FE590 – Statistical Learning

Country Risk Scoring/Risk Rating – the quantitative & qualitative approach

Introduction

- Country risk analysis is an important indicator as it is related to the risk of investing or lending (to companies) in a country
- Possible changes in the business environment may adversely affect operating profits, assets value and credibility/trustworthy of the country.
- Financial factors such as currency, bond yield, GDP/employment rate, regulatory changes, or politically stability factors such as mass riots, civil war and other potential events contribute to companies' operational risks (closely related ~ Political Risk)
- Objective is predict the **Country Risk scoring** (quantitative) and **Risk Rating group** (qualitative) using machine learning techniques

Data

Bloomberg

- Data of 83 countries collected (Brazil, Canada, China, Columbia, Egypt, Ghana, Japan, Kenya, Latvia, Luxembourg, Norway, Qatar, Singapore, South Korea, Sweden, Turkey, Vietnam, Ukraine, US etc.)
- Source: **Bloomberg**
- Bloomberg Country Risk Score & its country data:
  - Bond yield
  - GDP data / CPI data / Unemployment rate
  - Budget surplus
  - Import / Export, FDI etc.
- Risk rating group based on Fitch Ratings
  - AAA, AA+, AA, AA-, A+, A, A- (Good)
  - BBB+, BBB, BBB-, BB+, BB, BB- (Moderate)
  - B+, B, B-, CCC, CC etc. (Risky)

FitchRatings

Supervised Techniques (Quantitative)

- Data split into training (~70%) and testing sets (~30%)
- Linear regression in R language [using glm()] – select predictors based on smallest CP value on different subset selections

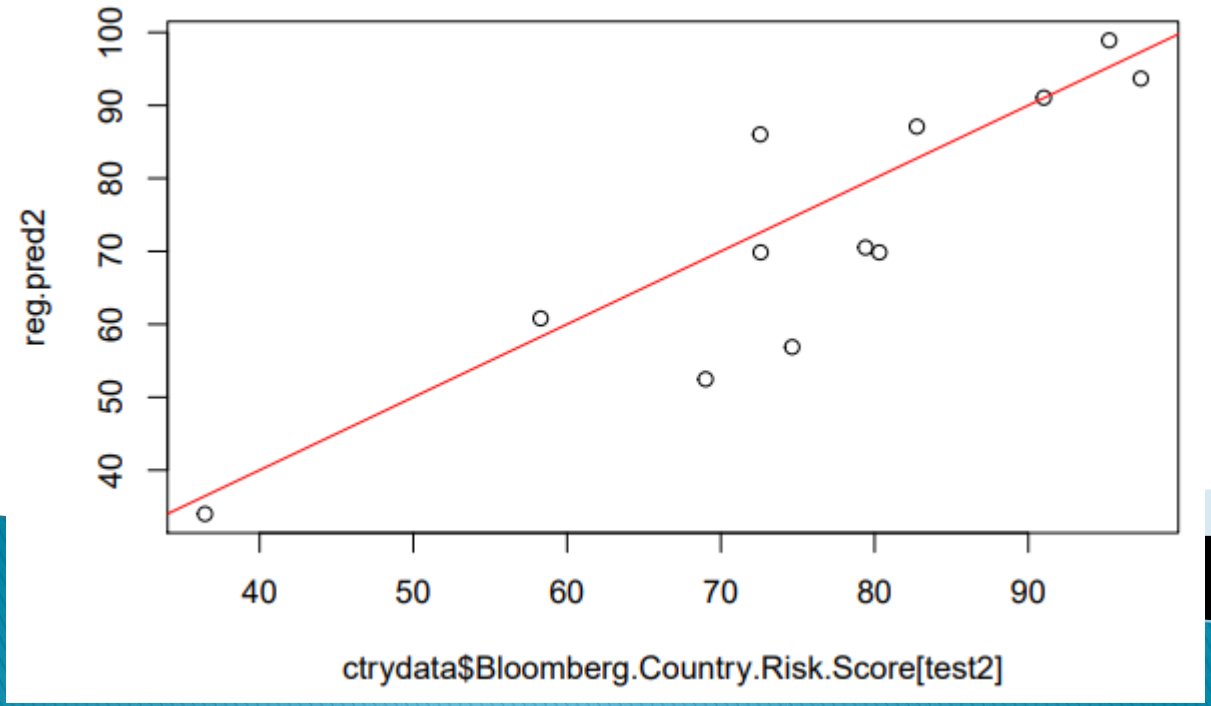
	1	2	3	4	5	6	7
(Intercept)	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
X5.Year.CDS	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE
X3.Month.Deposit.Rate	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Equity.Index.Price.Change.Perc	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE
Index.Returns.To.Global.Avg.Zscore	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
Historical.3.Month.Volatility	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE
Real.Effective.Exchange.Rate	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE
GDP.YOY.Perc	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
CPI.Actual	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Unemployment	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Total.External.Debt.PercGDP	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Imports.PercGDP	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Exports.PercGDP	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Currency.Reserves.PercGDP	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Foreign.Direct.Investment	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE
WGI.Control.Of.Corruption	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
WGI.Govt.Effectiveness	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
WGI.Regulatory.Quality	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
WGI.Rule.Of.Law	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE
Ease.Of.Doing.Business.Rank	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
Starting.A.Business.Rank	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Linear regression (Quantitative)

- Selected predictors:
  - 3 month deposit rate
  - Index returns to Global Avg. Z score,
  - Unemploymenet rate
  - FDI
  - World wide Governance Indictor (WGI) corruption index
  - WGI Gov. effectiveness
  - WGI Rule of Law
  - Ease of Doing Business ranking

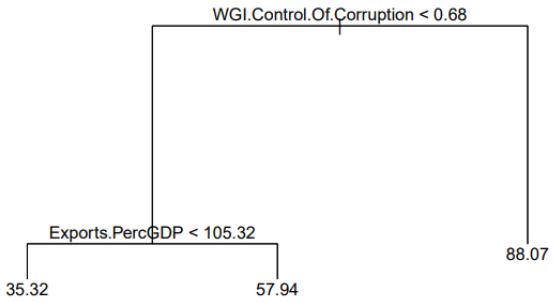
Linear regression (Quantitative)

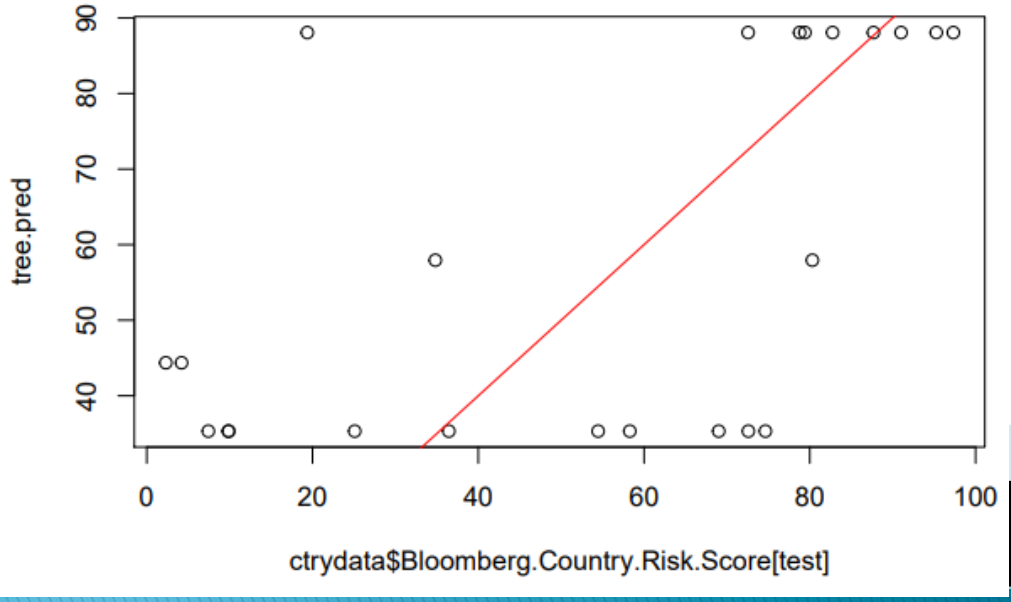
- Plot of predicted values vs. actual Bloomberg Risk Score



Tree regression (Quantitative)

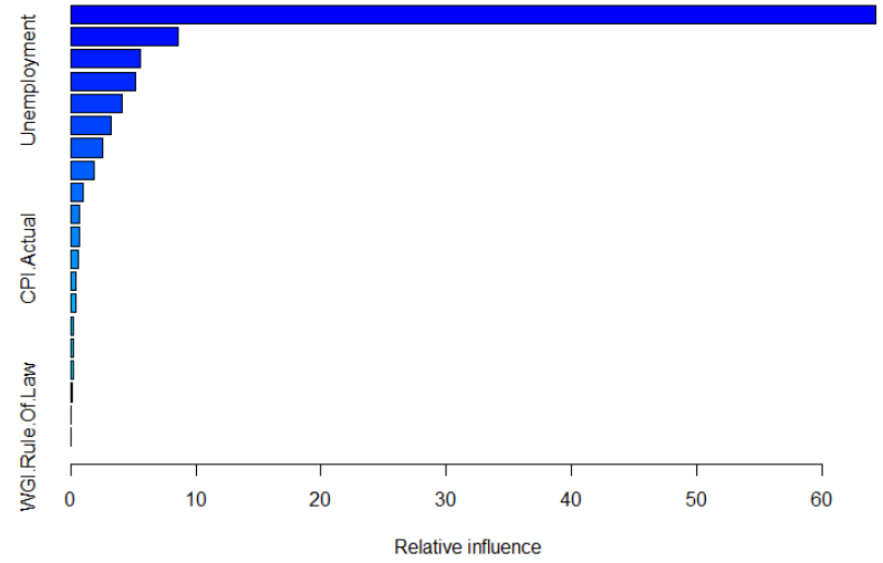
- Generated tree [using tree()] & predicted values vs. actual Bloomberg Risk Score





Boosting (Quantitative)

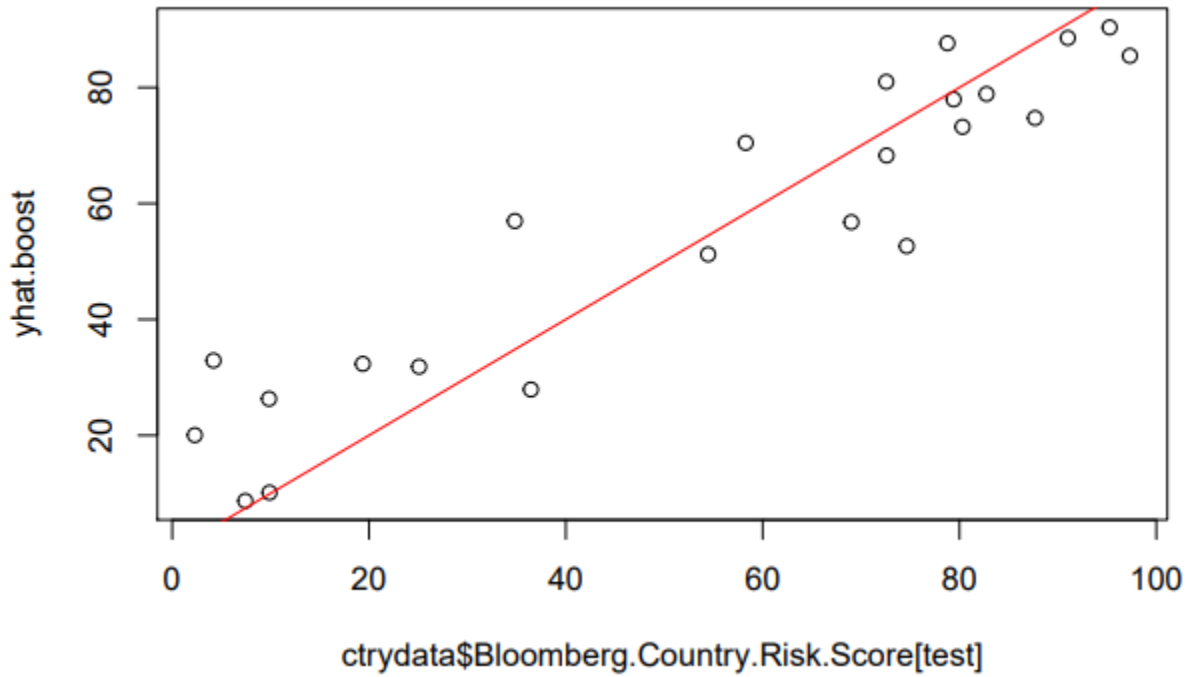
- Ensemble method to improve results of weak learners sequentially, using [gbm()]



	rel.inf
Equity.Index.Price.Change.Perc	64.31743852
Historical.3.Month.Volatility	8.58889305
X3.Month.Deposit.Rate	5.58321546
Unemployment	5.15688431
WGI.Govt.Effectiveness	4.09328425
Total.External.Debt.PercGDP	3.22075211
Ease.Of.Doing.Business.Rank	2.59328232
WGI.Regulatory.Quality	1.82447296
Exports.PercGDP	0.95530142
WGI.Control.Of.Corruption	0.70432854
Currency.Reserves.PercGDP	0.66449581
CPI.Actual	0.61352192
Imports.PercGDP	0.44300797
X5.Year.CDS	0.43755037
Index.Returns.To.Global.Avg.Zscore	0.24869574
Foreign.Direct.Investment	0.21865129
Starting.A.Business.Rank	0.21356855
Real.Effective.Exchange.Rate	0.10860436
GDP.YOY.Perc	0.01405102
WGI.Rule.Of.Law	0.00000000

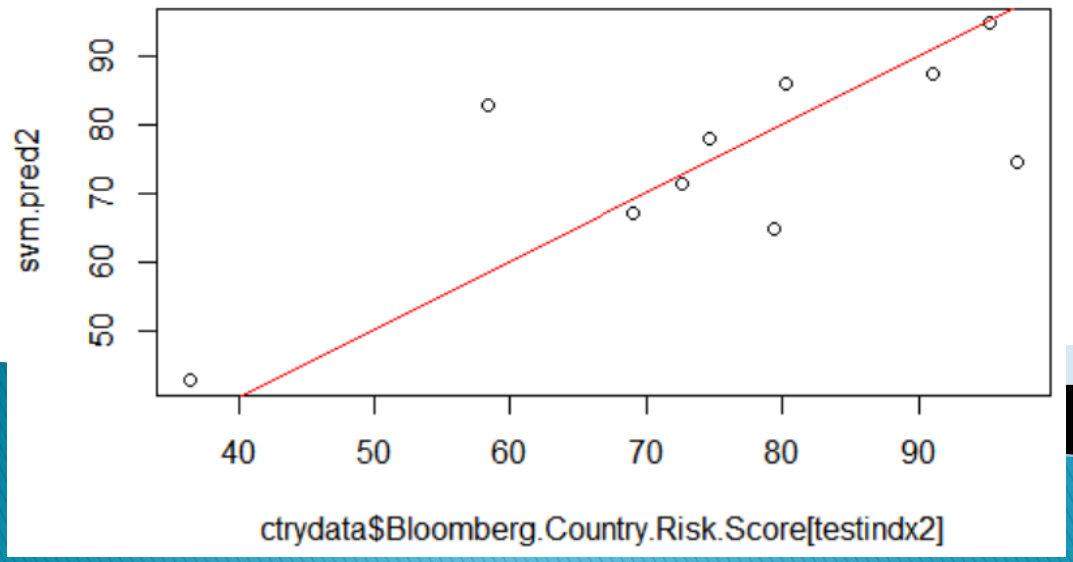
Boosting (Quantitative)

- Predicted values vs. actual Bloomberg Risk Score



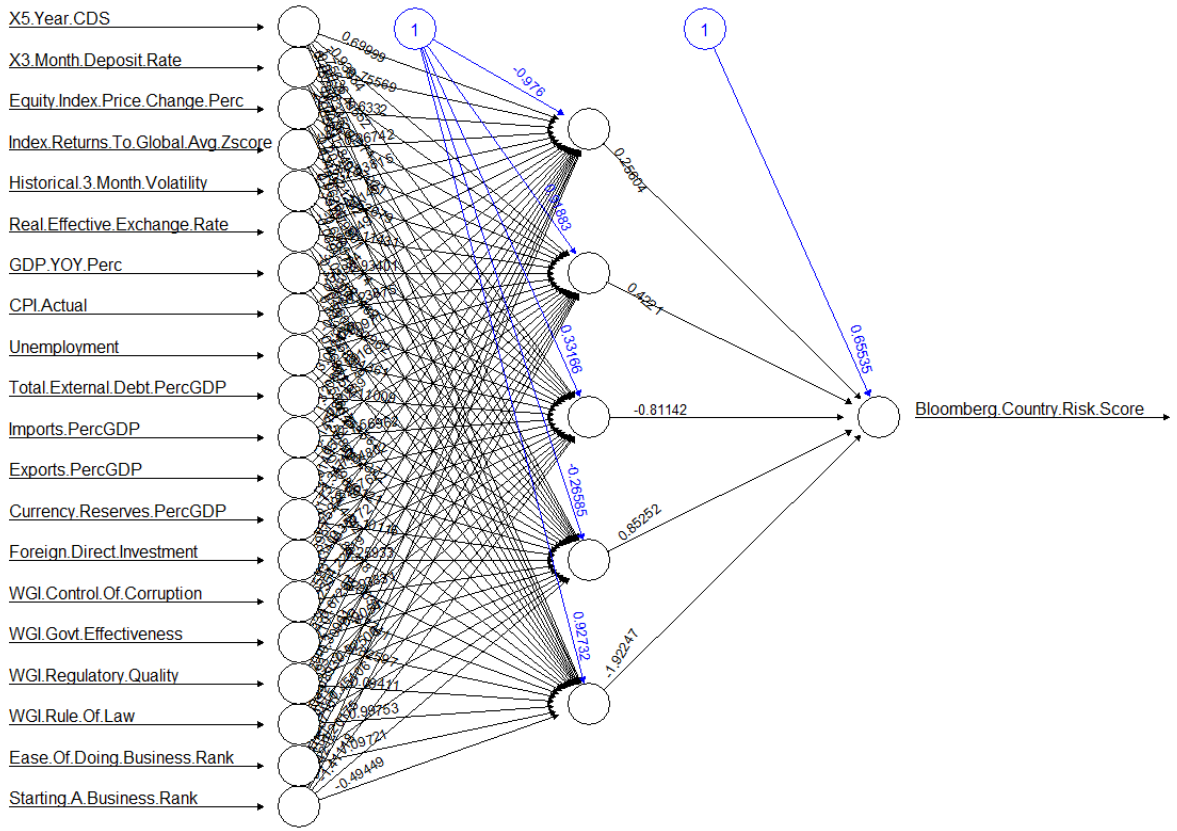
SVM regression (Quantitative)

- Support Vector Machine (SVM) regression to map max. number of points to decision boundary of a line
- Predicted values vs. actual Bloomberg Risk Score



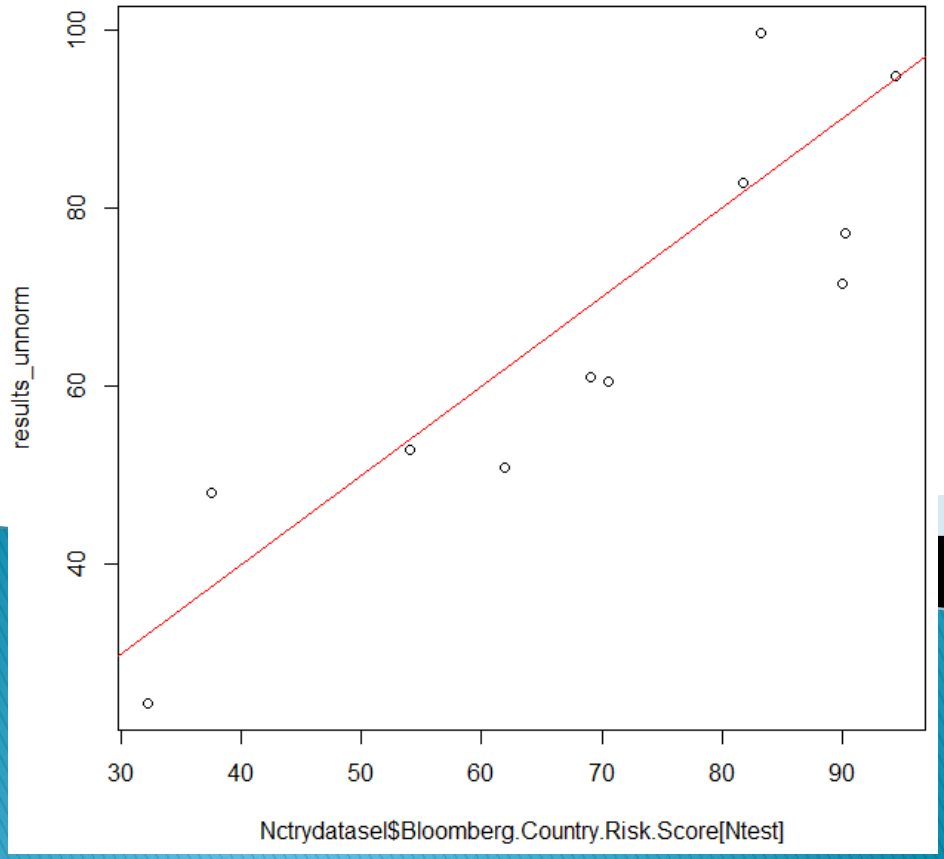
Neural Networks (Quantitative)

- Network diagram with 5 hidden layers [using neuralnet()]:



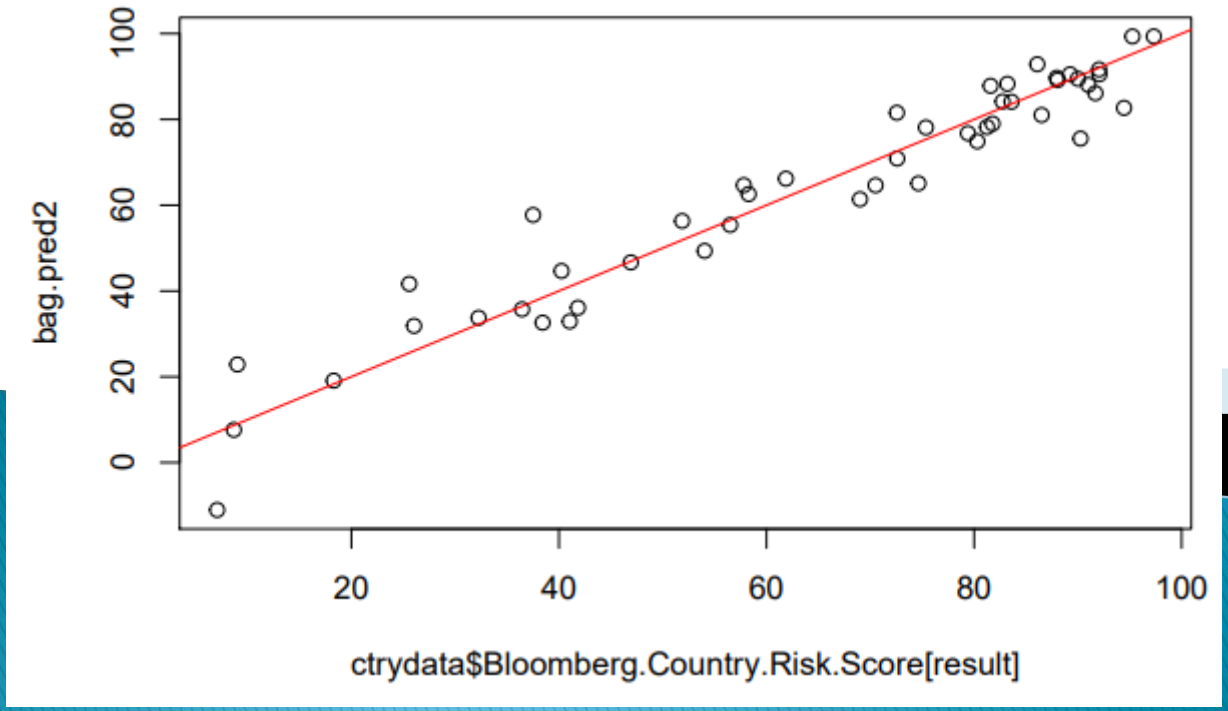
Neural Networks (Quantitative)

- Predicted values vs. actual Bloomberg Risk Score



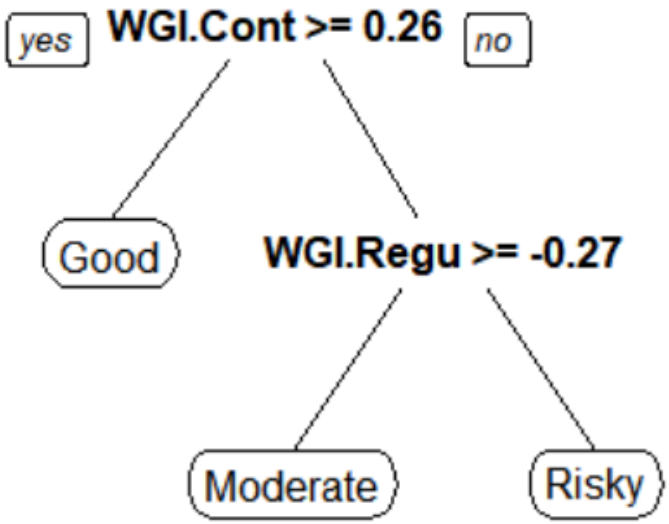
Bagging (Quantitative)

- Bootstrap aggregating (bagging) improves variance via model averaging technique. Bagging has the best results
- Predicted values vs. actual Bloomberg Risk Score



Decision Tree (Qualitative)

- Tree classification [using rpart()]
- Predicted Rating Grop vs. actual Risk Rating Group



	Good	Moderate	Risky
Good	10	0	1
Moderate	1	6	2
Risky	0	1	2

Confusion matrix | accuracy = 78.3%

LDA (Qualitative)

- Linear discriminant using linear combination of features for classification [using LDA()]
- Predicted Rating Group vs. actual Risk Rating Group

	Good	Moderate	Risky
Good	5	0	0
Moderate	2	3	2
Risky	0	0	0

Confusion Matrix | Accuracy = 80.0%

SVM (Qualitative)

- Support Vector Machine finds corresponding hyperplane that maximizes the separation between data points in classification [using SVM()]. Linear based & radial Kernel with best fit model.
- Predicted Rating Group vs. actual Risk Rating Group

*linear kernel	Good	Moderate	Risky
Good	5	1	0
Moderate	2	2	0
Risky	0	0	0

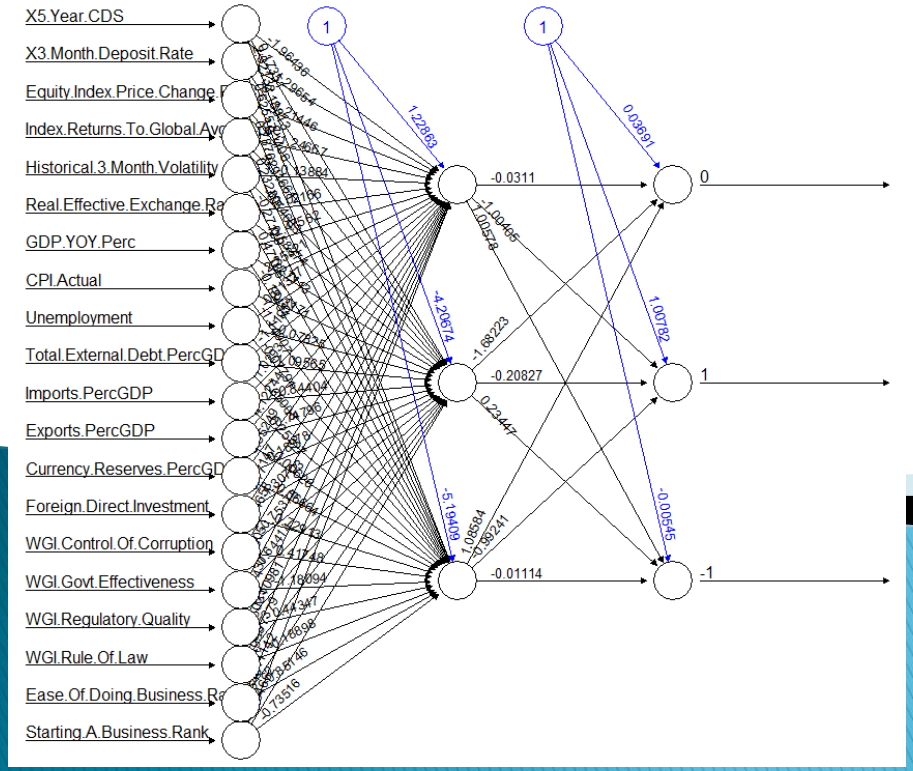
Confusion Matrix | Accuracy = 70.0%

*radial kernel	Good	Moderate	Risky
Good	5	0	0
Moderate	5	3	0
Risky	0	0	0

Confusion Matrix | Accuracy = 50.0%

Neural Networks (Qualitative)

- Network diagram with 3 hidden layers [using neuralnet()]:
- Predicted Rating Group vs. actual Risk Rating Group

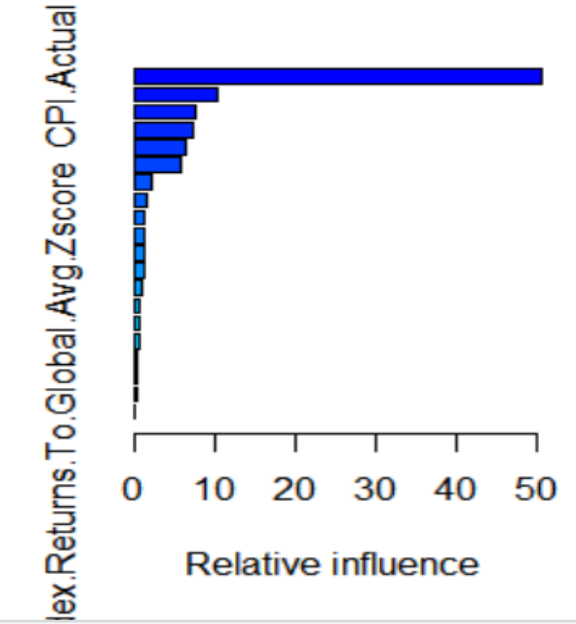


	Good	Moderate	Risky
Good	4	1	0
Moderate	2	2	2
Risky	0	0	0

Confusion Matrix | Accuracy = 54.5%

Boosting (Qualitative)

- Ensemble method to improve results of weak learners sequentially, using [gbm()]. Multinomial classification with specific learning rate.
- Predicted Rating Group vs. actual Risk Rating Group. Boosting has highest accuracy.



	Good	Moderate	Risky
Good	10	1	1
Moderate	1	6	1
Risky	0	0	3

Confusion Matrix | Accuracy = 82.6%

Conclusion

- Both bagging improves accuracy of quantitative results while boosting improves the accuracy of qualitative results
- The supervised learning techniques have fairly well accurate predictions > 80% for Country Risk Scoring / Risk Rating classification