Predicting Financial Crises

Sam Kamien
Sairajath
Chaitanya Shashi Kumar
Ram Kishore

The Problem

By leveraging machine learning approaches, can we predict various forms of financial crises (banking, inflation, gold standard, etc.) by examining various signifiers of economic health (GDP, debt, PPP, etc.)?

Data

We present data collected over many years by Carmen Reinhart (with her coauthors Ken Rogoff, Christoph Trebesch, and Vincent Reinhart). These include banking crisis data for more than 70 countries from 1800-present, exchange rate crises, stock market crises, sovereign debt growth and default, and many other data series.

Sample:

1	Case	CC3	Country	Year	Banking Cr	Banking_C	Systemic (	Gold Stanc	exch_usd	exch_usd_	exch_usd_	exch_usd_	conversion	national c	exch_prim	exch_sour	Domestic_	Domestic_	SOVEREIG	SOVEREIG	Defaults_E	GDP_Weig	<	Inflation, /	Independ	Currency C	Inflation Crises
1505	7	BOL	Bolivia	2000	0		0	0	6.39	6.36	6.39		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	4.6	1	0	0
1506	7	BOL	Bolivia	2001	0		0	0	6.82	6.809	6.82		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	1.6	1	0	0
1507	7	BOL	Bolivia	2002	0		0	0	7.49	7.4935	7.49		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	0.9	1	0	0
1508	7	BOL	Bolivia	2003	0		0	0	7.83	7.7955	7.83		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	3.3	1	0	0
1509	7	BOL	Bolivia	2004	0		0	0	8.05	8.022	8.05		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	4.4	1	0	0
1510	7	BOL	Bolivia	2005	0		0	0	8.04	7.95	8.04		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	5.4	1	0	0
1511	7	BOL	Bolivia	2006	0		0	0	7.98	7.995	7.98		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	4.3	1	0	0
1512	7	BOL	Bolivia	2007	0		0	0	7.62	7.53	7.62		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	6.7	1	0	0
1513	7	BOL	Bolivia	2008	0		0	0	7.02	7.0182	7.02		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	14	1	0	0
1514	7	BOL	Bolivia	2009	0		0	0	7.02	6.97	7.02		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	3.3	1	0	0
1515	7	BOL	Bolivia	2010	0		0	0	6.99	7.01	6.99		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	2.5	1	0	0
1516	7	BOL	Bolivia	2011	0		0	0	6.91	7.01	6.91		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	9.9	1	0	0
1517	7	BOL	Bolivia	2012	0		0	0	6.91	6.96	6.91		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	4.5	1	0	0
1518	7	BOL	Bolivia	2013	0		0	0	6.91	6.91	6.91		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	5.7	1	0	0
1519	7	BOL	Bolivia	2014	0		0	0	6.91	6.91	6.91		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	5.8	1	0	0
1520	7	BOL	Bolivia	2015	n/a		n/a	0	6.91	6.91	6.91		Series alre	Pre-1863-peso Spanis		Primary sc	0		0	0		0	n/a	4.1	1	0	0
1521	7	BOL	Bolivia	2016	n/a		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	4	1	0	0
1522	8	BRA	Brazil	1800	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	-5.1	0	0	0
1523	8	BRA	Brazil	1801	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	7	0	0	0
1524	8	BRA	Brazil	1802	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	5.7	0	0	0
1525	8	BRA	Brazil	1803	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	-3.4	0	0	0
1526	8	BRA	Brazil	1804	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	-1.2	0	0	0
1527	8	BRA	Brazil	1805	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	-12.8	0	0	0
1528	8	BRA	Brazil	1806	0		0	0					Series alre	1846-1942	BRXRXDE.	Primary sc	0		0	0		0	n/a	5.2	0	0	0

Types of Data

As seen in the previous slide, the dataset includes continuous, categorical, and binary data.

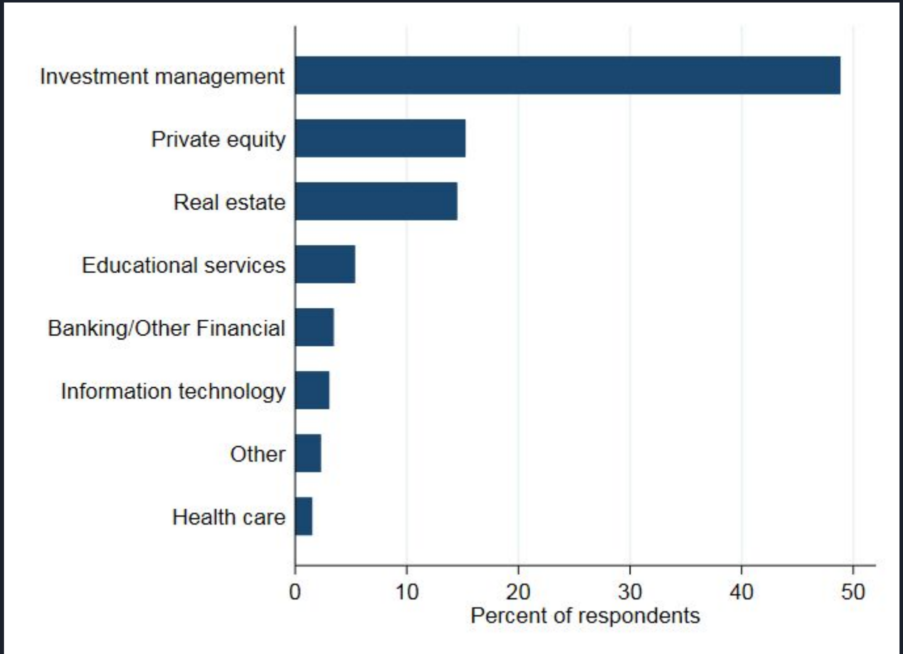
- Continuous Data example: US dollar exchange rate
- Categorical Data example: year
- Binary data example: banking crisis, domestic debt

Other types of data such as text also exist in the dataset which we will have to format for learning models

Survey Background

- In October of 2018, Harvard Business School hosted The Global Financial Crisis. In preparation for this event, the Behavioral Finance and Financial Stability Project (BFFS) at HBS conducted the 2018 Financial Risk and Regulation Survey.
- The survey was completed by 139 participants, with a variety of backgrounds and experiences, but most were financial industry professionals located in the U.S.. More than 80% of the participants listed investment management, real estate, private equity, or banking as their industry.

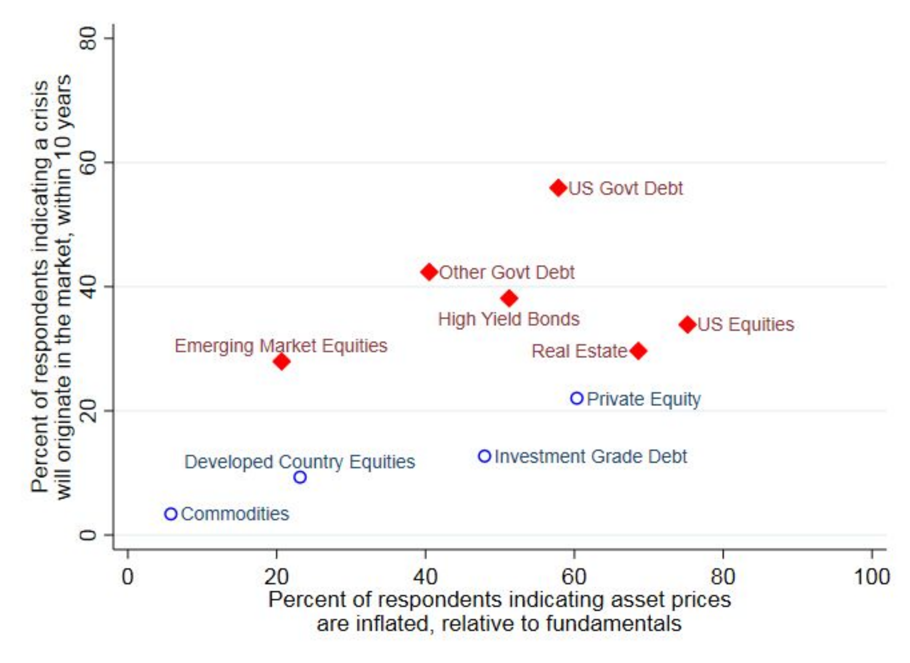
Share of Respondents by Industry:



Current Sources of Risk

- The survey asked the participants a number of questions regarding current sources of risk to the financial system.
- Participants identified asset prices in several markets as particularly inflated, relative to fundamentals, including U.S. government debt, other government debt, high yield bonds, real estate, and private equity. Generally, these same markets were identified as likely to originate a crisis in the next ten years.

Current Sources of Risk, by Market



Prior work

Machine learning as an early warning system to predict financial crisis

They employ structured financial networks and machine learning algorithms to enhance Early Warning Systems, identifying contagion risks with 98.8% accuracy. The findings emphasize the utility of financial network analysis for policymakers and investors, offering valuable insights for portfolio selection based on asset centrality.

Author: Aristeidis Samitas

Paper: https://www.sciencedirect.com/science/article/abs/pii/S1057521920301514

Predicting Financial Crises - The Role of Asset Prices

They examine a composite indicator, combining signals from asset price growth and volatility, as an effective early warning system for financial crises across 108 advanced and emerging economies from 1995 to 2017. The research demonstrates that elevated levels of this indicator, indicating rapid asset price growth coupled with low volatility, significantly enhance the predictive accuracy of future financial crises, outperforming credit-based metrics and providing critical information about systemic risk levels for policymakers.

Authors: Tristan Hennig, Plamen Iossifov, and Richard Varghese
Paper: https://www.imf.org/en/Publications/WP/Issues/2023/08/03/Predicting-Financial-Crises-The-Role-of-Asset-Prices-536491

Our Approach

Models:

- The problem is in the form of multiple binary classifications so we can train a number of models for each crisis type and combine them into an ensemble. In particular, we plan to use Naive-Bayes, decision trees, and linear regression models. If these have particularly poor accuracy, we can experiment with other models

Pre-Processing

- A challenge we anticipate is the pre-processing of the data. A number of columns are dominated by null values to signify unavailable data. Some columns have data that is not useful for us, such as the name of a country's currency. Data such as country or year can be digitized into categories or split into new columns via one-hot-encoding and we can trim columns that are not useful.