Predicting Sovereign Defaults

Andrew Huang {ahuang98}, Taresh Sethi {tksethi}
Department of Computer Science, Stanford University

Motivation

- Develop a model for predicting sovereign default
- Given publicly accessible economic data, predict whether or not a country will enter financial crisis.

Features and Data

- **Source of Data:**
  - International Monetary Fund (IMF) and Reinhart and Rogoff[1,2]
  - 1,334 observations, spanning 43 countries with economic data from the years 1980-2010
  - 50 features
  - Output generated by shifting "*Crisis this year*" feature to predict if in default the following year
- **Data was slightly imbalanced, 20.6% positives (countries in default)**
- **Preprocessing techniques**
  - Z-standardization and filling in averages
  - Removing chunks of sparse data

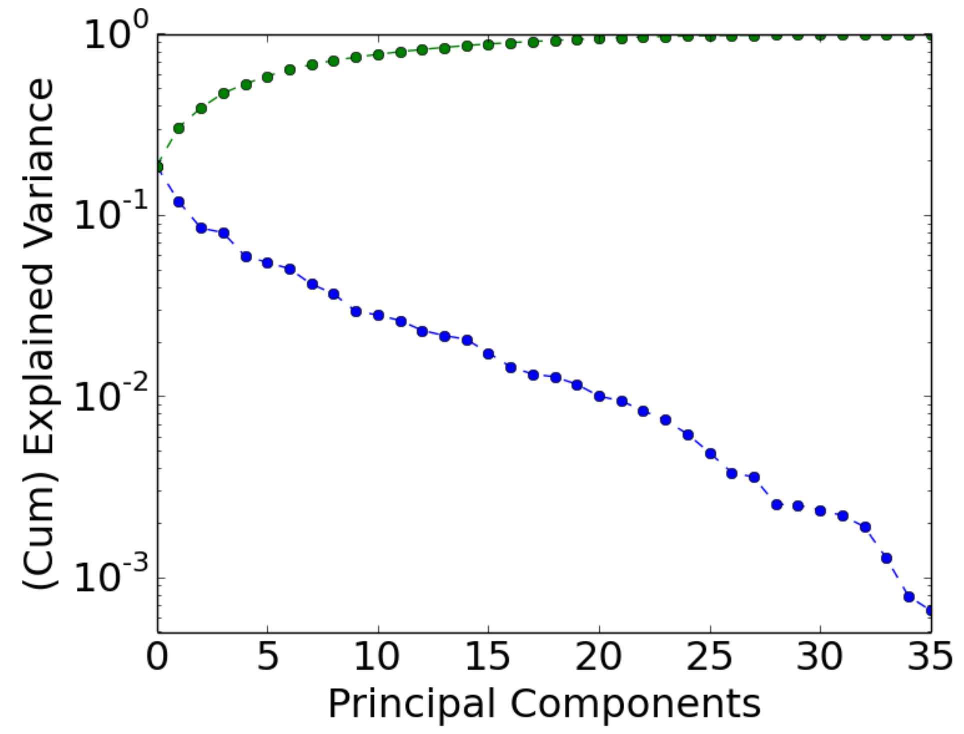
Future Steps

- Obtain diverse data (perhaps proprietary economic data, or sociopolitical data)
- Optimize performance on difficult predictions (coming in/out of crisis)
- Test different models

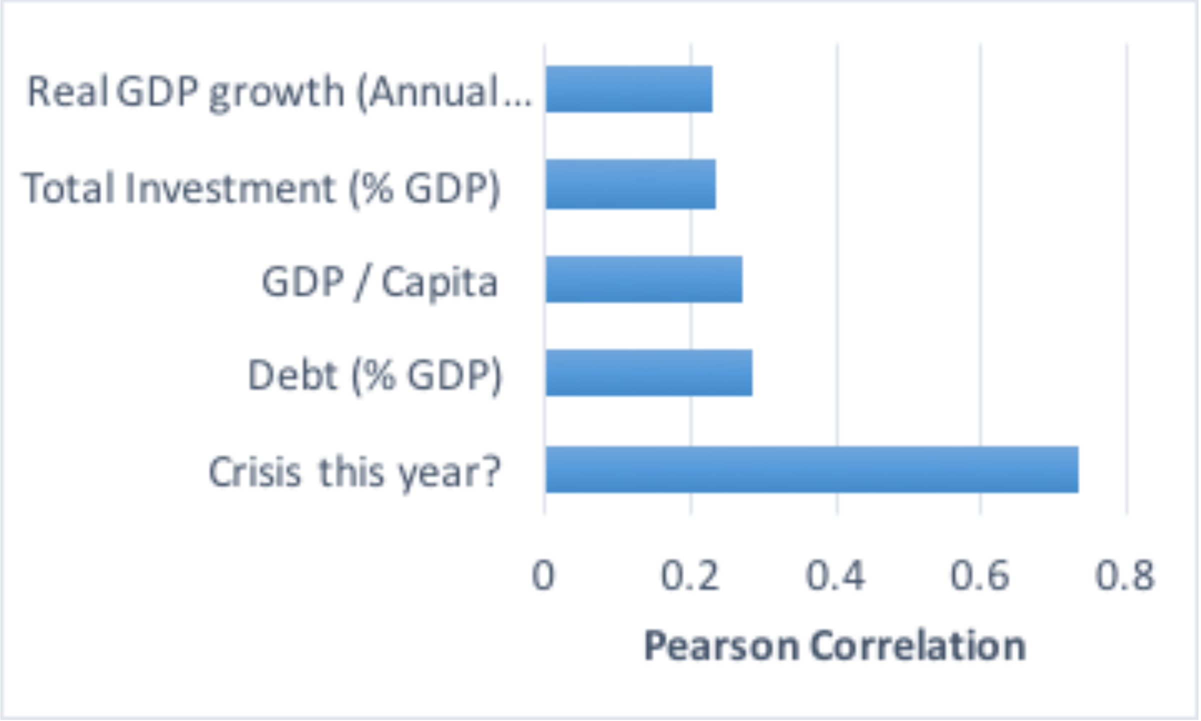
Model and Feature Selection

Feature Selection

- PCA
  - 30 features account for 99% of the variance



- RFE and Pearson Correlation w/ output
  - Tried selecting 30 highest-correlated features, but cross-validation scores were lower than model w/ all features
  - One feature very correlated ("Crisis this year")



Model Selection

- Logistic Regression
- SVM (Linear Kernel) with balancing
- Neural Network – momentum[3]:
- Random Forest

$$\begin{cases} v_{dW^{[\ell]}} = \beta v_{dW^{[\ell]}} + (1-\beta)\frac{\partial J}{\partial W^{[\ell]}} \\ W^{[\ell]} = W^{[\ell]} - \alpha v_{dW^{[\ell]}} \end{cases}$$

Discussion

We find an interesting result: our models, particularly random forest, were able to effectively predict crisis next year with and without the "*Crisis this year?*" feature. We speculate that the abundance of features allowed our models to down-weight the "*Crisis this year*" feature and allow for "smarter" predictions. Still, our models performed poorly on the addendum test set of meaningful predictions (coming in/out of crisis), which may suggest that our models may be over-separating the two classes. We were especially surprised by the performance we received given just publicly available data. We did not expect to see that feature selection did not aid the model, especially given initial perceived dependence among variables.

Model Evaluations and Results

- Performed k-fold cross validation
- Used F1 score as a performance metric
- 70-30 train/test split
- Evaluated w/ and w/o "*Crisis this year?*"
- Ground-truth: 78.3%

With "Crisis this year" column	K-fold Score	F1 Score (train)	F1 Score (test)	Train Accuracy	Test Accuracy
Logistic Regression	0.90	0.81	0.71	0.92	0.87
SVM (linear kernel) + balancing	0.92	0.82	0.74	0.93	0.89
MLP Neural Network	0.92	0.84	0.71	0.94	0.88
Random Forest (n=100)	0.93	1.00	0.76	1.00	0.91

Without "Crisis this year" column	K-fold Score	F1 Score (train)	F1 Score (test)	Train Accuracy	Test Accuracy
Logistic Regression	0.77	0.64	0.65	0.80	0.81
SVM (linear kernel) + balancing	0.85	0.63	0.62	0.87	0.86
MLP Neural Network	0.87	0.76	0.63	0.91	0.85
Random Forest (n=100)	0.91	1.00	0.71	1.00	0.91

- Picked out subset of testing set that represented difficult but meaningful predictions. Models did not perform well.

References

[1] International Monetary Fund. "World Economic Outlook Databases," October 10, 2017.
[2] C. Reinhart and K. Rogoff. "This Time is Different." http://www.reinhartandrogoff.com/data/
[3] A. Ng. "CS229 Lecture Notes: Deep Learning."