Queen Mary University of London

Assessed Midterm Assignment

Submit online via QMplus before 23:59 London Time on Monday 17th July 2023

ECOM193 Statistical Machine Learning in Finance

This assignment carries a maximum of 100 marks. A 20% weighting will be applied to your overall ECOM193 module score for this assignment.

You are required to submit a typed document in PDF or Word format containing written analysis together with supporting tables and graphics. It is not necessary to show any computer code you use to undertake your analyses but if you think it might be helpful, you may include a limited amount of annotated code in a clearly labelled appendix.

The assignment is a piece of data analysis detailed on the next page. Ideally, your analyses of the data should be in R. However, you are allowed to use any suitable scientific software, e.g. Python or Matlab, provided the output you produce tackles the problem appropriately.

You are not restricted in the type of research you undertake to help you complete the Assignment. However, it must be your own work.

Examiner: Dr R.A. Saldanha

Elucidating African Economic Crises

The `african_crises.csv` file contains annual data on economic and financial crises in 13 African Countries from 1860 to 2014. It is taken from the competition website kaggle.com; see *Africa Economic, Banking and Systemic Crisis Data*. The file contains the following 14 columns (variables):

`case`	number denoting specific country
`cc3`	country code
`country`	country name
`year`	observation year
`systemic_crisis`	systemic crisis (binary): 0 = no / 1 = yes
`exch_usd`	exchange rate (USD base)
`domestic_debt_in_default`	sovereign domestic debt default (binary): 0 = no / 1 = yes
`sovereign_external_debt_default`	sovereign external debt default (binary): 0 = no / 1 = yes
`gdp_weighted_default`	total debt in default as a proportion of total GDP
`inflation_annual_cpi`	annual consumer price inflation (CPI) rate
`independence`	country independence indicator (binary): 0 = no / 1 = yes
`currency_crisis`	currency crisis (binary): 0 = no / 1 = yes
`inflation_crisis`	inflation crisis (binary): 0 = no / 1 = yes
`banking_crisis`	banking crisis (character binary): "no_crisis" / "crisis"

1. Undertake a suitable exploratory analysis of the data explaining each of the variables. Do all the variables make sense? Are there any outliers?

[15 marks]

2. Which variables or factors are most associated with currency crises?

[20 marks]

3. What is the relationship between CPI and inflation crises (if any)?

[15 marks]

questions continue overleaf...

4. Can you find an effective model (or models) for systemic crisis in terms of a selection of the given variables? How well does your model work in-sample versus out-of-sample?

[25 marks]

5. What country conclusions (if any) do you draw from these data?

[12 marks]

6. Derive a classification model for systemic crisis. How well does your model work in-sample versus out-of-sample?

[13 marks]