Predicting Corporate Bond Ratings

Chadwick Balloo, Flatiron School

Scope and Purpose

- Moody's rates corporate bonds on a scale of Aaa to C (1-21) based on credit issue and borrower quality

- Using the issuer's financial data and relevant statistics, can we predict a bond's credit rating?

- Given current ratings, how can we use predicted ratings to profit?

The Data

-Bond Data - Markets.BusinessInsider.com - 6,130 corporate bonds

-Stock Data - iShares by BlackRock - ishares.com - 439 stocks from the Russell 1000

-Financial Data - Yahoo Finance - financial ratios and statistics

The Process

Data Cleaning and EDA > **Modeling** > **Refining and Evaluation**

Technologies:

-pandas

-numpy

-seaborn

-yellowbrick

-sklearn

Models:

-ordinal log reg

-random forest

-knn

-kmeans

Methods:

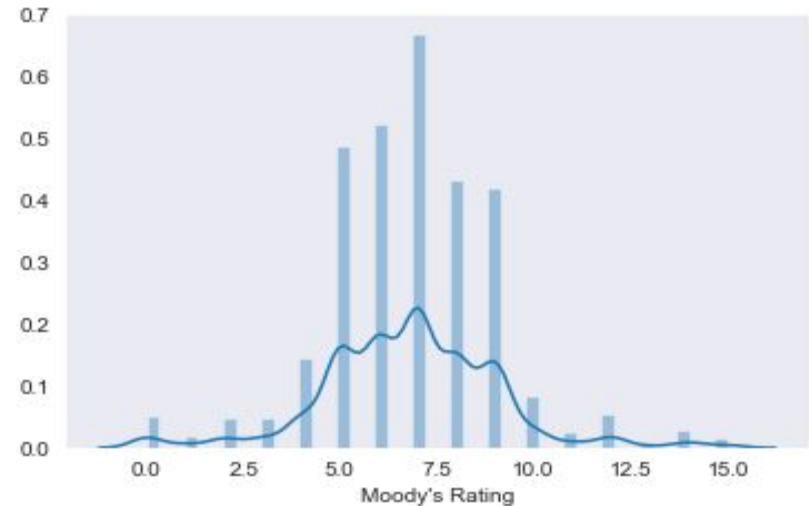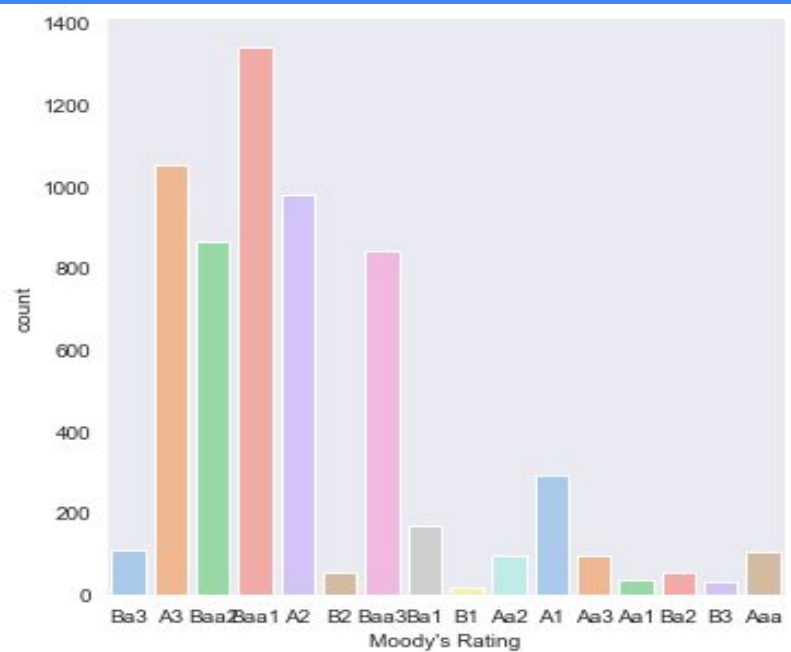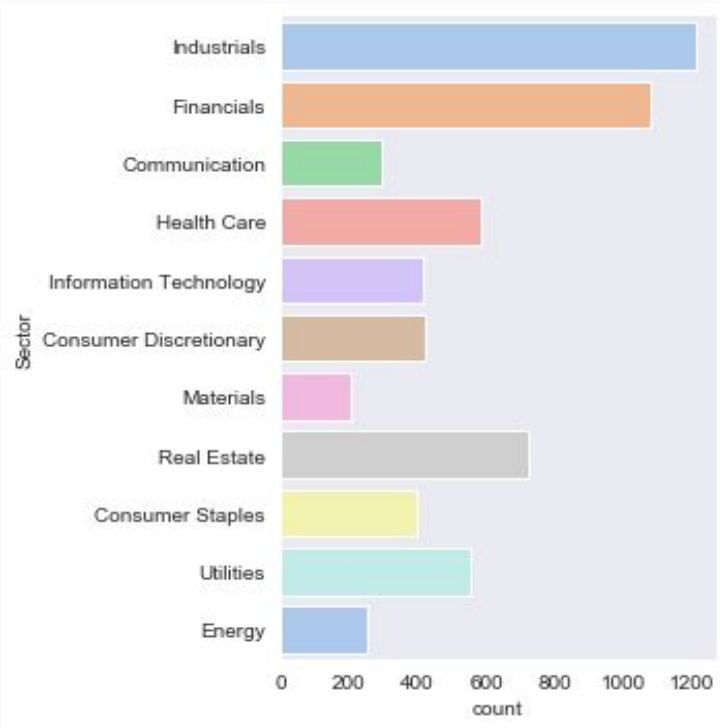-distance maps

-feature importance

-class balance

Exploring the Data

- Breakdown by rating

- Baa1, A2, A3, Baa2, Baa3,

  - "High ability or acceptable ability to repay short term debt"

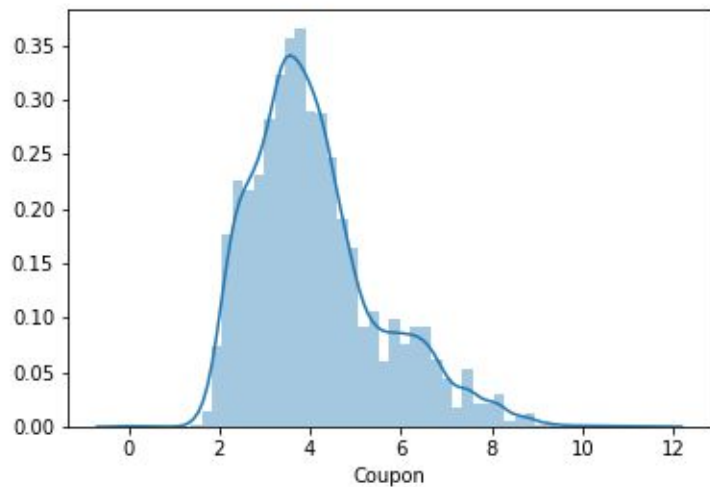 - Indicates shift toward higher debt levels and riskier debt

Sector Breakdown

- Industrials, Financials, Real Estate, HC, Utilities

- Typical of how these sectors conduct business

Coupon and Yield

Skewed to the right, higher coupons for riskier debt

Moody's Rating	Coupon	Yield
A1	3.472656	2.438255
A2	3.356087	2.469162
A3	3.928774	2.635681
Aa1	3.066303	2.310270
Aa2	3.579974	2.385795
Aa3	4.013421	2.621684
Aaa	3.511478	2.346516
B1	6.356944	3.900589
B2	6.495536	5.457689
B3	6.687500	7.535608
Ba1	4.983394	3.770574
Ba2	5.463113	3.719412
Ba3	5.622248	4.311754
Baa1	4.303829	2.817916
Baa2	4.492211	3.085166
Baa3	4.226217	3.060150

Ordinal Log Reg

- Performed poorly, likely due to number of classes

- MAE:
  - 1.75
  - 2.4 baseline

KNN

- Performed poorly for large dataset

  - Mean CV Score: .46

- Performed well for smaller dataset (~1800)
  - Mean CV Score: .83

KNeighborsClassifier Classification Report

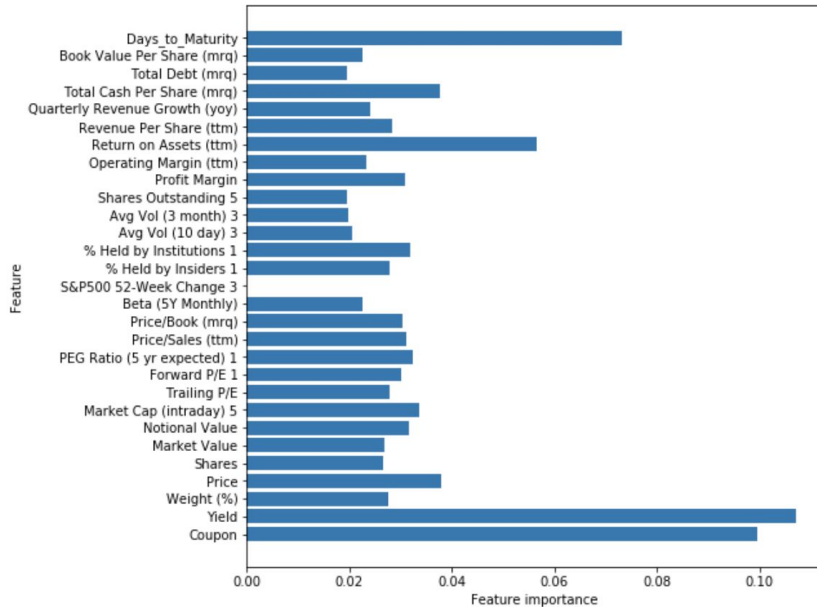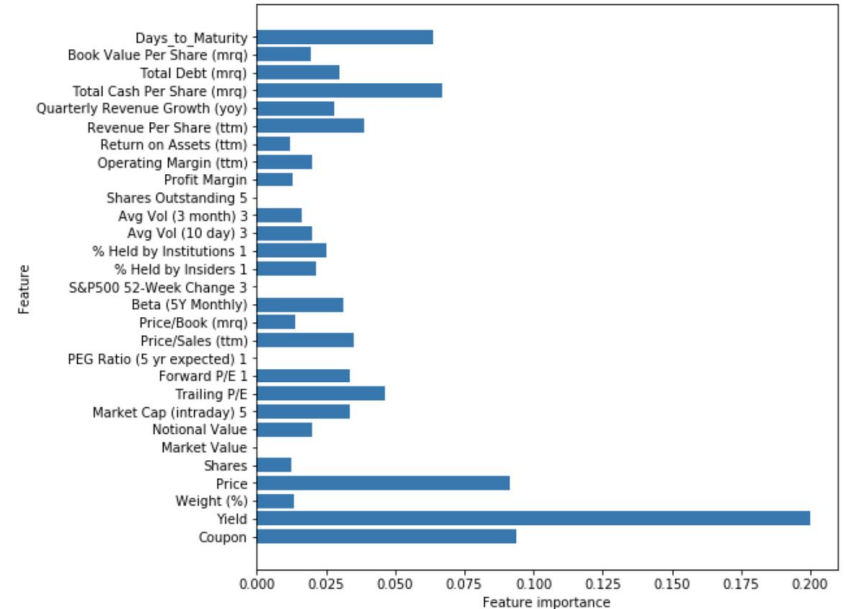|	precision	recall	f1	support
B1	0.615	0.505	0.554	218
Aa1	0.729	0.611	0.665	211
B3	0.647	0.576	0.609	356
Ba2	0.300	0.429	0.353	21
Aaa	0.500	0.700	0.583	10
Aa2	0.762	0.842	0.800	38
Aa3	0.500	0.111	0.182	9
A1	0.250	0.077	0.118	13
B2	0.200	0.250	0.222	4
Baa1	0.958	1.000	0.979	23
A3	0.560	0.438	0.491	32
Baa3	0.727	0.640	0.681	25
Ba1	1.000	1.000	1.000	7
Baa2	0.489	0.611	0.543	244
Ba3	0.498	0.595	0.542	252
A2	0.671	0.757	0.711	70

Random Forest

- Accuracy Score: .74
  - Baseline: .42
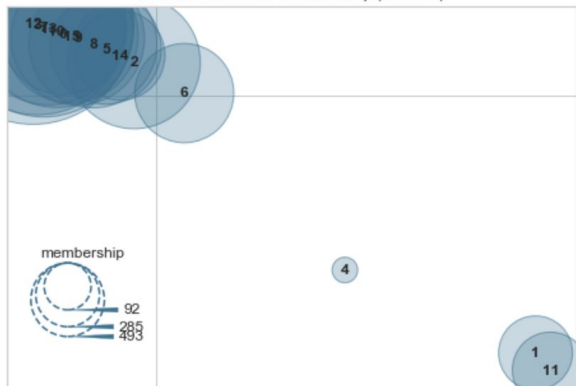- Also worked better on smaller data set

Forest

Tree

Credit Analysis

Select a ticker:

	VZ ✕	GOOGL ✕	PEP ✕		⊗

	Predicted Rating	Moody's Rating	Ticker	Name	Coupon
34	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	6.2500
35	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	3.8500
36	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	4.4000
37	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	2.9460
38	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	8.9500
39	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	7.7500
40	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	6.4000
41	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	5.0500
42	Aa2	Baa1	VZ	VERIZON COMMUNICATIONS…	4.9500
43	A3	Aa2	GOOGL	ALPHABET INC CLASS A	3.6250
44	A3	Aa2	GOOGL	ALPHABET INC CLASS A	3.3750

KMeans Clustering

KMeans Intercluster Distance Map (via MDS)

Predicted	Moody's Rating	Ticker	Name
A3	Aaa	JNJ	JOHNSON & JOHNSON
Ba3	A3	UTHR	UNITED THERAPEUTICS CORP
Baa1	Baa2	KMI	KINDER MORGAN INC
Ba1	A3	MS	MORGAN STANLEY
B3	A2	LLY	ELI LILLY

Predicted	Moody's Rating	Ticker	Name
Ba3	Baa3	DISCK	DISCOVERY INC SERIES C
Aa1	Aaa	MSFT	MICROSOFT CORP
B1	Baa2	NOC	NORTHROP GRUMMAN CORP
B1	Baa1	WM	WASTE MANAGEMENT INC
Aa3	A3	UAL	UNITED AIRLINES HOLDINGS INC

Moody's Rating	Predicted	Ticker	Name	Coupon	Yield
A3	Ba3	AMH	AMERICAN HOMES RENT REIT CLASS A	2.0995	1.910000
A3	Aa3	UAL	UNITED AIRLINES HOLDINGS INC	4.7000	2.498236
A3	Ba1	UPS	UNITED PARCEL SERVICE INC CLASS B	4.7000	2.498236
A3	Ba1	MS	MORGAN STANLEY	4.4310	2.920000
A3	Ba1	CAT	CATERPILLAR INC	2.5500	1.870000
A3	Ba3	AAL	AMERICAN AIRLINES GROUP INC	4.2000	2.290000
A3	Baa2	C	CITIGROUP INC	4.0440	2.640000
A3	Aa3	UAL	UNITED AIRLINES HOLDINGS INC	3.8875	2.498236
A3	Aaa	GILD	GILEAD SCIENCES INC	2.5000	2.090000
A3	Baa1	AIG	AMERICAN INTERNATIONAL GROUP INC	3.0000	2.260000

Implementation and Use

- Investors
  - Buying and selling opportunities
- Lenders
  - Monitoring borrowers and safety of loans
- Issuers
  - Financial health and reputation monitoring
  - Saving opportunity

Conclusion and Improvements/Future Work

- Better data
  - aka paid data
  - More data
  - More features
- More models
  - Better parameters
- More accurate data
  - Too many Nans
  - Timely data - trending data