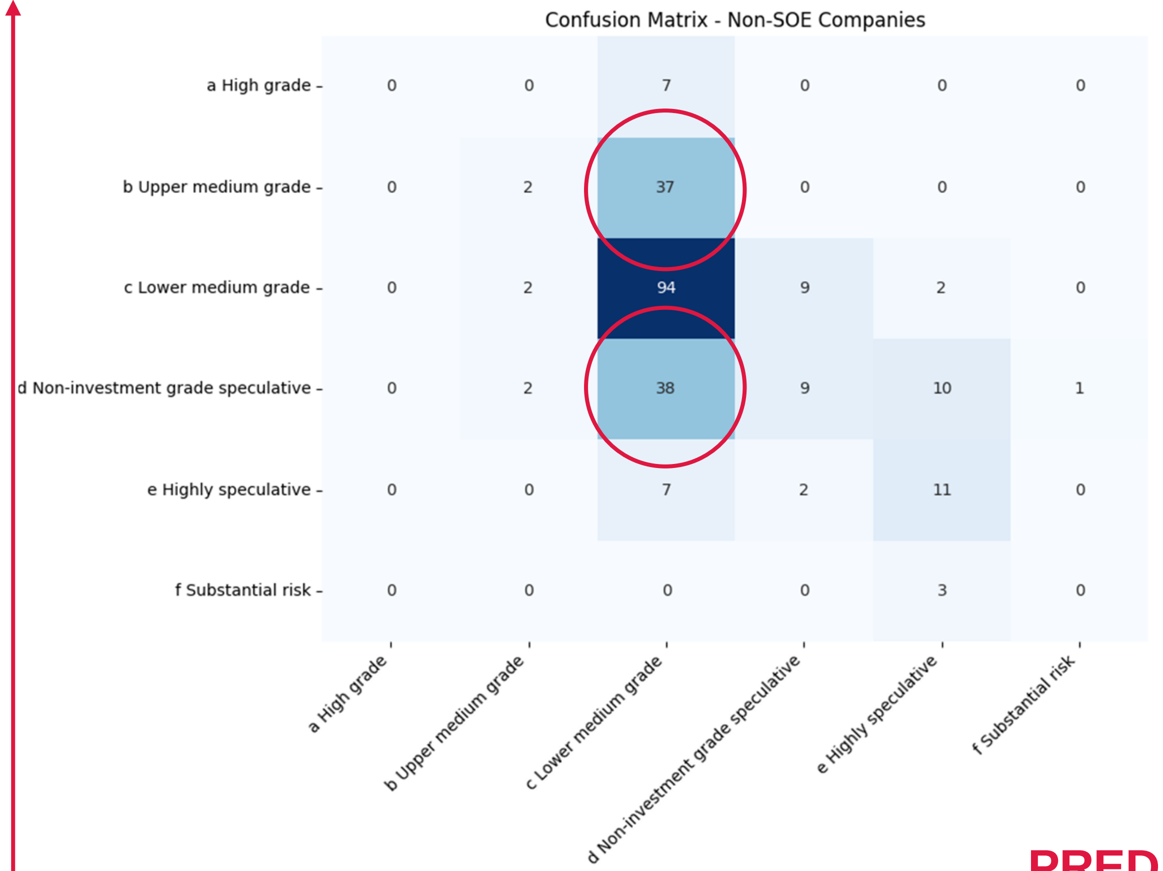
PREDICETD

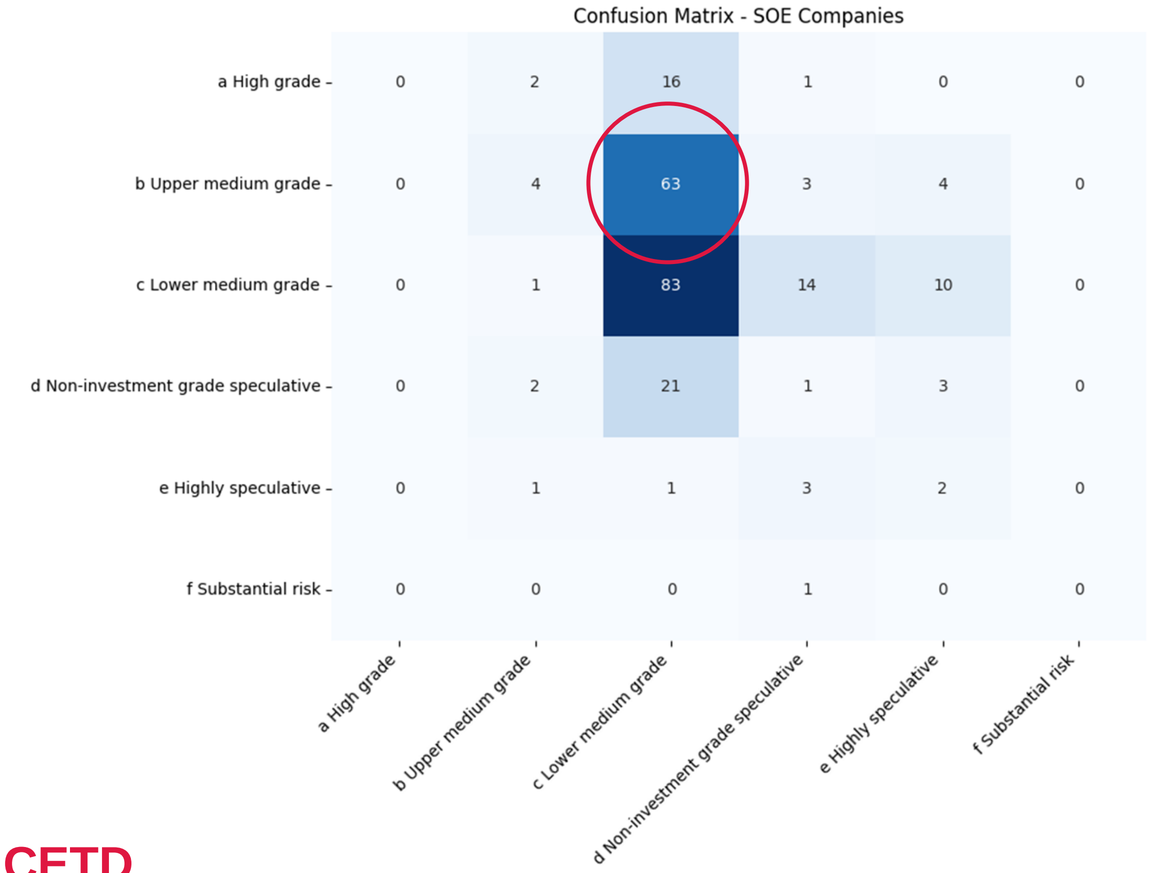
Confusion Matrix Logistic Regression: in the SOE a lot of mistakes were from C to B. With Penalty

Non-SOE

SOE

ACTUAL





PREDICETD

Data Management | Modeling (ML) | **Descriptive Analysis**

Model: Random Forest Model

Model Presentation

Random forests are built from decision trees:
"Trees have one aspect that prevents them from being the ideal tool for predictive learning, namely accuracy" (The Elements of Statistical Learning)

Random Forests are **not flexible** when it comes to classifying new samples because decision rules are hard-coded during training meaning they only work well when new data looks similar to the training data. It **struggles to adapt** to new, unseen patterns.

Why it might be better?

Credit ratings are based on **rule-like decisions**
Rating agencies use structured scorecards: if leverage is high, downgrade; if interest coverage is strong, upgrade.
Random Forests are well suited to replicate this rule-based logic.

Nonlinear interactions matter in credit: A company with high debt but massive cash flow is very different from a company with the same debt and no cash. **Random Forests naturally handle nonlinear combinations** of variables without you having to specify interactions manually.

Sources: Hastie, T., Tibshirani, R. and Friedman, J., 2009. The elements of statistical learning: Data mining, inference, and prediction. 2nd ed. New York: Springer.

Random Forest: Results

Results

We tested it on a random “virgin” sample of nongovernment stake

AUC: 88.25%
Acc: 57.20%

Then we tested it on companies with government stake

AUC: 72.28%
Acc: 43.22%

Description of the code

We used a **pipeline**, meaning all steps (data preparation + model) were packaged together for cleaner, repeatable analysis.

We used **class_weight='balanced'**, which automatically adjusts the importance of each class based on how frequent it is in the training data. This useful because our rating classes were not evenly distributed.

It tells the model:

- “Don't treat all mistakes equally. If you misclassify a rare class like CCC, that should count as a bigger mistake than misclassifying a common class like A.”

Number of trees: **(n_estimators) 2,000**

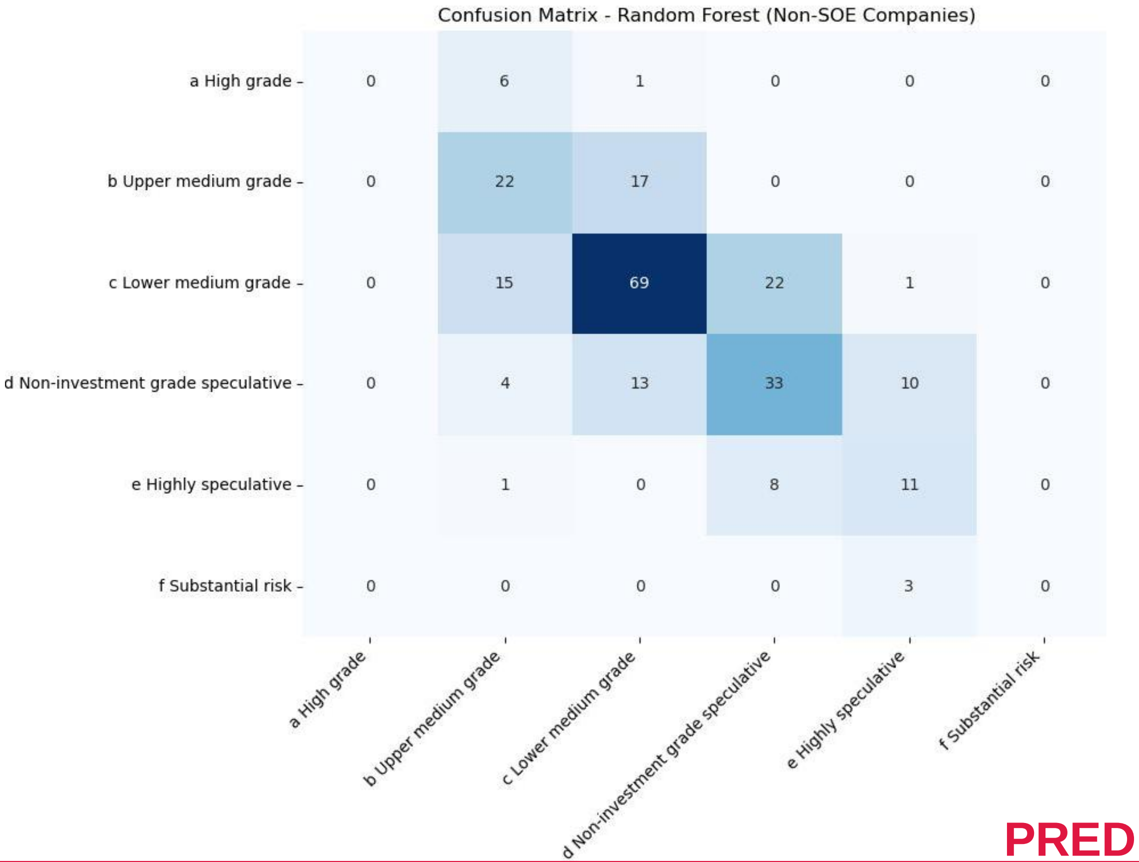
- Each tree is trained on a random subset of the training data (bagging).
- Considers a random subset of features at each decision split.

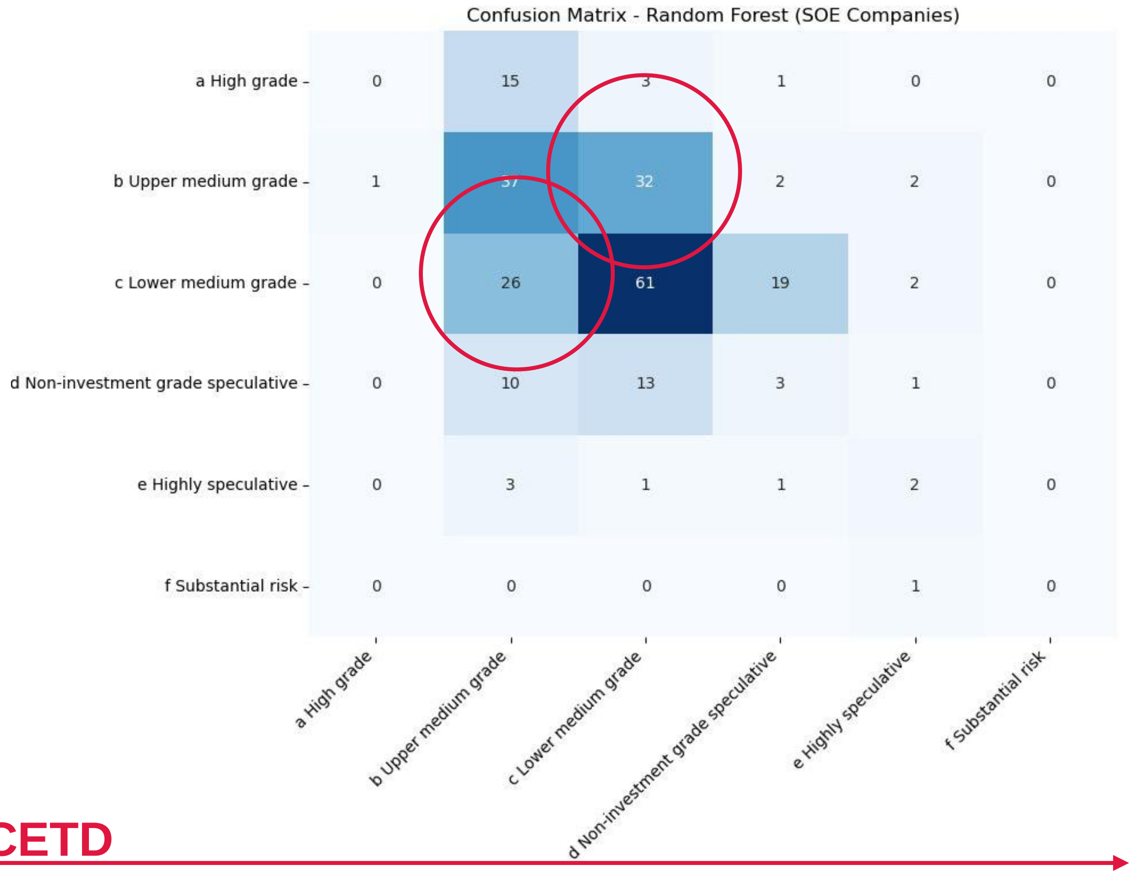
Confusion Matrix Random Forest: here the mistakes are more balanced.

Non-SOE

SOE

ACTUAL





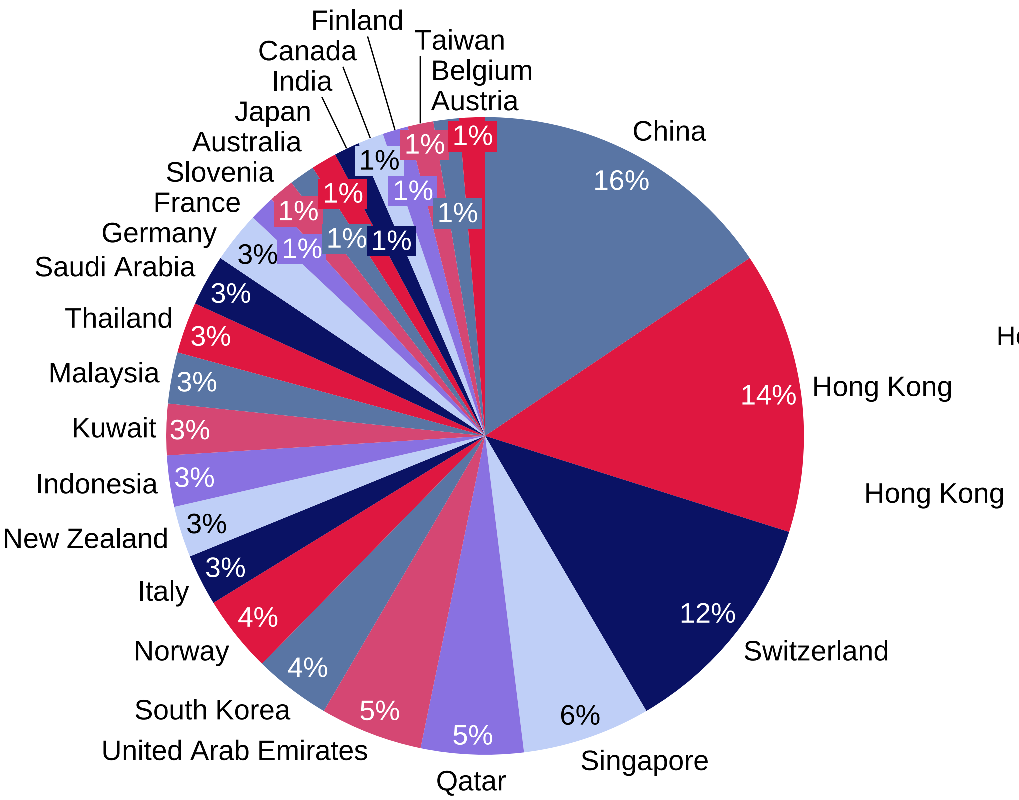
PREDICETD

Data Management | Modeling (ML) | **Descriptive Analysis**

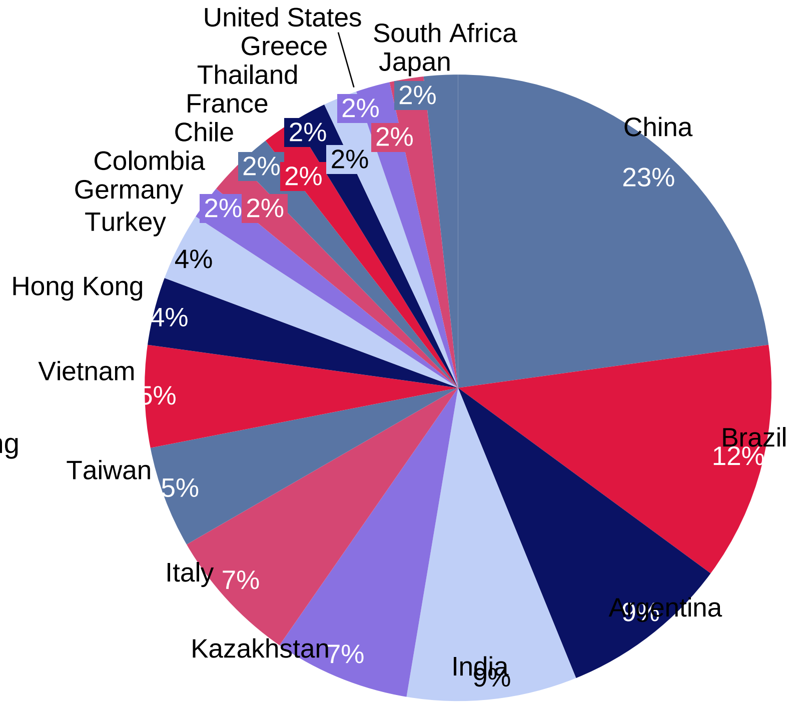
Model - Random Forest: which countries are uplifting the rating and which countries have a downgrading effect?

Hong Kong	AA+
Switzerland	AAA
Singapore	AAA

Countries in which the model predicts a lower rating than the actual rating for S-O-Es



Countries in which the model predicts a higher rating than the actual rating for S-O-Es

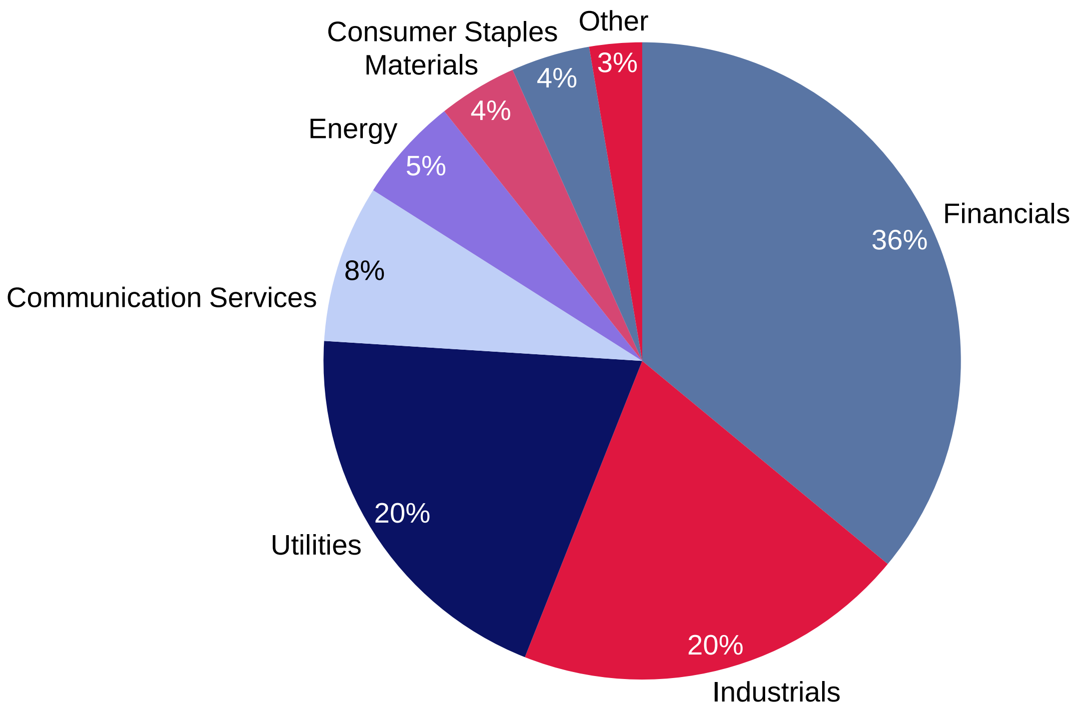


Brazil	BB
Argentina	CCC
India	BBB-

Data Management | **Modeling (ML)** | **Descriptive Analysis**

Model - Random Forest: which industries are uplifting the rating and which industries have a downgrading effect?

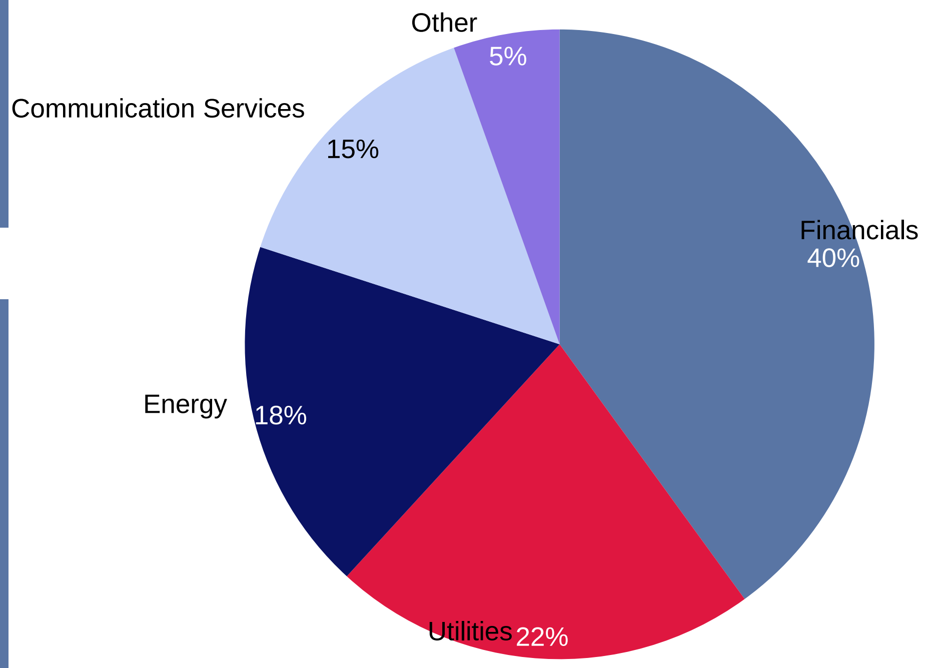
Sectors in which the model predicts a lower rating than the actual rating for S-O-Es



ΔIndustrials: 20%
ΔUtilities: -2%
ΔCommunications: -7%
ΔEnergy: -13%

Note: Industrials include transport & logistics, defense and construction.

Sectors in which the model predicts a higher rating than the actual rating for S-O-Es



Data Management | Modeling (ML) | **Descriptive Analysis**

FLASHBACK:
Are financials less important when determining the rating of State-Owned-Enterprises?

I. Data Retrieval

i) Bloomberg <EQS> we aim for 3000 companies (already have +- 2500 with only S&P) and 20 financials as well as the text of the ultimate parent, country and industry.
ii) Cleaning data duplicates between Moody's and S&P, outliers (in terms of financials)
iii) Text analysis (using GPT) identify S-O-Es extract the sample (say 300 companies and the same number randomly of private companies

II. Data Visualization

II. Data Visualization

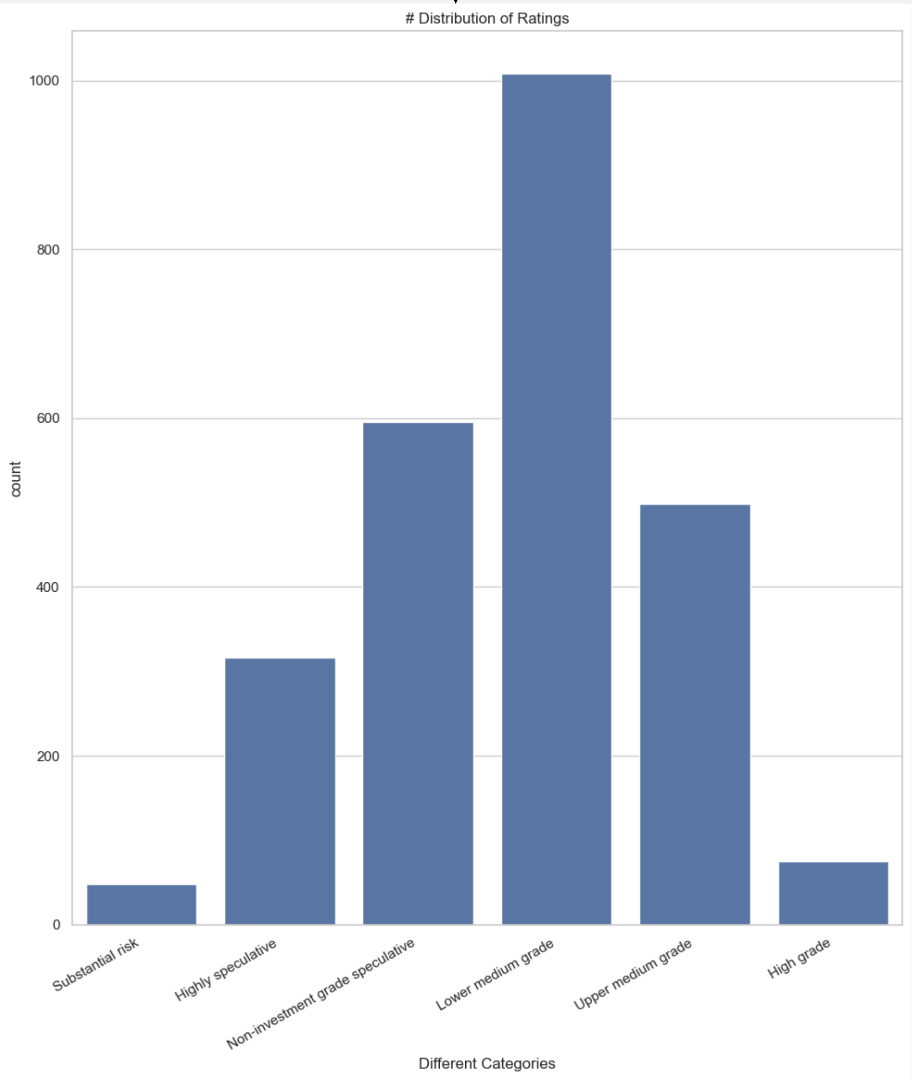
i) Government related companies' distribution per country
ii) Government related companies' distribution per company
iii) Results where the companies have been assigned compared to the real category – for government companies
iv) Results where the companies have been assigned compared to the real category – for private companies

=> Ideas for the 4 plots

III. Data Modelling

i) Logistic regression (multinomial (softmax) logistic regression). solver='lbfgs', 'saga' - K-th sample splitting
ii) Results: confusion matrix
iii) K-Means Clustering
iv) Conclusions: Are S-O-E treated similarly as private enterprises? Are they benefitting or not from the state and then **groupby.** (the mistakes in + or -) Country and Industry. Only plotting the errors (+ or -)

Current Data Distribution



Thank you!

Alessandro Fosselard
Antoine Ameye

May 2025