Predicting Electoral Violence
Final project report, March 2023

The D-Arch Team

1 Introduction

During the fall 2022 the D-Arch team has on the behalf of the Kofi Annan Foundation (KAF) worked to develop a system for forecasting electoral violence using state of the art machine learning techniques. The system uses data from the Varieties of Democracy (V-DEM) project (Coppedge et al., 2021b), as well as data from the Digital Societies Project (DSP) (Mechkova et al., 2022b), and the World Bank's World Development Indicators (WDI). Forecasts are made annually for each country with at least one scheduled election in the specific year, and made for two calendar years into the future from the last available data.

This report serves as a project report, briefly outlining the definition of electoral violence used for the forecasts as well as the data and methodology used and preliminary results for the forecasts of 2022 and 2023. The report will be followed by an academic article outlining the project in further detail.

2 Defining electoral violence

The first task of the D-Arch team was to develop an outcome measure for Electoral Violence which is conceptually sound and with properties which makes it practically useful for the KAF and which makes it possible to properly forecast. In discussions with KAF about what types of forecasts are most useful in practice, and internal discussions about what data are possible to use, we landed in agreeing that an ordinal outcome for electoral violence was the most useful. Based on a review of the existing sources of electoral violence in Fjelde et al. (2021) we decided to use the Varieties of democracy project indicators of electoral violence. To code this outcome we used two indicators measuring electoral violence by non-government actors (v2elpeace) and intimidation and harassment by government actors (v2elintim) from the VDEM election-level data. These two indicators are themselves ordinal on a 5-point scale from 0 (most electoral violence/harassment) to 4 (least electoral violence/harassment). These indicators are coded by country experts, and we use the mean value across all coded values (Coppedge et al., 2021a). The full description of these two indicators and their levels can be found in Appendix A.

In order to make the forecasts more stable and to make a more conceptually clear outcome, we re-coded these two indicators into a three-point ordinal scale measuring 'electoral violence and/or harassment'. We did this by first re-coding the 5 point ordinal scales to three point ordinal scales by merging the categories 0 and 1 to 'severe electoral violence or harassment', categories 2 and 3 to 'limited electoral violence or harassment' and category 4 to 'no electoral violence or harassment'. We then coded our outcome, 'electoral violence and/or harassment', as the most severe value across these two three-point indicators.

One important caveat with regards to this outcome measure is that this indicator captures *both* electoral violence *and* intimidation/harassment by the government. One consequence of this is that elections where the government repression has been so severe as to cause an artificially calm election with no outbursts of visible violence are also coded as having 'severe' levels of electoral violence since the government repression is so severe.

One alternative way of dealing with this potential problem is to separate the forecast into government harassment and/or intimidation as one category and electoral violence by other actors as a second category. Each country would, in that case, get a separate forecast for each of these two outcomes. This would make communication about the outcome more complex, but perhaps more in line with how electoral violence is conceptualized outwardly.

2.1 Forecasting the outcome

Since the outcome of electoral violence is ordinal the forecasts for the outcome show the probability that each election end up in the three categories of 'no electoral violence/harassment', 'limited electoral violence/harassment', and 'severe electoral violence/harassment'. In order to facilitate an easier interpretation of the results, two additional measures are also presented. First, the likelihood that *any* electoral violence occurs for the election, i.e. simply the sum of the probabilities of 'limited' and 'severe' electoral violence/harassment. Second, we also present a 'risk index' scaled from 0 to 1 where 0 indicates the least risk of electoral violence and 1 the highest risk of electoral violence. The risk index in constructed by simply taking the probability of 'severe' electoral violence and adding 0.5 x the probability of 'limited' electoral violence. We make these forecasts up to two years into the future.

3 Data and methodology

The forecasting methodology used for the project is anchored in the methodologies used by the Violence Early Warning System (ViEWS) to predict violence from armed conflict (Hegre et al., 2019; Hegre et al., 2021). To this end, train machine learning models on historical data with features (predictor variables) grouped into broad thematic *constituent models*. The predictions of these constituent models are then combined into an *ensemble* which produces the final forecasts. In total, for the December intermediary report, we have run tests with 37 constituent models grouped into five different overarching themes:

- Constituent models using features from election-level data from the last held election in the VDEM election-level data set
- Constituent models using features from the VDEM country year data set
- Constituent models using features from the digital societies project
- Constituent models using features from the World Bank's World Development indicators
- Constituent models using a combination of features from the above mentioned data sources

A list of all constituent models and which indicators each model contains can be found in Appendix B2, which also details which theme each constituent model corresponds to. The descriptions of the constituent models will be expanded for the final report.

For each model, the features are taken as the last value in the last available calendar year. I.e. for forecasts one calendar year into the future, the values are taken from December of the previous year and for forecasts two calendar years into the future, the values are taken from December two calendar years back in time. Most features are only updated annually, and in these cases the features are simply lagged one or two years respectively, but for features such as election-related variables the values are taken to be the last observed value in the last year with available data. Missing data are replaced by filling the last observed value forward.

We use all national level elections to either the presidency or the lower house (or combinations of elections which included at least one of these) in the period 1989-2021 (last available data) and coded by VDEM as the training data for the true forecasts, which we then produce for 2022 and 2023, i.e. two calendar years into the future. 2022 can thus be seen as a side-casted year since the model does not have access to the data from 2022 but we can evaluate how well the model did on the elections which have already been held. In the evaluation of the models (more below) we split the training data into different training and test periods in order to produce out of sample forecasts.

The process for producing the forecasts are detailed in Figure 1 below:

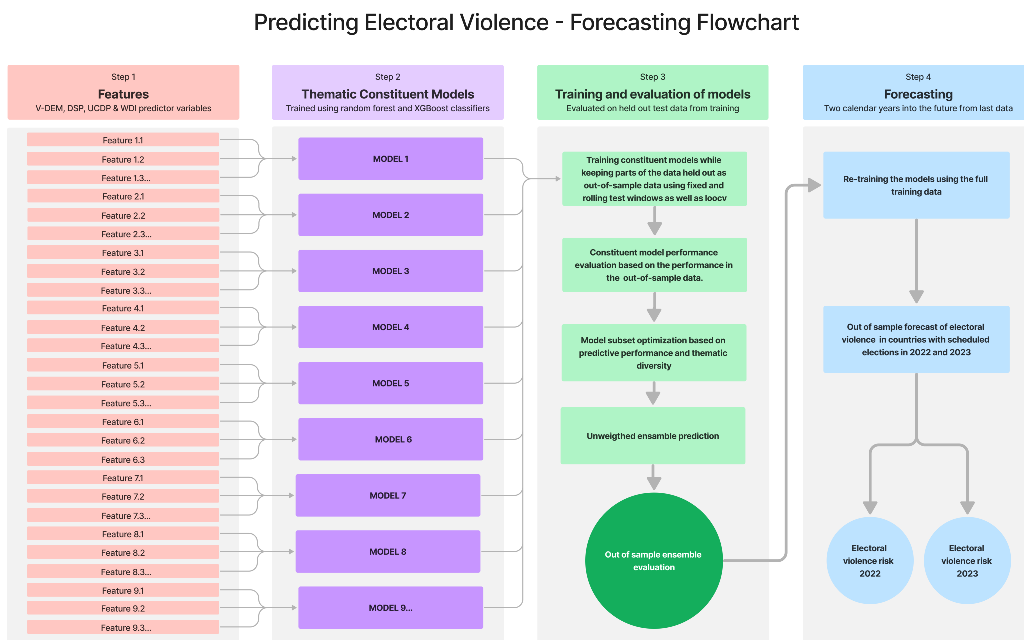

Figure 1. Forecasting flowchart

3.1 Forecasting models

Each constituent model is trained using two different types of machine learning models, a random forest classifier(Breiman, 2001), and an extreme gradient boosting (xgboost) classifier (Chen and Guestrin, 2016). Both of these models are developed to maximize the predictive performance of the model, rather than for optimizing the inferences which can be drawn from the models. This type of model can handle effects which are non-linear and effects which differ depending on other characteristics in the model (interactions). Models such as these have been shown to work well for predicting similar outcomes, for instance within the ViEWS project.

As the project uses a multitude of constituent models, these need to be weighted together in order to produce a final prediction. We do this by using an unweighted ensemble of a subset of the best performing models in the out of sample evaluation (see below). Currently we include nine models in this unweighted ensemble. These were selected among the best performing constituent models with consideration taken to which themes they relate to in order to not introduce bias in favour of specific themes and to ensure appropriate diversity among themes in the final ensemble. The models included in the nine model ensemble are discussed in Appendix B1. For the final forecasts we use an average of the random forest and xgboost predictions. We opted for using an ensemble of well-performing models over the use of the single best model for the final forecasts as an ensemble of models is likely to be more robust to overfitting when forecasting into the true future. We also wanted our final ensemble to span the feature space of available predictors in order to maximize the use of the available information.

3.2 Evaluation of models

The predictive performance of the constituent models and the ensembles were evaluated using a range of standard evaluation statistics. More specifically, the accuracy, brier, area under the precision recall curve (AUPR) and area under the receiver operator characteristic curve (AUROC) scores were computed. For the

AUPR and AUROC scores, the scores are the unweighted averages (i.e. macro) of the three one-against-all AUPR/AUROC scores. Accuracy measures the proportion of cases which are correctly classified on our three-level ordinal scale when the most likely level of electoral violence is taken as the prediction. The brier score, on the other hand, is a measure of the squared error, in terms of (decimal) probability for the model. AUPR measures the performance of the model when trading off the precision, i.e. the proportion of correctly classified cases among predicted positive cases, with recall, i.e. the proportion of all positive cases correctly predicted. AUROC is similar to AUPR but measures the performance of the model when trading off the recall with the false positive rate, i.e. the proportion of predicted positive cases which are in fact negative. Accuracy, AUPR and AUROC all theoretically range from 0 to 1 where 0 is the worst performing model and 1 is the best performing model. The Brier score also range from 0 to 1, but for this score a lower value indicate better performance (Zhou et al., 2021).

Crucial when evaluating machine learning models such as the random forest and xgboost models is that the evaluation happens on data which the model has not seen, since the models tend to overfit (i.e. perform artificially well) on data which are within the sample. To ensure that the constituent models were evaluated on out-of-sample data, we used two different methods to evaluate the performance. We evaluated the performance using:

- Fixed test window, where the period 2016-2021 was treated as out of sample test data. Consequently, the models were trained on data in 1989-2015 and predictions were made for 2016-2021 out of sample.
- Rolling test window where the period 2011-2021 was treated as the test period. For each year, the models were trained using all year prior, and predictions made one year into the future. I.e. for 2011 the period 1989-2010 was used as the training data, while for 2021 the period 1989-2020 was used as training data

3.3 Evaluation results

The results of the evaluation of the constituent model for our final ensemble can be seen in Appendix C1. In appendix C2 we also include all constituent models with all four performance metrics. The lists in Appendix C1-2 are sorted on the *brier-score* metric, where a lower value indicates a better performance. The results show that the best performing models are roughly on-par with one another with accuracies of 83-85%, brier scores of 0.12-0.15 and AUPR and AUROC scores of 0.8-0.85 and 0.9-0.93 for both the one and two year ahead forecasts.

Worth noting here is that the best performing models all include variables relating to the characteristics of the last held election, including the level of violence of that election. This is in line with results from other conflict forecasting efforts, such as the ViEWS project, where conflict history are usually the best predictors of future violence. The ensemble using a subset of nine well performing models also has good performance, with the highest AUPR and AUROC (tied) scores of all models.

The constituent models relating to the digital vulnerability aspect of the elections, i.e. the DSP models, are not among the best performing models. However, the combination model of DSP infrastructure, WDI structural factors, and VDEM mid level indicies do perform well, and the full DSP model also does relatively well. In order to incorporate the digital vulnerability aspects into the final forecasts, the full DSP model is also included in the ensemble of best performing models.

In figure 2 below are the performance metrics for the one year ahead forecasts in the fixed test window specification for the DSP full model and the nine model ensemble using the random forest and xgboost respectively. We also note that the random forest models and xgboost model perform roughly on par with one-another. The same figures for the rolling test window and the two year ahead fixed window can be found in appendix C1.

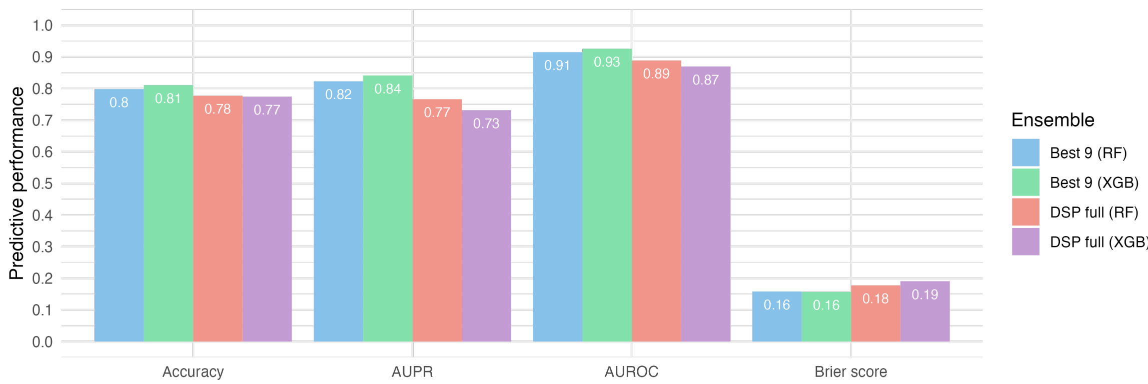

Figure 2. Performance comparison of xgboost and random forest for the one year ahead forecasts in the fixed test window

3.4 Digital vulnerability and electoral violence

One downside with using machine learning models such as the random forest and xgboost models are that it may be difficult to draw inferences about how specific factors affect the predictions. As this report has a specific interest in digital vulnerability we decided to do an exploration on how different indicators of digital vulnerability and infrastructure affect the predictions. To do this, we ran an analysis of the variable importance for the full DSP model. Variable importance measures the contribution of the individual variables on the performance of the model and can be used to identify which factors in a model are most important in the prediction. Running this on the random forest model shows that the features in the digital vulnerability model which has the largest impact on the performance of the model are v2smarrest, measuring the likelihood of arrests being made for political content, v2smgovdom, measuring the dissemination of misinformation by the government through digital channels, and v2smregcon, which measures what online content is regulated by legislation covering the digital sphere. For the xgboost model, the feature v2smdefabu measuring abuse of defamation and copyright law by elites also had a high importance (Mechkova et al., 2022a).

We also evaluated the impact of the DSP full model on the predictive performance of the ensemble by doing an ablation study, i.e. by comparing the performance of the ensemble when including compared to excluding the model. This analysis showed that excluding the DSP full model does not have a major impact on the ensmebles performance. In fact, the ensemble performs slightly better when the dsp full model is excluded. However, we still retain the DSP full model in the ensmble as we believe that it fills an important area of the feature space for the ensemble.

4 Forecasts for 2022 and 2023

In appendices D and E, the out of sample forecasts for 2022 and 2023 are presented for all countries with scheduled elections in these two years. The forecasts are presented for the nine model ensemble (appendix D) and the full DSP model (appendix E) averaged across the random forest and xgboost versions of the models and are sorted by the *risk index*.

The results show that the countries most at risk of electoral violence in 2022 were dictatorships such as Equatorial Guinea and Turkmenistan. The fact that they rank so high on the risk-index is related to the combination of electoral violence by non-government actors and the intimidation and harassment by government actors. For instance, in the country with the highest score on the risk index, Equatorial Guinea, no visible electoral violence was reported. Instead, the repression was so dense that no functioning opposition ran in the elections which were held.

For 2023, the countries most at risk of electoral violence contain Turkmenistan, Cuba, and Gabon. Comparing the forecasts from the nine model ensemble and the DSP full model no major differences can be spotted, although the ranking along the risk index has changed somewhat.

References

Breiman, Leo (2001). "Random forests". In: *Machine learning* 45, pp. 5–32.

Chen, Tianqi and Carlos Guestrin (2016). "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Coppedge, Michael et al. (2021a). *V-Dem Codebook v11*.

Coppedge, Michael et al. (2021b). *V-Dem Country-Year/Country-Date Dataset v11*.

Fjelde, Hanne et al. (2021). *The Deadly Electoral Conflict Dataset (DECO) Codebook*.

Hegre, Håvard et al. (2019). "ViEWS: A political violence early-warning system". In: *Journal of peace research* 56.2, pp. 155–174.

Hegre, Håvard et al. (2021). "ViEWS2020: revising and evaluating the ViEWS political violence early-warning system". In: *Journal of peace research* 58.3, pp. 599–611.

Mechkova, Valeriya et al. (2022a). *Digital Society Project Codebook*.

— (2022b). *Digital Society Project Dataset v4*.

Zhou, Jianlong et al. (2021). "Evaluating the quality of machine learning explanations: A survey on methods and metrics". In: *Electronics* 10.5, p. 593.

Appendix A: The outcome of electoral violence

Definition of v2elintim:

Question: In this national election, were opposition candidates/parties/campaign workers subjected to repression, intimidation, violence, or harassment by the government, the ruling party, or their agents?

Clarification: Other types of clearly distinguishable civil violence, even if politically motivated, during the election period should not be factored in when scoring this indicator (it is dealt with separately).

Responses:

0: Yes. The repression and intimidation by the government or its agents was so strong that the entire period was quiet.

1: Yes, frequent: There was systematic, frequent and violent harassment and intimidation of the opposition by the government or its agents during the election period.

2: Yes, some. There was periodic, not systematic, but possibly centrally coordinated — harassment and intimidation of the opposition by the government or its agents.

3: Restrained. There were sporadic instances of violent harassment and intimidation by the government or its agents, in at least one part of the country, and directed at only one or two local branches of opposition groups.

4: None. There was no harassment or intimidation of opposition by the government or its agents, during the election campaign period and polling day.

Definition of v2elpeace:

Question: In this national election, was the campaign period, election day, and post-election process free from other types (not by the government, the ruling party, or their agents) of violence related to the conduct of the election and the campaigns (but not conducted by the government and its agents)?

Responses:

0: No. There was widespread violence between civilians occurring throughout the election period, or in an intense period of more than a week and in large swaths of the country. It resulted in a large number of deaths or displaced refugees.

1: Not really. There were significant levels of violence but not throughout the election period or beyond limited parts of the country. A few people may have died as a result, and some people may have been forced to move temporarily.

2: Somewhat. There were some outbursts of limited violence for a day or two, and only in a small part of the country. The number of injured and otherwise affected was relatively small.

3: Almost. There were only a few instances of isolated violent acts, involving only a few people; no one died and very few were injured.

4: Peaceful. No election-related violence between civilians occurred.

Definition of electoral violence target:

Recoding of v2elintim and v2elpeace into three level ordinal variables with values 0-1 corresponding to 'Severe electoral violence/intimidation', values 2-3 corresponding to 'limited electoral violence/intimidation', and value 4 corresponding to 'no electoral violence/intimidation' perpetrated by government actors (v2elintim) and non-government actors (v2elpeace) respectively. The target 'electoral violence' is then taken as the max of the two three level ordinal variables measuring electoral violence/intimidation.

Appendix B1: Constituent models in the final ensemble

The final ensemble for the predictions consist of an unweighted ensemble of nine of the best performing constituent models in the evaluations outlined in Appendix C. The nine models span a wide range of the features and several of the major outlined themes of models and should therefore be relatively robust to potential changes in the causes of electoral violence. The nine models included in the ensemble are: While some of the individual models perform better than the final ensemble, we still believe that the ensemble is the appropriate choice for the final predictions since this ensemble spans a broader set of themes and should therefore be more robust to potential second degree overfitting on the test sets.

- **cmb vdem elcharirr excl wdistruct:** Constituent model with a combination of three themes: election irregularities in the last held election, vdem indicators on political exclusion (due to socio-economic, gender, urban/rural, political, or social group), and structural indicators from the world bank (including population, age composition, imr, life expectancy, and gdp)
- **lastdec elchar:** Constituent model containing variables on the characteristics of the last election, including irregularities, campaign finances, turnout, votebuying, outcome, etc
- **cmb vdem elcharirr cl wdistruct:** Constituent model with a combination of three themes: election irregularities in the last held election, vdem indicators on civil liberties, and structural indicators from the world bank (including population, age composition, imr, life expectancy, and gdp)
- **lastdec evhist:** Constituent model containing variables on the level of violence in the last election (both by government and by others)
- **cmb vdem mid wdistruc dspinfr:** Constituent model with a combination of three themes: vdem mid level democracy indicators (covering a wide range of dimensions of democracy), world bank structural indicators (same as above), and digital societies project infrastructure model (including cyber security and monitoring of social media, internet usage from wdi, and media usage by parties from vdem)
- **cmb vdem high wdistruc:** Constituent model with a combination of two themes: vdem high level democracy indicators (including broad indexes of electoral, liberal, egalitarian, and deliberative democracy), and world bank structural indicators (same as above)
- **dsp full model:** Constituent model with all variables from the digital societies project
- **vdem full model:** Constituent model with all variables from the vdem project
- **vdem mid:** Constituent model with all mid-level democracy indices from vdem

Appendix B2: Constituent models description

Table 1. Election history (VDEM-CD) constituent models

Model name	Description of features	included features
evhist long	Three features measuring the number of peaceful elections in a row, the number of violent (severe) elections in a row, and number of elections since last constitutional change	cons_elect, peaceful_streak, violent_streak
evhist short	Intimidation and harassment by the government in the last election and prevalence of other electoral violence in last election (basically lagged DV)	v2elintim_osp, v2elpeace_osp
elchar	Election (last election last year) characteristics features from vdem-cd	v2asuffrage, v2elcomvot, v2elgvsuflvl, v2eldonate, v2elpubfin, v2elembaut, v2elembcap, v2elmulpar, v2elrgstry, v2elvotbuy, v2elirreg, v2elintim, v2elpeace, v2elboycot, v2elfrcamp, v2elpdcamp, v2elpaidig, v2elfrfair, v2eldommon, v2elintmon, v2elmonden, v2elmonref, v2elaccept, v2elasmoff, v2elvaptrn, cons_elect
elchar_str	Election (last election last year) characteristics, structural features	v2asuffrage, v2elcomvot, v2elgvsuflvl, v2eldonate, v2elpubfin, v2elembaut, v2elembcap, v2elmulpar, v2elrgstry, v2elvotbuy, v2elfrcamp, v2elpdcamp, v2elpaidig, v2eldommon, v2elintmon, v2elvaptrn
elchar_irr	Election (last election last year) characteristics, irregularity related features	v2elembaut, v2elembcap, v2elmulpar, v2elvotbuy, v2elirreg, v2elintim, v2elpeace, v2elboycot, v2elfrfair, v2elmonden, v2elmonref

Table 2. DSP constituent models

Model name	Description of features	included features
DSP full model	All interval scale features from DSP	v2smgovdom, v2smgovab, v2smpardom, v2smparab, v2smfordom, v2smforads, v2smgovfilcap, v2smgovfilprc, v2smgovshutcap, v2smgovshut, v2smgovsm, v2smgovsmalt, v2smgovsmmon, v2smgovsmcenprc, v2smgovcapsec, v2smregcon, v2smprivex, v2smprivcon, v2smregcap, v2smregapp, v2smlawpr, v2smdefabu, v2smonex, v2smonper, v2smmefra, v2smorgviol, v2smorgavgact, v2smorgelitact, v2smcamp, v2smarrest, v2smpolsoc, v2smpolhate
dsp infra	Digital infrastructure features including media features from vdem-cy, emb capacity from vdem-cy, cyber security + monitoring and surveillance of social media from dspand internet use from WDI	v2smonex, v2elfrcamp, v2mecrit, v2merange, v2elembaut, v2smgovcapsec, v2smpolcap, v2smregcap, v2smgovfilprc, v2smgovsmmon, v2smgovsmcenprc, v2smarrest, it.net.user.zs
dsp smclimate	Social media climate from dsp including dissemination of disinformation, online polarization and hatespeach, and traditional use of social media by elites/political candidates	v2smgovdom, v2smpardom, v2smfordom, v2smonper, v2smmefra, v2smpolsoc, v2smpolhate, v2smorgelitact, v2smcamp
dsp monitoring	DSP features relating to government monitoring, surveillance and repression online	v2smregcap, v2smgovfilprc, v2smgovsmmon, v2smgovsmcenprc, v2smarrest
dsp disinform	DSP features relating to disinformation online	v2smgovdom, v2smpardom, v2smfordom
dsp media	DSP and VDEM-CY features on traditional and social media usage	v2smonex, v2elfrcamp, v2mecrit, v2merange

Table 3. VDEM-CY constituent models

Model name	Description of features	included features
vdem high	vdem-cy high level indices	v2x_polyarchy, v2x_libdem, v2x_partipdem, v2x_delibdem, v2x_egaldem
vdem full	vdem-cy full model	v2x_freexp_altinf, v2x_frassoc_thick, v2x_suffr, v2x_delibdem, v2xel_frefair, v2x_elecoff, v2xcl_rol, v2x_jucon, v2xlg_legcon, v2x_cspart, v2xdd_dd, v2xel_locelec, v2xel_regelec, v2xdl_delib, v2xeg_eqprotec, v2xeg_eqaccess, v2x_veracc, v2x_diagacc, v2x_horacc, v2x_ex_confidence, v2x_ex_direlect, v2x_ex_hereditary, v2x_ex_military, v2x_ex_party, v2xnp_client, v2xnp_pres, v2xnp_regcorr, v2x_clphy, v2x_clpol, v2x_clpriv, v2xpe_exlecon, v2xpe_exlgender, v2xpe_exlgeo, v2xpe_exlpol, v2xpe_exlsocgr, v2x_corr, v2x_execorr, v2x_pubcorr, v2x_gencl, v2x_gencs, v2x_genpp, v2x_rule, v2xdd_cic, v2xdd_i_ci, v2xdd_i_rf, v2xdd_toc, v2xdd_i_pl, v2xdd_i_or, v2xcs_ccsi, v2x_EDcomp_thick, v2xcl_disc, v2xcl_dmove, v2xcl_slave, v2xex_elecleg, v2xme_altinf, v2xps_party, v2x_divparctrl, v2x_feduni, v2xca_academ
vdem gender	vdem-cy gender features	v2x_gencl, v2x_gencs, v2x_genpp, v2x_gender
vdem mid	vdem-cy and cd mid level component indicies	v2x_api, v2x_mpi, v2x_freexp_altinf, v2x_frassoc_thick, v2x_suffr, v2x_delibdem, v2xel_frefair, v2x_elecoff, v2xcl_rol, v2x_jucon, v2xlg_legcon, v2x_cspart, v2xdd_dd, v2xel_locelec, v2xel_regelec, v2xdl_delib, v2xeg_eqprotec, v2xeg_eqaccess
vdem mid2	vdem-cy mid level indices	v2x_accountability, v2x_neopat, v2x_civlib, v2x_gender, v2x_corr, v2x_rule, v2xcs_ccsi, v2xps_party, v2x_divparctrl, v2x_feduni
vdem neopat	vdem-cy neopatriamonialism features	v2x_neopat, v2xnp_client, v2xnp_pres, v2xnp_regcorr
vdem party	vdem-cy features on party power, division and control	v2xps_party, v2x_divparctrl, v2x_feduni
vdem excl	vdem-cy features on exclusion of groups	v2xpe_exlecon, v2xpe_exlgender, v2xpe_exlgeo, v2xpe_exlpol, v2xpe_exlsocgr
vdem cl	vdem-cy features on civil liberties	v2x_clphy, v2x_clpol, v2x_clpriv, v2x_civlib
vdem corr	vdem-cy features on corruption	v2x_corr, v2x_execorr, v2x_pubcorr
vdem ac-countability	vdem-cy features on accountability	v2x_accountability, v2x_veracc, v2x_diagacc, v2x_horacc

Table 4. WDI Constituent models

Model name	Description of features	included features
wdi full	WDI full model	sp.pop.totl, ms.mil.xpnd.zs, ms.mil.xpnd.gd.zs, dt.oda.odat.pc.zs, nv.agr.totl.kn, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.le00.in, se.enr.prim.fm.zs, se.enr.prsc.fm.zs, se.prm.nenr, sh.sta.maln.zs, sh.sta.stnt.zs, sl.tlf.totl.fe.zs, sm.pop.totl.zs, sp.dyn.imrt.in, sh.dyn.mort.fe, sp.pop.0014.fe.zs, sp.pop.1564.fe.zs, sp.pop.65up.fe.zs, sp.pop.grow, sp.urb.totl.in.zs, se.xpd.totl.gb.zs, se.xpd.totl.gd.zs, sl.uem.neet.zs, ny.gdp.petr.rt.zs, ny.gdp.totl.rt.zs, it.net.user.zs
wdi structural	WDI structural factors, including population, age composition, imr, life expectancy, and gdp	sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs
wdi education	WDI education factors, including enrollment and expenditure	se.enr.prim.fm.zs, se.enr.prsc.fm.zs, se.prm.nenr, se.xpd.totl.gb.zs, se.xpd.totl.gd.zs
wdi resources	WDI factors on resources and gdp	ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, dt.oda.odat.pc.zs, ny.gdp.petr.rt.zs, ny.gdp.totl.rt.zs

Table 5. Combination Constituent models

Model name	Description of features	included features
cmb smclimate use	Combination of features of the smclimate model and social media usage	v2smgovdom, v2smpardom, v2smfordom, v2smonper, v2smmefra, v2smpolsoc, v2smpolhate, v2smorgelitact, v2smcamp
cmb smclimate secuse	Combination of features of the smclimate model and social media usage + cyber security	v2smgovdom, v2smpardom, v2smfordom, v2smonper, v2smmefra, v2smpolsoc, v2smpolhate, v2smorgelitact, v2smcamp, v2smgovcapsec, v2smpolcap
cmb smclimate disinformuse	Combination of features of the smclimate model and social media usage + disinformation	v2smgovdom, v2smpardom, v2smfordom, v2smonper, v2smmefra, v2smpolsoc, v2smpolhate, v2smorgelitact, v2smcamp
cmb vdemhigh wdistruc	Combination of features from vdem_high and wdi_structural	v2x_polyarchy, v2x_libdem, v2x_partipdem, v2x_delibdem, v2x_egaldem, sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs
cmb vdemmid wdistruc	Combination of features from vdem_mid and wdi_structural	v2x_accountability, v2x_neopat, v2x_civlib, v2x_gender, v2x_corr, v2x_rule, v2xcs_ccsi, v2xps_party, v2x_divparctrl, v2x_feduni, sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs
cmb vdemmid wdistruc dspinfr	Combination of features from vdem_mid and wdi_structural and dsp_infra	v2x_accountability, v2x_neopat, v2x_civlib, v2x_gender, v2x_corr, v2x_rule, v2xcs_ccsi, v2xps_party, v2x_divparctrl, v2x_feduni, sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs, v2smonex, v2elfrcamp, v2mecrit, v2merange, v2elembaut, v2smgovcapsec, v2smpolcap, v2smregcap, v2smgovfilprc, v2smgovsmmon, v2smgovsmcenprc, v2smarrest, it.net.user.zs
cmb elcharirr cl wdistruct	Combination of features from vdem_cl, elchar_irr, and wdi_structural	v2elembaut, v2elembcap, v2elmulpar, v2elvotbuy, v2elirreg, v2elintim, v2elpeace, v2elboycot, v2elfrfair, v2elmonden, v2elmonref, v2x_clphy, v2x_clpol, v2x_clpriv, v2x_civlib, sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs
cmb vdem elcharirr excl wdistruct	Combination of vdem_excl, elchar_irr, and wdi_structural	v2elembaut, v2elembcap, v2elmulpar, v2elvotbuy, v2elirreg, v2elintim, v2elpeace, v2elboycot, v2elfrfair, v2elmonden, v2elmonref, v2xpe_exlecon, v2xpe_exlgender, v2xpe_exlgeo, v2xpe_exlpol, v2xpe_exlsocgr, sp.pop.totl, ny.gdp.pcap.kd, ny.gdp.pcap.kd.zg, sp.dyn.imrt.in, sp.dyn.le00.in, sp.pop.0014.fe.zs, sp.pop.grow, sp.pop.65up.fe.zs
cmb full	Combination of all features above	

Appendix C1: Constituent model performance for models included in the ensemble

Table 6. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.84	0.93
2	elchar	Last Election characteristcs	0.86	0.12	0.83	0.93
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.85	0.12	0.83	0.93
4	ensemble best	Best 9 ensemble	0.84	0.13	0.85	0.93
5	evhist	EV History (last election)	0.83	0.13	0.82	0.92
6	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.84	0.14	0.81	0.91
7	vdem full	VDEM full model	0.83	0.14	0.82	0.92
8	ensemble all	Full ensemble	0.84	0.14	0.84	0.92
9	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.83	0.14	0.82	0.91
10	dsp full	DSP full model	0.82	0.14	0.81	0.91
11	vdem mid	VDEM mid incicies 1	0.82	0.15	0.83	0.92

Table 7. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	elchar	Last Election characteristcs	0.83	0.15	0.82	0.91
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.81	0.91
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.82	0.15	0.80	0.90
4	ensemble best	Best 9 ensemble	0.80	0.16	0.82	0.91
5	evhist	EV History (last election)	0.79	0.16	0.81	0.90
6	ensemble all	Full ensemble	0.79	0.17	0.79	0.90
7	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.78	0.18	0.76	0.88
8	vdem mid	VDEM mid incicies 1	0.77	0.18	0.79	0.90
9	dsp full	DSP full model	0.78	0.18	0.77	0.89
10	vdem full	VDEM full model	0.77	0.18	0.79	0.90
11	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.79	0.18	0.78	0.89

Table 8. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.83	0.93
2	elchar	Last Election characteristcs	0.85	0.12	0.82	0.92
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.84	0.13	0.83	0.92
4	ensemble best	Best 9 ensemble	0.85	0.13	0.83	0.93
5	evhist	EV History (last election)	0.82	0.14	0.81	0.92
6	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.82	0.14	0.80	0.91
7	vdem full	VDEM full model	0.83	0.14	0.81	0.92
8	ensemble all	Full ensemble	0.84	0.14	0.82	0.92
9	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.81	0.15	0.80	0.91
10	dsp full	DSP full model	0.82	0.15	0.80	0.91
11	vdem mid	VDEM mid incicies 1	0.82	0.15	0.81	0.92

Table 9. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.81	0.16	0.78	0.90
2	elchar	Last Election characteristcs	0.80	0.16	0.77	0.90
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.79	0.16	0.77	0.89
4	ensemble best	Best 9 ensemble	0.79	0.17	0.79	0.90
5	evhist	EV History (last election)	0.77	0.18	0.78	0.89
6	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.77	0.18	0.75	0.88
7	ensemble all	Full ensemble	0.78	0.18	0.76	0.89
8	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.18	0.75	0.87
9	vdem full	VDEM full model	0.76	0.19	0.75	0.89
10	vdem mid	VDEM mid incicies 1	0.77	0.19	0.76	0.88
11	dsp full	DSP full model	0.76	0.19	0.73	0.88

Table 10. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	evhist	EV History (last election)	0.86	0.12	0.85	0.94
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.85	0.94
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.85	0.12	0.85	0.94
4	elchar	Last Election characteristcs	0.85	0.13	0.85	0.94
5	ensemble best	Best 9 ensemble	0.86	0.13	0.86	0.94
6	vdem kitchen	VDEM kitchen sink	0.84	0.15	0.83	0.92
7	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.83	0.15	0.81	0.91
8	vdem mid	VDEM mid incicies 1	0.82	0.16	0.82	0.92
9	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.82	0.16	0.82	0.92
10	dsp kitchen	DSP kitchen sink	0.82	0.16	0.80	0.91

Table 11. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.85	0.93
2	evhist	EV History (last election)	0.83	0.15	0.84	0.92
3	elchar	Last Election characteristcs	0.82	0.15	0.83	0.92
4	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.81	0.15	0.83	0.92
5	ensemble best	Best 9 ensemble	0.81	0.16	0.84	0.93
6	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.78	0.18	0.79	0.90
7	vdem kitchen	VDEM kitchen sink	0.77	0.19	0.78	0.89
8	vdem mid	VDEM mid incicies 1	0.77	0.19	0.79	0.89
9	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.19	0.78	0.89
10	dsp kitchen	DSP kitchen sink	0.77	0.19	0.73	0.87

Table 12. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.83	0.92
2	ensemble best	Best 9 ensemble	0.84	0.15	0.83	0.92
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.82	0.15	0.82	0.92
4	elchar	Last Election characteristcs	0.82	0.16	0.82	0.91
5	evhist	EV History (last election)	0.82	0.16	0.80	0.91
6	vdem kitchen	VDEM kitchen sink	0.81	0.17	0.80	0.90
7	vdem mid	VDEM mid incicies 1	0.80	0.17	0.78	0.90
8	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.81	0.17	0.78	0.90
9	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.82	0.17	0.77	0.89
10	dsp kitchen	DSP kitchen sink	0.80	0.18	0.76	0.88

Table 13. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	ensemble best	Best 9 ensemble	0.78	0.19	0.80	0.90
2	elchar	Last Election characteristcs	0.78	0.19	0.78	0.89
3	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.76	0.19	0.77	0.89
4	evhist	EV History (last election)	0.78	0.19	0.77	0.88
5	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.76	0.20	0.77	0.88
6	vdem kitchen	VDEM kitchen sink	0.77	0.20	0.75	0.88
7	vdem mid	VDEM mid incicies 1	0.74	0.20	0.75	0.88
8	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.21	0.72	0.87
9	dsp kitchen	DSP kitchen sink	0.75	0.21	0.73	0.85
10	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.75	0.21	0.73	0.87

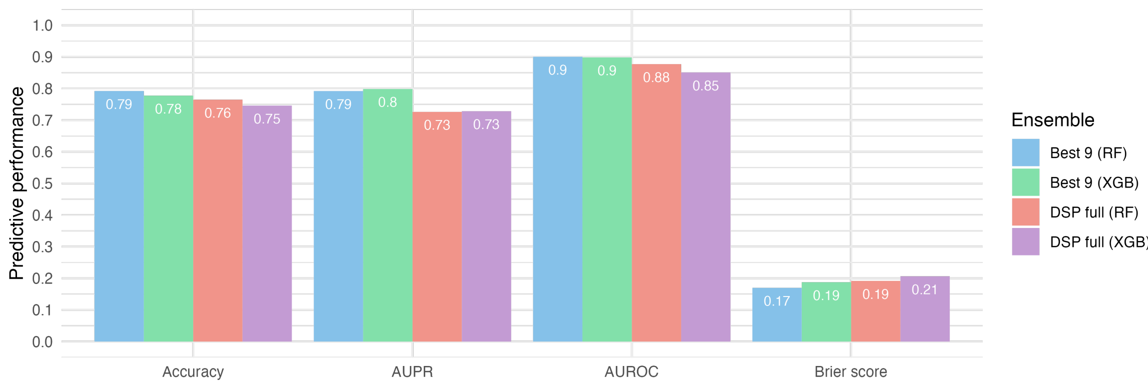

Figure 3. Performance comparison of xgboost and random forest for the two year ahead forecasts in the fixed test window

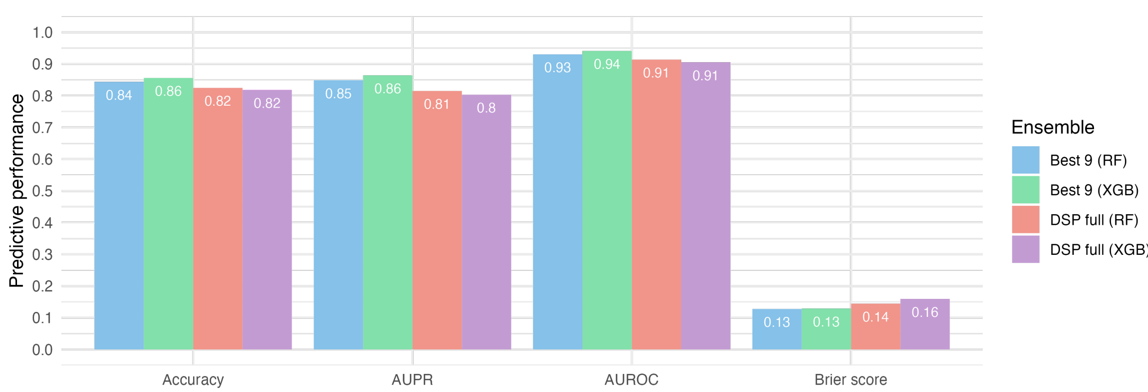

Figure 4. Performance comparison of xgboost and random forest for the one year ahead forecasts in the rolling test window

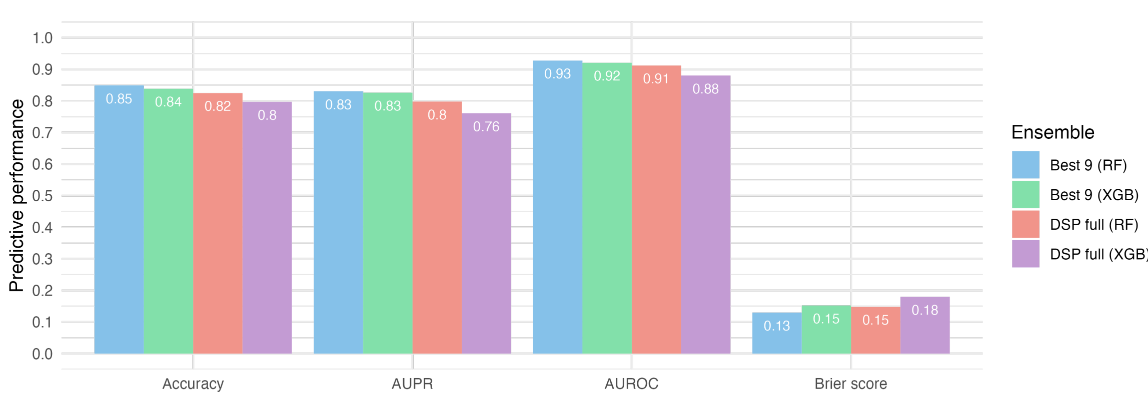

Figure 5. Performance comparison of xgboost and random forest for the two year ahead forecasts in the rolling test window

Appendix C2: All constituent model performance

Table 14. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.84	0.93
2	elchar	Last Election characteristcs	0.86	0.12	0.83	0.93
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.85	0.12	0.83	0.93
4	elcharirr	Last Election characteristcs (irregularities)	0.84	0.13	0.84	0.93
5	ensemble ablation	Ablation best 9	0.85	0.13	0.85	0.93
6	ensemble best	Best 9 ensemble	0.84	0.13	0.85	0.93
7	evhist	EV History (last election)	0.83	0.13	0.82	0.92
8	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.84	0.14	0.81	0.91
9	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.84	0.14	0.81	0.91
10	vdem full	VDEM full model	0.83	0.14	0.82	0.92
11	ensemble all	Full ensemble	0.84	0.14	0.84	0.92
12	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.83	0.14	0.82	0.91
13	dsp full	DSP full model	0.82	0.14	0.81	0.91
14	vdem mid	VDEM mid incicies 1	0.82	0.15	0.83	0.92
15	vdem excl	VDEM exclusion	0.82	0.15	0.83	0.92
16	evhist long	EV History (long)	0.84	0.15	0.80	0.91
17	elcharstr	Last Election characteristcs (structural)	0.82	0.15	0.78	0.90
18	vdem mid2	VDEM mid indices 2	0.82	0.15	0.80	0.90
19	dsp infra	DSP infrastructure	0.82	0.15	0.82	0.91
20	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.81	0.16	0.80	0.91

Table 15. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	dsp monitoring	DSP Social media monitoring	0.79	0.16	0.79	0.90
22	vdem high	VDEM high indices	0.80	0.16	0.79	0.89
23	cmb mega full	Comb. full model	0.79	0.16	0.70	0.83
24	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.80	0.16	0.79	0.89
25	cmb dsp smclimate use	Comb. sm climate + use	0.80	0.16	0.79	0.89
26	dsp smclimate	DSP Social media climate	0.80	0.17	0.79	0.89
27	vdem neopat	VDEM neopat	0.79	0.17	0.77	0.89
28	vdem account	VDEM accountability	0.78	0.17	0.77	0.89
29	vdem cl	VDEM civil liberties	0.76	0.19	0.74	0.86
30	vdem pp	VDEM parties and power	0.77	0.19	0.73	0.87
31	vdem corrupt	VDEM corruption	0.76	0.19	0.74	0.87
32	vdem gender	VDEM gender	0.75	0.19	0.73	0.86
33	wdi structural	WDI structural	0.75	0.19	0.71	0.86
34	dsp media	DSP traditional and social media	0.76	0.20	0.75	0.86
35	elcharout	Last Election outcome	0.73	0.20	0.71	0.86
36	wdi full	WDI full model	0.77	0.20	0.61	0.78
37	dsp disinform	DSP disinformation	0.73	0.20	0.73	0.86
38	wdi resources	WDI resources	0.69	0.24	0.65	0.81
39	wdi education	WDI education	0.68	0.25	0.60	0.77

Table 16. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	elchar	Last Election characteristcs	0.83	0.15	0.82	0.91
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.81	0.91
3	elcharirr	Last Election characteristcs (irregularities)	0.80	0.15	0.83	0.92
4	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.82	0.15	0.80	0.90
5	ensemble best	Best 9 ensemble	0.80	0.16	0.82	0.91
6	ensemble ablation	Ablation best 9	0.80	0.16	0.82	0.91
7	evhist	EV History (last election)	0.79	0.16	0.81	0.90
8	ensemble all	Full ensemble	0.79	0.17	0.79	0.90
9	evhist long	EV History (long)	0.80	0.18	0.77	0.89
10	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.78	0.18	0.76	0.88
11	vdem mid	VDEM mid incicies 1	0.77	0.18	0.79	0.90
12	dsp full	DSP full model	0.78	0.18	0.77	0.89
13	vdem full	VDEM full model	0.77	0.18	0.79	0.90
14	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.79	0.18	0.78	0.89
15	elcharstr	Last Election characteristcs (structural)	0.78	0.18	0.75	0.88
16	dsp infra	DSP infrastructure	0.78	0.18	0.78	0.89
17	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.78	0.18	0.73	0.87
18	dsp monitoring	DSP Social media monitoring	0.76	0.19	0.76	0.88
19	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.76	0.19	0.75	0.88
20	vdem excl	VDEM exclusion	0.76	0.19	0.77	0.88

Table 17. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	dsp smclimate	DSP Social media climate	0.75	0.19	0.75	0.87
22	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.75	0.19	0.74	0.87
23	cmb dsp smclimate use	Comb. sm climate + use	0.75	0.19	0.74	0.87
24	vdem account	VDEM accountability	0.75	0.19	0.75	0.88
25	vdem high	VDEM high indices	0.76	0.19	0.75	0.87
26	vdem mid2	VDEM mid indices 2	0.75	0.20	0.74	0.87
27	vdem neopat	VDEM neopat	0.76	0.21	0.70	0.85
28	elcharout	Last Election outcome	0.70	0.22	0.69	0.85
29	cmb mega full	Comb. full model	0.72	0.22	0.65	0.80
30	wdi structural	WDI structural	0.72	0.22	0.67	0.83
31	vdem pp	VDEM parties and power	0.71	0.22	0.70	0.84
32	vdem gender	VDEM gender	0.71	0.23	0.67	0.82
33	vdem corrupt	VDEM corruption	0.69	0.24	0.70	0.84
34	wdi full	WDI full model	0.71	0.24	0.61	0.77
35	vdem cl	VDEM civil liberties	0.69	0.24	0.65	0.81
36	dsp disinform	DSP disinformation	0.66	0.25	0.67	0.82
37	dsp media	DSP traditional and social media	0.69	0.25	0.67	0.81
38	wdi resources	WDI resources	0.64	0.28	0.61	0.77
39	wdi education	WDI education	0.61	0.30	0.53	0.72

Table 18. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.83	0.93
2	elchar	Last Election characteristcs	0.85	0.12	0.82	0.92
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.84	0.13	0.83	0.92
4	elcharirr	Last Election characteristcs (irregularities)	0.84	0.13	0.83	0.93
5	ensemble ablation	Ablation best 9	0.85	0.13	0.83	0.93
6	ensemble best	Best 9 ensemble	0.85	0.13	0.83	0.93
7	evhist	EV History (last election)	0.82	0.14	0.81	0.92
8	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.84	0.14	0.81	0.91
9	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.82	0.14	0.80	0.91
10	vdem full	VDEM full model	0.83	0.14	0.81	0.92
11	ensemble all	Full ensemble	0.84	0.14	0.82	0.92
12	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.81	0.15	0.80	0.91
13	dsp full	DSP full model	0.82	0.15	0.80	0.91
14	vdem mid	VDEM mid incicies 1	0.82	0.15	0.81	0.92
15	vdem excl	VDEM exclusion	0.81	0.15	0.81	0.92
16	vdem mid2	VDEM mid indices 2	0.81	0.15	0.80	0.91
17	dsp infra	DSP infrastructure	0.81	0.15	0.80	0.91
18	elcharstr	Last Election characteristcs (structural)	0.82	0.15	0.77	0.90
19	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.80	0.16	0.79	0.91
20	evhist long	EV History (long)	0.83	0.16	0.78	0.89

Table 19. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	dsp monitoring	DSP Social media monitoring	0.79	0.16	0.77	0.90
22	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.80	0.16	0.77	0.89
23	dsp smclimate	DSP Social media climate	0.81	0.16	0.78	0.89
24	cmb mega full	Comb. full model	0.80	0.16	0.71	0.83
25	vdem high	VDEM high indices	0.80	0.17	0.80	0.90
26	cmb dsp smclimate use	Comb. sm climate + use	0.80	0.17	0.77	0.89
27	vdem neopat	VDEM neopat	0.79	0.17	0.77	0.89
28	vdem account	VDEM accountability	0.78	0.17	0.78	0.89
29	vdem cl	VDEM civil liberties	0.76	0.19	0.74	0.87
30	vdem pp	VDEM parties and power	0.77	0.19	0.74	0.88
31	vdem gender	VDEM gender	0.76	0.19	0.74	0.87
32	vdem corrupt	VDEM corruption	0.76	0.19	0.75	0.87
33	wdi structural	WDI structural	0.76	0.19	0.72	0.87
34	elcharout	Last Election outcome	0.76	0.19	0.72	0.86
35	dsp media	DSP traditional and social media	0.75	0.20	0.75	0.87
36	wdi full	WDI full model	0.79	0.20	0.65	0.80
37	dsp disinform	DSP disinformation	0.74	0.20	0.73	0.87
38	wdi resources	WDI resources	0.71	0.23	0.66	0.82
39	wdi education	WDI education	0.71	0.24	0.64	0.79

Table 20. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.81	0.16	0.78	0.90
2	elchar	Last Election characteristcs	0.80	0.16	0.77	0.90
3	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.79	0.16	0.77	0.89
4	elcharirr	Last Election characteristcs (irregularities)	0.78	0.16	0.79	0.90
5	ensemble ablation	Ablation best 9	0.79	0.17	0.79	0.90
6	ensemble best	Best 9 ensemble	0.79	0.17	0.79	0.90
7	evhist	EV History (last election)	0.77	0.18	0.78	0.89
8	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.77	0.18	0.75	0.88
9	ensemble all	Full ensemble	0.78	0.18	0.76	0.89
10	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.18	0.75	0.87
11	vdem full	VDEM full model	0.76	0.19	0.75	0.89
12	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.79	0.19	0.73	0.87
13	vdem mid	VDEM mid incicies 1	0.77	0.19	0.76	0.88
14	evhist long	EV History (long)	0.78	0.19	0.75	0.87
15	dsp full	DSP full model	0.76	0.19	0.73	0.88
16	elcharstr	Last Election characteristcs (structural)	0.77	0.19	0.71	0.87
17	dsp infra	DSP infrastructure	0.77	0.19	0.74	0.88
18	vdem excl	VDEM exclusion	0.76	0.19	0.74	0.88
19	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.75	0.20	0.72	0.87
20	dsp monitoring	DSP Social media monitoring	0.74	0.20	0.72	0.87

Table 21. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the random forest model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	dsp smclimate	DSP Social media climate	0.75	0.20	0.72	0.86
22	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.74	0.20	0.72	0.86
23	cmb dsp smclimate use	Comb. sm climate + use	0.74	0.20	0.72	0.86
24	vdem mid2	VDEM mid indices 2	0.73	0.20	0.72	0.86
25	vdem high	VDEM high indices	0.75	0.21	0.72	0.85
26	vdem account	VDEM accountability	0.74	0.21	0.71	0.85
27	elcharout	Last Election outcome	0.73	0.22	0.67	0.84
28	vdem neopat	VDEM neopat	0.74	0.22	0.68	0.84
29	cmb mega full	Comb. full model	0.72	0.23	0.61	0.78
30	vdem pp	VDEM parties and power	0.70	0.23	0.68	0.84
31	vdem gender	VDEM gender	0.72	0.23	0.68	0.82
32	wdi structural	WDI structural	0.70	0.23	0.64	0.82
33	vdem corrupt	VDEM corruption	0.68	0.24	0.68	0.83
34	wdi full	WDI full model	0.71	0.24	0.57	0.76
35	vdem cl	VDEM civil liberties	0.68	0.25	0.64	0.81
36	dsp disinform	DSP disinformation	0.66	0.26	0.64	0.82
37	dsp media	DSP traditional and social media	0.69	0.26	0.66	0.80
38	wdi resources	WDI resources	0.64	0.28	0.60	0.76
39	wdi education	WDI education	0.64	0.30	0.54	0.72

Table 22. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	evhist	EV History (last election)	0.86	0.12	0.85	0.94
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.85	0.12	0.85	0.94
3	cmb mega kitchen	Comb. kitchen sink	0.85	0.12	0.86	0.94
4	elcharirr	Last Election characteristcs (irregularities)	0.85	0.12	0.85	0.94
5	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.85	0.12	0.85	0.94
6	elchar	Last Election characteristcs	0.85	0.13	0.85	0.94
7	ensemble ablation	Ablation best 9	0.86	0.13	0.87	0.94
8	ensemble best	Best 9 ensemble	0.86	0.13	0.86	0.94
9	evhist long	EV History (long)	0.86	0.14	0.81	0.91
10	vdem kitchen	VDEM kitchen sink	0.84	0.15	0.83	0.92
11	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.83	0.15	0.81	0.91
12	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.83	0.16	0.81	0.91
13	vdem mid	VDEM mid incicies 1	0.82	0.16	0.82	0.92
14	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.82	0.16	0.82	0.92
15	dsp kitchen	DSP kitchen sink	0.82	0.16	0.80	0.91
16	vdem mid2	VDEM mid indices 2	0.84	0.17	0.80	0.91
17	dsp infra	DSP infrastructure	0.80	0.17	0.79	0.90
18	elcharstr	Last Election characteristcs (structural)	0.80	0.17	0.78	0.89
19	vdem high	VDEM high indices	0.82	0.17	0.77	0.89
20	vdem account	VDEM accountability	0.80	0.18	0.79	0.90

Table 23. Constituent model performance for one calendar year ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.80	0.18	0.78	0.89
22	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.79	0.18	0.77	0.88
23	cmb dsp smclimate use	Comb. sm climate + use	0.79	0.18	0.77	0.88
24	dsp smclimate	DSP Social media climate	0.79	0.18	0.77	0.88
25	vdem excl	VDEM exclusion	0.80	0.18	0.77	0.90
26	vdem neopat	VDEM neopat	0.77	0.19	0.75	0.88
27	wdi kitchen	WDI kitchen sink	0.79	0.19	0.75	0.88
28	vdem cl	VDEM civil liberties	0.78	0.19	0.75	0.88
29	dsp monitoring	DSP Social media monitoring	0.79	0.19	0.74	0.88
30	vdem gender	VDEM gender	0.77	0.20	0.74	0.87
31	vdem corrupt	VDEM corruption	0.74	0.21	0.71	0.86
32	elcharout	Last Election outcome	0.74	0.21	0.70	0.87
33	wdi structural	WDI structural	0.75	0.21	0.68	0.86
34	wdi resources	WDI resources	0.75	0.22	0.70	0.85
35	dsp media	DSP traditional and social media	0.76	0.22	0.73	0.85
36	vdem pp	VDEM parties and power	0.76	0.22	0.72	0.86
37	dsp disinform	DSP disinformation	0.72	0.23	0.69	0.83
38	wdi education	WDI education	0.66	0.29	0.58	0.76

Table 24. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	elcharirr	Last Election characteristcs (irregularities)	0.82	0.15	0.84	0.92
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.85	0.93
3	cmb mega kitchen	Comb. kitchen sink	0.82	0.15	0.84	0.92
4	evhist	EV History (last election)	0.83	0.15	0.84	0.92
5	elchar	Last Election characteristcs	0.82	0.15	0.83	0.92
6	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.81	0.15	0.83	0.92
7	ensemble ablation	Ablation best 9	0.81	0.16	0.85	0.93
8	ensemble best	Best 9 ensemble	0.81	0.16	0.84	0.93
9	evhist long	EV History (long)	0.82	0.16	0.79	0.90
10	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.78	0.18	0.79	0.90
11	vdem kitchen	VDEM kitchen sink	0.77	0.19	0.78	0.89
12	vdem mid	VDEM mid incicies 1	0.77	0.19	0.79	0.89
13	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.19	0.78	0.89
14	dsp kitchen	DSP kitchen sink	0.77	0.19	0.73	0.87
15	dsp infra	DSP infrastructure	0.76	0.19	0.79	0.89
16	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.78	0.19	0.76	0.89
17	vdem account	VDEM accountability	0.76	0.20	0.77	0.89
18	vdem high	VDEM high indices	0.77	0.20	0.75	0.87
19	vdem mid2	VDEM mid indices 2	0.77	0.20	0.75	0.88
20	elcharstr	Last Election characteristcs (structural)	0.75	0.20	0.75	0.87

Table 25. Constituent model performance for one calendar year ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.76	0.21	0.71	0.85
22	cmb dsp smclimate use	Comb. sm climate + use	0.76	0.21	0.71	0.85
23	dsp smclimate	DSP Social media climate	0.76	0.21	0.71	0.85
24	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.75	0.21	0.71	0.85
25	vdem excl	VDEM exclusion	0.75	0.21	0.73	0.86
26	vdem neopat	VDEM neopat	0.71	0.22	0.70	0.85
27	wdi kitchen	WDI kitchen sink	0.75	0.22	0.68	0.85
28	elcharout	Last Election outcome	0.71	0.22	0.70	0.86
29	dsp monitoring	DSP Social media monitoring	0.74	0.22	0.71	0.84
30	vdem gender	VDEM gender	0.74	0.23	0.72	0.84
31	vdem cl	VDEM civil liberties	0.70	0.23	0.68	0.83
32	vdem corrupt	VDEM corruption	0.72	0.24	0.71	0.83
33	vdem pp	VDEM parties and power	0.71	0.24	0.69	0.84
34	wdi structural	WDI structural	0.72	0.25	0.63	0.83
35	wdi resources	WDI resources	0.71	0.25	0.65	0.82
36	dsp media	DSP traditional and social media	0.73	0.25	0.68	0.83
37	dsp disinform	DSP disinformation	0.67	0.26	0.64	0.80
38	wdi education	WDI education	0.62	0.32	0.54	0.73

Table 26. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb mega kitchen	Comb. kitchen sink	0.83	0.15	0.84	0.93
2	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.82	0.15	0.83	0.92
3	ensemble ablation	Ablation best 9	0.83	0.15	0.83	0.92
4	ensemble best	Best 9 ensemble	0.84	0.15	0.83	0.92
5	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.82	0.15	0.82	0.92
6	elcharirr	Last Election characteristcs (irregularities)	0.81	0.15	0.81	0.92
7	elchar	Last Election characteristcs	0.82	0.16	0.82	0.91
8	evhist	EV History (last election)	0.82	0.16	0.80	0.91
9	vdem kitchen	VDEM kitchen sink	0.81	0.17	0.80	0.90
10	vdem mid	VDEM mid incicies 1	0.80	0.17	0.78	0.90
11	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.81	0.17	0.78	0.90
12	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.82	0.17	0.77	0.89
13	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.80	0.18	0.77	0.89
14	dsp kitchen	DSP kitchen sink	0.80	0.18	0.76	0.88
15	elcharstr	Last Election characteristcs (structural)	0.79	0.18	0.78	0.89
16	vdem mid2	VDEM mid indices 2	0.80	0.18	0.77	0.89
17	evhist long	EV History (long)	0.80	0.18	0.73	0.87
18	dsp infra	DSP infrastructure	0.79	0.18	0.77	0.88
19	vdem high	VDEM high indices	0.81	0.18	0.75	0.88
20	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.80	0.18	0.78	0.88

Table 27. Constituent model performance for two calendar years ahead forecasts in the rolling test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.79	0.19	0.75	0.87
22	cmb dsp smclimate use	Comb. sm climate + use	0.79	0.19	0.75	0.87
23	dsp smclimate	DSP Social media climate	0.79	0.19	0.75	0.87
24	wdi kitchen	WDI kitchen sink	0.80	0.19	0.75	0.88
25	vdem cl	VDEM civil liberties	0.79	0.19	0.75	0.88
26	vdem account	VDEM accountability	0.77	0.19	0.75	0.88
27	vdem excl	VDEM exclusion	0.78	0.20	0.74	0.87
28	vdem neopat	VDEM neopat	0.76	0.20	0.72	0.87
29	vdem gender	VDEM gender	0.77	0.20	0.73	0.87
30	dsp monitoring	DSP Social media monitoring	0.77	0.20	0.72	0.86
31	vdem corrupt	VDEM corruption	0.74	0.22	0.71	0.85
32	wdi resources	WDI resources	0.75	0.22	0.68	0.85
33	wdi structural	WDI structural	0.74	0.22	0.67	0.84
34	dsp media	DSP traditional and social media	0.74	0.23	0.71	0.84
35	vdem pp	VDEM parties and power	0.74	0.23	0.68	0.85
36	dsp disinform	DSP disinformation	0.72	0.23	0.68	0.83
37	elcharout	Last Election outcome	0.72	0.23	0.67	0.85
38	wdi education	WDI education	0.67	0.28	0.60	0.78

Table 28. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
1	cmb mega kitchen	Comb. kitchen sink	0.77	0.19	0.78	0.89
2	ensemble best	Best 9 ensemble	0.78	0.19	0.80	0.90
3	ensemble ablation	Ablation best 9	0.77	0.19	0.78	0.89
4	elchar	Last Election characteristcs	0.78	0.19	0.78	0.89
5	elcharirr	Last Election characteristcs (irregularities)	0.76	0.19	0.77	0.88
6	cmb vdem elcharirr excl wdistruct	Comb. excl, wdi struct, elchar irr	0.76	0.19	0.77	0.89
7	evhist	EV History (last election)	0.78	0.19	0.77	0.88
8	cmb vdem elcharirr cl wdistruct	Comb. civ lib, wdi struct, elchar irr	0.76	0.20	0.77	0.88
9	vdem kitchen	VDEM kitchen sink	0.77	0.20	0.75	0.88
10	vdem mid	VDEM mid incicies 1	0.74	0.20	0.75	0.88
11	elcharstr	Last Election characteristcs (structural)	0.75	0.20	0.75	0.87
12	cmb vdem high wdistruc	Comb. vdem high, wdi struct	0.77	0.21	0.72	0.87
13	dsp kitchen	DSP kitchen sink	0.75	0.21	0.73	0.85
14	cmb vdem mid wdistruc dspinfr	Comb. vdem mid, wdi struct, dsp infr	0.75	0.21	0.73	0.87
15	evhist long	EV History (long)	0.76	0.21	0.69	0.84
16	cmb vdem mid wdistruc	Comb. vdem mid, wdi struct	0.75	0.21	0.71	0.85
17	dsp infra	DSP infrastructure	0.72	0.21	0.74	0.86
18	vdem high	VDEM high indices	0.76	0.22	0.70	0.85
19	vdem mid2	VDEM mid indices 2	0.75	0.22	0.71	0.86
20	cmb dsp smclimate secuse	Comb. sm climate + security + use	0.76	0.22	0.71	0.85

Table 29. Constituent model performance for two calendar years ahead forecasts in the fixed test window for the xgboost model

Rank	Constituent model	theme	accuracy	brier score	aupr	auroc
21	cmb dsp smclimate disinformuse	Comb. sm climate + use + disinform	0.73	0.22	0.70	0.83
22	cmb dsp smclimate use	Comb. sm climate + use	0.73	0.22	0.70	0.83
23	dsp smclimate	DSP Social media climate	0.73	0.22	0.70	0.83
24	vdem account	VDEM accountability	0.74	0.23	0.72	0.85
25	vdem cl	VDEM civil liberties	0.74	0.23	0.68	0.84
26	vdem excl	VDEM exclusion	0.72	0.23	0.66	0.83
27	wdi kitchen	WDI kitchen sink	0.72	0.24	0.67	0.83
28	vdem neopat	VDEM neopat	0.69	0.24	0.68	0.84
29	vdem gender	VDEM gender	0.71	0.24	0.67	0.83
30	vdem corrupt	VDEM corruption	0.69	0.25	0.68	0.83
31	elcharout	Last Election outcome	0.69	0.25	0.66	0.82
32	dsp monitoring	DSP Social media monitoring	0.69	0.25	0.66	0.81
33	wdi resources	WDI resources	0.70	0.25	0.63	0.81
34	dsp media	DSP traditional and social media	0.70	0.26	0.67	0.81
35	wdi structural	WDI structural	0.67	0.26	0.63	0.81
36	vdem pp	VDEM parties and power	0.66	0.27	0.61	0.81
37	dsp disinform	DSP disinformation	0.63	0.28	0.61	0.78
38	wdi education	WDI education	0.64	0.31	0.58	0.75

Appendix D: Forecasts for 2022 and 2023 using the nine model ensemble

Table 30. Risk of electoral violence in 2022 using the nine model ensemble (average of random forest and xgboost)

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
1	Equatorial Guinea	0.01	0.10	0.89	0.99	0.94
2	Turkmenistan	0.01	0.13	0.86	0.99	0.92
3	Congo - Brazzaville	0.01	0.29	0.70	0.99	0.84
4	Belarus	0.07	0.45	0.48	0.93	0.70
5	Bahrain	0.06	0.50	0.43	0.94	0.69
6	Papua New Guinea	0.05	0.58	0.37	0.95	0.66
7	Sudan	0.04	0.73	0.23	0.96	0.60
8	Angola	0.04	0.77	0.20	0.96	0.58
9	Algeria	0.03	0.81	0.16	0.97	0.56
10	Kenya	0.08	0.75	0.17	0.92	0.55
11	Kazakhstan	0.02	0.89	0.08	0.98	0.53
12	Lebanon	0.04	0.86	0.10	0.96	0.53
13	Nepal	0.06	0.85	0.09	0.94	0.51
14	Philippines	0.11	0.77	0.12	0.89	0.50
15	Fiji	0.05	0.89	0.06	0.95	0.50
16	Serbia	0.10	0.84	0.06	0.90	0.48
17	Malaysia	0.12	0.82	0.05	0.88	0.46
18	Hungary	0.17	0.76	0.08	0.83	0.45
19	Lesotho	0.14	0.81	0.05	0.86	0.45
20	Kuwait	0.21	0.69	0.10	0.79	0.45
21	Gambia	0.17	0.78	0.06	0.83	0.44
22	Bosnia & Herzegovina	0.15	0.82	0.03	0.85	0.44
23	Mexico	0.21	0.74	0.05	0.79	0.42
24	Senegal	0.25	0.72	0.03	0.75	0.39
25	Colombia	0.31	0.60	0.09	0.69	0.39

Table 31. Risk of electoral violence in 2022 using the nine model ensemble (average of random forest and xgboost)

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
26	Timor-Leste	0.24	0.75	0.02	0.76	0.39
27	Tunisia	0.35	0.59	0.06	0.65	0.36
28	Brazil	0.36	0.58	0.06	0.64	0.35
29	United States	0.49	0.50	0.01	0.51	0.26
30	São Tomé & Príncipe	0.64	0.35	0.01	0.36	0.19
31	Slovenia	0.67	0.31	0.02	0.33	0.17
32	France	0.73	0.25	0.02	0.27	0.14
33	Bulgaria	0.74	0.25	0.02	0.26	0.14
34	Israel	0.74	0.24	0.01	0.26	0.13
35	Vanuatu	0.78	0.21	0.01	0.22	0.12
36	Malta	0.79	0.20	0.01	0.21	0.11
37	Barbados	0.86	0.13	0.00	0.14	0.07
38	Cyprus	0.87	0.13	0.01	0.13	0.07
39	Czechia	0.90	0.10	0.01	0.10	0.05
40	South Korea	0.90	0.09	0.00	0.10	0.05
41	Uruguay	0.90	0.09	0.00	0.10	0.05
42	Switzerland	0.92	0.08	0.00	0.08	0.04
43	Latvia	0.92	0.08	0.00	0.08	0.04
44	Chile	0.93	0.07	0.00	0.07	0.04
45	Italy	0.92	0.08	0.00	0.08	0.04
46	Australia	0.93	0.07	0.00	0.07	0.04
47	Portugal	0.93	0.07	0.00	0.07	0.04
48	Denmark	0.93	0.07	0.00	0.07	0.04
49	Austria	0.93	0.07	0.00	0.07	0.03
50	Sweden	0.95	0.05	0.00	0.05	0.03
51	Japan	0.95	0.04	0.00	0.05	0.02
52	Costa Rica	0.97	0.03	0.00	0.03	0.02

Table 32. Risk of electoral violence in 2023 using the nine model ensemble (average of random forest and xgboost)

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
1	Turkmenistan	0.02	0.15	0.83	0.98	0.91
2	Cuba	0.07	0.33	0.60	0.93	0.76
3	Cambodia	0.02	0.45	0.53	0.98	0.75
4	Bangladesh	0.02	0.48	0.50	0.98	0.74
5	Congo - Kinshasa	0.02	0.52	0.47	0.98	0.73
6	Belarus	0.08	0.49	0.43	0.92	0.68
7	Myanmar (Burma)	0.04	0.61	0.35	0.96	0.66
8	Zimbabwe	0.01	0.71	0.28	0.99	0.63
9	Gabon	0.04	0.69	0.27	0.96	0.61
10	Eswatini	0.03	0.75	0.23	0.97	0.60
11	Thailand	0.06	0.70	0.24	0.94	0.59
12	Sudan	0.04	0.74	0.22	0.96	0.59
13	Togo	0.03	0.77	0.20	0.97	0.59
14	Benin	0.14	0.56	0.30	0.86	0.58
15	Pakistan	0.02	0.80	0.18	0.98	0.58
16	Rwanda	0.04	0.78	0.18	0.96	0.57
17	Turkey	0.06	0.74	0.20	0.94	0.57
18	Nigeria	0.04	0.79	0.16	0.96	0.56
19	Djibouti	0.01	0.88	0.11	0.99	0.55
20	Mauritania	0.02	0.86	0.11	0.98	0.55

Table 33. Risk of electoral violence in 2023 using the nine model ensemble (average of random forest and xgboost)

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
21	Kazakhstan	0.03	0.89	0.08	0.97	0.53
22	Guatemala	0.05	0.89	0.06	0.95	0.50
23	Liberia	0.08	0.86	0.06	0.92	0.49
24	Sierra Leone	0.08	0.87	0.05	0.92	0.49
25	Oman	0.12	0.80	0.08	0.88	0.48
26	Ukraine	0.12	0.82	0.06	0.88	0.47
27	Montenegro	0.18	0.75	0.07	0.82	0.45
28	Timor-Leste	0.29	0.68	0.03	0.71	0.37
29	Paraguay	0.30	0.66	0.04	0.70	0.37
30	Poland	0.38	0.56	0.06	0.62	0.34
31	Bhutan	0.65	0.30	0.04	0.35	0.19
32	Spain	0.85	0.15	0.00	0.15	0.07
33	Cyprus	0.88	0.12	0.01	0.12	0.07
34	Estonia	0.88	0.12	0.00	0.12	0.06
35	Greece	0.89	0.11	0.00	0.11	0.06
36	Czechia	0.90	0.10	0.01	0.10	0.05
37	Argentina	0.90	0.09	0.00	0.10	0.05
38	Switzerland	0.91	0.09	0.00	0.09	0.05
39	Finland	0.94	0.06	0.00	0.06	0.03
40	Luxembourg	0.95	0.05	0.00	0.05	0.03
41	New Zealand	0.95	0.05	0.00	0.05	0.03

Appendix E: Forecasts for 2022 and 2023 using the Digital Societies project full model

Table 34. Risk of electoral violence in 2022 by the Digital Societies Project full model

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
1	Equatorial Guinea	0.02	0.16	0.83	0.98	0.90
2	Turkmenistan	0.03	0.14	0.83	0.97	0.90
3	Belarus	0.03	0.46	0.51	0.97	0.74
4	Congo - Brazzaville	0.05	0.44	0.50	0.95	0.73
5	Bahrain	0.03	0.53	0.44	0.97	0.70
6	Sudan	0.03	0.60	0.38	0.97	0.67
7	Lebanon	0.06	0.66	0.27	0.94	0.61
8	Algeria	0.05	0.72	0.23	0.95	0.59
9	Philippines	0.06	0.71	0.23	0.94	0.59
10	Nepal	0.04	0.75	0.21	0.96	0.58
11	Colombia	0.05	0.74	0.21	0.95	0.58
12	Kazakhstan	0.03	0.78	0.19	0.97	0.58
13	Lesotho	0.07	0.81	0.12	0.93	0.52
14	Fiji	0.06	0.85	0.09	0.94	0.51
15	Brazil	0.16	0.67	0.17	0.84	0.51
16	Timor-Leste	0.03	0.92	0.04	0.97	0.50
17	Hungary	0.18	0.64	0.18	0.82	0.50
18	Serbia	0.15	0.71	0.14	0.85	0.50
19	Angola	0.09	0.86	0.06	0.91	0.48
20	Bosnia & Herzegovina	0.11	0.82	0.06	0.89	0.48
21	Kuwait	0.12	0.81	0.07	0.88	0.47
22	Malaysia	0.19	0.72	0.09	0.81	0.45
23	Papua New Guinea	0.27	0.60	0.13	0.73	0.43
24	Gambia	0.22	0.70	0.07	0.78	0.43
25	Senegal	0.29	0.61	0.10	0.71	0.41

Table 35. Risk of electoral violence in 2022 by the Digital Societies Project full model

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
26	Kenya	0.27	0.65	0.07	0.73	0.40
27	Tunisia	0.27	0.69	0.05	0.73	0.39
28	Mexico	0.30	0.64	0.05	0.70	0.38
29	Slovenia	0.37	0.55	0.08	0.63	0.35
30	United States	0.35	0.61	0.04	0.65	0.34
31	Israel	0.51	0.41	0.08	0.49	0.29
32	Bulgaria	0.49	0.49	0.03	0.51	0.27
33	France	0.47	0.52	0.01	0.53	0.27
34	Malta	0.53	0.45	0.02	0.47	0.24
35	Chile	0.62	0.34	0.04	0.38	0.21
36	Vanuatu	0.62	0.36	0.02	0.38	0.20
37	São Tomé & Príncipe	0.68	0.32	0.01	0.32	0.17
38	Cyprus	0.75	0.23	0.02	0.25	0.13
39	Australia	0.79	0.19	0.02	0.21	0.12
40	Czechia	0.80	0.18	0.02	0.20	0.11
41	South Korea	0.85	0.15	0.00	0.15	0.08
42	Uruguay	0.88	0.12	0.01	0.12	0.07
43	Italy	0.89	0.11	0.01	0.11	0.06
44	Barbados	0.89	0.10	0.01	0.11	0.06
45	Denmark	0.91	0.09	0.00	0.09	0.05
46	Austria	0.91	0.09	0.00	0.09	0.04
47	Costa Rica	0.96	0.04	0.00	0.04	0.02
48	Latvia	0.97	0.03	0.00	0.03	0.01
49	Sweden	0.98	0.02	0.00	0.02	0.01
50	Japan	0.99	0.01	0.00	0.01	0.01
51	Switzerland	0.99	0.01	0.00	0.01	0.01
52	Portugal	1.00	0.00	0.00	0.00	0.00

Table 36. Risk of electoral violence in 2023 by the Digital Societies Project full model

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
1	Cuba	0.01	0.16	0.83	0.99	0.91
2	Turkmenistan	0.03	0.12	0.84	0.97	0.91
3	Myanmar	0.06	0.34	0.60	0.94	0.77
4	Bangladesh	0.07	0.35	0.58	0.93	0.75
5	Cambodia	0.03	0.43	0.53	0.97	0.75
6	Belarus	0.02	0.49	0.49	0.98	0.74
7	Zimbabwe	0.02	0.50	0.48	0.98	0.73
8	Gabon	0.06	0.44	0.50	0.94	0.72
9	Eswatini	0.02	0.60	0.37	0.98	0.68
10	Sudan	0.05	0.59	0.36	0.95	0.66
11	Turkey	0.04	0.69	0.27	0.96	0.62
12	Kazakhstan	0.03	0.74	0.23	0.97	0.60
13	Togo	0.08	0.69	0.22	0.92	0.57
14	Thailand	0.05	0.76	0.19	0.95	0.57
15	Pakistan	0.03	0.83	0.14	0.97	0.56
16	Djibouti	0.03	0.83	0.14	0.97	0.56
17	Rwanda	0.03	0.85	0.11	0.97	0.54
18	Congo - Kinshasa	0.05	0.83	0.12	0.95	0.54
19	Ukraine	0.11	0.71	0.17	0.89	0.53
20	Nigeria	0.12	0.74	0.14	0.88	0.51

Table 37. Risk of electoral violence in 2023 by the Digital Societies Project full model

	country	p. no ev	p. limited ev	p. severe ev	p. violence	risk index
21	Sierra Leone	0.07	0.84	0.09	0.93	0.51
22	Oman	0.09	0.80	0.10	0.91	0.51
23	Guatemala	0.06	0.86	0.07	0.94	0.50
24	Mauritania	0.11	0.78	0.11	0.89	0.50
25	Bhutan	0.21	0.59	0.20	0.79	0.49
26	Timor-Leste	0.07	0.87	0.06	0.93	0.49
27	Benin	0.28	0.45	0.27	0.72	0.49
28	Montenegro	0.16	0.73	0.12	0.84	0.48
29	Liberia	0.32	0.57	0.11	0.68	0.40
30	Poland	0.36	0.60	0.04	0.64	0.34
31	Paraguay	0.54	0.40	0.06	0.46	0.26
32	Spain	0.76	0.23	0.01	0.24	0.12
33	Cyprus	0.78	0.20	0.02	0.22	0.12
34	Czechia	0.80	0.18	0.01	0.20	0.10
35	Argentina	0.85	0.14	0.01	0.15	0.08
36	Luxembourg	0.86	0.13	0.01	0.14	0.07
37	Greece	0.86	0.14	0.00	0.14	0.07
38	New Zealand	0.90	0.09	0.01	0.10	0.05
39	Estonia	0.92	0.08	0.00	0.08	0.04
40	Switzerland	0.99	0.01	0.00	0.01	0.01
41	Finland	0.99	0.01	0.00	0.01	0.00