**The Clock is Ticking for TikTok: how to explain social media ban in the Land of the Free**

Leonardo Stempfle de Azevedo

September 25, 2024

1. **Introduction**

In the 21st century, social media has reshaped various aspects of human life and social interaction. Intrinsically, the way we communicate online is different from how we speak, for it is shaped by the constraints of the digital medium, a dynamic in constant evolution and modification (Aleksic 2024). According to Freedom House’s estimates, over 5 billion people have access to the internet (Funk et al. 2024). Due to the development of technology, peoples around the world are capable of sharing information more quickly and widely than ever before, increasing transnational interconnectivity and globalization, not only of goods and services, but also of information, thoughts, and ideologies. These characteristics of social media allowed for the democratization of information and boosted freedom of speech, since anyone can speak (or message) whatever, wherever, and whenever they wish. Political mobilization and collective action efforts have been enhanced by the ability to message numerous individuals at once and take almost-immediate action when organizing protests. In autocratic[[1]](#footnote-2) countries, this ability prompts severe censorship and repression, to supress any sort of challenge to the government. In democratic countries, should free speech be unrestricted, unverified and distorted information – combined with ill-intentioned actors and broader and more efficient methods of communication – can devolve into the spread of hate speech and fake news, mobilizing masses to support anti-democratic ideologies, which also incentivizes censorship (Vidyarthi and Hulvey 2021). One example of this phenomenon is the American and Brazilian protests on January 6th, 2021, and January 8th, 2023, respectively, in which right-wing conservative protesters gathered around and stormed government facilities to challenge the democratic institutions of their respective countries (Cameron 2023). A movement generated by the exploitation of social media to promote anti-establishment and anti-democratic narratives of the world, advocating for and supported by the right for complete freedom of speech. In both cases, perpetrators were investigated and sentenced, leading to debates on whether there should be some censorship to free speech.

It is expected that undemocratic countries might ban[[2]](#footnote-3) social media. They limit or block citizens’ access to all kinds of mediums that spread anti-government ideas and any form of criticism or that could be used to do so, in an effort to keep incumbents unchecked and their populations alienated to the negative aspects of the government, therefore unable and unwilling to organize opposition, a dynamic which constitutes the core motivation for social media censorship (York 2022; Gehlbach et al. 2022). Afraid of losing power, authority, and potentially being overthrown, governments control how their population view them, by limiting unfavourable narratives and disseminating favourable ones (Gehlbach et al. 2022).[[3]](#footnote-4) Though not moved solely by state survival, closed, traditional, or fundamentalist nations might also block social media, striving to restrict the influence of foreign countries – of their traditions, cultures, and behaviours – on their own population, to protect their societies from these ideas that could corrupt their own customs and traditions.

What is not expected is that democratic countries ban social media, especially since social media usually originate from these countries (Facebook, Instagram, X[[4]](#footnote-5), and other mainstream platforms are all American) and because democratic nations should protect and foster democratic ideals – in particular, freedom of speech. For example, in April 2024, United States’ president Joe Biden signed the Protecting Americans from Foreign Adversary Controlled Applications Act, a bill that “prohibits distributing, maintaining, or providing internet hosting services for a foreign adversary controlled application (e.g., TikTok).” (Rep. Gallagher 2024, para. 1), namely mentioning TikTok, but applicable to all other foreign apps. The approval of the bill means that TikTok will (at least legally) either be bought by an American company or be banned from American territory altogether (House Committee on Energy and Commerce 2024). This is not the only case of democratic social media ban: in Brazil, in August 2024, Supreme Court Judge Alexandre de Moraes issued an order to block Twitter (Kahn 2024); WeChat and TikTok, two Chinese-owned platforms, have been banned in Canadian government officials’ devices (Collins 2023; Milmo 2024) and completely in India (A. Sharma 2024).

This thesis aims to answer the question of whether the reasons for blocking social media are the same for democratic and undemocratic countries or if they derive from different sources. To achieve this objective, I will provide a literature review on the minimalist and maximalist definitions of democracy and on the reasons for social media ban (and for censorship in general, in a broader sense) in both types of countries. Followed by an overview of TikTok bans around the world, to serve as examples of the dynamics driving the motives behind the social media bans on democratic and undemocratic countries, demonstrating the overlap and parallels between the two. Though not the only platform to be widely banned, TikTok is the most recent, fastest growing, and mainstream case of social media ban in recent years and serves as a clear example for all the drivers discussed in this paper. Subsequently, utilizing the Varieties of Democracy dataset (V-Dem) (Coppedge et al. 2024b; Pemstein et al. 2022), I will test three kinds of machine learning models to predict and also explain what variables influence social media ban: logistic regression, decision tree, and random forest. Three datasets will be used to train and test the models, one with all the data and two which are subdivisions: one with data only for democratic countries and one with data only for autocratic countries. To illustrate this process – and this illustration applies to all models – consider a logistic regression trained on democratic data. When tested on democratic data, it should perform well. Once having performed well, it will be evaluated against autocratic data to assess if the rules it apprehended from democracies can be accurately applied to autocracies as well. Likewise, one trained on autocratic data should perform well on this type of data and will be evaluated against democratic data to assess if autocratic bans follow the same rules as democratic bans.[[5]](#footnote-6) If either example is proven true, this analysis demonstrates that the underlying drivers of bans in democratic and undemocratic states are the same (or indistinguishable), shedding light on the issue that freedom of speech is not perfectly upheld only by maintaining democratic institutions, since democratic and autocratic institutions appear to act similarly on the matter of social media ban.

1. **Literature Review**

The literature review is subdivided into three subsections. The first two will provide conceptual definitions of how we understand democracy and discuss the motives and explanations for social media ban in both autocracies and democracies, how they are more similar than one might imagine, and what differentiates each type of ban. The third part will conduct an overview of cases where TikTok was blocked to exemplify the previously mentioned explanations for social media ban.

* 1. **Democracies and definitions**

Democracy – and, by extension, autocracy – is hard to define both conceptually and empirically. States on either side of the democratic-autocratic spectrum are easy to classify. Denmark, Sweden, Norway, Switzerland, New Zealand, and Belgium, to illustrate, can undoubtedly be classified as democracies, while Iran, North Korea, China, Saudi Arabia, Syria, and Afghanistan are clearly autocracies. But when it comes to countries such as Brazil, Mexico, and Poland, where do they fall on this scale? In the very middle, more to the left, or more to the right? Without a clear definition of what constitutes an autocracy and what is needed to become a democracy, middle-of-the-road countries cannot be properly classified. Why is it that Denmark is definitely a democracy and Saudi Arabia definitely an autocracy? One good explanation is the existence of free and fair elections, since it is understood that democracy is “the rule by the people” and, therefore, the people should have a say in how to compose the government and how it should act (Dahl and Shapiro 2024, para. 1). Thus, it is implied that autocracies are regimes which do not hold popular elections or do not do it properly, impeding the people from having relevant participation or any say at all in the government’s decisions.

Elections are the central pillar of electoral democracies, the understanding of democracy used in this paper and one V-Dem uses as a basis for all other high-level democracy indices, “an essential element of any other conception of representative democracy — liberal, participatory, deliberative, egalitarian, or some other.” (Coppedge et al. 2024a, 44). Though not the only one, electoral competition is the minimalist definition of democracy (Boese and Wilson 2022) and it is the most straightforward understanding to measure (and, for this reason, generally the most used), since it depends on elections, something material and observable; a concept adopted by the United Nations (UN) “the will of the people is the source of legitimacy of sovereign states” and present in the Universal Declaration of Human Rights “the will of the people shall be the basis of the authority of government” (United Nations, n.d., para. 3; 1948, art. 21). To be classified as a democracy, a regime must hold free and fair elections for the executive (they must not be tampered or frauded in any way to influence its results), allowing for multiparty participation and the opportunity for any party to compete in them; universal suffrage for all citizens of the state; and freedoms of expression and for organizations – civil and political – to operate freely (Coppedge et al. 2024a; Dahl 2005).

Apart from the electoral dimension, democracy can be understood as a regime in which the rights of individuals and minorities are protected against the tyranny of the state and of the majority (Coppedge et al. 2024a), a maximalist, more substantive definition that considers that there should be constraints to the action of the state, by enforcing the rule of law and establishing checks and balances, so as to create an environment where individuals are truly free and elections can be held and be free and fair in the first place (Boese and Wilson 2022). Broader definitions of democracy consider normative aspects of how a true democracy should be. It should, as mentioned, uphold the rights and freedoms of individuals and minorities, enable the participation of all eligible citizens in all political processes, ensure that the decisions reached by the majority derive from ample social debates and address pressing social issues, and guarantee that all of the above is done equally among individuals, ensuring that everyone enjoys equal access to the resources allocated by the state, to positions of power, and to their rights and freedoms (the state cannot protect the right of some more than it protects the rights of others, even if it is still protecting them) (Coppedge et al. 2024a; Sigman and Lindberg 2015).

After having attempted to define more clearly the definition of democracy, we can delve into why would such a system – structured entirely on the protection of individual freedoms, of access to information, and the liberty to share them – limit access to social media or even ban them in totality, contradicting its duty to protect free speech. One possibility is that it is precisely because these regimes must preserve the integrity of elections that they block social media, to restrain the spread of harmful information, which can twist individuals’ perceptions about these elections, leading them to distrust the government and its institutions, of which elections are one of them – once again, exemplified by the American and Brazilian January protests. This information could be intentionally spread by nationals attempting to promote an anti-establishment campaign, or by foreigners and foreign governments attempting to interfere with the election’s results, be it through third party social media, like Russia tampering with the 2016 US elections (an possibly the 2020 and 2024 elections), or China doing the same through its own social media apps, such as TikTok (US Department of Justice 2019, 1; Office of the Director of National Intelligence 2024, 12).

“National security concerns”, the most common reason democracies give to ban foreign social media concern (apart from election manipulation) data privacy of individuals and government officials, hacking, spying, and other kinds of cyberthreats. Usually, these cases revolve around Western countries banning Eastern technology, mostly Chinese, like TikTok, WeChat or Huawei. Since democracies cannot impede foreign investment in their own countries without proper justification, they utilize the subterfuge of national security to mobilize popular approval of the use of extraordinary powers (which should not usually be used in a democratic scenario) to answer the problem of foreign influence in the country (Roy 2017). There is a case to be made that Western social media do not raise national security concerns in Western countries due to the fact they are Western and would not serve the interests of neither China nor Russia. The data these companies collect remain stored in servers inside US borders and can be accessed by the state should the need for that ever arise. European and other US-allied countries do not worry about their data being collected by American companies as much as they worry about the same being done by Chinese companies, whose administrations are known to be subject to legal intervention by the Chinese Communist Party (Fung 2023). Meta, Google and TikTok collect the same data, using the same methods, but, because for one of these companies, data is being stored internationally (in a non-allied country), it becomes a national security risk.

The aforementioned reasons for democratic social media ban, while justifiable, raise concerns about the role of free speech in a world where its absolute protection can undermine other democratic institutions and, ultimately, its own integrity (because the remedy is censorship), and where foreign intervention bring about an environment of distrust of foreign social media, which, once banned, segregate the internet that once was a space of global interconnectivity (Funk et al. 2024; Supak 2022).

* 1. **Autocracies**

While democracies must be wary about not infringing individual freedoms, and ensuring civil rights, autocracies, on the other hand, do not have these limitations. They can be repressive with the opposition, arbitrary in their decisions, and completely illiberal, insofar as the regime and the incumbent’s safety is maintained, following the logic of state survival. Should a leader be too draconian and give rise to revolts and insurgencies, the future of the present regime is in jeopardy. Dictators need to maintain a tight balance between appeasing their winning coalition with private goods, so it does not depose them for another that does meet its demands (Bueno De Mesquita et al. 2003) – which consumes resources that could, otherwise, be directed at appeasing the population – and keeping the resource-deprived population at bay, in a constant state of not being too distressed to justify taking action against the government. Unlike democracies, autocracies do not have to provide goods equally between all members of the population, these regimes have the freedom to distribute them however they feel serves best their interests. And, once again, unlike democracies, there is no need to protect civilians, so repression can be applied as needed.

Autocracies aim to induce citizens to act in the government’s interest and, because they do not have the constraints democracies have, these regimes use censorship and state propaganda to control information access, resorting to repression if needed, since it is costly and raises dissatisfaction levels (Gehlbach et al. 2022). Sometimes banning a social media is necessary to censor opposition, sometimes it could be going a step too far and consequently inciting rebellion.

* 1. **Overview: The Case of TikTok**

Launched internationally in 2017, TikTok is a short-form video-sharing social media app created by the Chinese company ByteDance, owner of other popular apps such as Douyin (the domestic Chinese-controlled version of TikTok) and the video-editor CapCut. It was valued in USD $200 billion in 2023 and had over a billion users in July 2024, of which 157 million were located in Indonesia, 120 million in the United States, and 105 million in Brazil (Statista 2024).

Having said that, in the last few years, proposals to ban TikTok, on different levels, have appeared all around the world: in Indonesia (largest audience), the United States (2­nd largest), India, Europe, Australia, Pakistan (6th), Somalia, Kenya, Afghanistan, and more. In the beginning of 2024, the US House of Representatives passed a bill calling for ByteDance to divest from TikTok, demanding that a western – and preferably American – company owned it instead, alleging national security threats (Hale 2024), such as data collection malpractice, selling American citizens’ data to the Chinese government, manipulating public opinion, and spying on American government officials (Ceci 2024). About one month later, on April 24th, the bill was approved and signed by President Biden, meaning that TikTok had nine months to be sold, a decision for which ByteDance sought legal actions (Gerken and Singleton 2024). Previously, Donald Trump also tried to pass an executive order to ban TikTok and, though not a federal restriction, the state of Montana attempted to ban TikTok inside its jurisdiction in 2023. Both attempts claimed that the app threatened America’s national security, and both were unsuccessful in effectively banning the video-sharing platform (Kim 2024).

While the clock started ticking for TikTok in the US, it stopped in many other countries. India banned the app for a few days in 2019 and permanently in 2020 (alongside 50 more Chinese apps), alleging national security, sovereignty, and data privacy concerns (Clausius 2022; Ceci 2024), after a clash between Indian and Chinese troops at the disputed border of the Ladakh region (Abi-Habib 2020). NATO, the United Kingdom and the European Parliament, Canada, the US, Australia, New Zealand, and Taiwan have all banned TikTok on governmental devices, once more claiming security concerns regarding data leaks, spying, and monitoring from ByteDance and the Chinese government (Ceci 2024). Along with TikTok, France banned all recreational applications – which includes western apps like X, and Instagram – from public sector officials’ devices, all because of concerns about data security measures (AP News 2023). [[6]](#footnote-7) The US army has banned the app as early as 2019, after a national security review, due to privacy concerns (Anderson 2020). These cases show that using national security as a reason for banning social media, like TikTok, can prove successful. Huawei, another Chinese-owned company, suffers from the same fate: countries in Europe, Australia, New Zealand, Japan, Taiwan, and the US have all banned Huawei’s 5G technology from operating in their countries, at least partially, alleging, again, security concerns (Reuters 2023a; Buchholz 2020). France, once more, has raised national security concerns, this time to justify the blocking of TikTok in its overseas territory of New Caledonia, in light of the state of emergency triggered by indigenous Kanaks’ protests over proposed electoral reforms (Funk et al. 2024). The French government claimed that TikTok was being weaponized by China and Azerbaijan to spread disinformation, and the ban was made possible as a consequence of the exceptional state of emergency (Kirby 2024). The justification for the ban was national security concerns over interference from foreign governments, but the reasons behind the ban seem to be closer to authoritarian censorship aimed at quashing protests than at ending foreign intervention.

Having said that, national security is not the only reason that drives countries to ban TikTok, autocratic censorship (e.g. opposition silencing) is also a big contributor. In 2023 Senegal blocked access to TikTok (along with YouTube, Instagram, Facebook, and other social media) alleging that the app was used to promote hateful and subversive messages, threatening the stability of the country, and, to lift the ban, demanded the creation of a mechanism that allows the government to unilaterally remove accounts (*Reuters* 2023b). One month prior, the Senegalese government imprisoned opposition leader Ousmane Sonko, for supposedly “corrupting the youth and disturbing public order”, and dismantled his party, PASTEF, which triggered massive national protests (Dosunmu 2023). It is curious that Senegal is considered a democracy by V-Dem but acts, regarding censorship, just like an autocracy would. Jordan announced a temporary ban on TikTok after the death of a police officer during clashes with demonstrators protesting because of high fuel prices (Associated Press 2022). The temporary ban, months later, had not ended and kept the app out of the country, in an attempt to maintain public safety and order. In Armenia and Azerbaijan – most likely as a result of the ongoing tensions between them – TikTok malfunctioned on the latter and, on the former, the Azerbaijani government blocked not only TikTok, but also Facebook, X, LinkedIn, and YouTube, supposedly to prevent people from spreading unchecked, false, and unofficial information about the conflict and Azeri military (APA 2023; Acessnow 2022). Armenia crossed the threshold to become an electoral democracy in 2018, yet still banned TikTok a couple of years later. Russia criminalized what it deemed to be fake news about the Ukrainian invasion and, in response, TikTok only allowed old, state-approved content to be displayed on its app (Milmo 2022). China, with its “Great Firewall” is exceptionally effective at banning unwanted information inside its own territory (Supak 2022; Clausius 2022). All foreign companies have to adhere to Chinese Law, considerably limiting what Chinese users can consume when compared with international users. And even TikTok in China is not the same TikTok the rest of the world has access to, for inside Chinese borders ByteDance offers Douyin, the original version of TikTok, made to have only what is allowed by the government (Cueto 2023), which means technically TikTok is banned in the country.

Apart from national security and censorship, another reason why countries – both democratic and undemocratic – might seek to ban foreign-owned companies is to protect their populations from harmful content, material that could be contrary to tradition and cultural values or infringe upon human rights. In 2023 Somalia banned TikTok alleging indecent content, as tried Kenya for “‘eroding’ cultural and religious values” (Clynch 2024). The Taliban government in Afghanistan, in 2022, banned the app for not being consistent with Islamic Law, an act aimed to protect the younger generation and prevent them from being misled (*BBC News* 2022). Note that there can be censorship because of both political and cultural nature, such as in Iran, where social media is banned to control information, preventing foreign influence from harming tradition while also limiting the mobilization of protesters. Kyrgyzstan argued that its ban of TikTok was based on the protection of children from inappropriate content and behaviour, and lack of controls on time limits (RFE/RL’s Kyrgyz Service 2023), the same reasons Nepal utilized to support its ban (G. Sharma 2023) along with disturbing “social harmony and goodwill” (Reuters 2024); and so did France when threatening to block the app (Gerken and Singleton 2024).

These previously listed cases of attempts to ban TikTok (and foreign-owned companies in general), can be seen as a form of “net nationalism” in which countries are trying to get a tighter grip on their domestic digital space (largely controlled by big-tech companies) through the instrumentalization of the internet aimed at surveillance and thought control (Supak 2022). This means that countries, independent of their democratic leaning, have realized time and time again that they had weak control over their own digital space and could benefit considerably once they started to control it more effectively. This can hinder foreign intervention directed at damaging internal democratic institutions, or corrupting societies and their traditions. Stronger control of the internet also allows governments to directly influence access to information of any nature, enabling them to alter popular perceptions about the regime, contain the ability to rebel, track insurgents, and apply repression more efficiently.

The mechanisms and motives for democratic and autocratic states, though mostly distinct, more often than not converge and blur the lines between a democratic and authoritarian model of the internet. Both kinds of regimes do not have strongly conflicting ideas of the internet. “Democratic nations act as censors, too; they simply have different ideas of what should be censored” (York 2022, 4). In the end, what is driving social media ban if the distinctions between democratic and undemocratic motives are not clearly distinguishable and when the actual reasons behind public justifications for that are not apparent?

1. **Research Design**

To answer the question: “Are the drivers for social media ban in democratic countries and undemocratic countries the same?” I will apply three types of machine learning models, each more capable than the previous to predict social media ban, but also less interpretable[[7]](#footnote-8) in its ability to clearly explain what variables influence these bans. The models are: (i) logistic regression, the simplest of the three, since it assumes all relations among variables follow a linear pattern, but is the model with the clearest interpretation; (ii) decision tree, a model capable of dealing with non-linear relations that also provides clear rules for how it classifies each observation; and (iii) random forest which, unlike the previous models, is the hardest to interpret, since it utilizes countless randomized tree models at the same time (therefore, “forest”) to base its decisions, making it more capable to deal with complex variable interactions and make better predictions. Here, I utilize three different models for robustness, since the model with the clearest interpretation (linear regression) is expected to have the worst performance, while the model with the best performance (random forest) has the worst interpretability, and the decision tree serves as a middle ground for performance and interpretability. Should the three models provide the same results, in regard to which variables are the most important at explaining social media ban, it is safe to assume the conclusions of this paper are true.

After manipulating data from Surfshark’s Internet Shutdown Tracker (IST) project (Surfshark 2024) and combining it with selected variables from V-Dem’s dataset (Coppedge et al. 2024b; Pemstein et al. 2022), I will utilize three subsets of this whole dataset to train and test the models, one with all the data and two which are subdivisions: one with data only for democratic countries and one with data only for autocratic countries. In the end there will be three logistic regressions, three decision trees, and three random forests, one of each model for each dataset. A total of nine models. Seeing that this is a machine learning task, both the democratic and autocratic datasets will be divided each into a training set – data the models will use to extract patterns and rules – and a test set – data the models have not seen previously, on which those rules will be applied to make predictions and measure the accuracy of the models on unseen data (their capacity to generalize these rules). Every model will be trained and tested on its respective dataset and then tested again, this time on either the democratic or the autocratic datasets, to evaluate whether it is capable of explaining democratic and autocratic social media ban.[[8]](#footnote-9) The general model (trained on both kinds of data) serves to ascertain if a division between democratic and autocratic data is needed at all, seeing that a model trained on a more specific section of the data could produce a better model, tailored for this specific case. Nonetheless, creating two specific models entails dividing the data, which implies each model will have less observations to train with, and this could produce two weaker models than a single model with more training data. Likewise, a unified and bigger test set allows for a better measurement of the accuracy of the model.

**Table 1:** Schema of Train-Test-Evaluation model structure

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Train** | **Test** | **Evaluation**[[9]](#footnote-10) |
|  | Democracies | Democracies | Autocracies |
| **Logistic Regression** | Autocracies | Autocracies | Democracies |
|  | General (both) | General (both) | Aut. & Dem. (separately) |
|  | Democracies | Democracies | Autocracies |
| **Decision Tree** | Autocracies | Autocracies | Democracies |
|  | General (both) | General (both) | Aut. & Dem. (separately) |
|  | Democracies | Democracies | Autocracies |
| **Random Forest** | Autocracies | Autocracies | Democracies |
|  | General (both) | General (both) | Aut. & Dem. (separately) |

Given the research question “Are the drivers for social media ban in democratic countries and undemocratic countries the same?”, I argue that they are, indeed, similar, and can be explained by the same variables, as depicted in the TikTok ban overview. My hypotheses are:

1. *The variables that explain democratic social media ban are of the same nature as the variables that explain autocratic social media ban.*

Banning social media is undemocratic by nature, for the government is taking an action that infringes on its citizens’ rights without their consent, even if the action is intended to achieve a greater good. Bans have autocratic intentions behind them, be it with the objective of limiting free speech or thwarting foreign intervention – for which the case of France’s ban in New Caledonia is a perfect example, since it affected the free speech of protesters and was justified using national security against foreign intervention. Thus, social media bans in democracies will have the same drivers as bans in autocracies. The drivers might be any or all of: foreign influence in the country; level of religious freedom; how strong is government propaganda; and use of social media to organize mobilizations, their intensity, composition (mostly average people, mostly national elites), character (pro-democratic or pro-autocratic), and type (e.g. voting, protesting, rioting, rebelling).

1. *The variables that explain democratic social media ban are not of the same nature as the variables that explain autocratic social media ban.*

Even if there are some cases in which democracies and autocracies ban social media in similar fashions, most of them are still distinguishable in their motivations: democratic bans are driven more by foreign interference and threats to democratic institutions, while autocracies are driven more by political mobilization that can contest authority.

* 1. **Data**

The dependent (target) variable for this research is social media ban, a binary variable that assumes a value of 1 if, during a given year, for a given country, in respect to a given social media (either Facebook, Instagram, Telegram, TikTok, WhatsApp, X, or YouTube), it has been banned at least once, and assumes a value of 0 if it has not been banned. This information was acquired utilizing Surfshark’s Internet Shutdown Tracker (IST) project (Surfshark 2024), which compiled cases of internet shutdown for 196 political entities, ranging from 2009 to today, and was last updated on November of 2024. IST compiles cases of internet shutdown – which are not strictly cases of social media ban, for they target the internet as a whole – though both databases differentiate between service-based (focused on impacting social media) and other types of internet shutdown, which enables the analysis specifically on social media ban. The project structures cases in the following manner: each observation represents one instance of shutdown, with start and end dates, and there are separate dummy variables for identifying if Facebook, Twitter, YouTube, Instagram, Telegram, WhatsApp, or other social media have been affected by this shutdown. Because IST does not compile data for TikTok bans, I have done that manually (and for this reason, TikTok bans do not follow the exact same pattern as other social media). Whenever the service-based restriction variable is labelled as “True”, that instance of shutdown refers to a ban on (one or multiple) social media, with the banned social media being labelled as either 1 or 0. For this paper, to allow each observation to represent a single social media ban instead of a single ban for multiple platforms at once, the structure of IST has been altered by reshaping the table from wide to long format. For countries which have not banned any social media (e.g. Sweden, Australia, Germany), these observations are completely empty in IST. Therefore, considering the positive cases are differentiated by country, year, and social media, the negative cases should also be, meaning that if there has been no ban, the observation should be 0 and not empty. Multiple bans of the same social media in the same country in a single year are grouped together and considered as a single observation, and so did bans that lasted less than a year, unlike bans that span multiple years, which are considered as multiple observations, one for each year the social media remained banned. Thus, the final database[[10]](#footnote-11) used will be a combination table containing all possible combinations of 193 countries,[[11]](#footnote-12) 16 years (2009–2024), and 7 social media platforms, resulting in 15.815 datapoints – after impossible observations have been removed, such as South Sudan cases before its independence in 2011, and any social media case before its international release (e.g. TikTok prior to 2017 and Instagram prior to 2010). Of these 15.815 datapoints, only 667 (4.218%) are positive cases, implying that the methods utilized in this research will need to be able to handle unbalanced data.

The independent variables (features) derive mostly unaltered from V-Dem’s database (Coppedge et al. 2024b; Pemstein et al. 2022), with the exceptions of when data was missing, cases which have been imputed by applying the method of forward-filling. Variables were used in the ordinal version (which are denoted by the “\_ord” suffix), when available, to preserve interpretability, since V-Dem applies its own measurement model (Pemstein et al. 2022), which distorts the original variable codification. The following variables were collected:

* **v2clrelig\_ord**, measures the level of religious freedom, and was chosen because countries where citizens cannot choose their own religion are prone to censoring platforms that can be used to promote other religions.
* **v2exl\_legitideol\_ord**, measures to what extent the government promotes a specific ideology in order to justify the regime in place and was chosen because if the government strongly promotes an ideology, it will tend to ban platforms used to promote different ideologies.
* **v2exl\_legitideolcr**, measures the ideology character, and assumes binary values for each variable (a series of dichotomous scales), ending in:
  + \_0 if nationalist (v2exl\_legitideolcr\_0 and so on);
  + \_1 if socialist or communist;
  + \_2 if restorative or conservative;
  + \_3 if separatist or autonomist;
  + \_4 if religious.
* **v2cagenmob\_ord**, measures how frequent and large mass mobilizations have been and was chosen because autocracies will tend to ban social media if it can be used to induce mass mobilizations.
* **v2cademmob\_ord**, measures how frequent and large events of mass mobilization for pro-democratic aims have been and was chosen because autocracies will tend to ban social media if they can induce democratic mobilizations.
* **v2caautmob\_ord**, measures how frequent and large events of mass mobilization for pro-autocratic aims have been and was chosen because democracies will tend to ban social media if they can induce autocratic mobilizations.
* **v2smfordom\_ord**, measures how often foreign governments use social media to disseminate political misinformation in their country and was chosen because democratic and undemocratic regimes will tend to ban social media should it be weaponized by foreign governments.
* **v2smorgavgact\_ord**, measures use of social media by average people to organize offline political action and was chosen because autocracies will ban social media if they are used for political mobilization purposes.
* **v2smorgelitgact\_ord**, measures use of social media by elites to organize offline political action and was chosen because autocracies will ban social media if they are used for this reason, following Bueno de Mesquita’s (2003) winning coalition political survival theory.
* **v2smorgtypes**, measures the types of political mobilization generated through social media (a series of dichotomous scales), ending in:
  + **\_**0 if petition signing (v2smorgtypes\_0 and so on);
  + \_1 if voter turnout;
  + \_2 if street protests;
  + \_3 if strikes/labour actions;
  + \_4 if riots;
  + \_5 if organized rebellion;
  + \_6 if vigilante justice;
  + \_7 if terrorism;
  + \_8 if ethnic cleansing/genocide;
  + \_9 if other, but was not utilized in this research.
* **dem\_var**, although not a variable present in V-Dem, it was constructed based on the control variable v2x\_polyarchy as follows: the value of dem\_var for a given country in a given year for social media A is the difference between its democracy level (v2x\_polyarchy) and the democracy level for that given year of the country from which social media A originated. All the social media under analysis can only originate from the US (Facebook, Twitter, YouTube, Instagram, and WhatsApp), Russia (Telegram) or China (TikTok). This variable measures how distant two states are in terms of democracy levels. I expect high levels of dem\_var to be positively associated with social media ban.

Control variables (also called features) were also gathered from V-Dem, while data for population sizes and GDP *per capita* (in the models as PIB *per capita*) were collected and combined from the World Bank (2024a; 2024b),[[12]](#footnote-13) the International Monetary Fund (2024),[[13]](#footnote-14) and the United Nations Statistics Division (2024)[[14]](#footnote-15) datasets. Control variables such as country (to which ordinal encoding was applied), social media (to which one-hot encoding was applied), and year were already present in IST, while the following variables were collected from V-Dem:

* **v2x\_polyarchy**, measures a country’s democracy level and was used to construct “dem\_var”.
* **v2x\_regime**, also measures a country’s democracy level (even though its codification is ordinal, it does not have the “\_ord” suffix). It was used to separate the dataset into autocratic (values 0 and 1) and democratic (2 and 3) subdivisions – which created a 47%–53% split in the data – but was not utilized in model training, since it has a correlation of 0.93 with v2x\_polyarchy.[[15]](#footnote-16)
  + A value of 0 represents a closed autocracy, a regime that does not hold multiparty elections at all;
  + A value of 1 represents an electoral autocracy, a regime that holds elections that are neither free nor fair;
  + A value of 2 represents an electoral democracy, a regime that holds de-facto free and fair elections but does not meet minimum liberal institutional requirements, such as access to justice, transparent law enforcement, and rule of law;
  + A value of 3 represents a liberal democracy, a regime that holds free and fair elections, and satisfies liberal institutional requirements.
* **v2smgovfilcap**, measures the capacity of governments to censor the internet.
* **v2smgovshutcap**, measures the capacity of governments to completely shut down the internet.
* **v2smgovsmmon**, measures how comprehensive is governmental surveillance over social media.
* **v2smregcap**, measures the material capacity governments posess to regulate content on social media.
* **v2smpolsoc**, measures the level of polarization of society.
* **e\_regionpol**, a classification of political region.[[16]](#footnote-17)

With the addition of V-Dem’s variables to the dataset, the dimensions of the final database are 15.815 observations and 39 variables. [[17]](#footnote-18)

* 1. **Methodology**

A machine learning task is comprised of three main steps: feature engineering, training, and testing. Depending on the requirements of the models employed, features need to be prepared accordingly: for logistic regressions and tree-based models (decision trees and random forests), textual data must be transformed into numbers; since trees consider every variable separately, there is no need to rescale variables, but this step is necessary for regressions. Therefore, the variable with country names will be encoded with ordinal encoding (every country receives a unique ordered number) and the variable with social media names will be encoded with one-hot encoding (every category becomes a dummy column), since there are only 7 categories (a manageable number to apply one-hot encoding). And all other variables will be rescaled through robust scaling, a method fit to deal with possible outliers in the distribution, though V-Dem’s data is already scaled enough that there are no huge variations.[[18]](#footnote-19)

During training, to ensure I can get the best predictive capability from the models, I will apply the method of grid search to find the best combination of hyperparameters for each of the models (democratic, autocratic and general). And to verify the robustness of the results, to each grid search combination, I will implement cross-validation on 5 folds. This means that training data will be divided into 5 equal lots, and one will be separated from the rest to be the validation set (similar in nature to a test set). A model will be trained five times, each time with a different lot serving as the validation set, a rotation that ensures the model does not overfit to train data. After every combination has been cross-validated, the model with best performance will be tested against the test set and evaluation set, according to the schema in Table 1.

Due to the huge discrepancy in the balance of the target variable (way more 0’s than 1’s), the performance metric applied to the models cannot be the accuracy [[19]](#footnote-20) because a model that only predicts 0 every time would have high accuracy, since most values are indeed 0. I want a model less accurate, but better at predicting the positive cases, thus requiring a more robust metric. Since misclassifying 0’s and misclassifying 1’s would have the same impact (one type of mistake is not worse than the other), a good metric would be one that combines both precision (P) and recall (R): F1 Score, a harmonic mean of both metrics.[[20]](#footnote-21)

As a final step to ensure the robustness of the results, I added two random variables, one based on a normal distribution (random\_n) and a second one based on a uniform distribution (random\_u), to serve as a double threshold for the importance of features: if a feature has an importance as good as or worse than a random variable, its actual importance is practically null, for it brings as much predictive power as purely random numbers can (that is, none). This step is very useful for tree-based models, during the analysis of feature importances, and does not affect the logistic regression.

* + 1. **Logistic Regression**

Since the problem at hand is binary by nature, a simple linear regression is not fit for dealing with it, requiring the use of a logistic regression, which uses a different mathematical function, called sigmoid, capable of mapping values in an interval between 0 and 1 and approximating the results to be either 0 or 1, facilitating prediction. All regression models were cross-validated on the following grid search, with the best values for each of them highlighted.

**Table 2**: Regression Models’ Best Hyperparameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hyperparameters**[[21]](#footnote-22) | **Possible values** | **Democratic** | **Autocratic** | **General** |
| Penalty (regularization) | **None**, Lasso (L1), Ridge (L2), or Elastic-Net (a combination of L1 and L2) | None | None | None |
| Tolerance | , , , , |  |  |  |
| C (regularization) | **0.01**, 0.1, 1, 10, 100 | 0.01 | 0.01 | 0.01 |
| Class weight | None or **balanced** | Balanced | Balanced | Balanced |
| L1 ratio | **0.1**, 0.5, 0.9 | 0.1 | 0.1 | 0.1 |

The three models had the same best hyperparameter combination, all values that enable high levels of overfitting, by making the model very sensible to small changes in the data and limiting the constraints of penalties (even eliminating them completely) – though cross-validation, to some extent, serves to prevent excessive overfitting.

**Table 3**: Regression Models’ Performances (F1 Score)[[22]](#footnote-23)

|  |  |  |  |
| --- | --- | --- | --- |
| **Sets** | **Democratic** | **Autocratic** | **General** |
| Train | 0.036 | 0.397 | 0.266 |
| Test | 0.039 | 0.409 | 0.289 |
| Evaluation | 0.178 | 0.026 | |  |  | | --- | --- | | Democracy | Autocracy | | 0.075 | 0.342 | |

The results seem to be all over the place, without distinguishable trends and patterns – the model trained on democracies performs much better at autocracies than democracies. What they do show clearly is that the logistic regressions do not perform well with this data, denoting that the problem at hand may not have a linear patter after all, prompting the use of more complex models.

* + 1. **Decision Tree**

Decision trees are non-parametric supervised machine learning algorithms, more advanced than logistical regressions, that can divide observations into small homogenous groups after a series of subdivisions dependent on rules based on the independent variables provided (IBM 2021). The initial nodes of decision trees are easy to visualize and interpret, because of this step-by-step rule-based division, but the deeper you get into the tree, the less sense the rules make, since there are fewer and fewer observations per node, generating specific rules for those small sets of observations, leading to overfitting.[[23]](#footnote-24) This is why it is important to set limits to the depth of a tree and minimum amounts of cases per leaf. All decision tree models were cross-validated on the following grid search, with the best values for each of them highlighted.

**Table 4**: Decision Tree Models’ Best Hyperparameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hyperparameters**[[24]](#footnote-25) | **Possible values** | **Democratic** | **Autocratic** | **General** |
| Max depth | 1, 6, 11, 16, **21**, 26, 31, 50, None | 6 | 21 | 21 |
| Minimum number of samples per leaf | **2**, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 | 2 | 1 | 2 |
| Minimum number of samples to split a node | 1, **2**, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 | 2 | 2 | 2 |

The three models performed best when they did not limit the minimum size of the nodes and restricted the depth to a small or medium-sized tree. This might be the case because, should the tree only allow for larger numbers of samples per node, the unbalanced positive cases may drown in the middle of too many negative cases for them to be able to weigh in the decision of the final node. But then again, values that are too small might lead to overfitting.

**Table 5**: Decision Tree Models’ Performances (F1 Score)

|  |  |  |  |
| --- | --- | --- | --- |
| **Sets** | **Democratic** | **Autocratic** | **General** |
| Train | 0.627 | 1.0 | 0.926 |
| Test | 0.434 | 0.726 | 0.715 |
| Evaluation | 0.101 | 0.0 | |  |  | | --- | --- | | Democracy | Autocracy | | 0.666 | 0.719 | |

The results are definitely better, in comparison with the logistic regressions. While autocratic data is fairly well-predicted and democratic data shows a significant lack in predictability, the general model does a better job at predicting democratic ban than the democratic model and is on par, by a small difference, with the autocratic model at predicting autocratic ban.

* + 1. **Random Forest**

Random forests are like turbined decision trees. They create numerous randomized trees at the same time, each with random samples of the training data and of the available features and use all of them to predict new observations. The trick lies in the randomization: when trees are given the chance, they overfit immediately, but if the model allows for many overfitted trees to make specific predictions and even more non-overfitted trees to make more generalized predictions, the average of their answers results in generalizable yet accurate predictions. All random forest models were cross-validated on the following grid search, with the best values for each of them highlighted.

**Table 6**: Random Forest Models’ Best Hyperparameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hyperparameters**[[25]](#footnote-26) | **Possible values** | **Democratic** | **Autocratic** | **General** |
| Max depth | 5, 10, 15, 50, None | 10 | 50 | 50 |
| Minimum number of samples per leaf | 1, 2, 3, 5 | 1 | 1 | 1 |
| Minimum number of samples to split a node | 2, 3, 4, 5 | 3 | 2 | 2 |
| Maximum number of features per tree | Log2, Sqrt, 0.5, or None | Log2 | None | None |
| Number of trees | 100, 300, 500, 800, 1100 | 100 | 800 | 300 |

Once again the three models performed best when a small number of samples were allowed on the tree leaves, although the maximum depth is considerably higher for the autocratic and general models – which is expected, since random forests need to take different sized trees into consideration to make good predictions – unlike for the democratic model, which stayed very close to the dimensions of the decision tree model.

**Table 7**: Random Forest Models’ Performances (F1 Score)

|  |  |  |  |
| --- | --- | --- | --- |
| **Sets** | **Democratic** | **Autocratic** | **General** |
| Train | 0.892 | 1.0 | 1.0 |
| Test | 0.363 | 0.744 | 0.714 |
| Evaluation | 0.0 | 0.0 | |  |  | | --- | --- | | Democracy | Autocracy | | 0.434 | 0.733 | |

The results are not too different from what the decision trees offered. The random forests improved the results for autocracies and increased the gap between democratic train and test – a gap that might be a strong indicator of overfitting, since the performance for unseen data was so weak compared to observed data.

1. **Discussion of Results and Conclusion**

Firstly, let us remember this paper’s research question: “Are the drivers for social media ban in democratic countries and undemocratic countries the same?”, and its hypotheses:

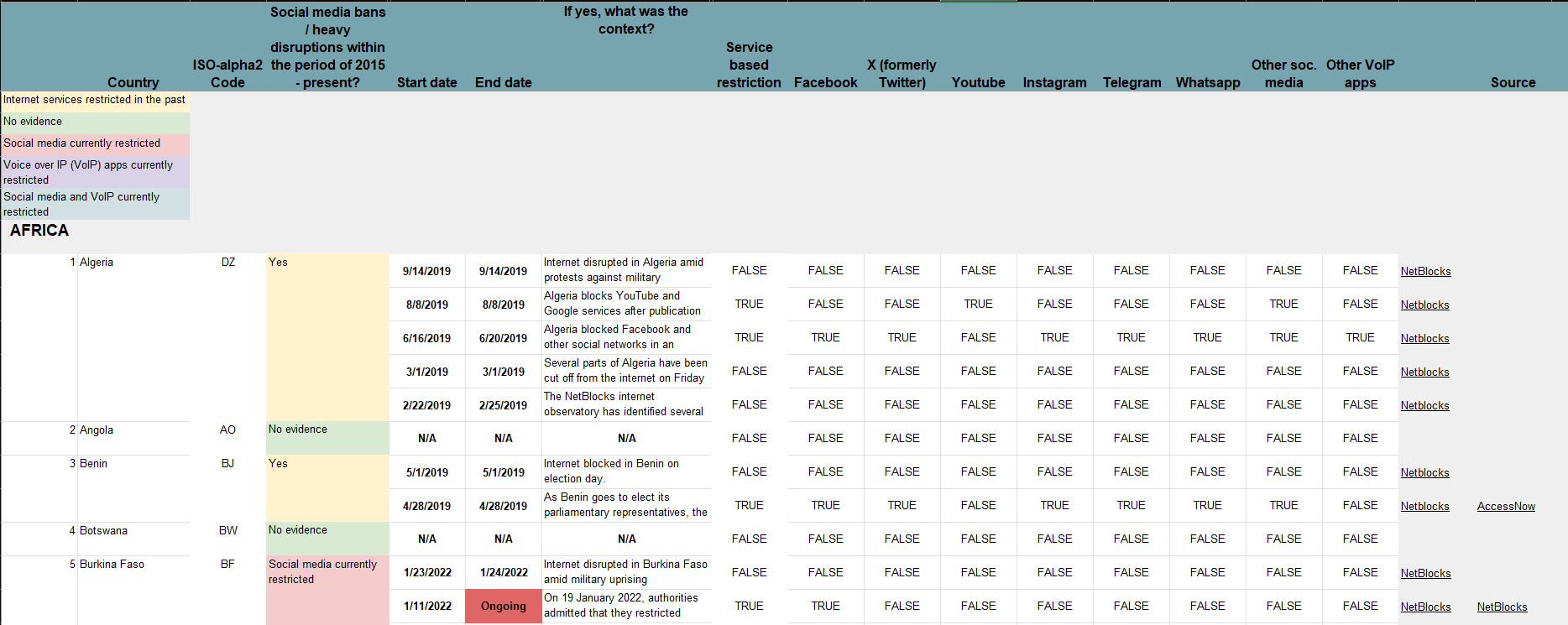
1. *The variables that explain democratic social media ban are of the same nature as the variables that explain autocratic social media ban.*
2. *The variables that explain democratic social media ban are not of the same nature as the variables that explain autocratic social media ban.*

To test the hypotheses, it is possible to utilize the tree-based model’s measure of feature importance to rank the most important/influential variables in predicting social media ban.[[26]](#footnote-27) All variables with higher importances than both random\_u and random\_n are definitely significant to explain social media ban; variables between them, could be significant, but is hard to justify; any variables below them are most definitely not relevant, since they have an inferior predictable power than random phenomena. Similarly, any variables ranked highly by models with poor performance in the test set are also likely to be irrelevant to the model, for they only explain cases they were tailored to explain and any case not seen during training is outside their tackling range. For this reason, results from logistic regressions are not going to be used in this analysis, since they underperformed for both train and test data. Likewise, all democratic models also underperformed and seem to be overfitted, especially when compared to the General Decision Tree, with a train F1 Score of 0.926, a test F1 Score of 0.715, and an evaluation F1 Score of 0.666 on democratic data, outperforming by almost 2 decimal points the F1 Score of the best models that attempted to predict democratic data. Thus, the models to be used to verify the claims of all other models regarding social media ban are Autocratic and General Decision Trees and Random Forests, since they have consistent performance.

The control variable “population size” appears on five instances ranking higher than random variables in the feature importances plots, as does “political region” once, “democracy index” four times, and “country” twice. The only independent variables that also rank higher than the random variables are “strikes/labour actions” once and “democratic variation” four times. In between the two random variables, “petition signing” appears twice and “religious ideology” once, as well as Telegram (Soc\_media\_4) also once.

If the general models performed well on their test set and, on average, both evaluation sets, there is an argument to be made that either social media ban is not properly explained through analysis of the democratic/autocratic dynamic – and, thus, there is a completely different reason that explains their occurrence, which can only be found when combining the two datasets, a reason completely detached from the protection of freedom of expression and from the dynamics of political survival (prompting future research) –, or there is a combination of variables that produces just the right rules to divide social media ban cases that only works when autocracies and democracies are scrutinized under a single lens. When analyzing the General Tree plot, the first division is based on democracy index, and so does the Autocratic Tree on its plot – with the exact same division at v2x\_polyarchy ≤ 0.21 – while the Democratic Tree does not utilize any variables of democracy on its first levels. How does the General Tree combine both trees to make predictions as good as a model trained solely on autocracies and even surpass the predictions of a model trained only on democracies? I argue that the missing piece to better predict democratic ban is the addition of rules that were based on autocratic bans, because, intrinsically to democratic ban, there is an autocratic component, exactly what my first hypothesis suggests. One could argue that none of the specific models were successful in their evaluation step, a finding that implies that the rules, the patterns found in either authoritarian or democratic social media ban cannot explain the bans in the other regime, effectively contradicting H1. And I counter this position, by reaffirming what I have just uncovered: indeed a model trained solely on democratic data could not efficiently predict bans in autocracies (and, apparently, not even in its own data), but it is when we combine the rules for democratic ban with the rules for autocratic ban that form the connection between these two, the motivations that unite these kinds of bans in a single thing emerge, and enhance the prediction capabilities of both models.

1. **Appendices**
   1. **IST Structure pre and post manipulation**



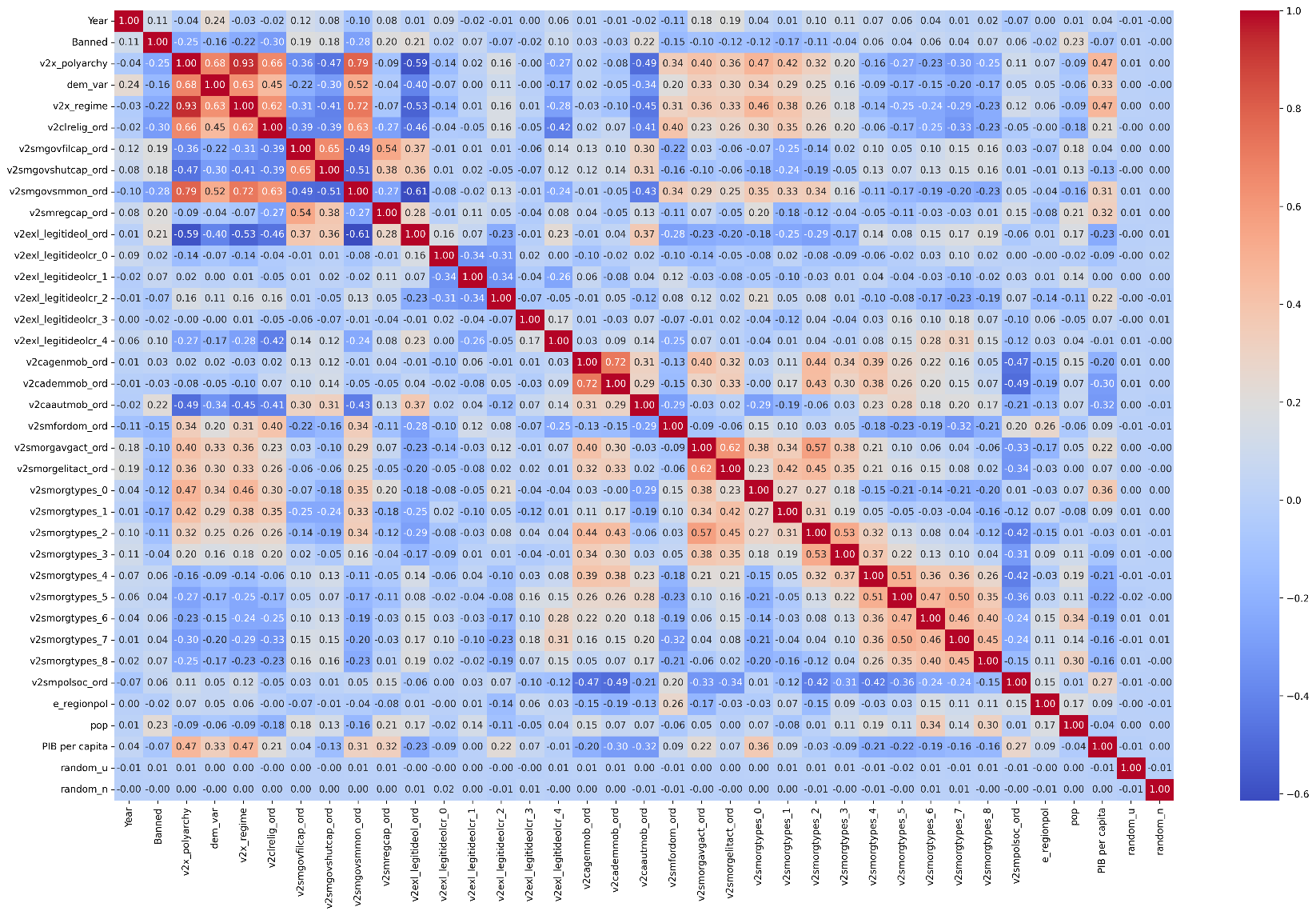
The structure of IST database pre-manipulation (columns hidden for clarity)

A screenshot of a computer

Description automatically generated

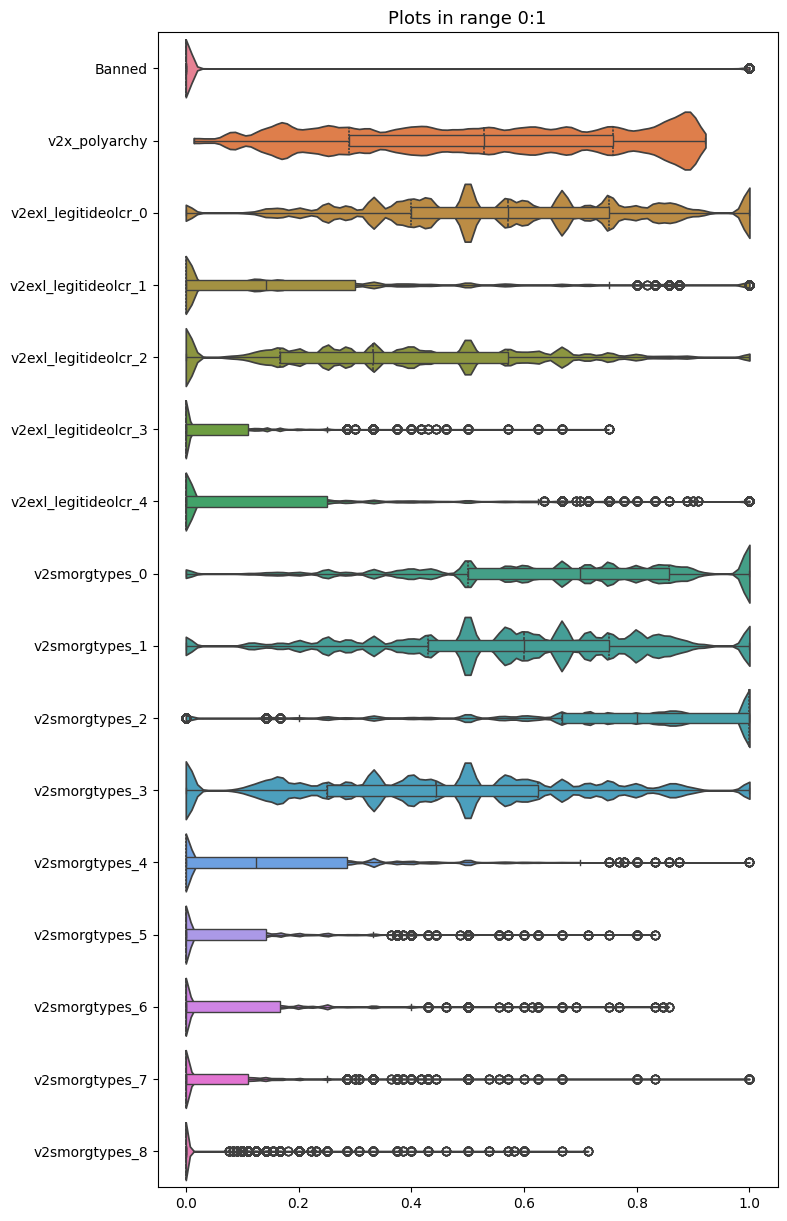
The structure of IST data post-manipulation (V-Dem data included; not showing all 41 columns)

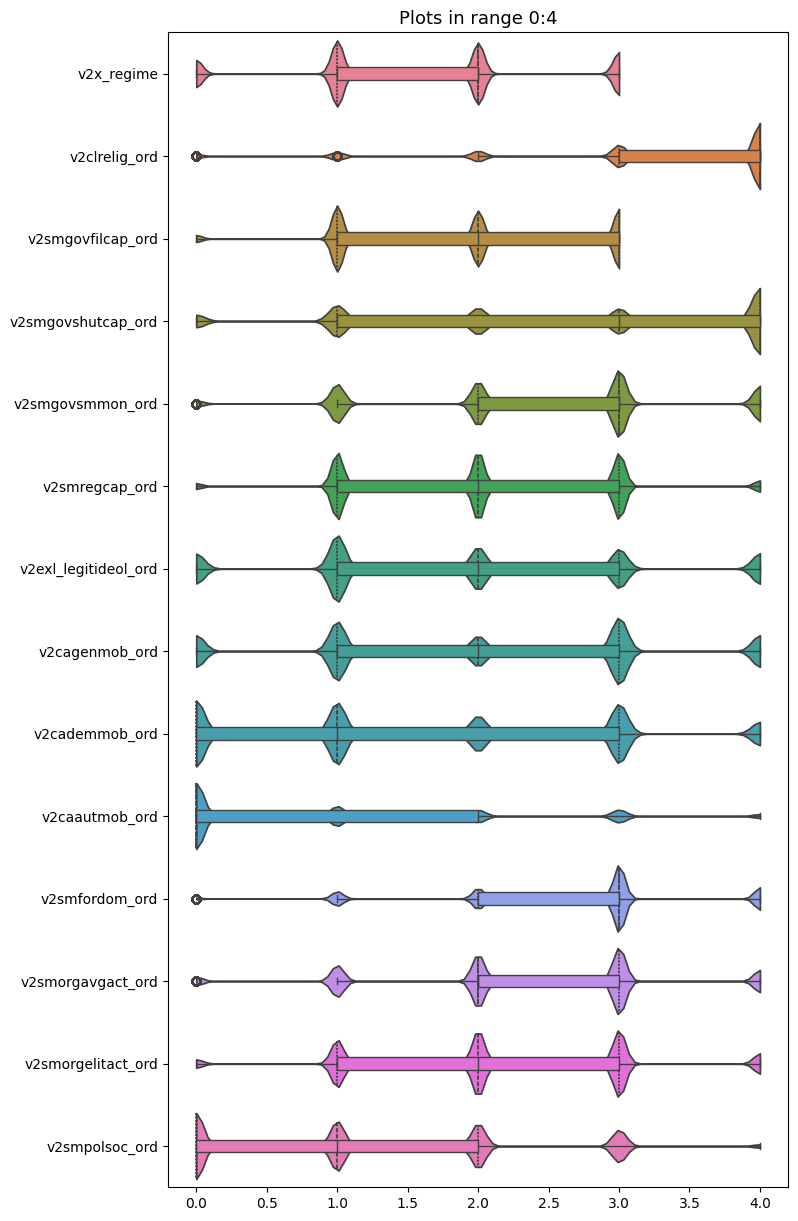
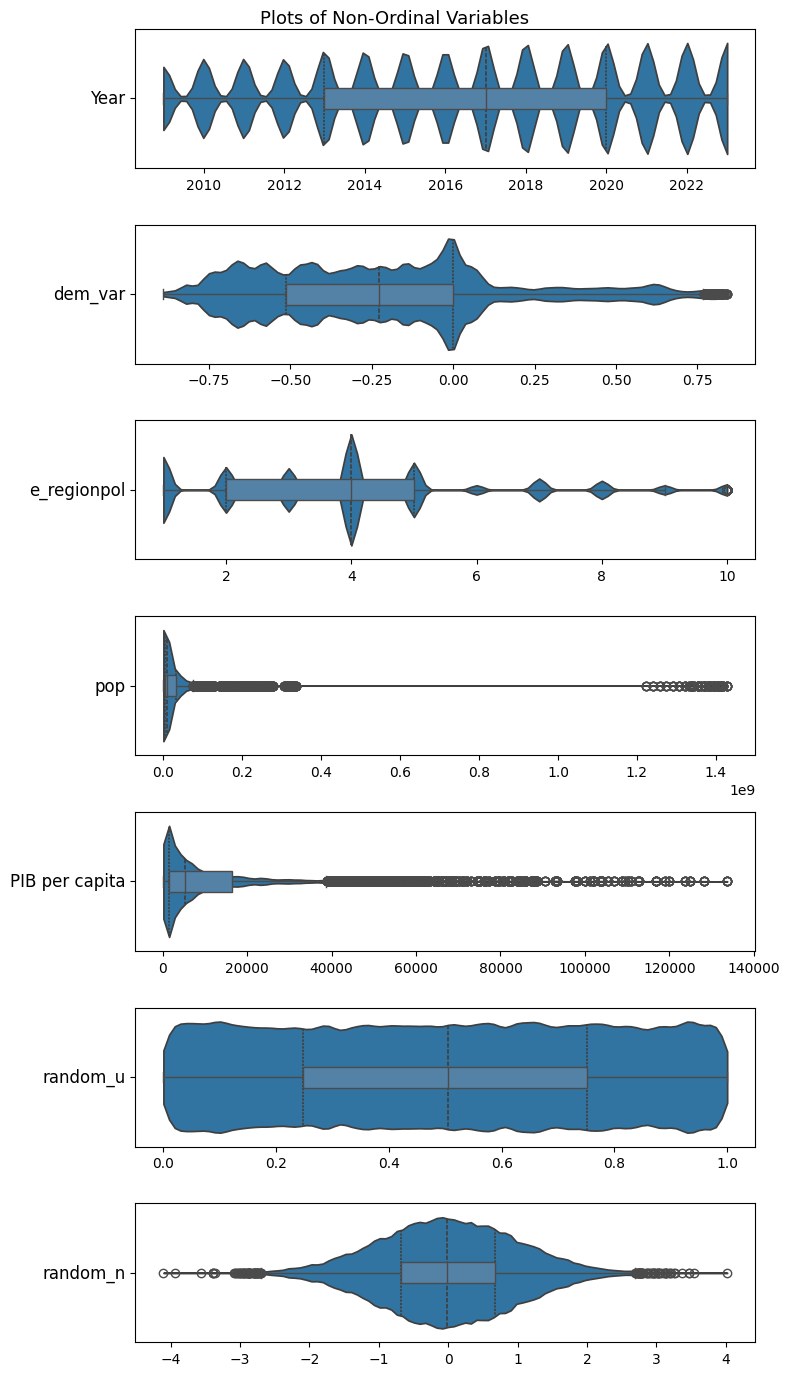
* 1. **Heatmap of correlations**

****

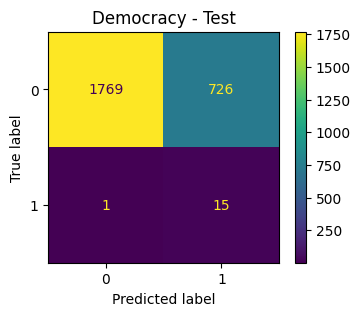
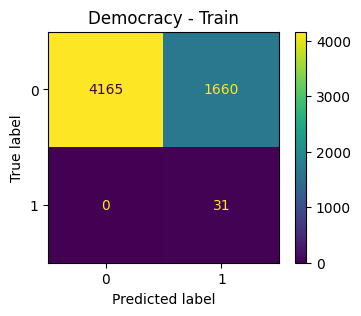
Correlations of all variables (excluding the textual variables, country and social media)

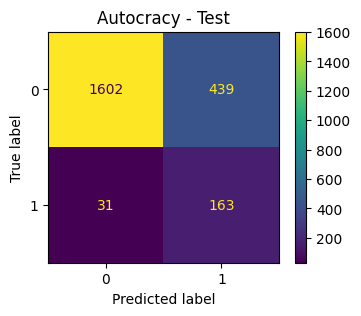
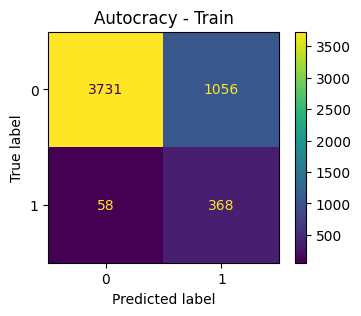
* 1. **Variable Violin Distributions & Boxplots**

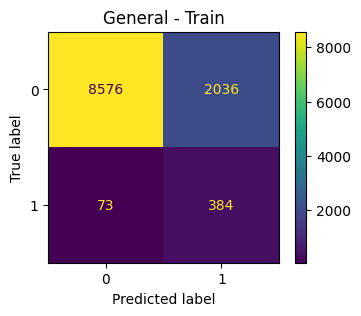


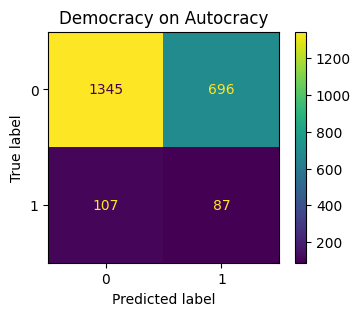
* 1. **Confusion Matrices**
     1. **Logistic Regression**



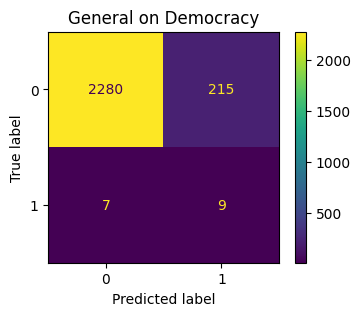


 A chart with numbers and labels

Description automatically generated with medium confidence

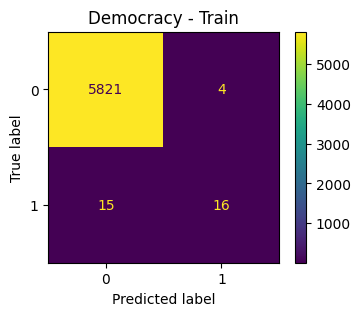
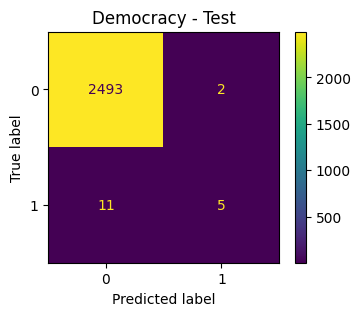
 A chart with numbers and labels

Description automatically generated

 A chart of a graph

Description automatically generated with medium confidence

* + 1. **Decision Tree**

A chart of a train

Description automatically generated A chart with numbers and a yellow and purple squares

Description automatically generated

A chart of a train

Description automatically generated A chart with yellow and purple squares

Description automatically generated

A chart of a political party

Description automatically generated with medium confidence A chart of a political party

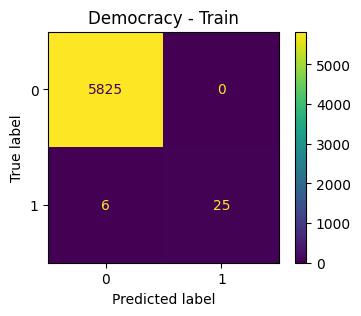
Description automatically generated with medium confidence

A chart of general on democracy

Description automatically generated A chart of a graph

Description automatically generated with medium confidence

* + 1. **Random Forest**

 A chart of a test

Description automatically generated with medium confidence

A chart of a train

Description automatically generated A chart with numbers and a few colored squares

Description automatically generated with medium confidence

A chart of a train

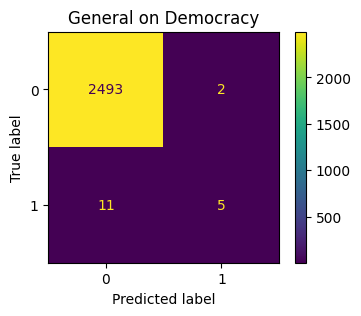
Description automatically generated A chart of a test

Description automatically generated

A chart of a political party

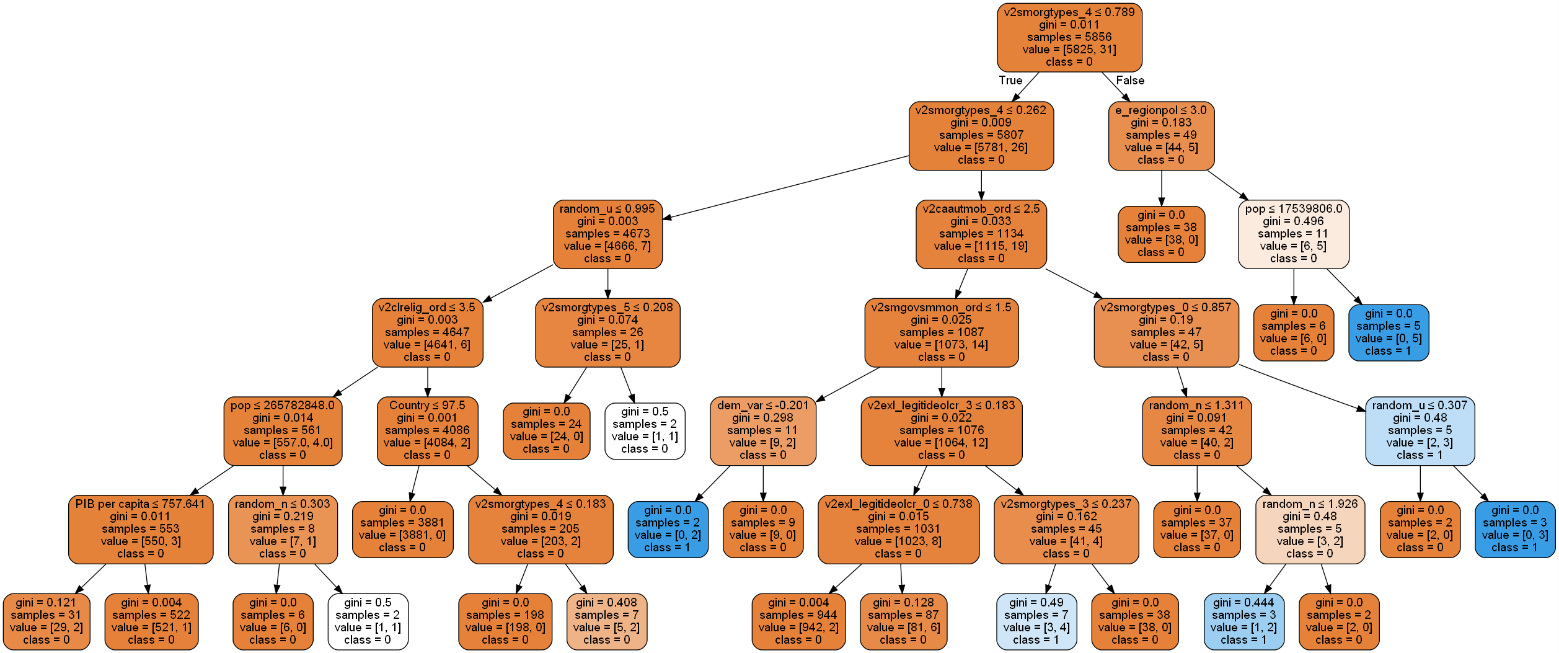
Description automatically generated with medium confidence A chart of a political party

Description automatically generated with medium confidence

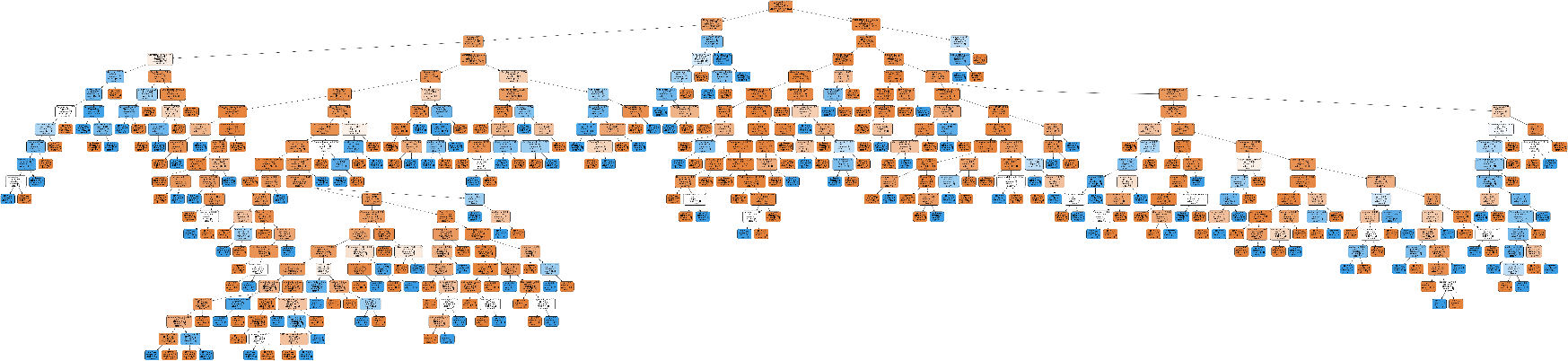
 A chart with numbers and a graph

Description automatically generated with medium confidence

* 1. **Decision Tree Graphs[[27]](#footnote-28)**
     1. **Democratic Tree**

****

* + 1. **Autocratic Tree**

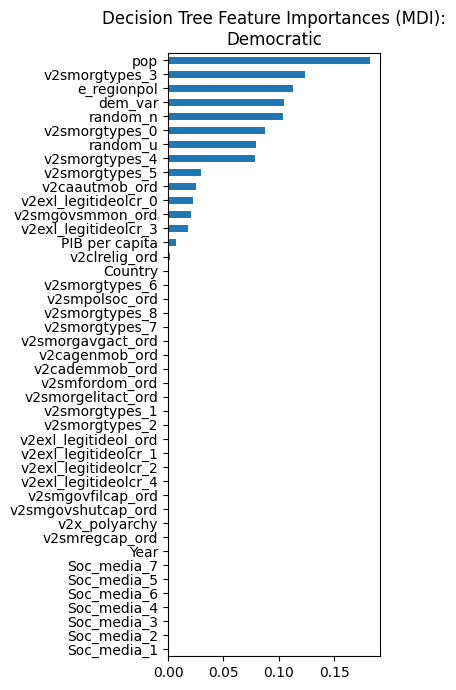
****

* + 1. **General Tree**

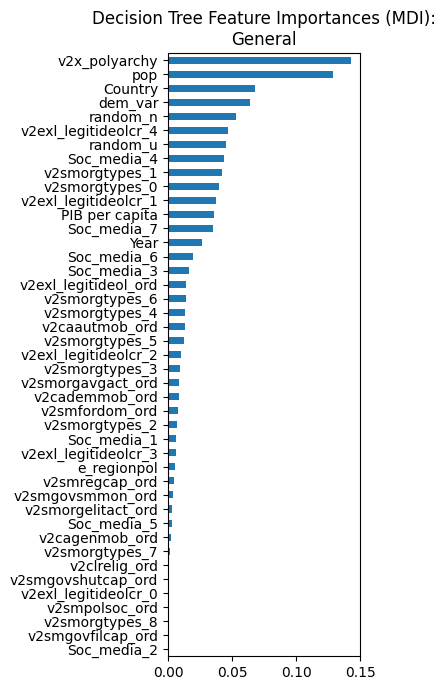
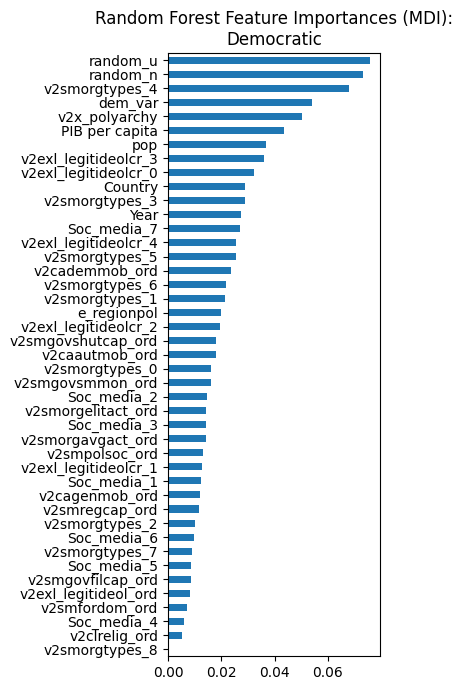
**A group of small squares

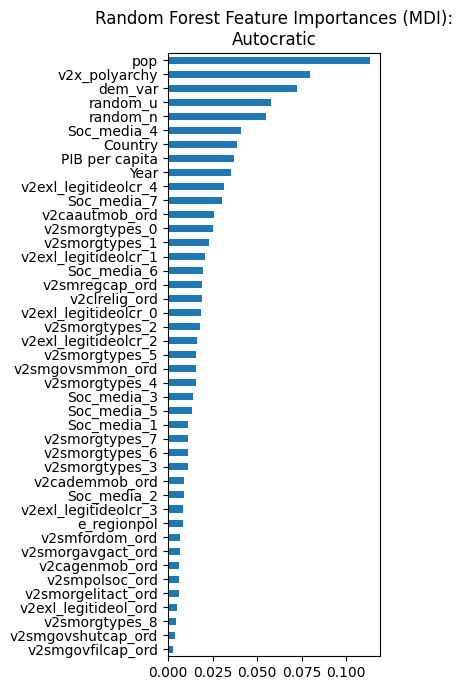
Description automatically generated with medium confidence**

* 1. **Feature Importances**

 A graph of a number of data

Description automatically generated with medium confidence

 A graph of a number of text

Description automatically generated with medium confidence

1. **Bibliography**

Abi-Habib, Maria. 2020. ‘India Bans Nearly 60 Chinese Apps, Including TikTok and WeChat’. *The New York Times*, 29 June 2020, sec. World. https://www.nytimes.com/2020/06/29/world/asia/tik-tok-banned-india-china.html.

Acessnow. 2022. ‘Turn TikTok on: Authorities in Azerbaijan and Armenia Must Not Restrict Access’. Access Now. 2022. https://www.accessnow.org/press-release/tiktok-azerbaijan-armenia/.

Aleksic, Adam. 2024. ‘Technology shapes our communication’. Instagram. 12 June 2024. https://www.instagram.com/reel/C8HvpHDA6Dv/.

Amazon Web Services. 2021. ‘Model Interpretability - Machine Learning Best Practices in Healthcare and Life Sciences’. Amazon Web Services. https://docs.aws.amazon.com/whitepapers/latest/ml-best-practices-healthcare-life-sciences/model-interpretability.html.

Anderson, Katie Elson. 2020. ‘Getting Acquainted with Social Networks and Apps: It Is Time to Talk about TikTok’. *Library Hi Tech News* 37 (4): 7–12. https://doi.org/10.1108/LHTN-01-2020-0001.

AP News. 2023. ‘France Bans TikTok, Twitter from Government Staff Phones’. AP News. 24 March 2023. https://apnews.com/article/tiktok-france-ban-cybersecurity-china-4c48564fbfe7b86bf44c30969902c293.

APA. 2023. ‘Access to TikTok Restricted in Azerbaijan’. Apa.Az. 19 September 2023. https://apa.az/incident/access-to-tiktok-restricted-in-azerbaijan-412160?locale=en.

Associated Press. 2022. ‘Jordan Bans TikTok after Police Officer Killed in Protests’. NBC News. 17 December 2022. https://www.nbcnews.com/news/world/jordan-bans-tiktok-police-officer-killed-protests-rcna62229.

*BBC News*. 2022. ‘Afghanistan: Taliban Orders TikTok, PUBG Ban for “misleading” Youths’, 22 April 2022. https://www.bbc.com/news/world-asia-61185931.

Boese, Vanessa Alexandra, and Matthew C. Wilson. 2022. ‘A Short History of Contestation and Participation’. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4022448.

Buchholz, Katharina. 2020. ‘Infographic: Which Countries Have Banned Huawei?’ Statista Daily Data. 30 January 2020. https://www.statista.com/chart/17528/countries-which-have-banned-huawei-products.

Bueno De Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. The MIT Press. https://doi.org/10.7551/mitpress/4292.001.0001.

Cameron, Chris. 2023. ‘The Attack on Brazil’s Seat of Government Resembles the Storming of the U.S. Capitol on Jan. 6, 2021.’ *The New York Times*, 9 January 2023, sec. World. https://www.nytimes.com/2023/01/08/world/americas/brazil-jan-6-riots.html.

Ceci, Laura. 2024. ‘TikTok - Statistics & Facts’. Statista. 2 June 2024. https://www.statista.com/topics/6077/tiktok/.

Clausius, Madison. 2022. ‘The Banning of TikTok, and the Ban of Foreign Software for National Security Purposes’. *Washington University Global Studies Law Review* 21 (2). https://journals.library.wustl.edu/globalstudies/article/id/8751/.

Clynch, Harry. 2024. ‘US TikTok Ban “Could Embolden African Governments”’. African Business. 15 March 2024. https://african.business/2024/03/technology-information/us-tiktok-ban-could-embolden-african-governments.

Collins, Benedict. 2023. ‘Canada Joins List of Nations to Ban Kaspersky and WeChat from Government Devices’. TechRadar. 31 October 2023. https://www.techradar.com/pro/canada-joins-list-of-nations-to-ban-kaspersky-and-wechat-from-government-devices.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Fabio Angiolillo, et al. 2024a. ‘V-Dem Codebook V14’. Varieties of Democracy (V-Dem) Project.

———. 2024b. ‘V-Dem [Country-Year/Country-Date] Dataset V14’. Varieties of Democracy (V-Dem) Project.

Cueto, José Carlos. 2023. ‘Qual a diferença entre o TikTok e o Douyin, versão do app usada na China’. BBC News Brasil. 4 April 2023. https://www.bbc.com/portuguese/articles/cq5zydp59j7o.

Dahl, Robert A. 2005. ‘What Political Institutions Does Large-Scale Democracy Require?’ *Political Science Quarterly* 120.

Dahl, Robert A., and Ian Shapiro. 2024. ‘Democracy - Factions, Parties, Politics | Britannica’. 21 November 2024. https://www.britannica.com/topic/democracy.

Dosunmu, Damilare. 2023. ‘The Politics of Africa’s TikTok Bans’. Rest of World. 7 September 2023. https://restofworld.org/2023/africa-tiktok-banned-politics/.

Fung, Brian. 2023. ‘TikTok Collects a Lot of Data. But That’s Not the Main Reason Officials Say It’s a Security Risk | CNN Business’. CNN. 24 March 2023. https://www.cnn.com/2023/03/24/tech/tiktok-ban-national-security-hearing/index.html.

Funk, Vesteinsson, Baker, Brody, Grothe, Agarwal, Barak, Loldj, Masinsin, and Sutterlin. 2024. ‘Freedom on the Net 2024’. Freedom House.

Gehlbach, Scott, Zhaotian Luo, Anton Shirikov, and Dmitriy Vorobyev. 2022. ‘A Model of Censorship, Propaganda, and Repression’.

Gerken, Tom, and Tom Singleton. 2024. ‘TikTok Vows to Fight “unconstitutional” US Ban’. BBC News. 24 April 2024. https://www.bbc.com/news/articles/c87zp82247yo.

Hale, Erin. 2024. ‘Why Has the US Passed a Bill to Ban TikTok, and What’s next?’ Al Jazeera. 14 March 2024. https://www.aljazeera.com/news/2024/3/14/why-has-the-us-passed-a-bill-to-ban-tiktok-and-whats-next.

House Committee on Energy and Commerce. 2024. ‘Chair Rodgers: “We Have Given TikTok a Choice: Divest or Face a Ban”’. House Committee on Energy and Commerce. 7 March 2024. https://energycommerce.house.gov/posts/energycommerce.house.gov.

IBM. 2021. ‘What Is a Decision Tree?’ 2 November 2021. https://www.ibm.com/topics/decision-trees.

International Monetary Fund. 2024. ‘GDP per Capita, Current Prices’. https://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEOWORLD.

Kahn, Gretel. 2024. ‘X Has Been Blocked in Brazil. Does It Matter for Journalism?’ Reuters Institute for the Study of Journalism. 16 September 2024. https://reutersinstitute.politics.ox.ac.uk/news/x-has-been-blocked-brazil-does-it-matter-journalism.

Kim, Juliana. 2024. ‘Trump Calls It an “honor” to Join TikTok after Previously Trying to Ban the App’. *NPR*, 2 June 2024, sec. Politics. https://www.npr.org/2024/06/02/nx-s1-4989327/trump-tiktok-ban.

Kirby, Paul. 2024. ‘New Caledonia TikTok Ban Lifted as State of Emergency Ends’. 2024. https://www.bbc.com/news/articles/c0dd94jv9jpo.

Milmo, Dan. 2022. ‘TikTok Users in Russia Can See Only Old Russian-Made Content | TikTok | The Guardian’. 10 March 2022. https://amp.theguardian.com/technology/2022/mar/10/tiktok-users-in-russia-can-see-only-old-russian-made-content.

———. 2024. ‘Canada Orders TikTok to Close Offices over “Security Risks”’. *The Guardian*, 7 November 2024, sec. World news. https://www.theguardian.com/world/2024/nov/07/canada-orders-tiktok-to-close-vancouver-office-over-security-risks.

Office of the Director of National Intelligence. 2024. ‘Annual Threat Assessment of the U.S. Intelligence Community’. https://www.odni.gov/files/ODNI/documents/assessments/ATA-2024-Unclassified-Report.pdf.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2022. ‘The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data’. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3595962.

Rep. Gallagher, Mike. 2024. ‘H.R.7521 - 118th Congress (2023-2024): Protecting Americans from Foreign Adversary Controlled Applications Act’. Legislation. 2024-03-05. 14 March 2024. https://www.congress.gov/bill/118th-congress/house-bill/7521.

Reuters. 2023a. ‘European Countries Who Put Curbs on Huawei 5G Equipment - ET Telecom’. ETTelecom.Com. 29 September 2023. https://telecom.economictimes.indiatimes.com/news/telecom-equipment/european-countries-who-put-curbs-on-huawei-5g-equipment/104048918.

*———*. 2023b. ‘Senegal Seeks Regulation Deal with TikTok after Ban’, 5 October 2023, sec. Technology. https://www.reuters.com/technology/senegal-seeks-regulation-deal-with-tiktok-after-ban-2023-10-05/.

———. 2024. ‘Protests, Court Appeals as Nepal’s TikTokers Decry Government Ban on App’. 24 November 2024. https://in-cyprus.philenews.com/international/protests-court-appeals-as-nepals-tiktokers-decry-government-ban-on-app/.

RFE/RL’s Kyrgyz Service. 2023. ‘Kyrgyz Authorities Ban TikTok Citing Effects On Child Development’. 30 August 2023. https://www.rferl.org/a/kyrgyzstan-tiktok-ban-children/32571122.html.

Roy, Enakshi. 2017. ‘Social Media, Censorship and Securitization in the United States and India’. Ohio University.

Sharma, Anchit. 2024. ‘How To Use WeChat In India After Ban’. Forbes Advisor INDIA. 12 January 2024. https://www.forbes.com/advisor/in/business/software/how-to-use-wechat-in-india-after-ban/.

Sharma, Gopal. 2023. ‘Nepal to Ban TikTok, Alleges Damaging Social Impact’. *Reuters*, 14 November 2023, sec. Technology. https://www.reuters.com/technology/nepal-govt-decides-ban-chinas-tiktok-ani-2023-11-13/.

Sigman, Rachel, and Staffan I. Lindberg. 2015. ‘The Index of Egalitarian Democracy and Its Components: V-Dem’s Conceptualization and Measurement’. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2727612.

Statista. 2024. ‘TikTok Users by Country 2024’. Statista. 2024. https://www.statista.com/statistics/1299807/number-of-monthly-unique-tiktok-users/.

Supak, Gabrielle. 2022. ‘Political Posturing or a Move Towards “Net Nationalism?” The Legality of a TikTok Ban and Why Foreign Tech Companies Should Be Paying Attention’ 22 (April).

Surfshark. 2024. ‘Internet Censorship Cases across the World\_last Update November 12th 2024’. https://docs.google.com/spreadsheets/d/122netE-cbusPi00GzXqfSvKPjFDxW\_9LKeVfPqWyv04/edit?gid=1765756421&usp=embed\_facebook.

United Nations. 1948. ‘Universal Declaration of Human Rights’. https://www.un.org/sites/un2.un.org/files/2021/03/udhr.pdf.

———. n.d. ‘Democracy’. United Nations. Accessed 23 November 2024. https://www.un.org/en/global-issues/democracy.

United Nations Statistics Division. 2024. ‘UNdata | Record View | Per Capita GDP at Current Prices - US Dollars’. https://data.un.org/Data.aspx?q=korea+gdp&d=SNAAMA&f=grID:101;currID:USD;pcFlag:1;crID:408,410.

US Department of Justice. 2019. ‘Report on the Investigation into Russian Interference in the 2016 Presidential Election’. *Washington D.C.* I.

Vidyarthi, Apratim, and Rachel Hulvey. 2021. ‘Building Digital Walls and Making Speech and Internet Freedom (or Chinese Technology) Pay for It: An Assessment of the US Government’s Attempts to Ban TikTok, WeChat, and Other Chinese Technology’. *Indian Journal of Law and Technology* 17 (1): 1–42.

World Bank. 2024a. ‘GDP per Capita (Current US$)’. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2023&start=2005.

———. 2024b. ‘Population, Total’. https://data.worldbank.org/indicator/SP.POP.TOTL?end=2023&start=2005.

York, Jillian C. 2022. ‘Understanding Internet Censorship in Democracies’. *Perry World House*.

1. For this paper, when terms such as “autocratic”, “undemocratic”, or “non-democratic” are used, they refer to either closed or electoral autocracies, following the definitions of V-Dem’s index of democracy. And whenever the term “democratic” is used, it refers to either electoral or liberal democracies, also in accordance with V-Dem’s definitions (Coppedge et al. 2024a), which can be found in Section 3.1: Data. [↑](#footnote-ref-2)
2. For this paper, terms such as “block” or “ban” will be used interchangeably and understood as a full/complete prohibition (be it *de jure* or *de facto*) of a social media inside the borders of a country, unless specified otherwise (for instance, censorship and partial or limited blocks). [↑](#footnote-ref-3)
3. According to Freedom House, “66% [of the 5 billion people with internet access] live in countries where authorities deployed progovernment commentators to manipulate online discussions.” (Funk et al. 2024) [↑](#footnote-ref-4)
4. Whenever X or Twitter are mentioned, they refer to the same social media platform. [↑](#footnote-ref-5)
5. See Table 1 for the visualization of the process. [↑](#footnote-ref-6)
6. A case that contradicts the argument that national security is not raised for Western companies. [↑](#footnote-ref-7)
7. “Interpretability is the degree to which a human can understand the cause of a decision. The higher the interpretability […] the easier it is to comprehend the model’s predictions” (Amazon Web Services 2021, 23). Also, see image in page 24 of Amazon’s whitepaper, comparing different models’ interpretabilities. [↑](#footnote-ref-8)
8. See Table 1 for the visualization of this structure and refer to the illustration in Section 1: Introduction. [↑](#footnote-ref-9)
9. Here, the term “evaluation” is used for clarity purposes. In practice, the evaluation step is the same as the test step (since it is the test set of another model), serving as a test on different data: after fitting the model on the training set, it predicts the result for the test set and then for the evaluation set. The test set is used to calculate the performance of the model on unseen data and test for overfitting, while the evaluation set serves to prove or disprove the research question. [↑](#footnote-ref-10)
10. See Appendix A.1. for an image of the manipulated and original IST databases. [↑](#footnote-ref-11)
11. Only 193 countries will be considered, to enable V-Dem compatibility. [↑](#footnote-ref-12)
12. Population and GDP *per capita*. [↑](#footnote-ref-13)
13. GDP *per capita*. [↑](#footnote-ref-14)
14. GDP *per capita* for North Korea. [↑](#footnote-ref-15)
15. See Appendix A.2. [↑](#footnote-ref-16)
16. The regions are separated as follows “1: Eastern Europe and post-Soviet Union (including Central Asia, Mongolia, and German Democratic Republic); 2: Latin America (including Cuba, Haiti, and the Dominican Republic); 3: North Africa and the Middle East (including Israel and Türkiye, but excluding Cyprus); 4: Sub–Saharan Africa; 5: Western Europe and North America (including Australia, New Zealand, and Cyprus, but excluding German Democratic Republic); 6: Eastern Asia (including Japan, excluding Mongolia); 7: South–Eastern; 8: Southern Asia; 9: The Pacific (excluding Australia and New Zealand); 10: The Caribbean (including Guyana and Suriname, but excluding Cuba, Haiti, and the Dominican Republic)” (Coppedge et al. 2024a). [↑](#footnote-ref-17)
17. 39 including two variables for ISO codes for country names (alpha-2 and alpha-3), that served as keys in the process of combining the different data frames, but not including the two posterior random variables. [↑](#footnote-ref-18)
18. See Appendix A.3. for the distributions of the variables and outliers. [↑](#footnote-ref-19)
19. Otherwise written as , where TP are True Positives, TN are True Negatives, FP are False Positive, and FN are False Negatives. [↑](#footnote-ref-20)
20. Otherwise written as , since precision is and recall is [↑](#footnote-ref-21)
21. Hyperparameters such as random state, solver, and maximum number of iterations had fixed values of 1234, saga, and 1.000, respectively. [↑](#footnote-ref-22)
22. See Appendix A.4. for the confusion matrices for all models. [↑](#footnote-ref-23)
23. See Appendix A.5. for the decision tree visualizations. [↑](#footnote-ref-24)
24. The only constant hyperparameter is the random state, fixed at 1234. [↑](#footnote-ref-25)
25. Again, the only constant hyperparameter is the random state, fixed at 1234. [↑](#footnote-ref-26)
26. See Appendix A.6. [↑](#footnote-ref-27)
27. The decision trees are too big to fit legibly inside a Microsoft Word document, therefore, I have uploaded them digitally, available at: https://github.com/LeonardoStempfle/Trabalho-de-Conclusao-de-Curso-TCC [↑](#footnote-ref-28)