“You could lose your job as a respectful director”

Empirical Study on Directors in US-Issued Companies

Group 7

Fatima Nurmakhamadova

Shuting Chen

Xiaolu Shen

Yumeng Xie

Content

- Research Methodology
- Descriptive Analysis
- Predictive Models
- Conclusions

Research Methodology

Research Steps

Research Assumptions & Hypotheses: identify the research scope and goal

Data Processing: feature engineering, data cleaning …

Descriptive Analysis: find out the patterns between departures and characteristics of individual directors

Predictive Model Analysis: compare the classification models' performances and conclude reliable statistical results

Research Assumptions & Scope

Assumption 1: all the controversial events recorded in the yearly total counts are negative events for the companies.

Assumption 2: if one director left the company in the year when his/her former company was recorded with one or more controversial events, his/her leave is regarded as a departure order from the company.

Assumption 3: If one director left the company in the year when his/her former company was recorded with NO controversial events, his/her leave is NOT regarded as a departure order from the company.

With our self-designed algorithm,
24,821 directors were successfully identified as “departed” or “non-departed”

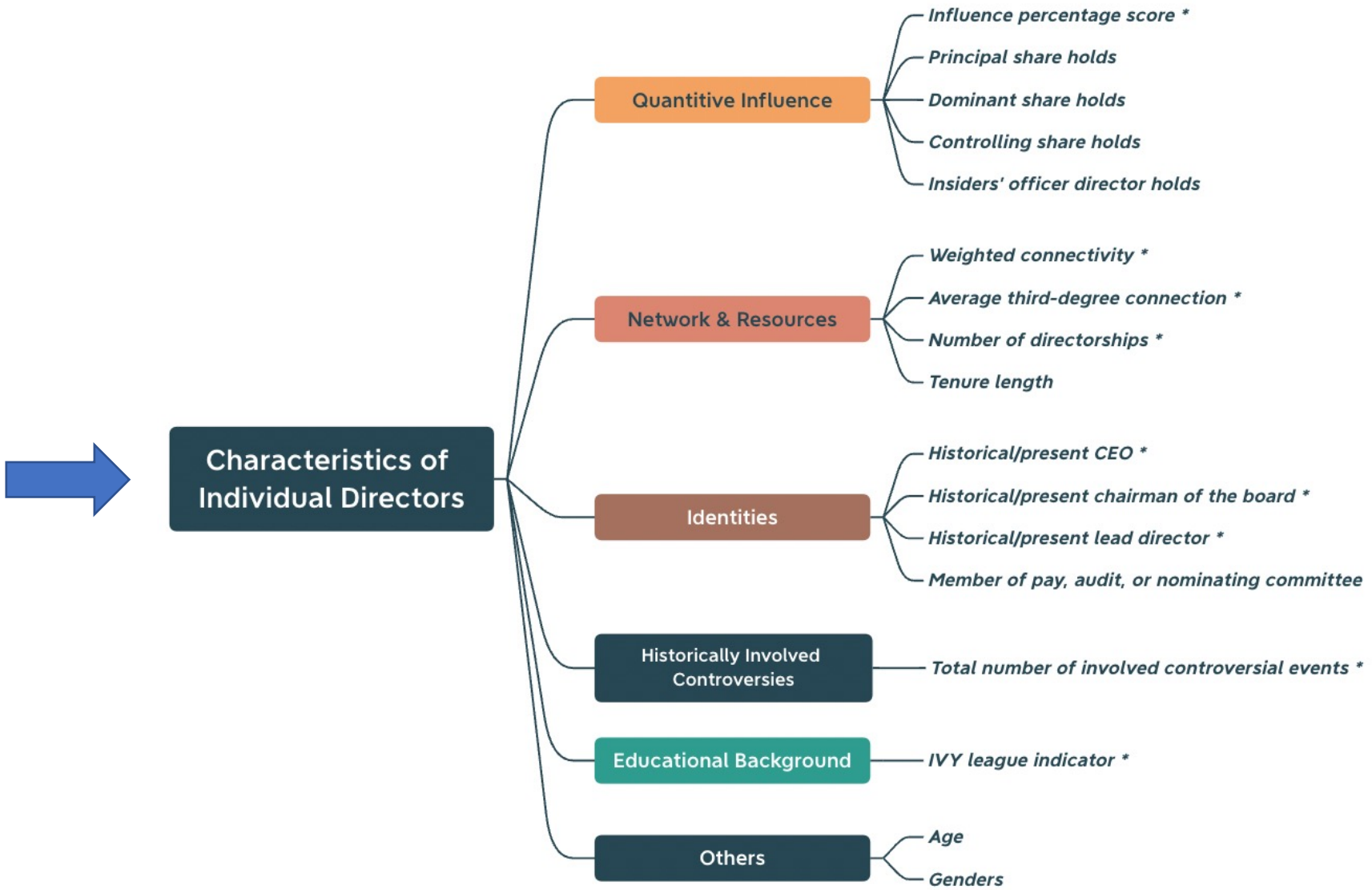
To obtain the best-balanced sample data with enough feature information,
4,155 observations of directors in US-issued companies are applied for further analysis
(361 positive targets and 3,794 negative targets)

Research Goal &Hypotheses

Project Goal:

How each characteristic of the individual director influences the departure when their companies involve controversial events.

16 Hypotheses were formed to be tested and answer:

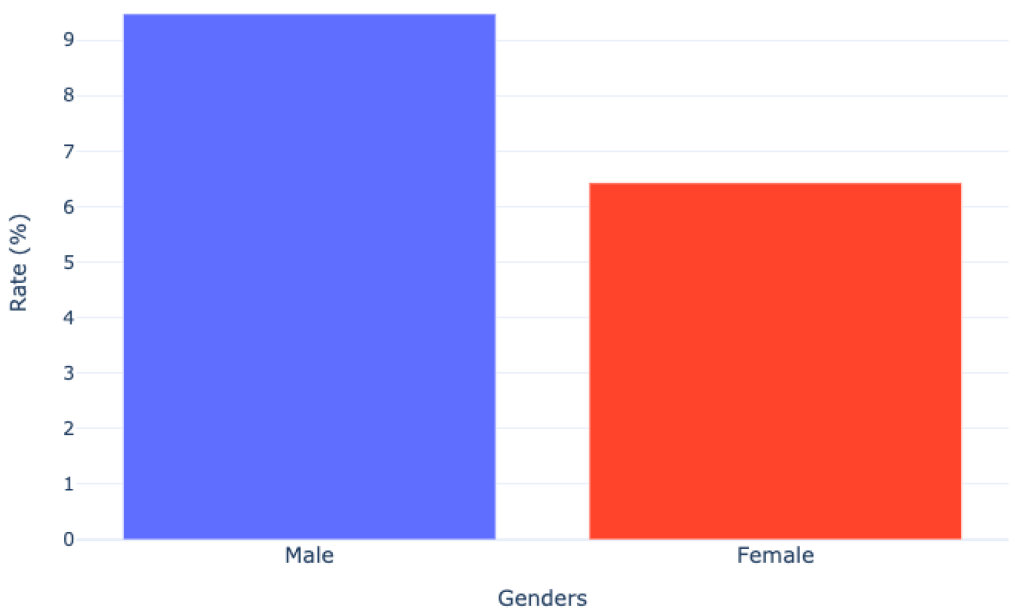


Note. Features with asterisks are generated from feature engineering processes.

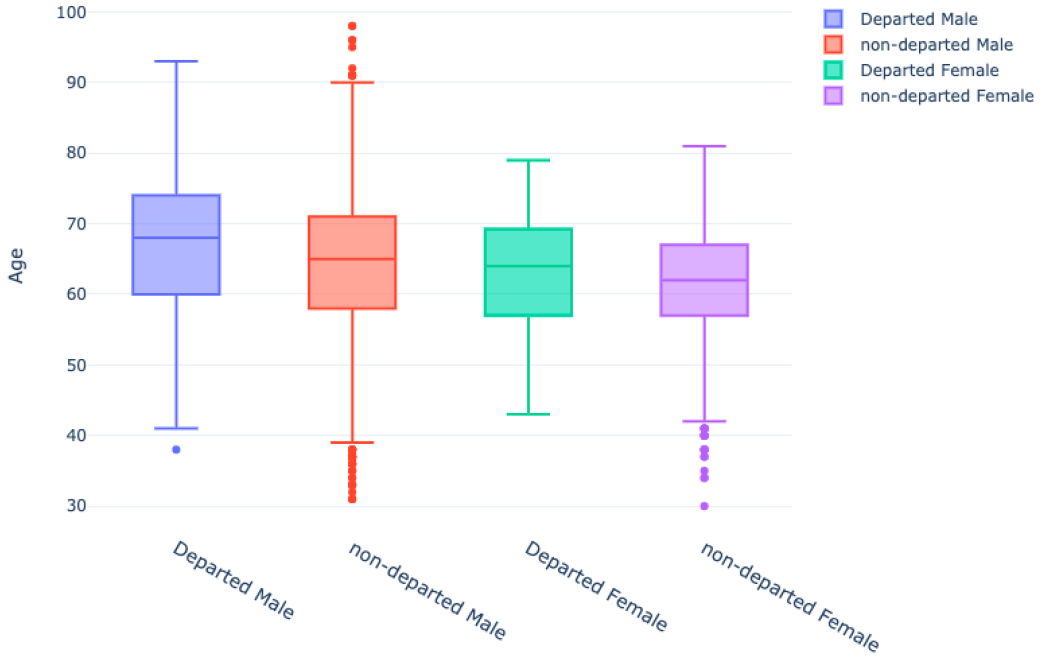
Descriptive Analysis

I. Departure VS Age – Genders

Departure Rates of Genders



Age Distributions Grouped by Genders & Departure



In US-issued companies (2017–2022),

- Male directors have a higher departure rate (9.47%) than female directors (6.43%);
- The age distributions of departed directors (in both genders) are higher than the non-departed directors;
- Directors under 40 years old were quite ”safe”.

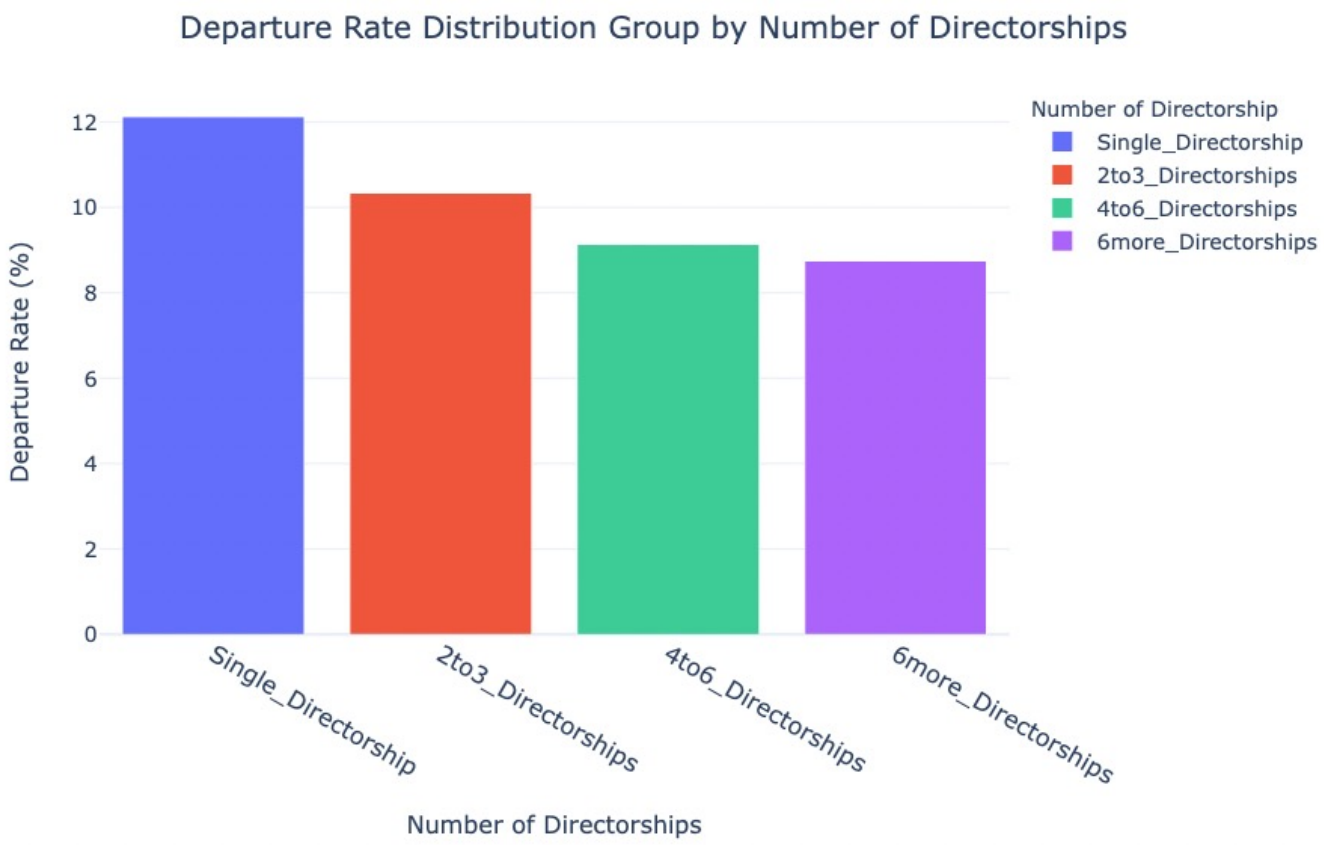
II. Departure VS Identities

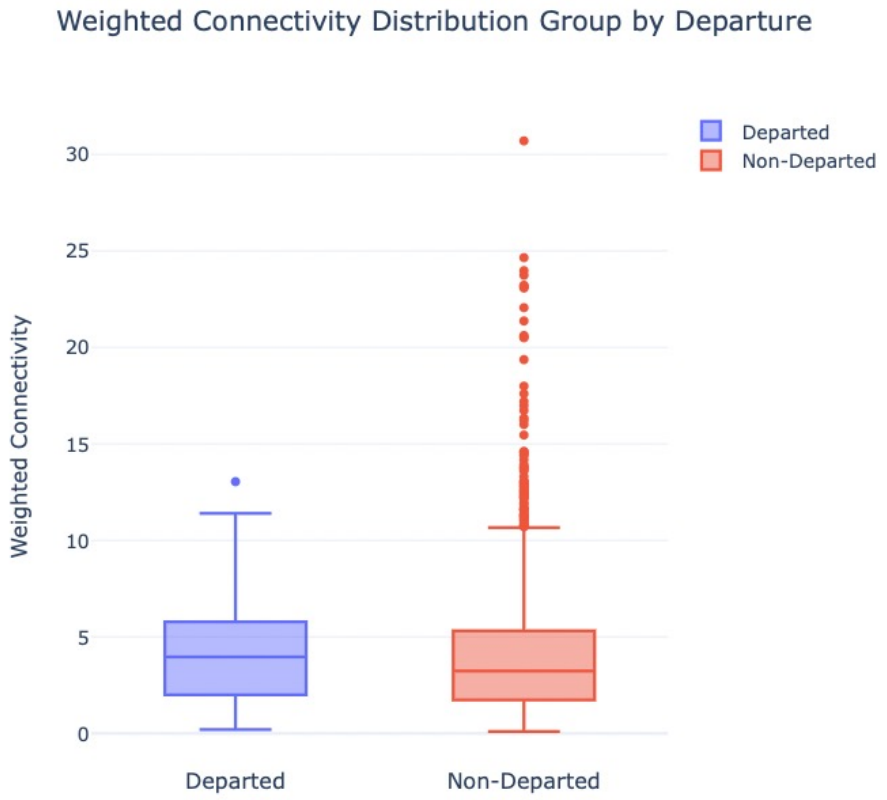
	Departure Rate (%)		
Identity	Yes	No	Difference
Pay Committee Member	0.10	11.48	-11.38
Nominating Committee Member	0.39	11.35	-10.96
Audit Committee	0	11.69	-11.69
Company Founder	6.54	8.74	-2.2
past or present CEO	10.30	8.47	1.83
past or present Chairman	9.83	8.53	1.3
past or present lead director	10.27	8.56	1.71

In US-issued companies (2017–2022),

- Directors having positions in the committees were very unlikely to be departed;
- Company founders are less likely to be kicked out (not a promise).
- Past or present important leaders of companies have more but slight chance (less than 2%) to be departed;

III. Departure VS Resources



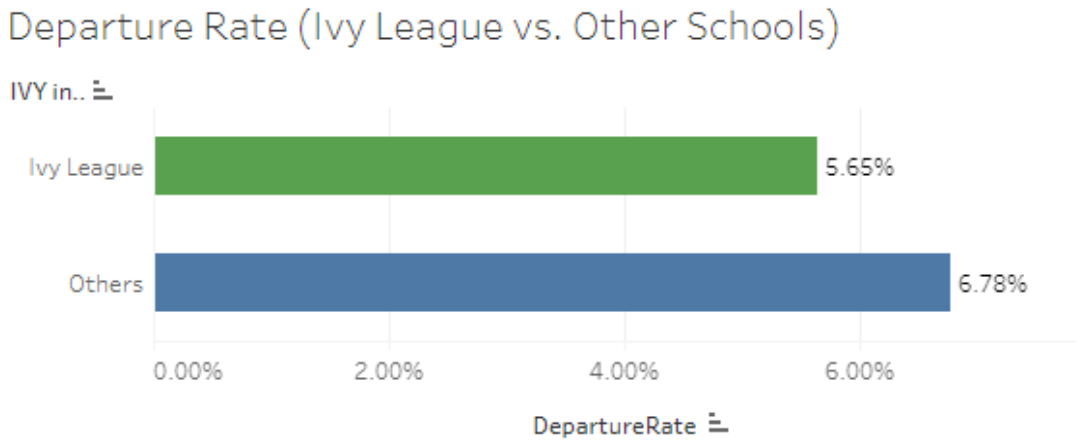


For the board directors in US-issued companies (2017—2022),

- More directorships, less likely to be departed – having more resources that the company needs;
- Outliers in weighted connectivity indicate that directors with extremely high weighted connectivity (greater than 11) are unlikely to be departed.

IV. Departure on Different Education Backgrounds

Of 151 people who were departed from the company, 38 are graduates from Ivy League Schools and 113 are not.



In US-issued companies (2017—2022):

- Ivy League graduates have slightly less departure rate (5.65%) than other school graduates have (6.78%)

Predictive Model Analysis

AUTO ML & Logistic Regression

AUTO ML

PYCARET

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
rf	Random Forest Classifier	0.9071	0.8620	0.1656	0.4536	0.2395	0.2021	0.2322	0.591
et	Extra Trees Classifier	0.8984	0.8692	0.1696	0.3668	0.2292	0.1827	0.1996	0.406
gbc	Gradient Boosting Classifier	0.8808	0.8584	0.2759	0.3186	0.2921	0.2280	0.2304	1.001
dt	Decision Tree Classifier	0.8762	0.6688	0.4161	0.3423	0.3735	0.3059	0.3088	0.053
ada	Ada Boost Classifier	0.8567	0.8512	0.4339	0.3001	0.3528	0.2759	0.2828	0.308
lr	Logistic Regression	0.7069	0.8453	0.8931	0.2213	0.3545	0.2459	0.3464	1.272
knn	K Neighbors Classifier	0.7054	0.6122	0.4301	0.1356	0.2059	0.0810	0.1024	0.072
lda	Linear Discriminant Analysis	0.6989	0.8416	0.9272	0.2216	0.3574	0.2484	0.3569	0.045
nb	Naive Bayes	0.6072	0.8230	0.9612	0.1818	0.3057	0.1821	0.3058	0.027
qda	Quadratic Discriminant Analysis	0.5698	0.7787	0.9656	0.1723	0.2917	0.1642	0.2870	0.031

AUTO ML


```
# Compares different models depending on their performance metrics. By default sorted by ac
#Exclude 'dummy', 'ridge','svm' models as they have the AUC score of 0.000 because it is no
best_model = compare_models(n_select = 5, exclude= ['dummy', 'ridge','svm','lightgbm']) #
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
rf	Random Forest Classifier	0.9083	0.8557	0.1779	0.4796	0.2549	0.2174	0.2497	0.630
et	Extra Trees Classifier	0.8991	0.8667	0.1701	0.3751	0.2312	0.1855	0.2035	0.678
gbc	Gradient Boosting Classifier	0.8808	0.8586	0.2759	0.3186	0.2921	0.2280	0.2304	1.179
dt	Decision Tree Classifier	0.8739	0.6641	0.4083	0.3365	0.3663	0.2975	0.3006	0.056
ada	Ada Boost Classifier	0.8567	0.8512	0.4339	0.3001	0.3528	0.2759	0.2828	0.329
lr	Logistic Regression	0.7066	0.8454	0.8931	0.2211	0.3541	0.2455	0.3461	0.833
knn	K Neighbors Classifier	0.7054	0.6122	0.4301	0.1356	0.2059	0.0810	0.1024	0.073
lda	Linear Discriminant Analysis	0.6989	0.8416	0.9272	0.2216	0.3574	0.2484	0.3569	0.084
nb	Naive Bayes	0.6072	0.8230	0.9612	0.1818	0.3057	0.1821	0.3058	0.028
qda	Quadratic Discriminant Analysis	0.5701	0.7812	0.9612	0.1719	0.2910	0.1634	0.2851	0.031

Results: Logit Summary

Model:	Logit	Pseudo R-squared:	0.494
Dependent Variable:	y	AIC:	3801.5974
Date:	2022-12-03 22:27	BIC:	4005.6466
No. Observations:	5336	Log-Likelihood:	-1869.8
Df Model:	30	LL-Null:	-3698.6
Df Residuals:	5305	LLR p-value:	0.0000
Converged:	0.0000	Scale:	1.0000

Features	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	1.0951	0.3426	3.1964	0.0014	0.4236	1.7667
AGE	0.0249	0.0052	4.8224	0.0000	0.0148	0.0351
Tenure	0.0139	0.0111	1.2474	0.2123	-0.0079	0.0357
DirectorshipCount	-0.3786	0.0492	-7.7017	0.0000	-0.4749	-0.2822
AVG_IPS	0.0439	0.0077	5.7395	0.0000	0.0289	0.0590
AVG_Weighted_Connection	-0.0686	0.0310	-2.2153	0.0267	-0.1292	-0.0079
INSIDERS_OFFICERS_DIRECTORS_HELD_PCT	0.0145	0.0044	3.3043	0.0010	0.0059	0.0232
PRINCIPAL_SHAREHOLDER_PCT	-0.0063	0.0042	-1.5200	0.1285	-0.0145	0.0018
CONTROLLING_SHAREHOLDER_PCT	0.0018	0.0036	0.4920	0.6227	-0.0053	0.0088
Avg_3rd_connect_17_22	0.0456	0.0183	2.4902	0.0128	0.0097	0.0815
Ttl_controv_17_22	0.0162	0.0149	1.0879	0.2766	-0.0130	0.0453
GENDER_Male	-0.1826	0.0999	-1.8291	0.0674	-0.3784	0.0131
COMPANY_FOUNDER_T	-1.2018	0.3999	-3.0056	0.0027	-1.9855	-0.4181
EXEC_OR_NON-EXEC_Non-Executive	0.8415	0.2784	3.0227	0.0025	0.2959	1.3872
HistIdentity_CEO_T	-2.4385	0.2149	-11.3447	0.0000	-2.8598	-2.0172
HistIdentity_Chairman_T	-1.1411	0.2148	-5.3129	0.0000	-1.5621	-0.7201
HistIdentity_LeadDirector_T	-1.5302	0.3137	-4.8783	0.0000	-2.1450	-0.9154
PayCommitteeMember_Positive	-6.0649	0.8003	-7.5780	0.0000	-7.6335	-4.4963
AuditCommitteeMember_Positive	-6.8575	0.8467	-8.0990	0.0000	-8.5170	-5.1980
NominatingCommitteeMember_Positive	-5.9681	0.8034	-7.4290	0.0000	-7.5427	-4.3936
INDEPENDENT_OF_MANAGEMENT_Yes	-0.8369	0.2325	-3.6003	0.0003	-1.2926	-0.3813
MULTIPLE_CLASSES_OF_VOTING_STOCK_Yes	-1.7029	0.2143	-7.9456	0.0000	-2.1230	-1.2829
OutsideRelatedReason_Positive	-2.2499	0.2995	-7.5130	0.0000	-2.8369	-1.6630
OutsideRelatedReason2_Positive	-0.2996	0.6590	-0.4546	0.6494	-1.5912	0.9920
CONTROLLED_VIA_STOCK_PYRAMID_Yes	0.5934	0.7041	0.8428	0.3993	-0.7866	1.9735
HAS_CORPORATE_PARENT_Yes	0.2339	0.5529	0.4231	0.6722	-0.8497	1.3176
IVY_indicator_True	-2.7702	0.2013	-13.7638	0.0000	-3.1647	-2.3757
IVY_indicator_UNKNOWN	-1.4904	0.0972	-15.3325	0.0000	-1.6809	-1.2999
CEO_leav_T	1.4923	0.2403	6.2098	0.0000	1.0213	1.9633
Chairman_leav_T	0.8665	0.2576	3.3632	0.0008	0.3615	1.3715
Lead_D_leav_T	-0.7514	0.4353	-1.7262	0.0843	-1.6046	0.1017

Logistic Regression Model

Performance Metrics:

Class	Precision	Recall	Accuracy
0 - not departed	0.95	0.75	
1 - departed	0.21	0.62	
			0.74

- only 22% of directors that were predicted as Departed are Departed
- 62% of all Departed directors were identified correctly

LR Feature Importance

- the odds of departures increase (positive impacts)
- the odds of departures decrease (negative impacts)

Name	Odds ratio
Constant/Intercept	2.98

Numeric Features

Name	Odds ratio
Avg 3rd degree connection	1.046
Avg Influence share %	1.044
Age	1.02
Insiders Officers Directors Held %	1.01
Avg Weighted Connection (edge)	0.93
Directorship Count *	0.68

* Departure likelihood gets 32% reduced with 1 more directorship

* Being the CEO and Chairman at the moment of controversial events increases their departure likelihoods to 2 - 4 times

Categorical Features

Feature	Odds ratio
CEO while leaving – True *	4.44
Chairman while leaving – True *	2.37
Non-Executive	2.31
Independent of Management - Yes	0.43
Was a Chairman while on the board - True	0.31
Company Founder - True	0.30
Was a Lead Director while on the board - True	0.21
Multiple Classes Of Voting Stock - Yes	0.18
Outside Related Reason - Positive	0.10
Was a CEO while on the board - True	0.08
Ivy League Alumni - True	0.06
Nominating Committee Member – Positive	0.0025
Pay Committee Member - Positive	0.0023
Audit Committee Member - Positive	0.001

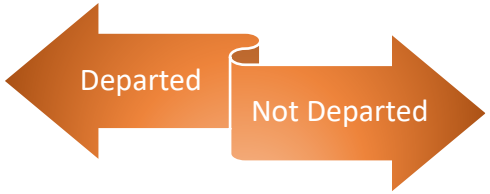
* * Directors with these characteristics (blue cells) are at lower risk to be departed

Predictive Model Analysis The Second Approach

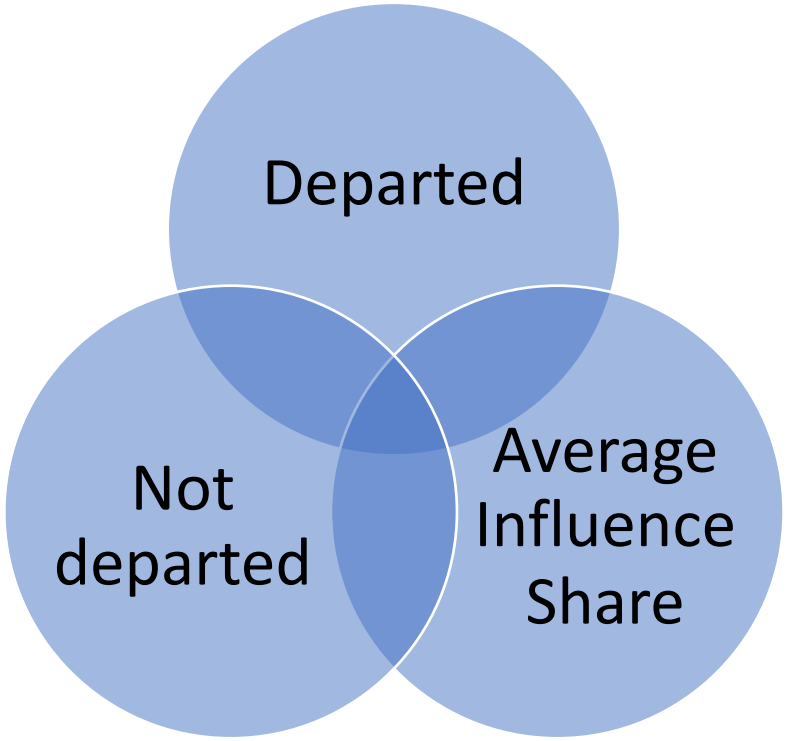
Models & SMOTE

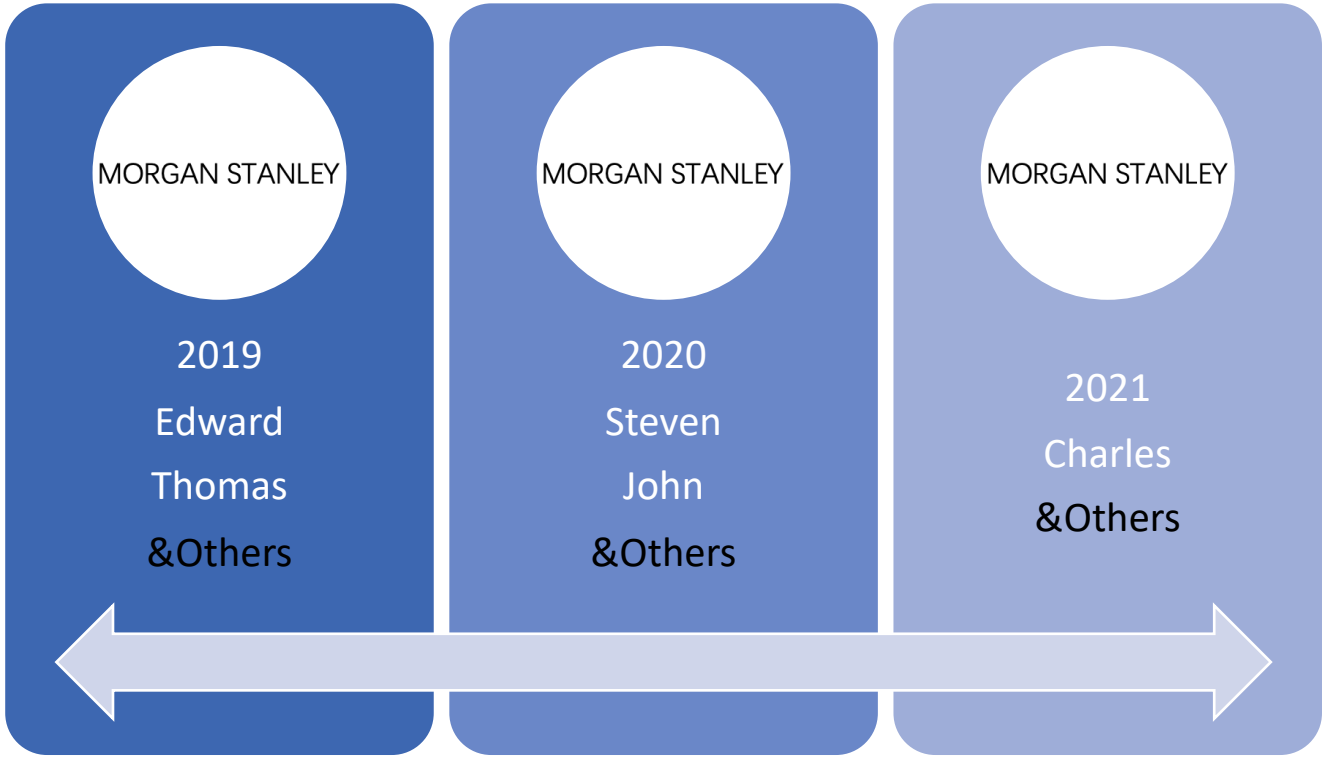
Data Structure

Study by case



- Each case is mutural exclusive.
- The influence share remains the year fraud happened.
- Explore more about the difference between two groups of people





Feature Selection

- 30380 rows × 28 columns

pct_share	EXEC_OR_NON-EXEC	OWNERSHIP_CATEGORY
that_influence_year_fraud_happened	OUTSIDE_RELATED_REASON	MULTIPLE_CLASSES_OF_VOTING_STOCK
that_year_share	OUTSIDE_RELATED_REASON_2	PRINCIPAL_SHAREHOLDER_PCT
AGE	INDEPENDENT_OF_MANAGEMENT	CONTROLLING_SHAREHOLDER_PCT
GENDER	HAS_CORPORATE_PARENT	GICS_SUB_IND
Tenure	CONTROLLED_VIA_STOCK_PYRAMID	Sector
PAY_COMMITTEE_MEMBER	DOMINANT_SHAREHOLDER_PCT	target
AUDIT_COMMITTEE_MEMBER	FAMILY_FIRM	3rd_median
NOMINATING_COMMITTEE_MEMBER	FOUNDER_FIRM	
COMPANY_FOUNDER	INSIDERS_OFFICERS_DIRECTORS_HELD_PCT	

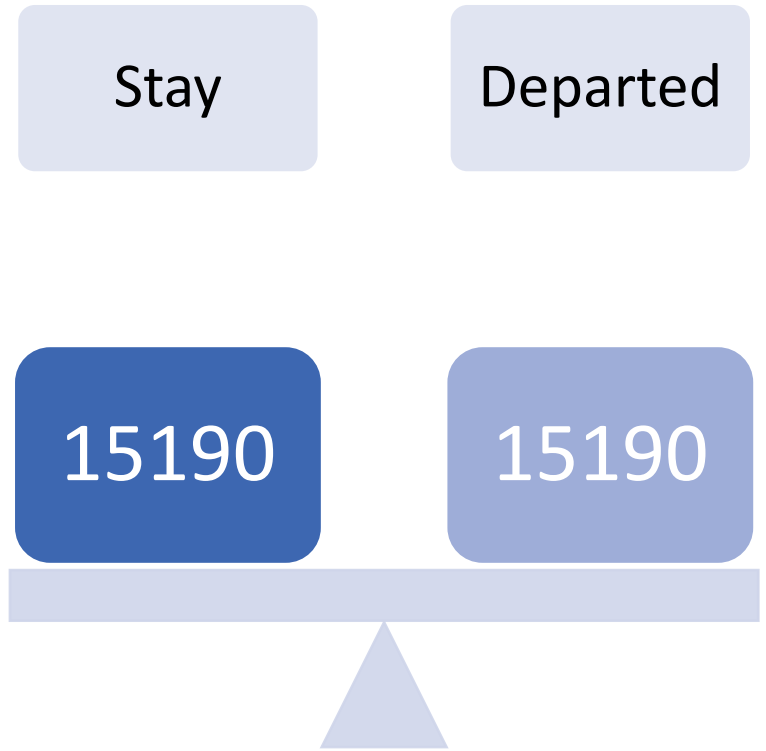
Balance data by SMOTE

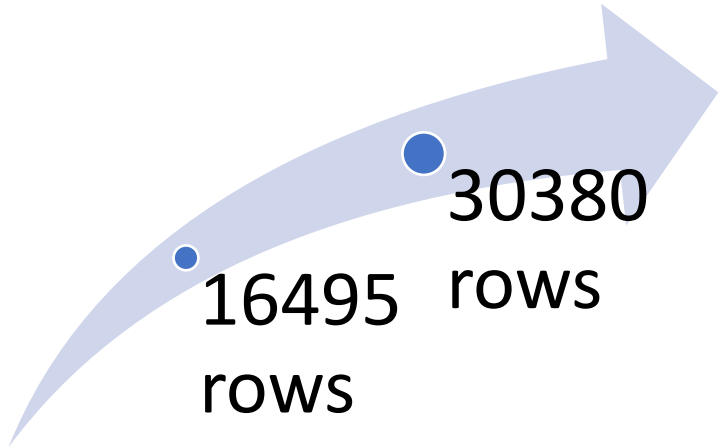
Synthetic Minority Oversampling Technique

Balance data by SMOTE

Synthetic Minority Oversampling Technique

30380 rows × 28 columns





Logistic Regression

- 79% accuracy after SMOTE
- 74% precision on predicting departed members
- PAY_COMMITTEE_MEMBER has highest coef

Before

	precision	recall	f1-score	support
0.0	0.93	0.99	0.96	5332
1.0	0.48	0.07	0.12	442
accuracy			0.92	5774
macro avg	0.70	0.53	0.54	5774
weighted avg	0.89	0.92	0.90	5774

After

	precision	recall	f1-score	support
0.0	0.87	0.67	0.76	5276
1.0	0.74	0.90	0.81	5357
accuracy			0.79	10633
macro avg	0.81	0.79	0.79	10633
weighted avg	0.80	0.79	0.79	10633

PAY_COMMITTEE_MEMBER	-4.3702
that_year_share	-4.1945
INDEPENDENT_OF_MANAGEMENT	3.7536
OUTSIDE_RELATED_REASON	3.0710
NOMINATING_COMMITTEE_MEMBER	-3.0643
EXEC_OR_NON-EXEC	-2.5090
3rd_median	2.4655
AGE	-2.2262
pct_share	1.9199
DOMINANT_SHAREHOLDER_PCT	1.3273
COMPANY_FOUNDER	-0.9227
Tenure	0.6717

Decision Tree

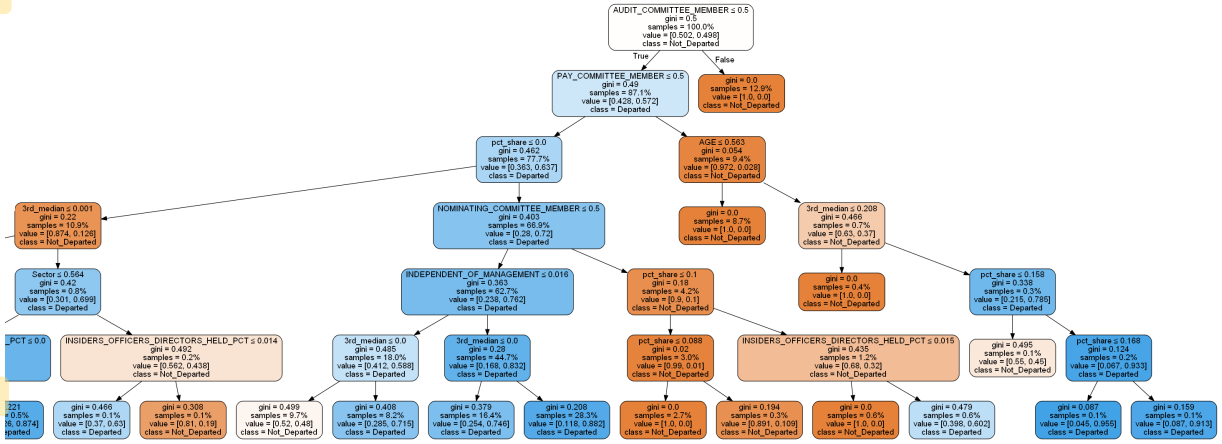
- 85% accuracy after SMOTE
- 82% precision on predicting departed members

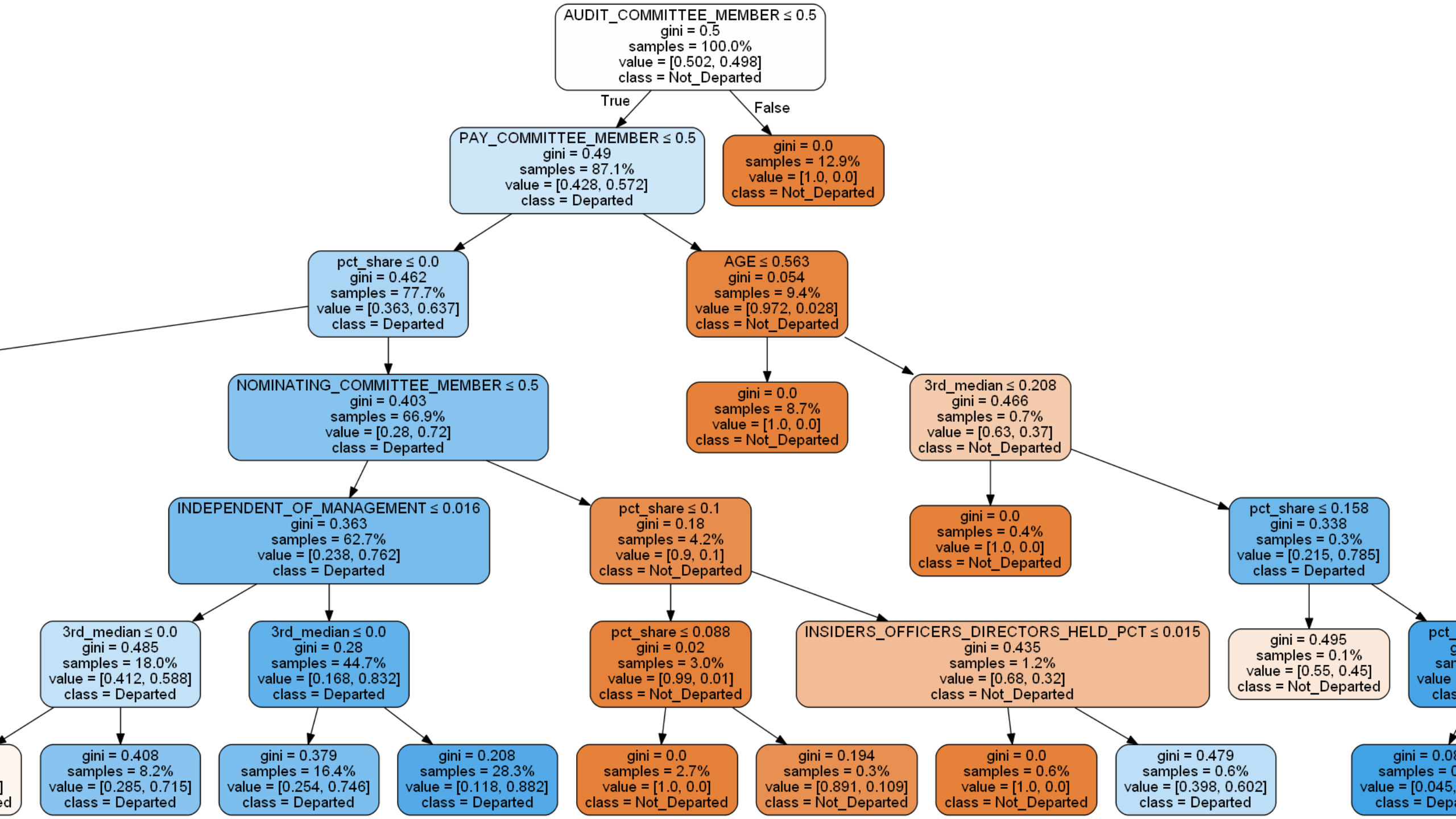
Before

	precision	recall	f1-score	support
0.0	0.93	0.99	0.96	5332
1.0	0.53	0.10	0.17	442
accuracy			0.92	5774
macro avg	0.73	0.55	0.57	5774
weighted avg	0.90	0.92	0.90	5774

After

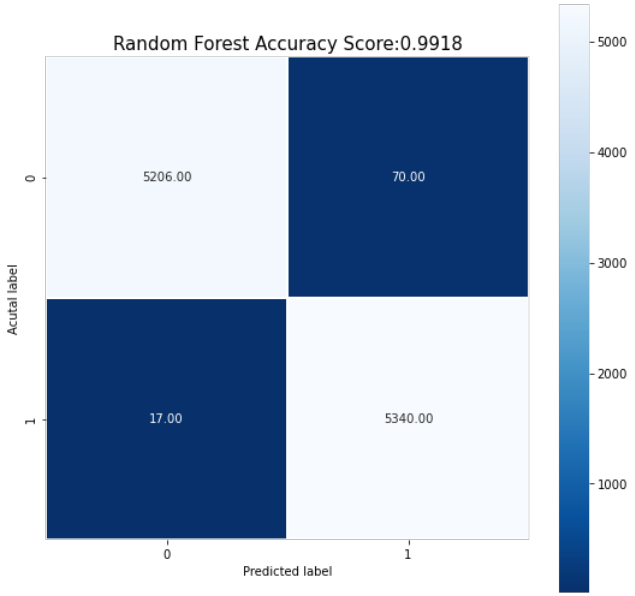
	precision	recall	f1-score	support
0.0	0.88	0.80	0.84	5276
1.0	0.82	0.90	0.86	5357
accuracy			0.85	10633
macro avg	0.85	0.85	0.85	10633
weighted avg	0.85	0.85	0.85	10633

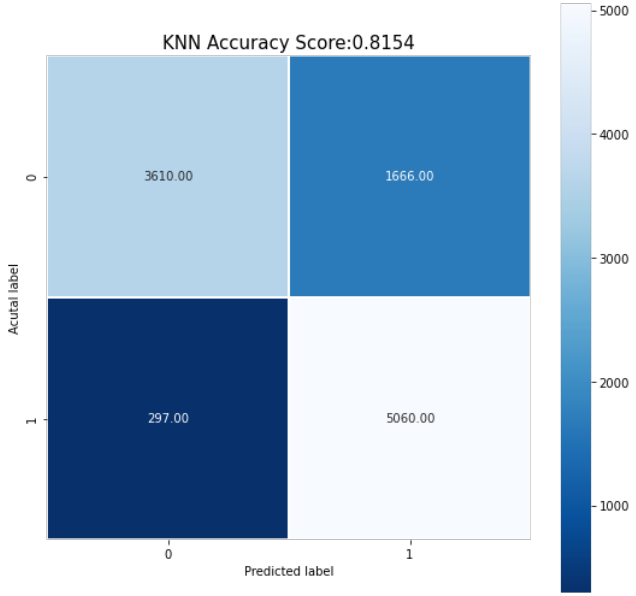


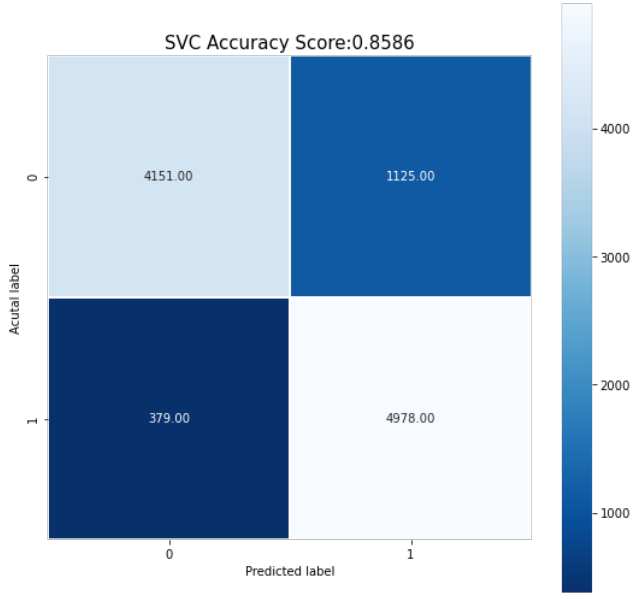

AUDIT_COMMITTEE_MEMBER ≤ 0.5
gini = 0.5
samples = 100.0%
value = [0.502, 0.498]
class = Not_Departed
True
False
PAY_COMMITTEE_MEMBER ≤ 0.5
gini = 0.49
samples = 87.1%
value = [0.428, 0.572]
class = Departed
gini = 0.0
samples = 12.9%
value = [1.0, 0.0]
class = Not_Departed
pct_share ≤ 0.0
gini = 0.462
samples = 77.7%
value = [0.363, 0.637]
class = Departed
AGE ≤ 0.563
gini = 0.054
samples = 9.4%
value = [0.972, 0.028]
class = Not_Departed
NOMINATING_COMMITTEE_MEMBER ≤ 0.5
gini = 0.403
samples = 66.9%
value = [0.28, 0.72]
class = Departed
gini = 0.0
samples = 8.7%
value = [1.0, 0.0]
class = Not_Departed
3rd_median ≤ 0.208
gini = 0.466
samples = 0.7%
value = [0.63, 0.37]
class = Not_Departed
INDEPENDENT_OF_MANAGEMENT ≤ 0.016
gini = 0.363
samples = 62.7%
value = [0.238, 0.762]
class = Departed
pct_share ≤ 0.1
gini = 0.18
samples = 4.2%
value = [0.9, 0.1]
class = Not_Departed
gini = 0.0
samples = 0.4%
value = [1.0, 0.0]
class = Not_Departed
pct_share ≤ 0.158
gini = 0.338
samples = 0.3%
value = [0.215, 0.785]
class = Departed
3rd_median ≤ 0.0
gini = 0.485
samples = 18.0%
value = [0.412, 0.588]
class = Departed
3rd_median ≤ 0.0
gini = 0.28
samples = 44.7%
value = [0.168, 0.832]
class = Departed
pct_share ≤ 0.088
gini = 0.02
samples = 3.0%
value = [0.99, 0.01]
class = Not_Departed
INSIDERS_OFFICERS_DIRECTORS_HELD_PCT ≤ 0.015
gini = 0.435
samples = 1.2%
value = [0.68, 0.32]
class = Not_Departed
gini = 0.495
samples = 0.1%
value = [0.55, 0.45]
class = Not_Departed
gini = 0.408
samples = 8.2%
value = [0.285, 0.715]
class = Departed
gini = 0.379
samples = 16.4%
value = [0.254, 0.746]
class = Departed
gini = 0.208
samples = 28.3%
value = [0.118, 0.882]
class = Departed
gini = 0.0
samples = 2.7%
value = [1.0, 0.0]
class = Not_Departed
gini = 0.194
samples = 0.3%
value = [0.891, 0.109]
class = Not_Departed
gini = 0.0
samples = 0.6%
value = [1.0, 0.0]
class = Not_Departed
gini = 0.479
samples = 0.6%
value = [0.398, 0.602]
class = Departed

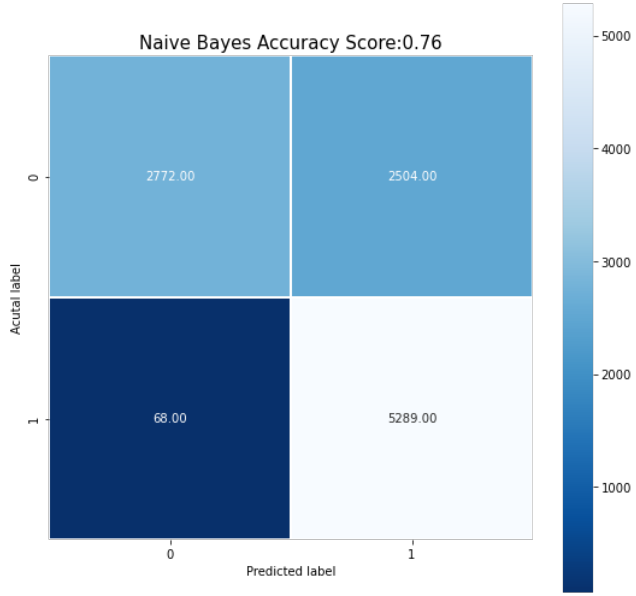
Random Forest & SVC & KNN & Naive Bayes

Random Forest	SVC	KNN	Naïve Bayes
99%	86%	82%	76%



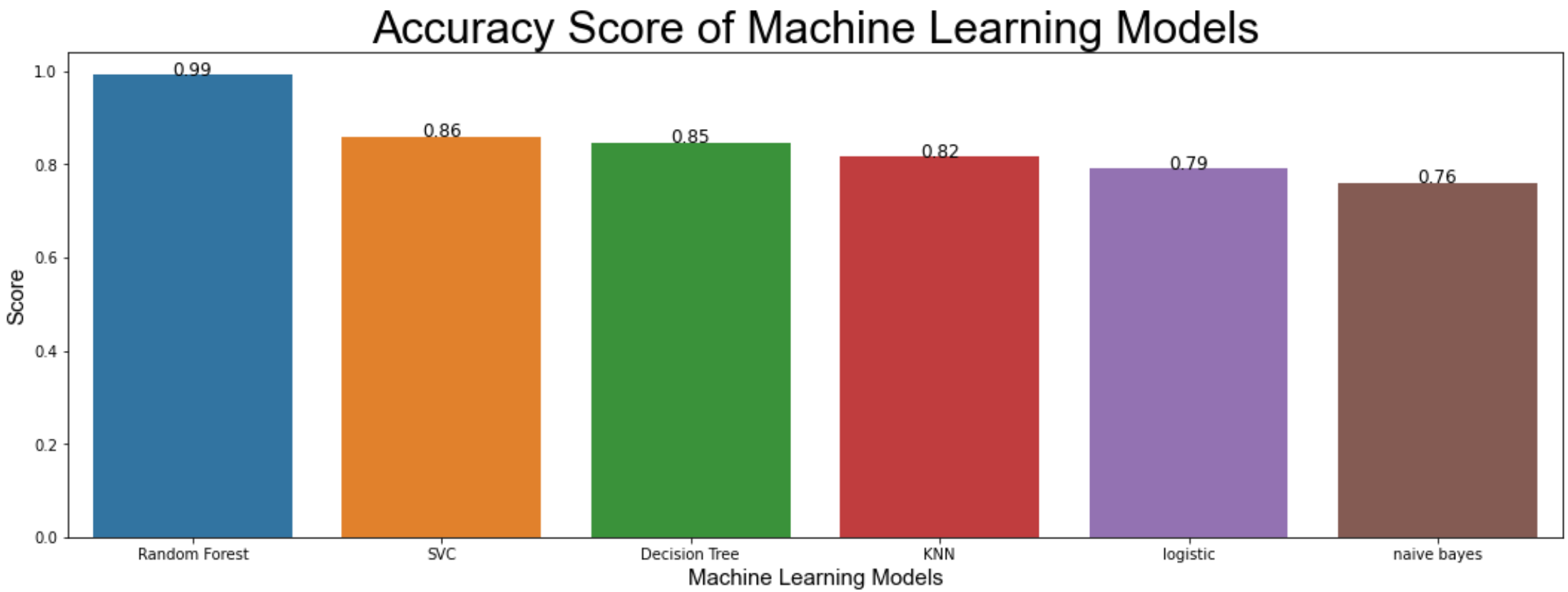


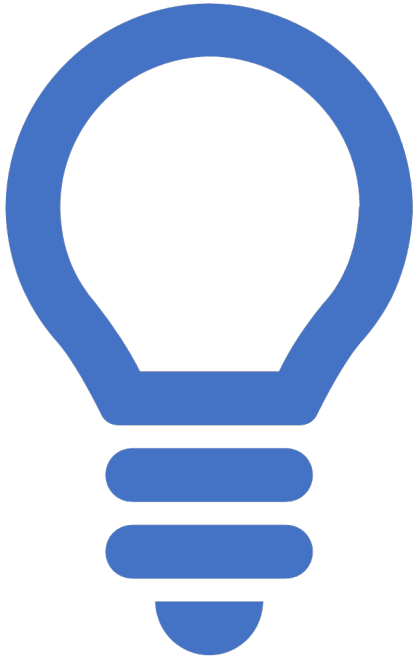




Bar Chart to Compare All Machine Learning Models

- 30380 rows × 28 columns





Conclusions

Identities – Departures

i. Being the CEO or Chairman of the board during controversial events have significantly **positive** impacts on departure results:

their departure likelihoods are 2 to 4 times compared to the directors who are not in the positions.

ii. Used to be CEOs, Chairmen, or lead directors before the controversial events have significantly **negative** impacts on departure results:

their departure likelihoods are reduced to less than 30% compared to those who had never been in these positions.

iii. Being a member of nominating, pay, or audit committees have significantly **negative** impacts on departure results:

their departure likelihoods are less than 1% of whom are not in these committees during the controversial events.

Note. The alpha level is at 0.05.

Network and Resources – Departures

i. The number of third-degree connections has significant **positive** impacts:

The departure likelihood gets 50% to 60% more when third-degree connection increases by 10.

ii. The number of directorships has significant **negative** impacts on departure results:

The departure likelihood gets around 30% reduced with one more directorship in different company.

Note. The alpha level is at 0.05.

Quantitive Influence – Departure

i. The influence percentage has significant **positive** impacts on departure results:

The departure likelihood increases about 40% with 10% additional influence percentage.

ii. Insider officers' directors holds have significant **positive** impacts on departure results:

The departure likelihood increases about 10% with 10% additional insider officer directors holds.

Note. The alpha level is at 0.05.

Educational Background – Departure

Graduating from IVY league colleges has significantly **negative** impacts on departure results:

Their departure likelihoods are less than 10% of the directors who have never been to these colleges.

Age – Departure

Age has significant **positive** impacts on the departure results:

The departure likelihood gets 20% increased with ten years older.

Note. The alpha level is at 0.05.

Predictive Modeling

The classification method of **Random Forest** turned out to have the highest accuracy in predicting the departures of directors when their companies involved controversial events regarding both data structures we applied.

Further improvement of the predictive model could be achieved by following aspects:

i. Overcome the imbalance distributions of independent variables;

ii. Variable control in sectors, capital levels, etc.;

iii. Combining the two strategies of processing IPS in different structures;

iv. Modifying the assumptions and target-identifying algorithm with additional and more detailed data sources.

Thank you for your listening

Open to any questions!

References:

- Ali, M. (2020). *Classification*. Classification - pycaret 3.0.0 documentation. Retrieved December 5, 2022, from https://pycaret.readthedocs.io/en/latest/api/classification.html

- Martin, K. G. (2022, October 14). *Why use odds ratios in logistic regression?* The Analysis Factor. Retrieved December 5, 2022, from https://www.theanalysisfactor.com/why-use-odds-ratios/

Appendix A: Research Hypotheses

- **H1.1**: the **influence percentage score** of the director has a negative influence on departures when controversial events happen (false, significantly positive influence)
- **H1.2**: the percentage of **dominant share holds**, **insiders' official directors' holds**, **principal share holds**, and **controlling share holds** all have a negative influence on departures when controversial events happen. (insiders' official directors' holds is true while others are false)

- **H2.1**: having positions in the **pay committee**, **audit committee**, and **nominating committee** negatively influences the likelihood of departures when controversial events happen compared to the situation of not having positions in these committees. (True)
- **H2.2:** directors who used to be or currently are **chairmen of the board**, **lead directors**, or **CEOs** have a lower likelihood to be departed than the directors who are not these characters in companies. (CEO & chairman true; lead directors false)
- **H2.3**: directors who are **company founders** or **executive directors** are less likely to be departed than the directors who do not have these identities. (True)

Appendix A: Research Hypotheses

- **H3.1**: the **weighted connectivity** of the director has a negative influence on the departure when the company involves controversial events. (false)
- **H3.2**: the **average third-degree connection** of the director has a negative influence on the departure when the company involves controversial events. (False, significant positive influence)
- **H3.3**: the **number of directorships** across different companies has a negative impact on the departure when controversial events happen. (True)
- **H3.4**: the **tenure length** has a negative impact on the departure when controversial events happen. (false)

- **H4.1**: the option of **multiple classes of voting stock** has negative impacts on the departure when controversial events happen. (True)
- **H4.2**: the option of **independent management** has negative impacts on the departure when controversial events happen. (True)

Appendix A: Research Hypotheses

- **H5**: having the **educational background of the IVY league colleges** has a negative influence on the departure when controversial events happen. (True)

- **H6**: the **male director** has a lower likelihood to be departed when controversial events happen than the female director. (False)