

Conflict Prediction with Spike and Slab prior

Anh Le

December 7, 2014

Introduction

Being able to predict conflict is very useful for governments and international organizations to deploy their limited resources effectively. Therefore, this project will try to predict the (binary) presence of four events (i.e. insurgency, rebellion, dpc-domestic political crisis, and erv-ethnic and religious violence). I use monthly data of 167 countries from 2001 to present (dimension = $27,000 \times 550$). The feature variables include each country's politics, economic performance, financial status, and their 2-month lags.

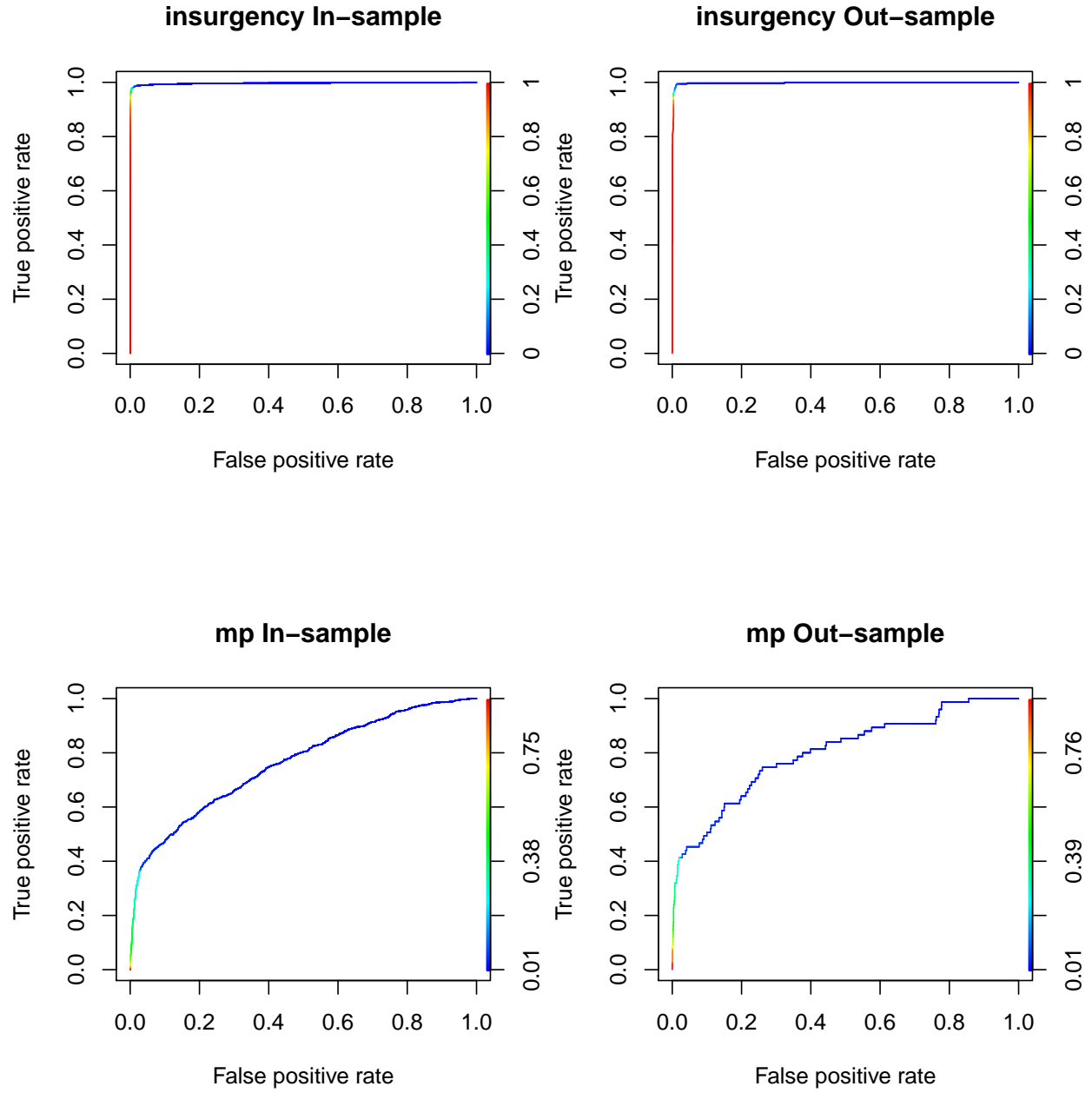
The project will use logistic regression with spike-and-slab prior to choose a small subset of variables. This will isolate important predictive factors, facilitating model interpretation. Given high correlations between temporally lagged variables, sparse regression will also avoid over-fitting. I will compare the predictive accuracy of this model against others currently used in our lab.

Analysis

I pre-process the data by removing extraneous variables (country names, time ID, etc.) and, for each country in the dataset, add a binary variable that indicates whether an observation belongs to a country. This allows the model to have different intercepts for each country, essentially adding country fixed effect.

I fit the logistic model with spike-and-slab prior using the package `BoomSpikeSlab`. To simultaneously fit four models for my four response variables, I use packages `doMC` and `foreach`.

The model is trained on March2001-June 2013 data and tested on July2013-Sept2014 data. The results below come from a MCMC chain with iterations = 1000 and burn-in = 100. Even though the MCMC chain is not that long, due to the size of the dataset the computational time is already over 1.5 hour.



	insurgency	rebellion	dpc	erv	mp
brier	0.005	0.006	0.042	0.008	0.036
auc.C	0.996	0.999	0.927	0.989	0.764
precision	0.981	0.957	0.789	0.961	0.548
recall	0.767	0.769	0.410	0.681	0.068

Table 1: In-sample predictive performance

	insurgency	rebellion	dpc	erv	mp
brier	0.008	0.020	0.097	0.033	0.024
auc.C	0.998	0.930	0.865	0.975	0.801
precision	0.976	0.907	0.544	0.907	0.647
recall	0.946	0.789	0.548	0.490	0.147

Table 2: Out-sample predictive performance