

# Conflict Prediction and Machine Learning

Anh Le

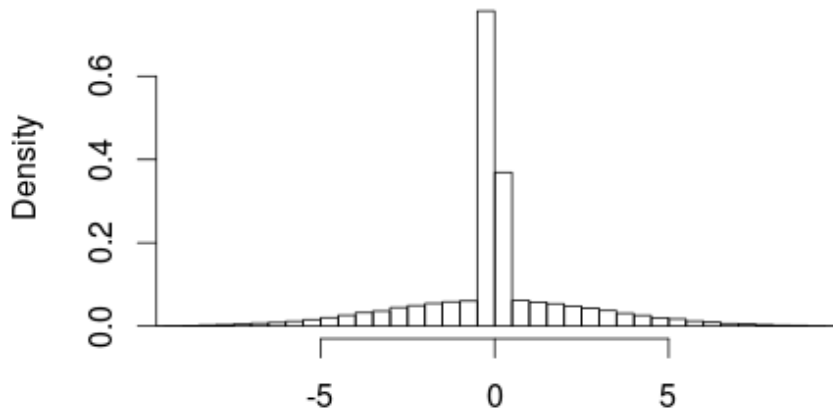
Duke University

*anh.le@duke.edu*

January 21, 2015

# Overview

## Spike and Slab Prior



# Model performance (with country dummies)

	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

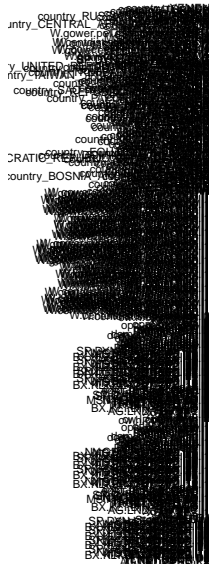
Table: Spike and Slab (out-sample)

	insurgency	rebellion	dpc	erv
brier	0.06	0.03	0.12	0.03
auc.C	0.94	0.97	0.78	0.93

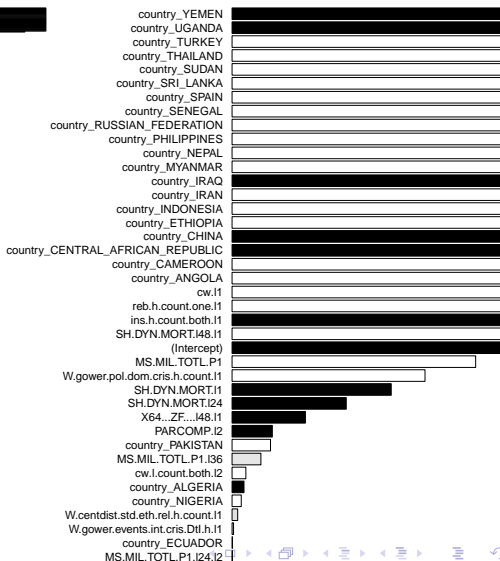
Table: EBMA (out-sample)

# Variable selection (rebellion)

rebellion : all variables



rebellion : variables with inclusion prob > 0.0



# Model performance (without country dummies)

	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: With dummies (out-sample)

	insurgency	rebellion	dpc	erv	mp
brier	0.080	0.057	0.167	0.037	0.049
auc.C	0.886	0.895	0.688	0.922	0.704
precision	0.519	0.575	0.162	0.508	0.097
recall	0.778	0.586	0.458	0.504	0.033

Table: Without dummies (out-sample)

# Boosted classification tree

- Fit an initial tree
- Get the residuals, fit another tree to the residual
- Add (part of) the new tree to the existing tree
- Tune 1) the number of trees, 2) how much of the new tree to add back to the old tree, 3) the complexity of each tree

So the algorithm can learn *slowly*

## Boosted tree result (for 2-month out...)

	insurgency	rebellion	dpc	erv
brier	0.043	0.023	0.057	0.016
auc.C	0.956	0.984	0.875	0.979
precision	0.781	0.818	0.611	0.833
recall	0.633	0.749	0.261	0.769

Table: Boosted tree (in-sample)

	insurgency	rebellion	dpc	erv
brier	0.046	0.030	0.065	0.020
auc.C	0.956	0.978	0.813	0.978
precision	0.694	0.815	0.649	0.837
recall	0.659	0.650	0.224	0.606

Table: Boosted tree (out-sample)